HARDNESS RECOGNITION IN SYNTHETIC SOUNDS

Bruno L. Giordano, Karin Petrini

Department of General Psychology, University of Padova, Italy
bruno.giordano@unipd.it

ABSTRACT

Sound source recognition investigates recovery of different features of the objects, whose interaction lead to the generation of the acoustical signal. Among them material type have received particular attention, while recovering of material properties, such as hardness, have been scarcely considered. Hardness plays a significant role in the musical field too, especially for percussion instruments, where resonating objects of variable hardness are struck with mallets of variable hardness. Comparison of previous results on hardness recognition point toward the perceptual independence of the resonator and exciter properties. This issue was addressed in four experiments conducted on stimuli synthesized with a physical model, which allowed independent manipulation of the exciter and resonator properties. Free identification and forced choice tasks have been used to investigate the ability of listeners to discriminate variations in the exciter from variations in the resonator. Scaling tasks have been used to investigate the relationship between the synthesis parameters and the hardness estimates of the exciter and of the resonator. Free identification and forced choice data reveal a bias toward the interpretation of the acoustical signals in terms of features of the resonating object. Hardness scaling results reveal the perceptual dependence of exciter and resonator properties, although strong individual differences are found.

1. INTRODUCTION

A relatively recent field in auditory perception investigates recognition of the features of the sounds source [1]. The most studied class of signals is that of impact sounds [2], [3], generated by the interaction between a highly damped object, the exciter or hammer, and a vibrating object, called resonator or sounding object. Particular interest have received the study of categorization or discrimination of the material type of the resonator [4], [5], [6], [7]. Wildes and Richards [8] hypothesized that material could be uniquely specified, at the acoustical level, by means of the $tan\phi$ coefficient, which measures the degree of anelasticity of the material of the resonator, defined as

$$tan\phi = 1/\pi f t_e \tag{1}$$

where f is the frequency of vibration and t_e is the time it takes for the amplitude to decay to $1/e$ of its starting value. Different perceptual studies found material categorization to be strongly influenced by this acoustical parameter (rubber and wood chosen for higher $tan\phi$ values than glass and steel)[5], [4]. Other acoustical parameters, modulated by physical features extraneous to the material type of the resonator, were found to influence recognition of this sound source feature: the decay time of signal amplitude, as modulated by external damping of the resonator, and signal frequency, as modulated by the size of the resonator [6]. In a further experiment Giordano [6] demonstrated recognition of material type of the resonator to be independent of the material of the exciter, although the range of variation in the elastic properties of the used exciters was too low to induce a relevant timbral variation in the generated signals. Scaling of a mechanical material property, namely mallet hardness, was investigated by Freed [9]. Different cooking pans of variables size were struck with mallets of variable hardness. Participants were found to properly scale the hardness of the mallets, independently of the size of the resonator. Hardness estimates were found to be influenced by signal amplitude, spectral centroid, and amplitude decay velocity. On the basis of these results, recognition of the material properties of the resonator can be hypothesized to be independent of the features of the exciter, and viceversa. This was first assessed by investigating the ability of listeners to distinguish variations in the resonator from variations in the exciter features. Both free identification and forced choice tasks were used. Second, scaling of the hardness of both the exciter and the resonator was conducted, in synthetic stimuli generated by manipulating the properties of the resonator as well as those of the exciter. These topics are of interest to the musical field too, particularly for the field of percussion instruments, where resonators made of a large variety of materials (e.g., metallic and wooden instruments, membranophones, lithophones, glass and ceramic instruments) are struck with objects of variable hardness (e.g., felt and wooden mallets, metallic sticks, hands) [10]. Then research on hardness scaling could partially highlight the problem of the perceptual quality of percussion instruments, scarcely considered in timbre perception research.

2. SYNTHESIS PARAMETERS

Stimuli were generated with the impact model by Avanzini et al. [11]. Three parameters of the model were investigated. The first two were related to the resonator: the frequency of the lowest mode f, and $\tau = 1/\pi tan\phi$. The first parameter can be thought as modelling the geometrical properties of the resonator (e.g., length). The τ parameter models the material of the resonator. High τ values correspond to hard or stiff materials (steel, glass). Low τ values correspond to soft or elastic materials, such as rubber. The third synthesis parameter was related to the interaction between the exciter and the resonator: the force stiffness coefficient k, directly related to the elastic properties of the exciter. Low k values may represent rubber or felt mallets. High k values may represent stiff mallets (metal, hard woods).

