INVESTIGATIONS WITH THE SONIC BROWSER ON TWO OF THE PERCEPTUAL AUDITORY DIMENSIONS OF SOUND OBJECTS: ELASTICITY AND FORCE

Laura Ottaviani

Dipartimento di Informatica
Università degli Studi di Verona
Strada Le Grazie, 15 - 37134 Verona, Italy
ottaviani@sci.univr.it

Mikael Fernström

Interaction Design Centre
Dept. of Computer Science and Information Systems
University of Limerick, Ireland
Mikael.Fernstrom@ul.ie

ABSTRACT

The Sonic Browser is a software tool developed especially for navigating among sounds in a 2-D space, primarily through listening. It could be used for managing large collections of sounds, but now it is turning out to be useful also for conducting psychophysical experiments, aiming at investigating perceptual dimension scaling of sounds.

We used it for analyzing the relationship between the physical parameters involved in the sound synthesis and for studying the quality of the sounds generated by the SOb models. Some experiments in this direction have been already reported [1, 2], examining real and model generated sounds of impacts and bounces of objects made with different materials.

In this paper, we introduce our further investigations, by analyzing perceptually the impacts and bounces sounds from a different perspective, focusing on other two perceptual dimensions, i.e elasticity of the event and the force applied to the dropped object. We will describe the new experiment we conducted and we will report the collected data, by analyzing the resulting perceptual evaluation spaces.

1. INTRODUCTION

The Sonic Browser [3, 4], a software tool developed especially for navigating among sounds in a 2-D space, primarily through listening, is turning out to be useful, not only for managing large collections of sounds, but also for psychophysical experiments, aiming at investigating perceptual dimension scaling of sounds.

This task is particularly important for the Sounding Object (SOb) project [1]. Its aim consists in developing physically-based sound models for generating sounds, that are controllable in real time through physical parameters. In particular, the Sound Objects can be integrated into artifacts or interfaces which are directly controllable by human gestures [5].

In this context, the Sonic Browser can be exploited in order to analyze the relationship between the physical parameters involved in the sound synthesis and to study the quality of the sounds generated by the SOb models. Some experiments in this direction have been already reported [1, 2] and the use of the Sonic Browser for conducting psychophysical experiments was helpful in collecting data, because, besides allowing the subjects to evaluate the perceptual scaling in a direct and natural way by navigating through the bi-dimensional plot by means of the mouse and moving the sound objects according to their estimations, it gives to the subjects the option of comparing two or more evaluations by listening simultaneously to all the sounds they want. In fact, one of the most distinctive features of this application is the *aura*, a circle surrounding the cursor, that can be resized and which defines the range of sounds to be played simultaneously.

Examining real and model generated sounds of impacts and bounces of objects made with different materials, previous experiments concerned the relationship between perceived height of the object drop and perceived size of dropped objects. In this paper, we introduce our further investigations, which continue the previous work, by analyzing perceptually the impacts and bounces sounds from a different perspective, focusing on the relationship between other two important perceptual dimensions: perceptual elasticity of the impact/bounce and perceptual force "throwing" the object, as well as looking at the judgments about the quality and the realism of the synthesized stimuli. With perceptual force "throwing" the object we mean both the perception of the object just dropped without any force and the perception of the object thown by applying some force.

We will present the experiment conducted and the data collected, and we will comment the results obtained, comparing them with those from the previous experiments.

2. THE EXPERIMENT

The experiment was conducted with the same method used for the previous ones [1, 2], collecting both data logging, that is the object positions in the 2-D space of the Sonic Browser, and verbal comments, through the *Thinking-Aloud Protocol* [6]. The Thinking-Aloud Protocol, which consists in asking the users to express aloud what they are thinking while performing the task required by the experiment, was very useful for the analysis of the data collected by the previous experiments. As the previous setting, all the participants' sessions were video-taped.

2.1. Participants

The participants were 10 volunteers who were studying at the University of Limerick. Four of them referred to have a musical training for about 10 years, while five referred to have less than 3 years of musical training, including two that said to have never practiced music. One subject referred to practice music for 5 years. Nobody referred to have hearing problems, and four required glasses.

[1] http://www.soundobject.org

2.2. Stimuli

Since we obtained quite good results in comparing recorded and synthetic sounds in previous experiments [1, 2], at this stage we preferred to involve only sounds synthesized with the sound models.

