Fall 9-2014

A design exploration on the effectiveness of vocal imitations

Stefano Delle Monache

Stefano Baldan

Davide A. Mauro
Marshall University, maurod@marshall.edu

Recommended Citation

Delle Monache, S., Baldan, S., Mauro, D. A., & Rocchesso, D. "A design exploration on the effectiveness of vocal imitations." 40th International Computer Music Conference (ICMC) joint with the 11th Sound & Music Computing conference (SMC), Athens (Greece), 14-20 September, 2014.

A design exploration on the effectiveness of vocal imitations

Stefano Delle Monache **Stefano Baldan** **Davide A. Mauro** **Davide Rocchesso**
Department of Architecture and Arts, IUAV University of Venice
stefano.dellemonache, stefano.baldan, davide.mauro, davide.rocchesso @iuav.it

ABSTRACT

Among sonic interaction design practices a rising interest is given to the use of the voice as a tool for producing fast and rough sketches. Goal of the EU project SkAT-VG (Sketching Audio Technologies using Vocalization and Gestures, 2014-2016) is to develop vocal sketching as a reference practice for sound design by (i) improving our understanding on how sounds are communicated through vocalizations and gestures, (ii) looking for physical relations between vocal sounds and sound-producing phenomena, (iii) designing tools for converting vocalizations and gestures into parametrized sound models. We present the preliminary outcomes of a vocal sketching workshop held at the Conservatory of Padova, Italy. Research through design activities focused on how teams of potential designers make use of vocal imitations, and how morphological attributes of sound may inform the training of basic vocal techniques.

1. INTRODUCTION

Any design activity starts from sketching. By means of sketches the designer produces, verifies, selects, communicates, and refines ideas [1, 2]. Over the last ten years, the research community on sonic interaction design has been focusing on the development of sketching approaches to enable, communicate and preserve the expressive and performative qualities of sonic interactive artifacts [3]. A sonic interactive sketch can have several manifestations, depending on the level of detail and the particular aspects that the designer is investigating: Demonstration methods include vocal sketching [4], foley-based and electroacoustic Wizard-of-Oz sound making [5], film and theatre-based strategies [6], basic design of sensorized mockups exploiting procedural audio tools [7, 8].

Vocal sketching, in particular, is a promising approach which exploits the sound-producing tool par excellence, the human voice, to act out sound design ideas. Indeed, notable vocal imitators and comedians, such as Michael Winslow [1] and Fred Newman [9], have developed basic techniques to imitate the sounds of everyday appliances and sound effects to compose complex narratives just by using their vocal capabilities and a microphone. The immediacy of communication, the inherent enactive and embodied nature of vocal imitations potentially make the human voice the ideal tool for sound designers, similarly to what the use of pencil (i.e., drawing) actually represents for visual designers. Indeed, vocal sketching is promptly available to the sound designer, does not require particular skills to enact meaningful vocalizations, and is highly performative, especially when sketches are made for tightly coupled interactions.

Lemaitre and colleagues investigated the effectiveness of non-conventional vocal imitations (i.e., creative utterances intended to be acoustically similar to the sound of the thing to which they refer) and verbalizations to communicate sound in recognition tasks [10, 11]. It was shown that (i) vocal imitations are as effective as verbalizations in communicating identifiable sounds, (ii) humans maximize the use of vocal imitations when the referent sound (i.e., the sound-producing event) is not clearly identifiable, and (iii) the imitated sounds are effectively communicated when the spectral and especially temporal information (envelopes, patterns, etc.) of the referent sound are retained. With a complementary approach, Merer and colleagues investigated the existence of a typo-morphology of spectro-temporal signal properties (i.e., sound attributes) capable of evoking motion, or in other words gestures. Drawing tasks were exploited to study the association between unidentifiable sounds and basic visual trajectories, in order to extract intuitive attributes for sound design and synthesis. The resulting list of attributes included direction, randomness, continuity, shape, and size of the morphological profile of a given sound [12, 13]. Given these premises, it is straightforward to hypothesize a future scenario wherein vocal utterances to sketch sonic interactions are exploited to inform digital sound models, control the sound synthesis, and refine the sonic sketches iteratively.

