GROWING THE PRACTICE OF VOCAL SKETCHING

S. Delle Monache, D. Rocchesso, S. Baldan, D. A. Mauro

Iuav University of Venice
Department of Architecture and Arts
Dorsoduro 2196, 30123, Venezia, Italy
{sdellemonache, roc, sbaldan, dmauro}@iuav.it

ABSTRACT

Sketch-thinking in the design domain is a complex representational activity, emerging from the reflective conversation with the sketch. A recent line of research on computational support for sound design has been focusing on the exploitation of voice, and especially vocal imitations, as effective representation strategy for the early stage of the design process. A set of introductory exercises on vocal sketching, to probe the communication effectiveness of vocal imitations for design purposes, are presented and discussed, in the scope of the research-through-design workshop activities of the EU project SkAT-VG.

1. BACKGROUND

In many everyday life situations, people find more effective and economical to support their reasoning by backing or even replacing verbal communication with sketches. In most cases, this inherently knowledge-based activity takes the form of drawings, that is visual representations of direct percepts or mental images, aimed at quick recording and processing of information, and its communication.

The etymology of the word "sketch" goes back to the ancient Greek term $\sigma\chi\acute{\epsilon}\delta\iota o\varsigma$[1], which means "done or made offhand, on the spur of the moment, temporary, done extempore". The main characteristic of sketching, both as a process and a product, is its capability to elicit meaning and interpretation, even unintended. This is a peculiar, reflective activity which characterises the early stage of any inventive process, whenever individuals engage in the conception of a not yet existing artefact, being it an artwork, a physical thing, or even a concept. In this early stage, creative professionals (e.g., artists, designers, architects, composers) produce a wide variety of sketches (e.g., studies, diagrams, schematics, etc.) and annotations, which serve to generate and fix ideas, compare, select, communicate and refine them, and eventually set and solve problems that are often ill-defined. These rough displays are provisional, ephemeral representational acts which are known to be embodied means of design thinking [1, 2]. As reported by Goldschmidt [3, p. 80], the origin of sketching as a practice can be dated back to the late-fifteenth century in Europe, with the widespread availability of good quality paper. The incomplete and often minimal drawings, often collected in sketchbooks, were called "pensieri", meaning "thoughts" in contemporary Italian.

The central role of sketching in contemporary design has been advocated by many thinkers, educators, designers, and researchers [4, 5]. Although being traditionally associated to architecture and visual design [6], sketching as a design practice has nowadays breached several disciplines and domains. Designers are called to shape artefacts apparently immaterial, experiences, emotions, interactive behaviours in space and time [7, 8, 9].

Indeed, the knowledge and technological advancement in HCI-related domains has led to the widespread diffusion of computational artefacts in our everyday life. In the era of the *disappearing, yet ubiquitous computing*, this is reflected in the concern of how disciplines traditionally considered humanistic, such as architecture and design, could benefit from the use of computational tools in the practice and education [10], and conversely in the methodological issue of how grounding HCI research in design practice and theory [11, 12].

In the eighties, the first WIMP-based (Window, Icon, Menu, Pointer) CAD (Computer-Aided Design) tools became available on the market (Autodesk's AutoCAD[2], above all), and architectural computing became a dreamland, the gate to revolutionising the art of manual graphics and design (thinking). A new generation of "pensieri" was expected to emerge. Currently, post-WIMP interfaces and applications, such as Autodesk[3]'s Sketchbook Pro, Wacom[4]'s tablets, Anoto[5]'s digital pens, or Fiftythree's Pencil and Paper[6] among the others, are attempting to make their way among design practitioners. Certainly, CAD tools sped up the later stages of the design processes, by improving the efficiency, the development of the technical documentation of designs, and the overall product quality.

However, CAD criticism acknowledges that little support is still given by CAD tools to the early phases of conceptual design, and that the growing complexity of these systems demands an increasingly high cognitive load of operation. Seen from a telescopic standpoint, challenges and contradictions in the established usage of CAD tools are reflected in a kernel of activities which are still based on geometric and or structural modelling and representations; in a corresponding inborn caesura occurring in the design process flow, deriving from the effort required to transfer concept

[1] http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.04.0057%3Aentry%3Dsxe%2Fdios.

[2] http://en.wikipedia.org/wiki/AutoCAD.

[3] http://www.autodesk.com/products/sketchbook-pro/.

[4] http://www.wacom.com/.

[5] http://www.anoto.com/.

[6] https://www.fiftythree.com/.

designs in CAD specifications; and in the resulting reluctance of stakeholders (i.e., industrial developers and users) towards radical innovations [13] (e.g., ubiquitous computing approaches, consider Ishii's Luminous Table, a well-known *research-through-design* exemplar of augmented reality for urban planning [14]). For this reason, research on computational support for design started focusing on the purpose, practices and behaviours incident to the early stages of the design process, as they are carried out without computation, in order to effectively support them by means of computation [15].