We decided to design them all with the PD-modules modeling impact interactions of two modal resonators [7], simplified returning only one mode, because it was the model that gave us less spread data in the aforementioned works. As we aimed at investigating the relationship between elasticity and force applied to the objects, we decided for objects of two materials with completely different elasticity properties: wood and rubber.

Moreover, in designing the stimuli set, we paid attention to change slightly no more than two parameters simultaneously, in order to be able to make some observations on the influence of the parameters' values on the scaling and estimation results. In particular, we worked on the following parameters of the model: elasticity of the contact, force damping, gravity force, strike velocity, frequency, decay time. For the meaning of each parameter and the details of the model structure, we suggest to consult [7].

The stimuli included in the sounds set were 18, consisting of 11 sounds of wood and 7 of rubber. All of them were sounds of bouncing events, excluded 2 for each material who consisted in single impact events.

2.3. Procedure

The experiment was conducted in a quite, but not acoustically isolated room of the Interaction Design Centre, at the University of Limerick. The procedure was the same as the one applied in the previous studies: The participants had to navigate in the bidimensional plot of the Sonic Browser, where they were represented by the cursor and free to decide about the aura size, to listen to the stimuli through headphones, by placing the mouse on the objects representing the sounds, and to move the objects according to the two axes of the plot, i.e. perceptual elasticity of the impact/bounce and perceptual force "throwing" the object.

As in the previous experiments, the resulting data coordinates have been normalized between 0 and 1, for being able to compare the objects' locations estimated by each user, who preferred to arrange the sounds according to their own scales, and not paying attention to the screen boundaries.

Besides collecting data logging, we also recorded each session on video-tapes, in order to keep the data coming from the Thinking-Aloud Protocol.

After the scaling task, the subjects were asked to tag the sounds which they judged unrealistic and, in the debriefing phase that concluded each session, each participant filled out a 7-point Likert scale questionnaire, from a "poor" evaluation (0) to a "excellent" evaluation (6). The questionnaire was similar to that used in the previous experiments, but contrary to those experiments, we preferred to enhance the fact that each session comprised two tasks, i.e. scaling and tagging, by asking the subjects to evaluate the difficulty of the two tasks separately.

2.4. Results and Observations

In fig. 1 (a) and fig. 1 (b), we report the representation of the individual perceptual scaling and tagging information sorted by stimuli, and grouped according to the material: wood and rubber respectively.

We can see that the sounds locations, according at least to one dimension, are slightly spread and some objects are placed in the 2-D plot uniformly in both the dimensions. In particular, the sounds *6-rubber* and *7-rubber* have only 2 outliers for the force axis, while *5-rubber* has 3 outliers for both the axis and *9-wood* has 4 and 5 outliers respectively for the elasticity and the force dimension.

Anyway, we can observe from the data collected that most of the users agreed in scaling the stimuli, at least in one dimension. In particular, a part from some outliers, the sounds *5-rubber*, *6-rubber*, *7-rubber* and *9-wood* were estimated uniformly in both the dimensions, while all the other sounds were judged uniformly in at least one dimension. We can observe that the sounds *6-rubber* and *7-rubber* are the only two stimuli of the rubber-set that consisted of a single impact event.

Looking at the tagging task, we can notice that there is one sound, *2-wood*, that was considered by all participants to be realistic, and 8 sounds were defined to be realistic by at least 7 users. It is interesting to underline that, among these 8 sounds, the sound *7-rubber* belongs to the best uniformly scaled stimuli.

The better results of the tagging task, achieved with this experiment rather than those conducted previously [1, 2], could be due to the use of a sound set designed with the same sound model and of a model parameters' setting that is more value-centered.

From the verbal protocol arose that the scaling task was difficult for some participant, because of the influence of other dimensions involved in the event, such as size or weight of the objects, loudness or height of the drop. In particular, five subjects referred to be influenced on their scaling task performances by a pitch variation in the sounds and two by the objects size.