The EU project SkAT-VG [2] (Sketching Audio Technologies using Vocalizations and Gestures, 2014-2016) is aimed at finding ways to exploit voice and gestures in sonic interaction design. Research in SkAT-VG proceeds along three directions: (i) improving our understanding on how sounds are communicated through vocalizations and gestures; (ii) looking for relations between vocal/gestural primitives and the physical characteristics of sound-producing phenomena; (iii) designing tools for converting vocal and gestural actions into parametrized sound models.

[1] https://www.youtube.com/user/michaelwinslowtv.

[2] http://www.skatvg.eu/

In this paper we describe the research through design activities carried out by SkAT-VG researchers in the workshop on "Sonic sketches, voice and gestures for sonic interaction" [3], organized by the Sound and Music Processing Lab SaMPL [4] of the Conservatory of Music "C. Pollini" of Padova (Italy), in collaboration with the Iuav University of Venice. Typically, in research through design activities, raw models and sketches of sonic interactive scenarios and artifacts become means of embodied design thinking, that allow to study both the designer's cognition and behavior, and to assess the goodness of experimental designs, tools, and approaches. The workshop is the natural venue wherein research meshes with design activities. Shared doing, reflective practices and inter-observation are instrumental to enable understanding through designing sonic and interactive objects [14]. The workshop was aimed at investigating the effectiveness of vocal imitations for design purposes and setting the playground for the development of a workshop format in the scope of SkAT-VG researches.

The paper is organized as follows: in section 2 the workshop is described in terms of learning and research objectives; in section 3 the preparatory exercises on vocalization techniques are described and assessed; section 4 presents the outcomes of the design session; finally we draw our conclusions.

2. THE WORKSHOP

The workshop took place on March, 24th, 2014 at the Conservatory of Music "C. Pollini" of Padova (Italy). Twenty-four students of sound and music technologies took part to the workshop. The participants were split in four teams of six participants, tutored by the authors. The workshop was planned as a tight schedule of "learning by doing" activities. The first part of the workshop (morning) was devoted to sensitize participants to the topics of sonic interaction design and vocal sketching. After a brief presentation of the SkAT-VG project, warming-up and preparatory exercises were paced in order to incrementally engage and practice vocalizations. Ear cleaning, practice of vocal imitations and their Foley reproduction, vocalization of referent sounds according to given typo-morphological profiles had a double objective: For the participants they were means to learn some basic techniques to communicate basic sound events and effects; for the researchers the proposed exercises were instrumental to (i) observe the imitator's behavior, in terms of strategies for both imitation and recognition, and natural use of gestures, under the design perspective; and (ii) having a first insight on the effectiveness of use of typo-morphological profiles for vocal sketching purposes. In the second half of the workshop (afternoon), participants were engaged in a design session: The assignment was to conceive the sonic interaction with an artifact, within the framework of well-defined objectives and constraints. The resulting sketches were presented and discussed collegially. All the activities were video-recorded for further analyses.

[3] http://skatvg.iuav.it/?page_id=196.
[4] http://sampl-lab.sme-ccppd.info/.

3. PREPARATORY ACTIVITIES

The goal of preparatory exercises is to foster students to start discussing and thinking about the importance of sound in real and mediated environments. Arguments (e.g., about ecological and musical listening, auditory perception in interaction) are passed through playful activities such as sound walks and ear-cleaning exercises. In addition, sensitizing exercises are aimed at fostering engagement and breaking the ice of the typical social discomfort that may characterize the exposure of own body (and voice) in public.

3.1 Exercise 1: Ear cleaning and basic vocalization techniques

This simple exercise is presented as recognition task, the acousmatic listening of samples of sonic interactions is aimed at improving analytical skills. Samples are played in order to introduce the taxonomy of everyday sounds [15], the concepts of affordance, structural and transformational invariants, and to reflect on the informative, yet deceptive nature of sound. The presentation of sound samples is paced in order to gradually move the discourse from the simple description of sound to the understanding of the sound-action relations. The second part of the exercise introduces basic vocalization techniques [9] instrumental to imitate the basic sound phonomena previously presented. For example the *inhaled fry* technique, can be used to imitate a large variety of frictions, while the *palate grind*, shown figure 1, allows to reproduce a large variety of grinding and scraping sounds, as the name itself suggests. Participants listen to the moderator's example and try to repeat it. The progressive exploration of the technique enacts the control over one's own voice, nourishes the emergence of semantic associations, and facilitates self-confidence.

