ICSV13 - Vienna
The Thirteenth International Congress
on Sound and Vibration
Vienna, Austria, July 2-6, 2006

Relationship between sound classification of xylophone-like bars and wood species properties

Mitsuko Aramaki[(1)], Henri Baillères[(2)], Loïc Brancheriau[(2)], Richard Kronland-Martinet[(1)] and Sølvi Ystad[(1)]

[(1)] CNRS - Laboratoire de Mécanique et d'Acoustique
31, chemin Joseph Aiguier 13402 Marseille Cedex 20, France

[(2)] CIRAD-Forêt, TA 10/16, avenue Agropolis, 34398 Montpellier Cedex 5, France

Abstract

In this paper, we propose a methodology associating an analysis-synthesis process and a perceptual classification test of impacted wooden bars to point out the relationship between sound quality and physical parameters characterizing wood species. Statistical analysis have made it possible to point out the influence of the pitch in the classification, and the importance of two timbre descriptors, i.e. the frequency dependent damping and the spectral bandwidth. These descriptors can be linked to physical and anatomical characteristics of wood species, leading to new clues in the choice of attractive wood species from a musical point of view.

INTRODUCTION

This study is the first step towards a better understanding of what makes the sound produced by an impacted wooden bar attractive for xylophone makers. Sounds from a wide variety of wood species have been recorded and their sound quality has been compared and linked to the wood properties. To this aim, an original methodology associating analysis-synthesis processes and perceptual classification analysis is used and tested with a renowned instrument maker.

The xylophone maker community seems to agree in the choice of wood species, and most commonly uses the (*Dalbergia* sp.) which fulfulls certain sound quality criteria together with properties like robustness and esthetic aspects. In spite of this specific choice of wood specie in xylophone manufacturing, other tropical wood species fulfill the conditions for high sound quality and could replace traditional species. Several authors have already addressed the problem of understanding which physical characteristics are of importance for the generated sound. In particular, [1] has concluded that an “ideal” xylophone wood type bar is characterized by a specific value range of density, Young modulus and damping factors. Studies on auditory cues of natural or synthetic sounds associated to geometry and material properties of vibrating objects, has already been presented by authors like [2, 3, 4]. These studies revealed the existence of perceptual cues allowing the source of the impact sound to be identified merely by listening. In particular, the perception of material correlated mainly with the damping factors of the spectral components [5, 6]. Nevertheless, it has not been established whether the nature of the perceptual cues highlighted in the distinction of different materials are pertinent to explain the subjective classification of different species of wood.

The perceptual differences reported in the literature are linked to subtle changes in the timbre of the sound. Actually, timbre is usually defined as “the perceptual attribute that distinguishes two tones of equal, pitch, loudness and duration” (American National Standards Institute). This points out the importance of comparing sounds with similar loudness, duration and pitch. What loudness and duration are concerned, the sounds of interest in this experiment can easily be adjusted in intensity by listening, and have approximately the same duration, since they correspond to the very narrow category of impacted wooden bars. What the pitch is concerned, it differs depending on the wood specie, since it is determined by physical characteristics of the material, such as the Young modulus and the mass density. To tune the sounds to the same pitch, we therefore propose to digitally process the sounds recorded using bars of equal length. Synthesis models can be used for that purpose, allowing a virtual tuning by altering the synthesis parameters. Such an approach, combining sound synthesis and perceptual analysis, has already been proposed by several authors [7, 4], [8, 9, 10]. Nevertheless, the models used for this purpose do not easily allow for an analysis-synthesis process, implying the generation of a synthetic sound perceptually similar to an original one. To overcome this drawback, we propose an additive synthesis model based on the physics of vibrating bars, the parameters of which can be estimated from the analysis of natural sounds.

DESIGN OF THE SOUND DATA BANK

In this experiment, we use tropical and sub-tropical wood species that are uncommon to xylophone makers to compare their physical and perceptual properties to commonly used species (such as *Dalbergia* sp. or *Pterocarpus* sp.). A set of 59 different species presenting a large variety of densities (from 206 to 1277 kg/m^3) have been chosen from the huge collection of species (about 8000) possessed by the CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement, Montpellier, France). By manufacturing wooden bars with the same geometry (as prismatic as possible with dimensions L=350 mm $\times$ W=45 mm $\times$ T=20 mm, without singularities and cut in the grain direction) and boundary conditions, the intrinsic quality of the species can be compared. We assume that the growth rings are parallel to the tangential wood direction and that their curvatures are negligible. The longitudinal direction is collinear to the longitudinal axis of the bars which were stabilized in room conditioning.