For instance, a participant referred that "it's really hard to judge about force, because they sound like they are different sized objects. So you don't know whether the impact sound that they make is because they are thrown, or because they are bigger". Four subjects were "confused from them (some sounds) thrown from a different height". Some participants were biased by the loudness, the hardness of the dropping objects, their weight or their material. In particular, one subject reported to hear not only the dropping object but also the object where it is dropping on, its shape and material, by saying that "the object where has been dropped on sounds quite strange. It is a kind of complex thing being dropped on and there's something else vibrating as well " and, moreover, "it's like it has been dropped inside an object, where there is a kind of rim that is vibrating", and, as far as the material is concerning, "that's a completely different material, this one, where it is dropping on". Another participant noted that some dropping objects " hit on something that has a different density or dropped from a lower height".

As we already underlined, all the sounds were scaled quite uniformly, at least in one dimension, and, moreover, we can observe that the four sounds of single impact events, i.e. *7-wood*, *11-wood*, *6-rubber* and *7-rubber*, were all judged to have minimum elasticity, as it could be expected, and all were scaled very uniformly, probably because the scaling of the single impact events is not affected by the bouncing pattern. For example, in a bouncing sound "the second and the third (bounces) go higher, louder . . . I think it's higher in elasticity" and "the sound of it seems to bounce for too long and the bouncing thing is very, very small . . . too small for it". In addition, we can see that, among them, the rubber objects converge to the zero-elasticity point less than the wood objects, probably due to the characteristic elasticity of the materials

