SHORT REPORT

Brightness scaling of periodic tones

Andre Almeida¹ · Emery Schubert¹ · John Smith¹ · Joe Wolfe¹

Published online: 27 July 2017
© The Psychonomic Society, Inc. 2017

Abstract Brightness is an attribute often used by musicians when describing timbral characteristics. It is related to the spectral distribution of energy, as is sharpness, studied by Zwicker (*Psychoacoustics: Facts and Models*, 1990). In the current work, subjects adjusted the spectral slope and thus the spectral centroid (SC) of one of a pair of sounds to make it twice as bright as the other, so as to build a perceptual scale. The ratio of SC required to double brightness is a little less than 2 and decreases as the SC of the tones increases. For these tones, the ratio of brightness is statistically different from the ratio of sharpness calculated from published models.

Keywords Brightness · Sharpness · Timbre · Stationary tones · Spectral centroid · Psychoacoustics

Timbre is a broad term that encompasses all the possible perceived differences between two tones playing at the same pitch, duration, and loudness. It is multidimensional (Grey, 1977) and includes spectral and temporal features. For a stationary sound, an important feature of timbre is the distribution of energy across the frequency spectrum. One descriptor of timbre used by musicians is brightness, which increases when more energy is distributed to higher frequencies. Some studies quantify a psychophysical variable called sharpness,

which is also related to the spectral distribution on energy (Zacharakis, Pastiadis, & Reiss, 2014; von Bismarck, 1974a). Von Bismarck (1974b) has also proposed a numerical scale for sharpness, which has later been assigned a unit: the acum (Zwicker & Fastl, 1990), and a model to estimate the perceived sharpness for a given spectrum. Ilkowska and Miśkiewicz (2006) found that sharpness and brightness are usually equivalent, with slight differences when the sound has high-frequency content (around 6 kHz), but to the authors' knowledge no psychoacoustic scale exists for brightness. For subjects with knowledge of music, the nomenclature *sharpness* to describe timbre is confusing because *sharper* and *flatter* are used to compare and to describe pitch.

Perceived brightness is strongly correlated with the spectral centroid (SC; Schubert & Wolfe, 2006). SC describes the distribution of power at different frequencies in a way that is analogous to the definition of the centre of mass, a point around which the mass of an object is distributed. The value of SC is defined by an analogous equation:

$$SC = \frac{\int_0^{\infty} fP(f)df}{\int_0^{\infty} P(f)df} \quad (1)$$

where f is frequency and $P(f)$ the power spectrum distribution of the pressure signal.

The present study aims to relate perceived brightness to SC and to other established timbral descriptors in the context of musical notes by varying the spectral slope of test tones while keeping all spectral content below 3 kHz so as to maintain all the spectral content in a region where the ear's response is somewhat flat. It also investigates whether the perception of brightness is similar to existing psychoacoustic models of sharpness. This study was included in a larger study

Electronic supplementary material The online version of this article (doi:10.3758/s13414-017-1394-6) contains supplementary material, which is available to authorized users.

✉ Andre Almeida
a.almeida@unsw.edu.au

¹ The University of New South Wales, Sydney, NSW 2052, Australia

administered to students in a university course, with a significant proportion of the subjects being musicians, presumably a subset of the population that has more extensive ear training in some aspects. This sample allowed us to compare groups with different extents of musical training.

Relevant investigations in the past have studied the perceptual scaling of sharpness and compared it with brightness in some cases, in particular in the work of Ilkowska and Miśkiewicz (2006). Zwicker and Fastl (1990) and Aures (1985) proposed sharpness models based on von Bismarck's findings by using newer auditory data for the calculation of specific loudness, which forms the basis of loudness and sharpness models (among other psychoacoustic features).

Procedure and methods

An experiment was performed using a computer graphical interface (see Fig. 1), which presented a reference tone and allowed users to adjust a second, target tone using a slider control until they judged it to be 'twice as bright'. The tones were presented diotically with closed circumaural ear headsets (Sennheiser HD280 Pro). The sound samples were presented at levels of 46 to 49 dBA. The sound level inside the headphone with no signal from the computer, due to room ventilation, was 39 dBA, according to a Bruel and Kjaer model 2250-S sound level meter. Initial tests at a higher level of 55 dB were reported to be uncomfortable by subjects for the 30-

minute duration of the experiments. Separate tests for harmonic distortion in the computer-headphone reproduction system were made in a quiet room using a pure sine wave, showing that any harmonic components due to distortion were at least 40 dB below the fundamental. Consequently, any differences in spectral centroid due to distortion would be less than 14% for a pure tone at 500 Hz and less than 1.7% for the tone with SC = 1160 Hz.

Stimuli

All the sound samples were complex periodic tones of 600 ms duration generated from the six lowest harmonics of a tone with fundamental frequency of $f_0=500$ Hz. All sounds were normalised so that they had the same RMS amplitude. They had linear starting and finishing transients with durations of 50 and 20 ms, respectively.