3. FREE IDENTIFICATION

A free identification experiment was conducted in order to assess how variations in the resonator or in the exciter parameters were described or interpreted by listeners. A set of 16 stimuli was synthesized combining two values for the f parameter (50 and 800 Hz), two values for the τ parameter (30 and 200), and four values for the k parameter (10^2, 10^4, 10^6, and 10^8). Stimuli were presented in eight series of four sounds. In half of them only the resonator parameters were varied, while the k parameter remained constant, in the other half the k parameter varied and the resonator parameters were kept constant. Stimuli were presented via headphones in a silent room. The order of the series within the experiment, as well as the order of the stimuli within each series, was randomized. Nine participants were presented each series once and were asked to describe what was varying within each series. No information concerning the nature of the stimuli was given.

3.1. Results

Descriptions were categorized in three groups: generic descriptions (e.g.,"four sounds together"), descriptions based on the perceptual features of the sounds (e.g.,"high pitch", "bright"), descriptions which focused mainly on sound source features (e.g., "finger tapping on wood"). This latter was divided in two subgroups: those which described a constant resonator (e.g., "a xylophone sound with different decay velocities"), and those which described different resonators (e.g., "a drum, a bell..."). Table 1 shows the proportion of occurrence of each category.

Varied parameter	**Category** G	P	Cr	Vr
Resonator	0.166	0.027	0.111	0.694
Exciter	0.027	0.111	0.583	0.278

Table 1: *Proportion of choosing each description category, for the variable resonator and variable exciter series. G = generic, P = perceptual, CR = constant resonator, VR = variable resonator.*

Overall listeners tended to describe sounds mainly in terms of sound source features (83% of the descriptions), rather than in terms of perceptual features alone. Notably the constant object descriptions where more frequent when the k parameter was varied, while the variable object descriptions were more frequent when the resonator parameters where varied. This indicated that a variation in the exciter features did not generate a variation in the recognized resonator features, and viceversa. This conclusion can be biased by the fact that, differently from what observed for the resonator, the exciter was rarely mentioned in the descriptions. When a constant object description was given, listeners tended to describe the variation within the series either in terms of perceptual qualities (e.g., differences in the attack quality) or in terms of a change in the action exerted over the resonator (e.g., a variation in the external damping).

4. FORCED CHOICE CATEGORIZATION

A forced choice task was used to asses wether listeners were able to distinguish variations in the resonator from variations in the exciter. Stimuli were selected from a synthesis space defined by the combination of five equally log-spaced levels for f, τ, and k. Table 2 shows the levels used for each parameter.

Sixteen series of stimuli were built by varying only one of the three investigated parameters, either in ascending or descending order. In half of them $[(f_a, \tau_2, k_2)$, (f_a, τ_4, k_4), (f_2, τ_a, k_2), $(f_4, \tau_a, k_4)]$, a resonator parameter was varied, in the other half $[(f_2, \tau_2, k_a)$, (f_2, τ_4, k_a), (f_4, τ_2, k_a), $(f_4, \tau_4, k_a)]$ the k coefficient was varied (the subscript a denotes the use of all the available parameters within the series.). All series were judged once by each participant, and were presented in randomized order. Two types of instructions were given. In both cases participants were given a definition of resonator and of exciter, and an example of exciter/resonator interaction was described verbally (a baked clay dish struck with a wooden ball). Additionally a sub-group of participants was presented auditory examples of variations in the resonator and in the exciter properties, using real rather than synthetic sounds. They were asked to tell wether the presented series was generated by a variation in the resonator or in the exciter. Twelve participants were given verbal-only instructions; nine participants were presented the verbal+auditory instructions.

Parameter	**Levels** 1	2	3	4	5
f (Hz)	50	100	200	400	800
τ	10	20	40	80	160
k	$5e^6$	$3.3e^7$	$2.24e^8$	$1.495e^9$	$1e^{10}$

Table 2: *Synthesis parameters values used for the investigated stimuli.*

4.1. Results

Table 3 shows the proportion of choosing the response "variation of the resonator" for the variable exciter and variable resonator series, and for the two types of instructions.