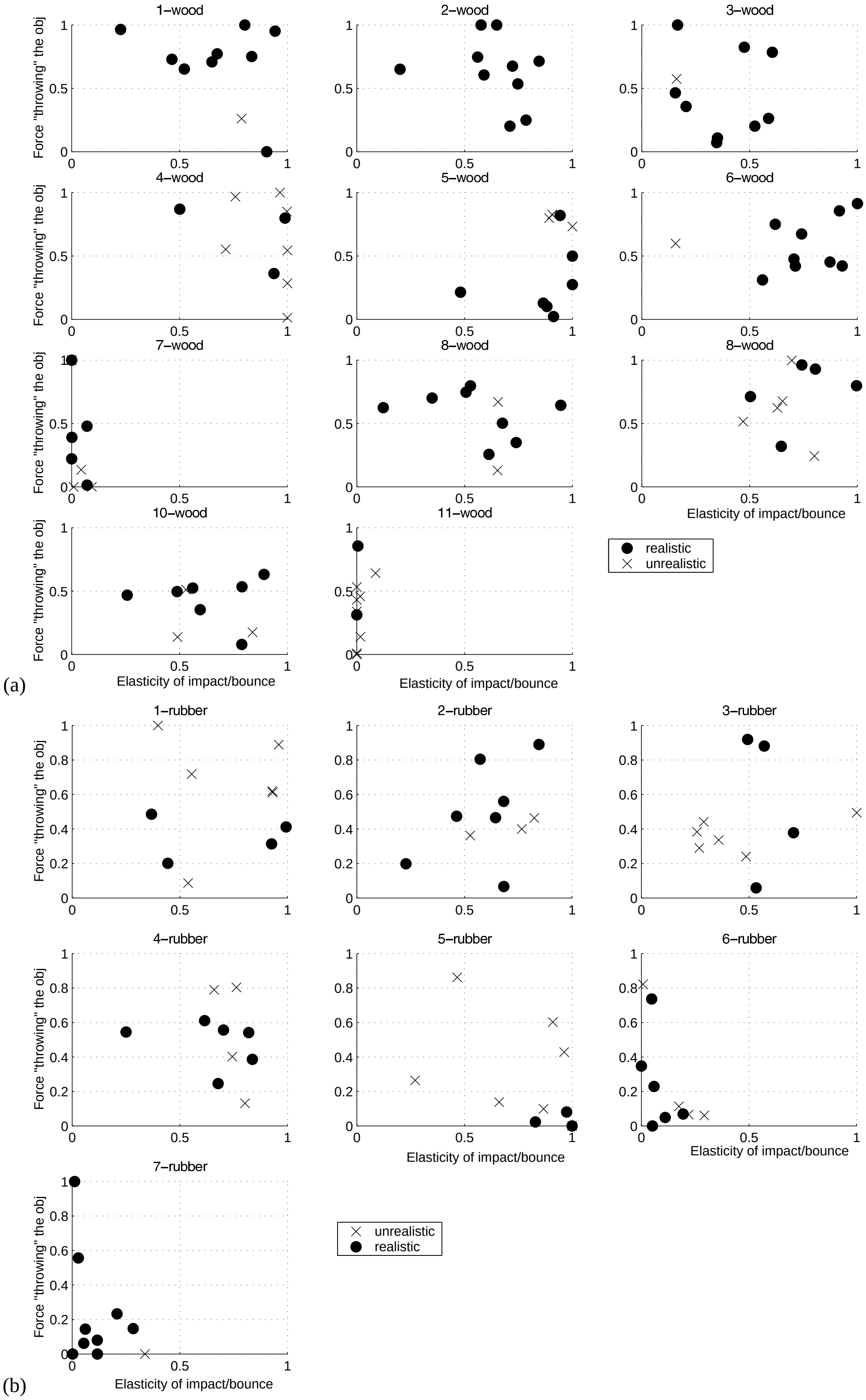

Figure 1: *Representation of the individual perceptual scaling and tagging information sorted by stimuli. (a) Only wood objects. (b) Only rubber objects.*

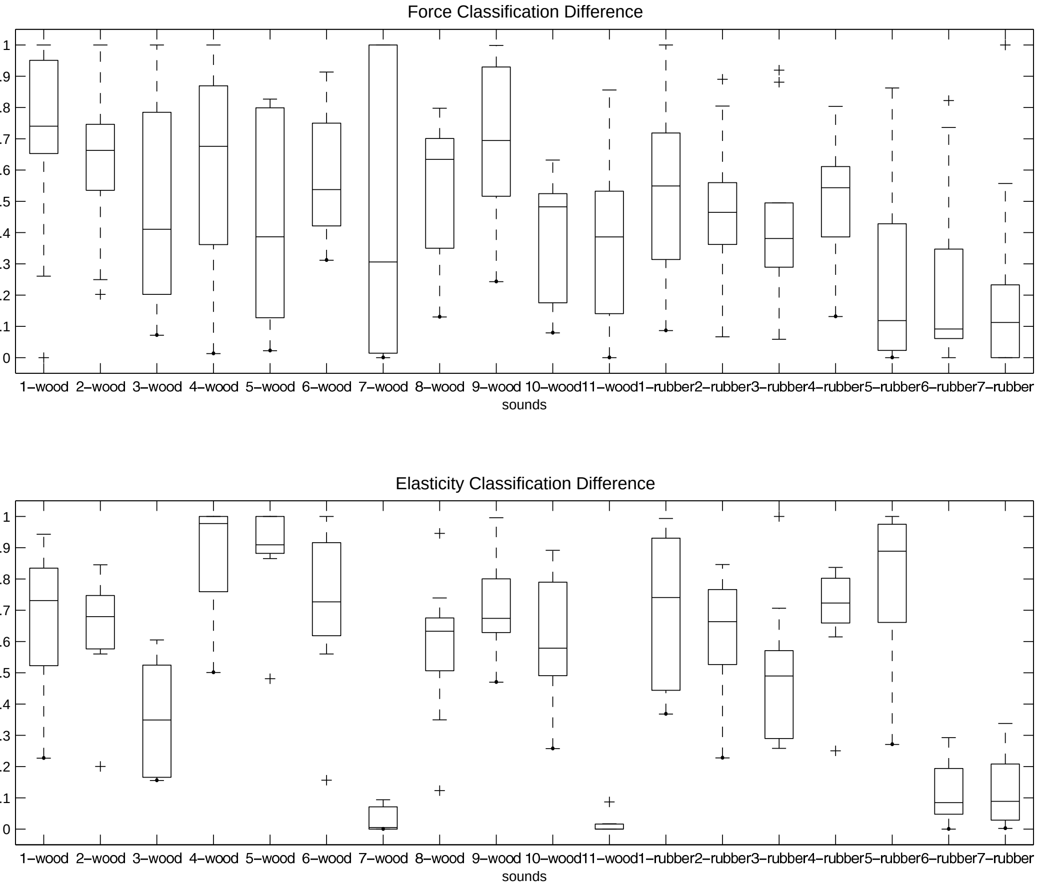

Figure 2: *Representation, by a box plot, of the perceptual scaling of the elasticity and force of the impact/bounce event sorted by stimuli.*

involved. We can see this by looking at fig. 3, where we plot the centroid of the scaling of each stimulus. These centroids are computed by keeping all the values, including the outliers. Although the barycentres would be more accurate by excluding the outliers, the plot represented in fig. 3 can give a general view of the stimuli mean positions within the perceptual space.