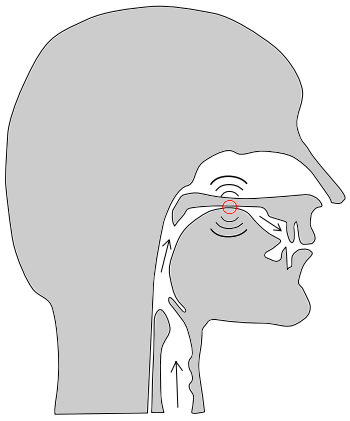

Figure 1: Arching the tongue on the soft palate while exhaling produces the *palate grind*, which is useful for the imitation of hand saws, trains, coffee grinders and so on.

3.2 Exercise 2: What is that makes *tictic*, *vvrrrrr*, *psssst*?

The goal of this exercise is to practice vocal skills by imitating a set of referent sounds. The exercise is proposed as competitive cards game, suitable for small teams of players.

3.2.1 Rules and pacing

Each team is provided with a deck of cards of referent everyday sounds, and a set of Foley-oriented ordinary, sounding objects such as clips, strings, pens, marbles, dice, boxes of various shapes and materials, elastic bands, brushes, and pipes. Within each team, the first player (i.e., the performer), placed behind the other components' back, takes a card and tries to mimic the described referent sound (e.g., an explosion). The other players (i.e., the listeners) should guess the referent sound. The listeners can manipulate the sounding objects at their disposal to facilitate the guess. When someone guesses correctly the team receives one point, and the turn is passed to the nearest player on the left. Discarded cards entail one point of penalty. The amount of cards collected in the preallocated time slot determine the winning team.

The set of referent sounds was chosen in order to encompass a mixture of the categories of sounds used by Lemaitre and Rocchesso [11], that is (i) identifiable complex events, (ii) elementary mechanical interactions, and (iii) unidentifiable mechanical sounds. We did not use the category of artificial sound effects, which would result too difficult to describe univocally on the card and consequently difficult to mimic. Table 1 shows the list of referent sounds present in our deck of cards:

sharpening blade	bouncing ball	hammering
shutting squeaking door	rolling ball	explosion
blender	buzzer	blowing
bottle pop	sea waves	wooshing
electric fan	engine	typing
horse trot	crumpling paper	whipping
scissors	pouring liquid	hitting
suction cup	wooden impact	sawing
wind	throwing dice	drilling
gas stove	spring	puffing

Table 1: List of referent sounds.

3.2.2 Results and discussion

In the scope of the SkAT-VG project, this exercise was aimed at observing the behavior of the potential designer. The players were free to move and use free gestures to accompany and support their mimicking, no constraints were given except for the ban on using onomatopoeia and, of course, speech. For this purpose, a laptop was placed in front of the performer in order to document the imitation. Each team had forty minutes to mimic a total amount of thirty cards (i.e., sounds). All the teams managed to play at least ten cards, which gives a hint of the initial difficulty and reticence of vocal sketching as a practice, and the effort needed to pinpoint a specific sound through vocalizations. The analysis made on edited video-recordings stressed a number of interesting issues.

Familiarity: A first observation is that the familiarity with the referent sound, in terms of internalized experience, largely affects both the imitation and guessing stages. Cards like typing, shutting squeaky door, scissors, explosion or pouring liquid took a short time to be communicated. On the other side, cards like drilling and whipping required several iterations for the teams to agree on the guessing.

Temporal patterns: Similarly, it was observed that referent sounds with dense and irregular temporal patterns require more effort to be reproduced with the voice. For instance, the compound sonic gesture of throwing dice was mimicked as a combination of shaking and bouncing. However, the listeners initially interpreted these two sound events as throwing and falling of rounded objects, such as spinning tops. Only after several tries of Foley exploration of the imitated sound (dice were among the available sounding objects), performer and players agreed on the imitation.

Direct and indirect sounds: A further consideration concerns the ease of vocal imitation of direct and indirect sounds, being the former the sound of tightly coupled manipulative actions (e.g., pouring liquid, scissors), and the latter the sound produced by a process caused by either a manipulative or triggering action (e.g., throwing dice, drilling). Direct sounds seemed to be easier to imitate and guess compared to indirect sounds, which implies more levels of mediation and result in less embodied interaction.