Varied parameter	**Instructions** Verbal	Verbal+Auditory
Resonator	0.604 (0.673–0.535)	0.694 (0.770–0.619)
Exciter	0.573 (0.643–0.503)	0.403 (0.483–0.323)

Table 3: *Proportion of choosing the "variable resonator" response category for the variable resonator and exciter series, and for the verbal and verbal+auditory instructions. 95% confidence intervals are given in parentheses.*

An overall tendency of participants to attribute the change in the auditory stimuli to a variation in the resonator was observed. This tendency was particularly strong when only verbal instructions were given, where the proportion of choosing the response "variable resonator" in the variable exciter series was higher than chance level and not significantly different from the same proportion for the variable resonator series. This tendency was weakened by the use of auditory examples. These results are consistent with those collected with the free identification procedure. Listeners revealed a bias toward the interpretation of the perceptual variations in terms of features of the resonator, rather than in terms of features of the exciter. The effect of the auditory examples supports this conclusion, as it shows that listeners do not pay attention to the perceptual variations associated to a change in the features of the exciter until are instructed in this sense.

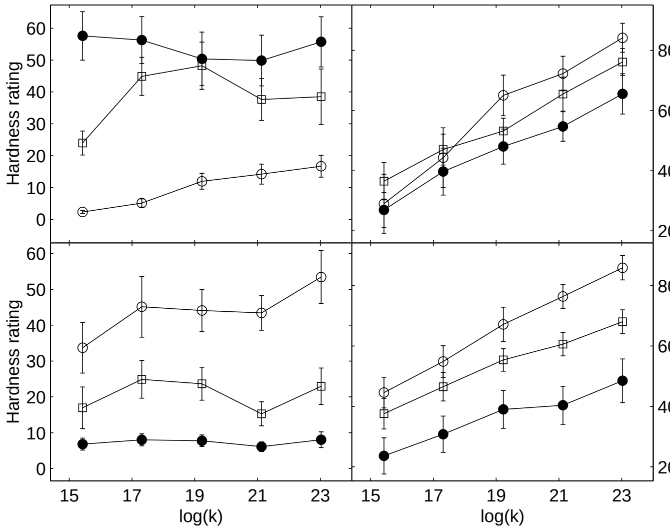

Figure 1: *Average exciter hardness ratings as a function of the k coefficient. Variable f set: top row, variable τ set: bottom row. Filled circles: f_1 (top) τ_1 (bottom), empty squares: f_3 (top) τ_3 (bottom), empty circles: f_5 (top) τ_5 (bottom). Left column: cluster 1, right column: cluster 2. Error bar =* ± 1 *SE*

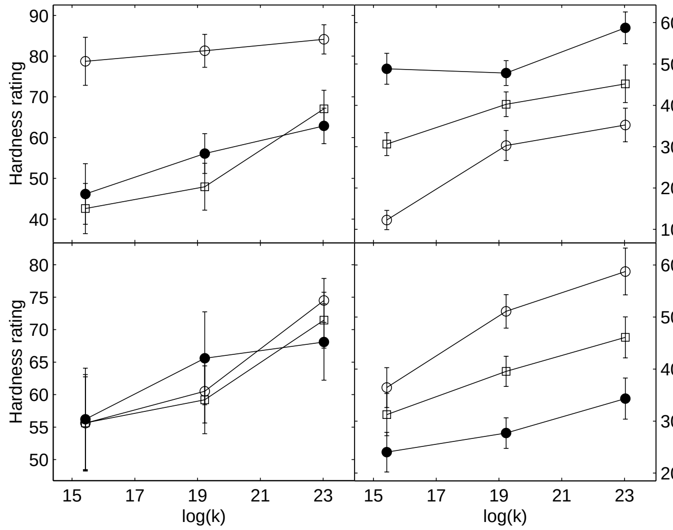

Figure 2: *Average resonator hardness ratings as a function of k. Filled circles: f_1 (top row) τ_1 (bottom row), empty squares: f_3 (top row) τ_3 (bottom row), empty circles: f_5 (top row) τ_5 (bottom row). Left column: cluster 1, right column: cluster 2. Error bar =* ± 1 *SE*