Observing the locations of the single impact sounds, we can notice a distinction of their perceptual estimations from the rest of the stimuli set and, moreover, that the material characteristics of these events appear to be distinct on both the scaling dimensions. In fact, the wood objects with a single impact sounds are identified to have in average less elasticity and more force rather then the rubber objects.

In the plot of the centroid three more sounds, that are *4-wood*, *5-wood* and *5-rubber*, are judged to have a high level of elasticity, resulting to be on the right hand of the graph, quite separated from the other stimuli. We can observe that these three sounds were the only in the set to have the strike velocity parameter with a lower value, rather than the others, which could mean that the changing in the strike velocity in the model is perceived as an increase in elasticity.

As far as the other stimuli locations are concerned, we can see the complex relations that connect the model parameters with the stimuli, although the centroid positions includes the outliers.

By the sessions task performances, it is clear the utility of the aura for conducting the scaling task. Usually, at the beginning of the experiment, the subjects took “the aura smaller. I will listen to them (the sounds) all together when they are organized a bit more. I slow down the aura”, while they used the aura for comparing the sounds and evaluating their judgments afterwards.

Moreover, it is interesting to notice that some participant preferred to judge the two dimensions together, while some others preferred to start by scaling one dimension and then moving to scale the other, because “it’s really hard to judge the two together”. In fact, “if I will try to think of the two things I will come confused”.

In fig. 4 we report the results of the questionnaire filled out by each participant during the debriefing phase, by representing the cumulative participant response with a bar chart.

The most interesting result, arisen from it, is that, even if the users found both the tasks difficult, they judged the scaling task harder then the tagging task, probably because of the influence of other dimensions in the event perception, as resulting from the verbal protocol as well. Despite this difficulty, the sounds were scaled consistently, at least in one dimension.

Moreover, the sounds were judged to be realistic and of good quality and these positive results could be connected to those obtained from the tagging task, as we have already reported.

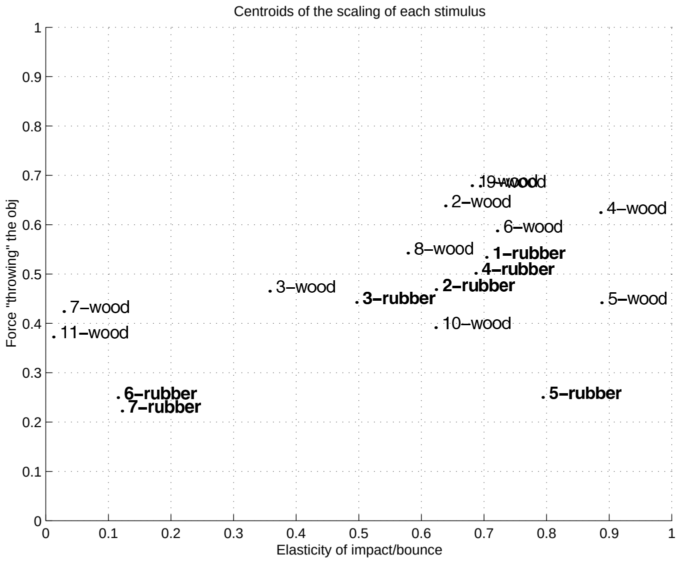

Figure 3: *Centroids of the scaling of each stimulus, sorted by material (Rubber objects written in bold).*