The independent variable was spectral slope (hereafter referred to as slope). This was the (constant) difference in level between successive harmonic amplitudes, measured in decibels (see Fig. 2). In principle, this variable can change the spectral centroid (SC) over a range from $f_0=500$ Hz (fundamental), when no higher harmonics are present, to 3000 Hz (only the sixth harmonic present). In practice, values of SC close to 3000 Hz were not used because they are heard as having a different pitch due to the dominance of a single harmonic (de Cheveigne, 2005).

Procedures

Four reference tones with different spectral centroid (SC) values were presented to each subject. The adjustment of the target tone was done with a GUI slider that could be adjusted

UNSW Tone adjustment study

Fig. 1 The interface used by participants to adjust a target tone (Tone 2) to twice the perceived brightness of the reference (Tone 1). A confidence rating (lower left corner) was mandatory whereas a comment was optional. The interface ran on a web browser

Fig. 2 Examples of the amplitude of harmonics in a complex tone as a function of the slope parameter. RMS amplitude is constant across each slope parameter signal. Tones corresponding to values of spectral centroid can be consulted in the [Supplementary Materials](#)

in 100 equal steps that produced equal changes in SC at constant total RMS power. (Numerical values were not displayed to the subject.) The slider position is proportional to SC so that the leftmost corresponds to SC = 500 Hz (all energy in the first harmonic) and the rightmost to SC = 2750 Hz. The four reference tones used had SC = 500, 720, 940, and 1160 Hz. The highest of the reference values ensured that there was enough slider range for subjects to adjust the last value to more than twice the spectral centroid SC. (The distribution of responses was later examined for any evidence of bias due to the limited range.) The four reference tones were presented to each participant in a different random order.

For each reference sample, the users were allowed to adjust the target tone and to play the sounds as many times as desired and in the order they wished. They were asked to rate their confidence in their answer before continuing to the next trial. The confidence rating was converted to a numerical scale, from 0 (*Guessing*) to 4 (*Certain*)—see Fig. 1. Additionally, participants could leave comments about the trial if they wished. At the end of the tasks, demographic data were collected, including self-rated musical background (Ollen, 2006).

Participants

Participants were 176 students attending the music psychology course at a university in Australia, comprising 49 music students, 84 others with music experience, and 43 without music experience. One hundred and seven were female and 69 male, and the age range was from 18 to 23 years for 169 of the participants; seven were older than 23. They completed the exercise in return for course credit. The subject of timbre or brightness had not been addressed in their curriculum at the time of the experiment.

Results

Analysis of spectral centroid and comparison with sharpness models for ‘twice as bright’ adjustment

The adjusted values of SC for the targets corresponding to each of the reference sounds are grouped together. Sharpness values were calculated for the reference tone and for the final selected target tone for each response by each participant, according to Zwicker’s original model (ZS) and a modified model by Aures (AS), taking into account the loudness of the sound. These values were obtained using PSYSOUND 3 (Cabrera, 1999) and are plotted in Fig. 3 for comparison with the ratios of SC. If brightness and sharpness were the same quantity, then all values of ZS and AS would be expected to be close to two (twice as sharp equals twice as bright). These are plotted in Fig. 3.

Fig. 3 Mean (thick line) and distribution (box plots) of the ratio of spectral centroids of two tones, one being adjusted to be twice as bright as its reference tone. The dotted and dashed lines show the ratios of sharpness of the mean adjusted tone to the reference tone, calculated from the spectra using the Zwicker and Aures algorithms

The values of ZS and AS are not normally distributed (Kolmogorov-Smirnov test of normality with Lilliefors significance correction returned $p < .05$, not shown graphically.) Consequently, nonparametric testing was applied. For each set of ZS and AS corresponding to a single reference tone, a one-sample Wilcoxon signed rank test (Conover, 1971) was conducted using a null hypothesised median value of 2 (double brightness). All of the eight tests (4 trials $\times$ 2 models) were statistically significant (all p values less than .001), meaning that neither model was a good predictor of the participant’s judgments of twice as bright, regardless of the value of spectral centroid of the reference tone SC.

The distributions of SC chosen as twice as bright are represented in the box plot in Fig. 3. Whiskers extend to 1.5 times the interquartile difference. The solid line represents the mean value of each distribution fitted to the cumulative distribution of measured values. The distributions are slightly skewed towards low SC values but closer to normal for higher SC reference tones.

The prediction of the ratios of sharpness based on ZS and AS for the sample with an SC value corresponding to the mean of the SC distribution are shown in Fig. 3 as dashed and dot-dashed lines, respectively.

Figure 3 shows that, on average, a doubling of perceived brightness requires somewhat less than a doubling of the spectral centroid, a difference that is statistically significant (p values $< .001$) for all but the first reference centroid ($t = -1.1, -4, -12$, and -20 , successively). Moreover, the spectral centroid ratio needed to double perceived brightness decreases monotonically as spectral centroid increases. The equation $SC \text{ ratio} = 2.2 - \frac{SC}{2000 \text{ Hz}}$ approximates the double brightness judgment in the measured range, 500–1160 Hz (but should not be extrapolated to tones with a higher number of harmonics since this has not been

tested, and an SC ratio < 1 would be meaningless in this context).