Use of gestures: From the viewpoint of a future SkAT-VG tool, this exercise was aimed at observing in which way a potential designer, asked to produce, for instance, a crumpling sound, would make use of his/her body and gestures to support vocal sketching. Recent studies showed that humans mainly mimic the sound-producing action when the causal action can be identified. Conversely, when the sound-action causality is lost, humans trace contours related to the acoustic features of sound [16]. From the design viewpoint, we observed that the more extroverted participants made use of mimicking gestures when the referent sounds were directly linked to causal actions. For example, whooshing sound was mimicked as acting with a sword, pouring liquid was accompanied by the referent action. Scissors sound was described either by representing the motion of the sharpened edges with the index and middle fingers, or by representing the opposable movement of the thumb and the index to operate the handles of the scissors. Head movements were often used to trace the sound dynamics and accompany the gestures.

Use of Foleys: when conceiving this exercise, we introduced the use of Foleys as means to reinforce the guess by the listeners. In particular, given the competitive nature of the exercise and the constraint of time, we expected that the listeners would have exploited the objects manipulation to facilitate the guess. For this purpose, we were

interested in observing the guessing strategies used in objects manipulation, whether focused on finding the right object to reproduce the sound producing action rather than the spectro-temporal characteristics of the sound to guess. Nonetheless, this happened very few times. Listeners generally preferred to enable discussion through sound and action mimicking. On the other hand, we should account retrospectively the fact that the use of Foley in the scope of the exercise was somewhat ill-posed.

3.3 Exercise 3: Play it *repeated*, *irregular*, *passing by* and *approaching* again, Sam!

The goal of this exercise is to enact and fine-tune the control over single vocal techniques, by modifying a given vocal sound according to some typo-morphological attributes. The exercise is proposed as competitive cards game, suitable for small teams of players. Under a design viewpoint, this exercise represented a first exploration on the effectiveness of such attributes as vocal sketching patterns.

3.3.1 Rules and pacing

Each team is provided with two decks of cards containing the verbal description of the typo-morphological attributes of sound, listed in table 2. This exercise leans on the list of high-level descriptors of spectro-temporal dynamics, associated to categories of motion by Merer and colleagues [13, 12]. They were proposed as they are, to test their validity in the context of abstract vocalizations. Each team is provided with the same predefined vocal sound. The sound is representative of a specific vocal technique (e.g., glottal fry), and should be abstract and short enough to not be associated to any sound-producing event. Within each team, the first player picks up a card from the first deck and tries to vocalize the assigned sound, according to the resulting attribute. The other players (i.e., the listeners) have to guess which attribute was vocalized among the ones available in their second deck. When someone guesses correctly the team receives one point, and the turn is passed to the nearest player on the left. Discarded cards entail one point of penalty. The amount of cards collected in the preallocated time slot determine the winning team.

For this exercise, we also asked the players to draw, at each round, their personal, visual interpretation of the attributes' verbal description. Each team had forty minutes to go through the whole deck.

3.3.2 Result and discussion

An observed limitation of vocal sketching as a design practice is that it is quite complicated to intentionally vocalize sounds that have not a direct referent in the physical world [17]. Sonic interaction design workshops' experiences showed that designers mostly make a metaphorical use of vocal sketching. In other words, its use is not only constrained by one's own vocal capabilities, but also by the difficulty to detach from the personal repertoire of mental images of sonic events. Even when abstract, electronic sounds are concerned, they are mostly related to specific sci-fi imagery and effects. This exercise is conceived as a possible training method to go beyond this limitation. For this purpose, we assigned a low-pitched, "aaah" glottal fry sound. This technique is used to describe a particular series of clicking sounds created in the glottis.

Shape	Size	Randomness	Direction	Continuity
Linear	Small	No irregularities	Passing by	Stable
Repeated	Medium	Low irregularities	Approaching	Impulsive
Rotating	Large	High irregularities	Going away	
			Rising up	
			Falling down	

Table 2: List of categories (shape, size, randomness, direction and continuity) and their morphological attributes.