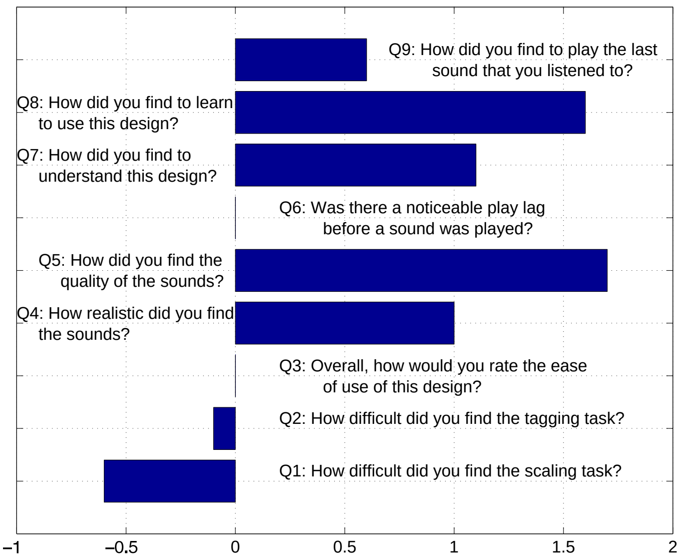

Figure 4: *Results of the questionnaire filled out during the debriefing phase: Cumulative Participant Response.*

As far as the software application is concerned, the ease of use of the Sonic Browser was judged to be on average, the slight delay of up to .3 seconds for playing a sound was noticed but accepted, and it didn't affected the tasks performances, since the question about the difficulty in playing the last sound listened was answered positively. Finally, good evaluations were achieved for the questions regarding how subjects found the interface to be understandable and learnable.

3. CONCLUSIONS

We can see that the Sound Objects are judged to be quite realistic and, even if they are tagged as unrealistic, because they are cartoonification of the reality, they still convey information and the physical properties of the events are still perceived by the listeners.

In particular, we have seen that the participants distinguished quite clearly among materials and about event identity. We have noted a parameter, the strike velocity, that could be particularly involved in the elasticity scaling. Nevertheless, some other investigations are needed in order to confirm this hypothesis.

As in previous experiments, the results are affected by the influence of other dimensions, which weren't examined in this case. The sound objects could be located in a multidimensional physical space and in a multidimensional perceptual space, connected to each other by a complex relationship.

Nevertheless, these experimental results were more uniformly estimated rather than those from the previous results. Therefore, we can state that, by providing to users sound objects synthesized with the same model, the listeners could scale the stimuli more easily and clearly. In this way, the sounds could convey information to the listeners. Moreover, even if the sounds are judged to be not realistic, what it is important is not to introduce distractors, such as a buzz tail, that could turn the user listening attitude.

4. ACKNOWLEDGMENTS

This work has been supported by the European Commission under contract IST-2000-25287 ("SOb - the Sounding Object"). The Authors thank the users, who participated voluntarily to the experiment. They are also grateful to Eoin Brazil for the contribution and installation of the Sonic Browser and to Colm McGettrick for his help in the room experiment setting.

5. REFERENCES

[1] L. Ottaviani, E. Brazil, and M. Fernström, "Psychoacoustic experiments for validating sound objects in a 2-D space using the sonic browser," in *Proceedings of the XIV Colloquium on Musical Informatics (XIV CIM 2003)*, Firenze, Italy, 2003, pp. 90–94.

[2] E. Brazil, M. Fernström, and L. Ottaviani, "Psychoacoustic validation and cataloguing of sonic objects: 2D browsing," in *The Sonic Object*, D. Rocchesso and F. Fontana, Eds., pp. 257–294. Mondo Estremo, 2003, Freely distributed under the GNU Free Documentation License. Available at `http://www.soundobject.org/SObBook/`.

[3] E. Brazil and M. Fernström, "Let your ears do the browsing - the Sonic Browser," *The Irish Scientist*, 2001.

[4] E. Brazil, M. Fernström, G. Tzanetakis, and P. Cook, "Enhancing sonic browsing using audio information retrieval," in *Proceedings of the 2002 International Conference on Auditory Display*, Kyoto, Japan, 2002.

[5] D. Rocchesso, R. Bresin, and M. Fernström, "Sounding objects," *IEEE Multimedia*, vol. 10, no. 2, pp. 42–52, 2003.

[6] M.T. Boren and J. Ramey, "Thinking aloud: Reconciling theory and practice," *IEEE Transactions on Professional Communication*, vol. 43, no. 3, pp. 261–278, September 2000.

[7] F. Avanzini, M. Rath, D. Rocchesso, and L. Ottaviani, "Low-level models: resonators, interactions, surface textures," in *The Sonic Object*, D. Rocchesso and F. Fontana, Eds., pp. 119–148. Mondo Estremo, 2003, Freely distributed under the GNU Free Documentation License. Available at `http://www.soundobject.org/SObBook/`.