As a matter of fact, design professionals still prefer to engage in "analog" media, when approaching the conceptual phase of a new project. Whiteboards, paper, pen and pencil, ready-mades and gestures are means to give shape and immediately enact ideas in space and time. Conceptual reasoning is embodied in placeholders (i.e., the sketches), that progressively mark the boundaries of the design space. In interaction design practices, "Wizard of Oz" [16] techniques, bodystorming [17], and paper prototyping [18] are exploited as sketching approaches and tools. To some extent, sensing and actuating technologies, and microprocessors represent raw materials at the designer's disposal, as far as they are exploited to generate fast, explorative and tentative representations.

Sonic interaction design (SID) is a brand new area of design science which leverages the design culture on the world of the audible and vibrations. The design challenge is concerned with how to create meaningful, engaging, and aesthetically pleasing sonic interactions, that is affecting through sound design an overall shape aspect of things, their appearance, identity, and experience of use [19]. In this respect, research in SID is strongly committed to 1) constructing solid foundations for the development of the design discipline, and 2) grounding the research activity in the design practice [20].

A large variety of approaches and techniques are available, not to mention the plethora of computational tools to model and generate sound. Techniques such as model-based sonification, audification, earcons, auditory icons, sound editors, plug-ins, and programming languages do exist and offer infinite possibilities to sound designers [21], however their use implies that specific design choices had already been made, in terms of problem-solving. In other words, there is a whole part of the design process which is still missing appropriate knowledge, skills, and tools for shaping the acoustic behaviour of artefacts. Sketching sound, at least in acceptation established in design studies, is a practice mainly neglected in the sound creation process. Yet, sound designers are usually more prone to show advanced proposals, with a high degree of refinement, than to expose themselves with early drafts of sound ideas.

In this context, a recent line of basic research has been focusing on devising meaningful practices to inform and develop computational supports for the conceptual stage of the sound design process [22]. In particular, voice and gestures, and the innate human ability to communicate acoustic phenomena by means of vocal imitations, are being investigated and systematised as main strategy to trigger voice-driven sketching of synthetic sounds [23].

In this work, we report the experiences collected in several workshops on vocal sketching. Exercises and design assignments are reviewed, in the perspective of developing a literacy and a propaedeutics on sonic sketching [24, 25] .The paper is organised as follows: In the next Section we trace the salient aspects of sketching and highlight some critical research themes in sketch-based interaction; in Section 3 we move our discourse in the aural domain; in Section 4 we describe the workshop format, and discuss three introductory exercises aimed at probing the effectiveness of vocal imitations; finally, we draw our conclusions.

2. SKETCHING FOR DENSITY AND AMBIGUITY

In the last years, Buxton's book "Sketching user experiences" [5] had the undeniable merit of bringing back to the fore the question of sketching in the context of interactive experiences and technologies. Designing is about getting the right design, by distilling between many ideas, and getting the design right, by transforming and refining the selected best idea. Sketching is that peculiar activity that enables the interpretation and emergence of lateral and vertical transformations of ideas towards the prototype.

By bearing on some seminal studies on the relationship between drawing and sketching in architecture and design [26, 27, 2], Buxton elaborates a non-exhaustive set of features which marks sketching from what is not. Whatever the tool or technique used are, the resulting self-generated displays are *quick*, since they provide impressions; are *timely*, *economical*, and especially *disposable*, they can be provided at a glance, and thrown away, since their investment is in the concept and not in the execution; they are *dense* and *ambiguous*, in that they do not exist in isolation, their variations carry a substantial semantic depth, and facilitate the emergence of new perceptual relationships between their distinctive elements.

It can be argued that ambiguity and density represent the least common multiple of any sketch, the quintessential qualities which characterise the use of representational displays in the early phase of a project [28]. Goldschmidt contends the extreme value of the sketch as means to modulate the design problem space [3], in that provisional representations allow the addition of overlays and the creation of new relationships, perceptual and cognitive. As the sketching activity proceeds from the first rudimentary idea to a certain number of more detailed variations, the boundaries of the design space expand and shrink in a process of progressive understanding and resolution of the problem at hand.

Sketching skills and expertise certainly affect the effectiveness of the search process, and fluency is one major component. A fluent sketcher is focused on manipulating the representations without actually sparing attention on the production process. The "sketchiness" of the tool at hand is transparent, embodied in the sketcher's kinaesthetic creativity, and progressively mastered through training and rehearsal, whether it is about drawing, paper folding, bodily acting, making, coding or vocal mimicking [29]. As a consequence, a second major expertise is the choice of appropriate methods and techniques to represent displays [4]. There is a significant basic design education dimension associated to the acquisition of skills and expertise that facilitates perceptual understanding, and the development of cognitive abilities in incorporating perceptual factors in works and recognise them in the work of others [20].