In general, this kind of exercise required more engagement, in terms of body effort and exposure, compared to the previous one. We observed a limited use of free gestures, except when they accompanied attributes which needed some reference point. For example the size attribute, which refers to the perceived size of the sound profile, is highly subjective. In some cases, the players exploited their open arms to accompany their utterances in order to convey small, medium or large size. Similarly head and body movements were exploited for supporting passing by, approaching, and going away direction, and rotating shape. Major confusions were also observed regarding the meaning of linear shape, stable continuity, and no irregularities (randomness), which were often perceived as interchangeable. On the other side, rising up and falling direction, low and high irregularities, impulsive continuity resulted almost immediate to perform and guess. These observations are reflected in the visual representations made by participants. The size attribute is represented with various shapes of different sizes. Passing by, approaching and going away visualizations always use a referent sign, like a dot representing the listener's head. As a general remark, the exercise proved to be useful to foster the exploration of the given vocal technique, but failed in providing a coherent typo-morphological strategy. This confirms Merer and colleagues' observation about the possible bottleneck occurring in the perceptual overlap between some elementary patterns of their proposed typology. Further iterations are needed to improve both the exercise and the deck of cards.

4. THE DESIGN SESSION

The afternoon activity was devoted to the design session. The assignment for each team was to conceive the sonic interaction with an artifact within the following constraints; (i) use the glottal fry sound practiced in exercise 3; (ii) produce at least four variations of the glottal fry sound; (iii) the sound is coupled to the actions needed to use the artifact;

(iv) no speech. Additionally, we asked the teams to try to synthesize the vocalized sounds with the help of their tutor. For this purpose, we prepared a custom GUI to control the synthesis parameter of the friction sound model, available in the Sound Design Toolkit (SDT) [8]. Nonetheless, only one team managed to work on the SDT. The remaining approached the sound synthesis too marginally to be evaluated. Rather, that was a cue of the fact that time-consuming activities such as sound programming need to be properly paced and organized. Indeed, the teams had two hours to develop and prototype their design ideas, produce and rehearse their sonic sketches, and prepare their presentation.

4.1 The Maiden Iron

Sebastiano Aleo, Stefano Ceri, Andrea Fregnan, Matteo Girardi, Elisa Leonori, Elia Zupin.

Ironing is an activity of rich multisensory interaction, involving vision, proprioception, touch, heat perception, and sound. An expressive sound can make continuous ironing a more engaging and effective activity. Sound can help avoiding accidents by signaling presence of the hot iron. The amount of vapor can be, and already is in some extent, conveyed through sound.

The design process:

In the brainstorming several everyday activities where considered for sound augmentation. Ironing was unanimously recognized as an activity where sounds can be profitably used both in continuous manipulations and in signaling. The construction of a mockup iron that could be actually manipulated allowed extensive vocal sketching accompanied by the gestures of ironing. The final mockup is shown in figure 2.

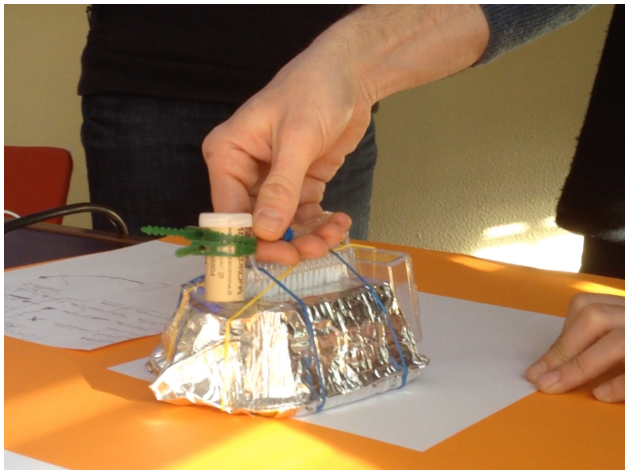

Figure 2: The Maiden Iron.

Discussion: It is interesting to notice that the prop materials were not much used for experimenting with different sounds. Instead, they were used for making a credible mockup artifact the group could play with. Gestures were exerted directly on the artifact as simulated control action, and the voice was used as exclusive sound sketching tool. The team coordinator preferred to use most of the time for practicing the stirring actions and the associated vocal sound simulations. The group was exposed at the Sound Design Toolkit only at the end of the session, thus getting only the impression on how the vocal imitations could be actually turned into synthetic sounds and further refined by manipulation of the model parameters.

4.2 ElectroSonar ByCycle

Simone Arrighi, Alessandro Aruffo, Enrico Baccarin, Matteo Michelutti, Matteo Pilotto, Nicola Sanguin.