An experimental setup combining an easy way of generating sounds with a relative precision insuring the repeatability of the measurements was designed. For this purpose, a steel ball of 12 g and a 14 mm diameter, tightened by a 30 cm long string was used to struck the bars close to one of their extremities, allowing high frequencies to develop. An omni-directional microphone (Neumann KM183mt) was placed in the close soundfield at the opposite of the impact location to measure the sound radiated pressure.

Figure 1 shows the temporal signal, the spectral representation and the time-frequency representation of a typical sound obtained experimentally.

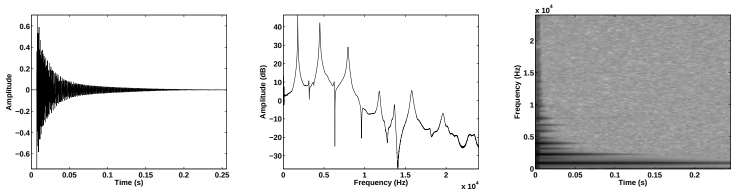

Figure 1: (a) Temporal signal, (b) spectral representation and (c) spectrogram (amplitude in logarithmic scale) of a typical sound obtained by impacting a wooden bar.

The temporal signals are characterized by a short onset and a strong decay. Consequently, their durations generally do not exceed 1 s. Their spectra are composed of some emergent resonances which do not much overlap. As shown by the time-frequency representation, the damping of these spectral components is frequency dependent, the high frequency components being more heavily damped than the low frequency components.

TUNING THE SOUND DATA BANK

As mentioned in the introduction, the pitches of the impact sounds produced by the different species vary due to their material specificities (e.g. Young modulus and mass density). To compare timbre properties produced by different wood species, the pitches have to be equalized. Since the geometry of the bars should not be modified in this study, a sound synthesis model is used to virtually tune the sounds. The use of a synthesis model allows accurate sound transformations according to the physical phenomena, which could not have been possible with other signal processing approaches such as pitch shifting.

Synthesis model An additive synthesis model is used to tune the sounds. The model takes into account the main characteristics of the vibrations produced by an impacted bar to exhibit the most important properties of the radiated sound. The model is based on the well-known Euler-Bernoulli equation:

$$EI\frac{\partial^4 y(x,t)}{\partial x^4} + \rho S\frac{\partial^2 y(x,t)}{\partial t^2} = 0 \tag{1}$$

where E is the longitudinal Young modulus, I the quadratic moment, ρ the mass density and S the cross section area. It can be shown that the general solution of this equation can be written as a sum of elementary components consisting of exponentially damped monochromatic signals:

$$y(x,t) = \sum_n Y_n(x) e^{i\omega_n t} e^{-\alpha_n t} \quad \text{with} \quad \begin{cases} \omega_n \approx \sqrt{\frac{E_d I}{\rho S}}(2n+1)^2 \frac{\pi^2}{4L^2} \\ \alpha_n \approx \frac{\delta}{2}\omega_n = \frac{\omega_n^2}{2}\tau \end{cases} \tag{2}$$

Hence, the frequency of these elementary components is inversely proportional to the square of the length of the bar and their damping is proportional to the square of the frequency. Hence, the signal measured at a fixed location is considered to be well represented by the expression (2).

Estimation of synthesis parameters Before the tuning process, the recorded sounds are analyzed and then resynthesized with the synthesis model described above. For that purpose, the synthesis parameters are directly estimated from the analysis of the recorded sounds. The estimation of the parameters defining the sound is obtained by fitting the recorded signal with the expression given in relation (2). For that purpose, we used a signal processing approach which consists in identifying the parameters of a linear filter by an ARMA (Auto Regressive and Moving Average) analysis using classical techniques such as Steiglitz-McBride [11].