For these samples, sharpness and brightness are not the same quantity, as the predicted values of sharpness ratio differs from that of the brightness distribution. Statistical tests also show that the ratios of the sharpness prediction are also significantly different from 2 (all p values $< .001$, according to one-sample t tests returning $|t| > 10$ in all but one case where $t = 3$ and Wilcoxon tests had all p values < 0.001). Figure 3 shows that SC ratio approximates brightness ratio rather better than the modelled sharpness ratios.

Reliability of answers

Figure 3 shows a larger spread of results for the darkest sound (SC = 500). This is possibly due to a feature of the interface: for a lower reference SC, the perceived brightness seems to vary more quickly for a small slider displacement. The error in adjustment could therefore be partly due to a difficulty in adjusting the slider precisely with the mouse. For the three highest SC tones, this does not seem important. For each of the different reference tones, at least half the subjects chose SC ratios within 20% of the mean.

Differences due to musical experience

The responses were also analysed in five subgroups according to the musical experience of the subjects. These data are shown in Fig. 4, which shows, for each of the subgroups of subjects, the subgroup's ratio of SC divided by the ratio of SC for the entire population of subjects.

Despite a small visible trend showing professional musicians' selection of a proportionally higher SC ratio than non-musicians, there was no significant difference in the ratio of SC to mean SC ratio as a function of the self-declared musical expertise according to an ANOVA, $F(4, 172) = 1.463$, $p = .216$, partial eta squared = .033.

Discussion

In a complex tone composed of a small number of harmonics, brightness is judged to double when spectral centroid—SC—is increased by a ratio of 1.6 to 2.0. The ratios are highest when the reference tone has a low SC and decrease monotonically but slowly with increasing SC. Judgement of 'twice as bright' is reliable to an uncertainty of about 20% in the ratio of SC. The experimental data show that the ratios of sharpness predicted by the Zwicker and Aures algorithms are significantly smaller than the ratio of brightness measured here (except for the 500 Hz SC reference tone), even though Ilkowska's study suggests that the two are well correlated. There may be several reasons for this:

- Sharpness and brightness are different percepts.
- Ilkowska's study was performed using tones with complex spectra, whereas the tones used here have very well-defined frequency peaks. There may be more significant differences between the two percepts when the spectra consist of well-defined tones.
- Sharpness models have been developed for random signals with wide-band spectra. They may not be applicable for tonal sounds.

Fig. 4 Box plots showing distribution of the ratios of spectral centroids SC by music experience groups (numbers of participants in parentheses): NO-musicians (43), AMateur (57), ReGular player (21), music STudent (49), PRofessional (6)

The empirical investigation of the psychophysics of timbre perception has only sporadically received attention, despite its significance to musicians and audiences (Donnadieu, 2007). The current data suggest, however, that there is still much work to be done to obtain reliable models of brightness and sharpness perception, applicable over a wider range of sound types and frequencies.

Acknowledgements The support of the Australian Research Council is gratefully acknowledged.

References

- Aures, W. (1985). Berechnungsverfahren für den sensorischen Wohlklang beliebiger Schallsignale. *Acta Acustica united with Acustica*, 59(2), 130–141.
- Cabrera, D. (1999). PSYSOUND: A computer program for psychoacoustical analysis. *Proceedings of the Australian Acoustical Society Conference*, 24, 47–54.
- de Cheveigne, A. (2005). Pitch perception models. In C. Plack & A. Oxenham (Eds.), *Pitch* (pp. 169–233). New York: Springer Verlag.
- Conover, W. J. (1971). *Practical nonparametric statistics*. New York: Wiley.
- Donnadieu, S. (2007). Mental representation of the timbre of complex sounds. In J. Beauchamp (Ed.), *Analysis, synthesis, and perception of musical sounds: The sound of music* (pp. 272–319). New York: Springer.
- Grey, J. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, 61(5), 1270–1277.
- Ilkowska, M., & Miśkiewicz, A. (2006). Sharpness versus brightness: A comparison of magnitude estimates. *Acta Acustica United with Acustica*, 92(5), 812–819.
- Ollen, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings* (Unpublished doctoral dissertation, The Ohio State University).
- Schubert, E., & Wolfe, J. (2006). Does timbral brightness scale with frequency and spectral centroid? *Acta Acustica United with Acustica*, 92(5), 820–825.
- von Bismarck, G. (1974a). Sharpness as an attribute of the timbre of steady sounds. *Acta Acustica United with Acustica*, 30(3), 159–172.
- von Bismarck, G. (1974b). Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acta Acustica United with Acustica*, 30(3), 146–159.
- Zacharakis, A., Pastiadis, K., & Reiss, J. D. (2014). An interlanguage study of musical timbre semantic dimensions and their acoustic correlates. *Music Perception: An Interdisciplinary Journal*, 31(4), 339–358.
- Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and models*. New York: Springer-Verlag.