The ElectroSonar ByCycle, shown in figure 3, was conceived as a sonic augmentation of the pedaling, similarly to what children do when they tape the card on the frame, so that it touches the spokes. Rising up and falling down profiles informed the evolution of the sound feedback, according to speed. Breaking sound, horn, and an anti-theft system were designed.

The design process: Vocalizations and gestures were used to produce fast and rough sketches of the envisaged functions. Several and extensive iterations were needed in order to translate the vocalizations in synthetic sounds, on the SDT environment. In practice, the tutor acted like an interface between the software tool and the designers, manipulating control parameters, proposing sounds and waiting for further inputs from the participants. For this purpose, the participants were asked to refine the synthetic sound, by using their voice and gestures.

Figure 3: The ElectroSonar ByCycle.

Discussion: The main concept is to hide from the view of the designers all the software complexity and let them "play" with the sonic material. In the context of the project, the approach pursued by the team highlights the potentialities of vocal sketching, but also shows how, generally speaking, designers want to have full access and complete control over the tool they are interacting with. In this case, the participants were curious to learn about the sound synthesis parameters and their perceptual contribution to the resulting sound.

4.3 The Warmefresher

Elio Bizzotto, Matteo Boscolo, Mirko Brigo, Marco Parlante, Lorenzo Pasquotti, Mattia Piovani.

The Warmefresher is a heat adjuster for liquids and beverages. It is controlled by turning the bottle as one would do with a common tap mixer. Continuous sound feedback informs the user about the difference between the current heat of the liquid and the desired one: rattling if the bottle is too cold, harsh if the bottle is too hot, smooth and stable if the bottle is at the right heat. The Warmefresher was conceived to ease the process of warming milk for babies.

The design process: After trying potential interactions with found objects, the team chose a plastic bottle, and explored vocal associations of direct manipulations on the object, and various liquid sound qualities. Several rehearsals were exploited to evaluate and refine the design idea. Each participant in turn played the roles of the user, the vocal performer, and observer. The final sketch is shown in figure 4, and is composed of a plastic cup, which serves as a warming up/cooling down docking base for the bottle.

Figure 4: The Warmefresher.

Discussion: The team made use of metaphorical associations of sound, so that they would mimic, to a certain extent, instinctive vocal reactions which occur when feeling cold or touching hot objects. For this purpose, high irregularities in the sound intensity were used to represent the concept of "too cold", as they somewhat reminded the sound of rattling teeth and shivering. On the other side, an increase in pitch and roughness of the timbre was used to suggest the concept of "too hot", as if the device was screaming after being burnt by something. The last part of the discussion involved the manipulation of the friction model available in the Sound Design Toolkit, and the proposal of configurations of parameters to synthesize the sounds.

4.4 The Crushingnoise

Matteo Polato, Alessandro Perillo, Daniele Pezzi, Niccoló Granieri, Laura Nao, Devid Bianco

The Crushingnoise, shown in figure 5, is an expressive device which reacts to simple manipulations like pumping, crushing, rotating, tilting and their combinations. The stand-by sound is modulated according to the actions exerted. The Crushingnoise was conceived as a playful and relaxing tool. Two or more Crushingnoise allow users to engage in collaborative non-verbal interactions.

The design process: The brainstorming focused on building relationships between improvised configurations of found objects and continuous manipulative actions. The team explored intuitive associations between sound and gestures. Blindfolded interaction allowed to better concentrate on the sound-action links. Each member either performed the vocal sketch or the physical interaction. In this way, the sound and interaction designers could enable mutual comparison and exchange observations. The final sketch is composed of two small wooden bases between whom a small piece of foam is placed.

Figure 5: Two Crushingnoises in interaction.

Discussion: After several iterations, the team decided to abandon their initial musical approach, to focus on the use of some sound attributes explored in the morning. They used the attributes "rising up", "falling down" and "randomness" to sketch the sonic interaction. Rising up and falling down were associated to the upwards and backwards pushing, while randomness was linked to the tilting. However, together with the expressive potential of the device, several other attributes emerged, depending on the manipulation exerted (e.g., fast irregular pumping, impulsive tilting, rotating pushes). The public demonstration showed a collaborative use of vocal sketching, in terms of social interaction mediated by the use of two devices.