In addition to the synthesis model described above, the attack time has been preserved by multiplying the synthetic signal by an adequate fading function. Synthesis sounds have been evaluated by informal listening tests confirming that their original sound qualities were conserved. The synthesis quality has been further confirmed by results obtained from the professional instrument maker showing similar classification of original and synthetic sounds.

Tuning process What the tuning process is concerned, some assumptions have been made in accordance with the vibratory behavior of the bar. Firstly, the pitch of the sound produced by an impacted wood bar is related to the frequency of the first vibration mode, which is correlated to the length of the bar (cf. Equ. (2)). Actually, if the length L changes to βL, then ω_n changes to ω_n/β^2. As a consequence, a change in pitch corresponds to a dilatation of the frequency components. Secondly, according to the mechanical model of a wooden bar, the damping is proportional to the square of the frequency. Thus, from the expression (2), a damping law can be defined by a parabolic function. As a consequence, when the pitch is changed, the damping coefficient of each tuned frequency component has to be evaluated according to the damping law measured on the original sound.

These assumptions made a virtual equalization of the pitches of the recorded bank of sounds possible. To minimize the pitch deviation, the whole set of sounds has been tuned by transposing the fundamental frequencies to 1002 Hz, which is the mean fundamental frequency of all the sounds. Once again, no precise listening test has been performed, but synthesis sounds have been described by our colleagues as conserving the specificity of the material. The whole sound data bank is available at http://www.lma.cnrs-mrs.fr/∼kronland/JASA_Xylophone/sounds.html.

EXPERIMENTAL PROTOCOL

Sounds from different wooden bars were evaluated by one instrument maker specialized in xylophone manufacture with a free classification test. The instrument maker is placed in front of a computer screen on which the sounds (all represented as identical crosses) are randomly distributed. The task consists in placing the sound as a function of its musical quality according to a (virtual) horizontal axis from the left of the screen (best quality) to the right (worst quality). He can listen to the sounds as often as he wants by simply clicking on the cross. The tests have been effectuated on a laptop Macintosh with a Sony MDR CD550 headset. The instrument maker evaluated the original sounds (recorded sounds with different pitches) and the set of tuned sounds (with the same pitch). The obtained classifications are respectively called C2 and C3. The xylophone maker did not know that the sounds were synthesized, and was told that the same pieces of wood had been sculpted in order to tune the sounds to the same fundamental frequency.

RESULTS

Statistical analysis of the classifications made by the instrument maker has highlighted the importance of salient descriptors. Two descriptors were of great significance: the damping of the first partial and the spectral bandwidth of the sound, indicating the prospection of highly resonant and crystal clear sounds. Moreover, the comparison of the classification made on both the original and processed sounds has shown how the pitch influences the instrument maker. This result is in accordance with the preponderance of the damping and reinforces the importance of the pitch manipulation to better dissociate the aftereffect of the wood species from the one of the bar geometry. Finally, the results obtained allowed to better understand some of the manufacturers' strategies and to point out important mechanical and anatomical characteristics of wood species used in xylophone manufacturing.

We briefly summarize the main results of this study and refer the reader to the article [12] for more information. We have chosen to focus the scope of this current paper on the relationship between descriptors and classification of tuned sounds and on the influence of pitch on the classification.

Relationship between the descriptors and the acoustic classification of tuned sounds

Descriptors To characterize the sounds from an acoustical point of view, we calculated several timbre descriptors which have been proposed in the literature [13, 14, 15, 16]: attack time TA (the time it takes for the signal to deploy its energy from 10% to 90% of the maximum), spectral bandwidth SB (spectrum spread), spectral centroid CSG (brightness) and spectral flux SF (spectral envelope with respect to time).

In addition to these well-known timbre descriptors, we have also considered various acoustical parameters which have been chosen as function of the specificities of the impact sounds: the amplitude ratio between the first two frequency components of the sound, noted $A_{2/1}$, the damping coefficients α_1 and α_2 of components 1 and 2 and the parameters linked to the inharmonicity. As the wooden species used for this study have been intensively examined at CIRAD (their anatomical and physical characteristics are well-known), we have also considered some mechanical descriptors such as the mass density MV, the longitudinal modulus of elasticity EL, the transverse shear modulus GLX, the specific longitudinal modulus ELMV and the specific shear modulus GLXMV. All these parameters are more precisely described in [12].