5. HARDNESS SCALING

Hardness scaling experiments were conducted in order to test the perceptual independence of the resonator and of the exciter properties. Two experiments were conducted. Stimuli were selected from the synthesis space investigated in the previous experiment. Scaling of the hardness of the resonator was investigated with a set of 27 stimuli, given by the combination of the first, third and fifth level of each parameter. Scaling of the hardness of the exciter was investigated with two sets of 15 stimuli. In both cases all the five levels of k were used. In the variable f set τ was fixed at the third level, and three f values were used (f_1, f_3, f_5). In the variable τ set f was fixed at f_3 and τ assumed the following values: τ_1, τ_3, τ_5. Scaling of the hardness of the resonator and of the exciter was performed by two different groups of 18 participants. The variable f and variable τ sets were presented in counterbalanced order across participants. Stimuli were presented in randomized order within each session. Every stimulus was judged three times by each participant.The verbal+auditory instructions of the previous experiment were given. Participants were asked to rate the hardness of the exciter/resonator on a 1 (very soft)–100 (very hard) scale.

5.1. Results

In order to evaluate the presence of different response profiles, average hardness ratings were subjected to cluster analysis. An agglomerative hierarchical clustering procedure, with average linkage between clusters as grouping criterion and a quadratic Euclidean distance measure was used. This analysis was performed separately on the variable τ, on the variable f, and on the resonator hardness data sets. Two clusters of participants were extracted from each data set, by considering only the two furthest ones, joined in the last stage of the agglomerative procedure. This analysis individuated two clusters of nine participants in the variable f data set, one cluster of seven participants and one of eleven in both the variable τ and in the resonator hardness data sets. Despite the cluster analysis was performed separately for the variable τ and f data sets, the division of participants among clusters was almost the same for both these sets. In particular all the seven participants of the first cluster in the variable τ data set were included in the first cluster of the variable f set. Not surprisingly their response profiles for both the sessions had strong similarities, as shown below.

Figures 1, and 2 show average hardness estimates as a function of the k coefficient, for the different data sets, and for all the extracted clusters. Average ratings were analyzed with a repeated measures ANOVA. A separate analysis was performed for each cluster of participants.

Hardness estimates for the first cluster in the variable f set were significantly influenced by f ($F_{2,16} = 38.67$, $p < 0.001$), by its interaction with k ($F_{8,64} = 5.70$, $p < 0.001$), but not by the one-way effect of k ($F_{4,32} = 2.28$, $p = 0.083$). Analysis of the simple effects of the k variable revealed that its effect on the hardness estimates was significant only for the intermediate f level ($F_{4,32} = 5.80$, $p = 0.001$). For this reason it can be concluded that hardness estimates were determined almost only by signal frequency, higher frequency leading to higher hardness estimates. Responses, for the second cluster of the same data set, were influenced only by k ($F_{4,32} = 55.88$, $p < 0.001$), higher k values being associated to higher hardness estimates. Signal frequency, in other words, did not influence exciter hardness estimates. In the variable τ set, hardness estimates for the first cluster were influenced by τ $F_{2,12} = 34.94$, $p < 0.001$) and by its interaction with k ($F_{8,48} = 2.19$, $p = 0.045$). As with the variable f set, the one-way effect of the k coefficient was not significant ($F_{4,24} = 2.24$, $p = 0.095$). Analysis of the simple effects of the k parameter revealed that it slightly influenced hardness estimates only for the highest τ level ($F_{4,24} = 2.91$, $p = 0.043$). For this reason it can be concluded that participants belonging to the first cluster focused mainly on the τ parameter, discarding almost completely the k parameter as with the variable f set. Within this cluster hardness estimates increased as a function of the τ coefficient. On the contrary responses in the second cluster were influenced by the one-way effect of both τ and k variables ($p < 0.001$ in both cases), and by their interaction ($F_{8,80} = 2.17$, $p = 0.038$). As in the first cluster an increase in the τ parameter was associated to an increase in the exciter hardness estimates. The same effect was found for the k parameter. Overall participants belonging to the first cluster of both data sets ignored the k coefficient when estimating the hardness of the exciter, focusing on the parameters related to the resonator, although instructed otherwise. The second cluster of participants, instead, gave results highly consistent with those collected by Freed [9] on real sounds, where hardness estimates were found independent on signal frequency. Hardness estimates of the exciter in both clusters were found influenced by the material of the resonator, as modelled by the τ coefficient.