5. CONCLUSIONS AND FUTURE WORKS

The workshop described in this paper is the first of its kind organized by the EU project SkAT-VG. Its format was very compact (one day), the scope quite limited (minimal use of computer-based tools), and the participants biased by their field of studies (computer music and sound engineering). Nevertheless, most of the participants recognized the effectiveness and power of vocal imitations, especially when they are coupled with manipulations of physical objects. At this early stage, the workshops are especially means to distill and fine-tune the exercises. Several followup workshops are being planned with different kinds of audience (product design students, high-school students, etc.), span-

ning a few days, and with extensive use of sound synthesis models.

Acknowledgments

The authors are pursuing this research as part of the project SkAT-VG and acknowledge the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 618067.

6. REFERENCES

[1] B. Verplank, "Interaction design sketchbook," *Unpublished paper for CCRMA course Music 250a*, 2003. [Online]. Available: https://ccrma.stanford.edu/courses/250a-fall-2004/IDSketchbok.pdf

[2] B. Buxton, *Sketching User Experiences: Getting the Design Right and the Right Design*. Morgan Kaufmann, 2007.

[3] D. Rocchesso, S. Serafin, and M. Rinott, "Pedagogical approaches and methods," in *Sonic Interaction Design*, K. Franinovć and S. Serafin, Eds. MIT Press, 2013, pp. 125–150.

[4] I. Ekman and M. Rinott, "Using vocal sketching for designing sonic interactions," in *DIS '10: Proceedings of the 8th ACM Conference on Designing Interactive Systems*. New York, NY, USA: ACM, 2010, pp. 123–131.

[5] D. Hug and M. Kemper, "From foley to function: A pedagogical approach to sound design for novel interactions," *Journal of Sonic Studies*, vol. 6, no. 1, 2014. [Online]. Available: http://journal.sonicstudies.org/vol06/nr01/a03

[6] S. Pauletto, "Film and theatre-based approaches for sonic interaction design," *Digital Creativity*, vol. 25, no. 1, pp. 15–26, 2014.

[7] D. Rocchesso, P. Polotti, and S. Delle Monache, "Designing continuous sonic interaction," *International Journal of Design*, vol. 3, no. 3, pp. 13–25, December 2009.

[8] S. Delle Monache, P. Polotti, and D. Rocchesso, "A toolkit for explorations in sonic interaction design," in *AM '10: Proceedings of the 5th Audio Mostly Conference*. New York, NY, USA: ACM, 2010, pp. 1–7.

[9] F. Newman, *MouthSounds: How to Whistle, Pop, Boing and Honk for All Occasions... and Then Some*. New York: Workman Publishing Company, 2004.

[10] G. Lemaitre, A. Dessein, P. Susini, and K. Aura, "Vocal imitations and the identification of sound events," *Ecological Psychology*, vol. 23, no. 4, pp. 267–307, 2011.

[11] G. Lemaitre and D. Rocchesso, "On the effectiveness of vocal imitations and verbal descriptions of sounds," *The Journal of the Acoustical Society of America*, vol. 135, no. 2, pp. 862–873, 2014.

[12] A. Merer, S. Ystad, R. Kronland-Martinet, and M. Aramaki, "Semiotics of sounds evoking motions: Categorization and acoustic features," in *Computer Music Modeling and Retrieval. Sense of Sounds*. Springer, 2008, pp. 139–158.

[13] A. Merer, M. Aramaki, S. Ystad, and R. Kronland-Martinet, "Perceptual characterization of motion evoked by sounds for synthesis control purposes," *ACM Transactions on Applied Perception (TAP)*, vol. 10, no. 1, p. 1, 2013.

[14] S. Delle Monache, "Sonic interaction: Research through basic design," Ph.D. dissertation, IUAV University of Venice, Venice, Italy, 2012. [Online]. Available: http://rice.iuav.it/372/

[15] W. W. Gaver, "What in the world do we hear? An ecological approach to auditory event perception," *Ecological Psychology*, vol. 5, pp. 1–29, 1993.

[16] B. Caramiaux, F. Bevilacqua, T. Bianco, N. Schnell, O. Houix, and P. Susini, "The role of sound source perception in gestural sound description," *ACM Trans. on Applied Perception (in review)*, 2013.

[17] W. Hatch and A. Pirhonen, "Designing alarm sounds for the control of a hydraulic platform," in *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound*. ACM, 2011, pp. 116–121.