Quantitative analysis The classification made by the xylophone maker is correlated to several descriptors. Among these, three descriptors related to the time evolution of the sound (α_1, α_2 and SF), and two descriptors related to the spectral content of the sound (CSG and SB) play an important role. One may notice that the physical descriptors linked to the wood properties don't explain by themselves the classification of the instrument maker, even though ELMV seems to be the most pertinent one.

The multivariate linear regression analysis used to give the high level of correlation between the descriptors highlighted two main descriptors: α_1 and SB. These descriptors are non-correlated and give rise to a linear predictor $\hat{C}3_{Linear}$ of the classification which explain 77% of the variance:

$$\hat{C}3_{Linear} = -3.82 \times 10^{-2}\alpha_1 - 1.32 \times 10^{-4}SB + 1.752 \quad (3)$$

This model is of great importance what the choice of specie is concerned. Actually, it emphasizes the fact that the xylophone maker looks for a highly resonant sound (the coefficient of α_1 is negative) containing a few spectral components (the coefficient of SB is also negative). Such a seek for a crystal-clear sound could explain the general choice of *Dalbergia* sp. which is the most resonant specie and the most commonly used in the design of xylophone bars. Actually, xylophone makers use a specific way of carving the bar [17] by removing substance in the middle. This operation tends to minimize the importance of partial 2, decreasing both the CSG and the SB. The importance of α_1 in the model is inline with several studies showing that the damping is a pertinent cue in the perception of impacted materials [6, 5]. What the parameter SB is concerned, it has been shown that the spectral distribution of energy is also an important cue, especially for categorization purposes. Previous studies [18] showed that the inharmonicity plays an important role for such tasks, but in our case, where sounds belong to the same category of material, such a descriptor doesn't seem to be of much interest. The linear classification prediction has been improved by taking into account non linear phenomena. Thus, we proposed a non-linear model which explains 82% of the variance [12].

One may finally notice that no correlation has been evidenced between the classification and the wood density, although it is well known that the wood density is of great importance for the instrument maker. This lack is due to the way we designed our experimental protocol, focusing on the sound itself and minimizing multi-sensorial effects. Actually, in a situation where the instrument maker has access to the wood, bars with weak density are rejected for manufacturing and robustness purposes, no matter their sound quality.

Influence of the fundamental frequency (pitch) upon the classification

As discussed previously, the timbre is a key feature for appreciation of sound quality and makes it possible to distinguish tones with equal pitch, loudness and duration (ANSI, 1973). Since this study aims at better understanding which timbre descriptor is of interest for wood classification, one expected differences in the classification of the tuned and the original sound data banks. The difference between the classifications C2 (various pitches) and C3 (same pitches), represented in Figure 2 as a function of the original fundamental frequency of the bars, shows a clear linear tendency.

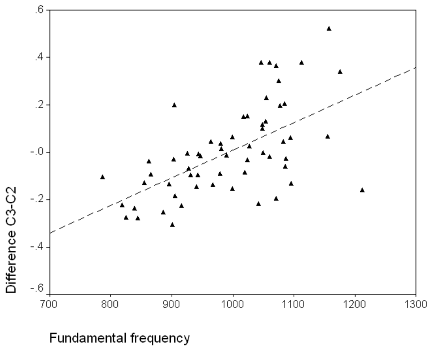

Figure 2: Linear relationship between fundamental frequency and arithmetic difference C3-C2 (R = 0.59, N = 59).