Resonator hardness estimates for the first cluster (seven participants) were slightly influenced by f ($F_{2,12} = 4.01$, $p = 0.046$), and by k ($F_{2,12} = 4.18$, $p = 0.042$). The other effects were non-significant ($p > 0.08$ in all cases). Higher frequencies were given higher resonator hardness estimates. As found with the previous experiment, an increase in k was associated to an increase of the hardness estimates. The absence of a significant effect of the τ parameter for this cluster is counterintuitive, and contradicts the already known effects that this parameter has on material type categorization. The low significance of the effects f and k variables, makes more probable the hypothesis that this group of participants gave unreliable estimates. In the second cluster (eleven participants) responses were significantly influenced by the one-way effect of all the synthesis parameters ($p < 0.002$ in all cases), as well as by the two-way interactions between k and f, and between k and τ ($p < 0.02$ in both cases). An increase in both k and τ was associated to an increase in the resonator hardness estimates. The same response pattern was found in the second cluster for the variable τ set. The effects of f and τ on resonator hardness scaling, are consistent with results gathered on categorization of material type [4], [5], [6], and establish a link between results collected with largely differing experimental techniques (scaling and categorization).

6. CONCLUSIONS

Three different experimental techniques were used to address the perceptual independence of the resonator and exciter features. Free categorization and forced choice tasks revealed a bias toward the interpretation of the variation among stimuli in terms of a variation in the resonator. Confusion among these two objects was investigated by means of hardness scaling of both the exciter and the resonator. Across almost all clusters, independent of wether the exciter or the resonator hardness was scaled, the k and τ parameter were found to have similar effects. Both these synthesis parameters model material properties, so that it can be concluded that material properties of the exciter and of the resonator are confused each other. A large variability was associated to the effects of the f parameter, which modelled the geometrical properties of the resonator. This was associated either to an increase or decrease of the resonator hardness, and to an increase or, consistently with results by Freed [9], to no effects on the exciter hardness. Acoustical analyses will have to be performed to highlight the dependence of hardness estimates of the exciter and of the resonator on the material properties of both. Plausibly the equivalence of the effects of these sound source features on the recognized hardness is due to similar effects on relevant acoustical signal properties.

7. ACKNOWLEDGMENTS

This work was supported by the European Commission, under contract IST–2000–25287.

8. REFERENCES

[1] McAdams, S., Recognition of sound sources and events, in McAdams, S. and Bigand, E., eds., Thinking in sound: the cognitive psychology of human audition, pages 146-198, Oxford University Press, 1993.

[2] Kunkler-Peck,A.J.,Turvey, M.T., “Hearing shape”, Journal of Experimental Psychology: Human Perception and Performance, 26(1):279-294, 2000.

[3] Lakatos, S., McAdams, S., Causse, R., “The representation of auditory source characteristics: simple geometric form”, Perception & Psychophysics, 59(8):1180-1190, 1997.

[4] Klatzky, R.L., Pai, D.K., Krotkov E. P., “Perception of material from contact sounds”, Presence: Teleoperators and Virtual Environment, 9(4):399-410, August 2000.

[5] Rocchesso, D.,Ottaviani, L., Fontana, F., Avanzini, F., Size, shape, and material properties of sound models, in Rocchesso, D., and Fontana, F., eds., The sounding object, Mondo Estremo, Firenze, 2003.

[6] Giordano, B.L., Material categorization and hardness scaling in real and synthetic impact sounds, in Rocchesso, D., and Fontana, F., eds., The sounding object, Mondo Estremo, Firenze, 2003.

[7] Lutfi, R. A., Oh, E. L., “Auditory discrimination of material changes in a struck clamped bar”, J. Acoust. Soc. Am., 102(6):3647-3656, 1997.

[8] Wildes, R., Richards W., Recovering material properties from sound, In W. Richards, editor, Natural Computation, pages 356-363. MIT Press, Cambridge, MA, 1988.

[9] Freed, D.J., “Auditory correlates of perceived mallet hardness for a set of recorded percussive events”, J. Acoust. Soc. Am., 87(1):311-322, 1990.

[10] Rossing, T.D., Science of Percussion Instruments, World Scientific, New Jersey, 2000.

[11] Avanzini, F., Rath, M., Rocchesso, D., Ottaviani, L., Low level models: resonators, interactions, surface textures, in Rocchesso, D. and Fontana, F., eds., The sounding object, Mondo Estremo, Firenze, 2003, available at `http:\\www.soundobject.org`