The difference is negative for frequencies below the mean frequency and positive in average for high values of the frequencies. The Pearson coefficient associated to the linear relationship between the arithmetic difference of the classification and the fundamental frequency leads to an important observation stating that *a wooden bar with a low fundamental frequency tends to be up graded while a wooden bar with a high fundamental frequency tends to be down graded*. This is in accordance with our linear prediction model that expects weakly damped sounds to be better classified than highly damped ones. Actually, sounds with low (respectively high) fundamental frequency have been transposed toward high (respectively low) frequencies during the tuning process, implying α_1 to increase (respectively decrease). As an important conclusion, one may say that the instrument maker cannot judge the wood itself independently of the bar dimensions, since the classification is influenced by the pitch changes, favoring wood samples generating low fundamental frequency sounds. Once again, it is worth noticing the good reliability of our instrument maker who kept unchanged the classification for sounds which fundamental frequency was close to the mean fundamental frequency of the data bank (i.e. sounds with nearly unchanged pitch). Actually, the linear regression line passes close to 0 at the mean frequency 1002 Hz. Moreover, the *Dalbergia* sp. has been kept at the first position after the tuning process, suggesting that no dramatic change in the value of α_1 had been made. Actually, this sample was transposed upwards by 58 Hz, changing α_1 from $13.6\mathrm{s}^{-1}$ to $14.28\mathrm{s}^{-1}$ which still was the smallest value of the tuned data bank.

References

[1] D. Holz, *Acoustically important properties of xylophon-bar materials: Can tropical woods be replaced by european species?*, Acustica/acta acustica, 82(6), (1996), 878–884.

[2] R. Lutfi, E. Oh, *Auditory discrimination of material changes in a struck-clamped bar*, Journal of the Acoustical Society of America, 102(6), (1997), 3647–3656.

[3] S. McAdams, A. Chaigne, V. Roussarie, *The psychomechanics of simulated sound sources: material properties of impacted bars*, Journal of the Acoustical Society of America, 115(3), (2004), 1306–1320.

[4] F. Avanzini, D. Rocchesso, *Controlling material properties in physical models of sounding objects*, La Habana, Cuba (2001), pp. 91–94.

[5] R. P. Wildes, W. A. Richards, *Recovering material properties from sound*, MIT Press, Cambridge (1988) pp. 356–363.

[6] R. L. Klatzky, D. K. Pai, E. P. Krotkov, *Perception of material from contact sounds*, Presence: Teleoperators and Virtual Environments, 9(4), (2000), 399–410.

[7] J. M. Adrien, *The missing link: modal synthesis*, MIT Press (1991) pp. 269–297.

[8] A. Chaigne, V. Doutaut, *Numerical simulations of xylophones. i. time-domain modeling of the vibrating bars*, Journal of the Acoustical Society of America, 101(1), (1997), 539–557.

[9] V. Doutaut, D. Matignon, A. Chaigne, *Numerical simulations of xylophones. ii. time-domain modeling of the resonator and of the radiated sound pressure*, Journal of the Acoustical Society of America, 104(3), (1998), 1633–1647.

[10] I. Bork, *Practical tuning of xylophone bars and resonators*, Applied Acoustics, 46, (1995), 103–127.

[11] K. Steiglitz, L. E. McBride, *A technique for the identification of linear systems*, IEEE Trans. Automatic Control, AC-10, (1965), 461–464.

[12] M. Aramaki, H. Baillères, L. Brancheriau, R. Kronland-Martinet, S. Ystad, *What makes the sound produced by an impacted wooden bar attractive for a xylophone maker ?*, submitted to Journal of the Acoustical Society of America.

[13] A. Caclin, S. McAdams, B. K. Smith, S. Winsberg, *Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones*, Journal of the Acoustical Society of America, 118(1), (2005), 471–482.

[14] S. McAdams, S. Winsberg, S. Donnadieu, G. D. Soete, J. Krimphoff, *Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes*, Psychological Research, 58, (1995), 177–192.

[15] J. W. Beauchamps, *Synthesis by spectral amplitude and "brightness" matching of analyzed musical instrument tones*, Journal of the Audio Engineering Society, 30(6), (1982), 396–406.

[16] J. Marozeau, A. de Cheveigné, S. McAdams, S. Winsberg, *The dependency of timbre on fundamental frequency*, Journal of the Acoustical Society of America, 114, (2003), 2946–2957.

[17] N. H. Fletcher, T. D. Rossing, *The physics of musical instruments*, Springer-Verlag (1991).

[18] M. Aramaki, S. Ystad, R. Kronland-Martinet, M. Besson, *Perception and categorization of sounds from impacted materials: Acoustical and cognitive neuroscience approaches*, submitted to Journal of the Acoustical Society of America.