2.1. Sketching by computing

It is a fact that computers entered the design practice as indispensable tool for information search and retrieval, and creative production stages. However, computers are still far from being used as proper conceptual tool. In most cases the digital sketch is actually a "scanned" sketch, a digitised version of the analogue original, whether it is a drawing scanned for manipulation in software such as Photoshop, Rhinoceros, SketchUp, and so forth, or a recorded sound imported in any kind of sound editor and processor. Johnson and colleagues provided an exhaustive overview of the state of the art on computational support for sketching in design, analysing several sketch-based system exemplars, from early pen-based interfaces such as GRAIL and SketchPad, diverse recognition approaches and toolkits (e.g., on/off line, stroke-oriented, feature-based), to diverse application environments [15]. This overview nicely traces the contours of several research themes in sketch-based interaction. Although the analysis is carried out with respect to the visual domain, the research agenda set by the authors can be generalised to any kind of design:

Nature and ubiquity of traditional sketching, and opportunities for computational support: The design research is not only concerned with the understanding of the perceptual and cognitive processes involved in this problem-setting activity, but also with framing its semantics in basic elements to effectively inform recognition strategies (e.g., when and what recognising), and filling that application gap between early and later stages of the design process;

Sensing technologies for physical input devices: The most advanced sketch-based systems exploit stylus-based interaction, finger-tracking, multitouch sensing, electronically enhanced paper and ink. Depending on the relative size, the input devices (e.g., PDAs, smartphones, tablet, whiteboards) may show a trade-off between accuracy and convenience of drawing. However, other interaction techniques (e.g., exploiting voice and gestures) and devices are likely to emerge, not only to disambiguate recognition errors without interrupting the workflow, but also to strengthen directness in the representation of interactive sketches;

Strategies and techniques for sketch recognition: Despite the availability of a sizeable literature on recognition-based interaction, the authors point out the lack of empirical data on what accuracy rate is actually acceptable in *sketch* recognition, whether it is a drawing, a sound or a gesture. Hence, the development of machine learning techniques and control strategies should take in account the specific aspects of sketching as an activity, and not the mere use of the tool at hand;

Human-computer communication, emerging interaction styles and user interfaces: sketching practices are taking root in the internet of things, and new or alternative input hardware may support diverse interaction styles for the manipulation of representations, and promote collaborative activities. As pen-based interaction exploits manual drawing as means for enabling visual thinking, similarly new interaction styles and user interfaces should support the direct, embodied interaction with the perceptual characteristics of the design dimension at hand. For instance, Fasciani and Wyse proposed an automatically adaptive system that exploits the human voice as acoustic control source for the manipulation of perceptually-relevant features of sound synthesis algorithms [30]. In addition, future interactive sketching tool should envisage the ambiguity of the sketch input as a resource, leaving the user potentially free to manage errors and conflicts.

Computers are not only tools, but also proper design materials. Computing technologies are physically embedded in designs, and are provided with perceptual and expressive capabilities (i.e., sensors and actuators/displays) to manifest computed effects on the environment. From a complementary viewpoint, microcontrollers platforms such as Arduino, Lilypad, Phydgets, and the plethora of toolkits increasingly made available on the market, represent a further way of sketching by computing. The sketching activity concerns both the control/display dimension and the rough crafting of the electronic circuitry. Paper sketches are quick to produce, economical, and disposable, and can be easily provided with interactive controls and displays [18]. Paper circuits, robust enough to survive a few manipulations, can be easily implemented by means of conductive inks[7], layers of graphite, movable electric contacts, and other affordable solutions [31]. Paper sketches of sonic interactive artefacts can be quickly realised by embedding piezo loudspeakers and driving digital sound models by means of augmented paper mechanisms, thus enhancing the co-location of action and feedback [32].

3. SKETCHING SOUND

Sound professionals are more prone to decision-making by verbal thinking, and advanced prototyping approaches towards their clients. Traditional paper and pencil sketching hardly captures the inherent salient aspects of sonic information and interaction. For this purpose, the design research community has been developing and proposing novel approaches to sound sketching and design which could potentially lead to a prototypical process exploitable in educational and commercial contexts. Hug and Misdariis in [33] proposed a conceptual framework, which integrates designerly and scientific sound design methods, based 1) on the morphological characteristics and degree of abstraction of the functional relationship between sound and artefact (i.e., the degree by which a sound closely relates to the object, its use and functioning, or is a virtual placeholder for other kinds of information, like the displays in smartphones; 2) on situational categories which may represent the relationship between the sonic interactive artefact and its context of use, such private/public, intimacy level, causal/professional use, etc.; 3) on a structured set of semantic abstract themes and attributes, such as qualities of use and control, transformation processes, mood, structural states, and dramaturgy. The conceptual framework is meant to support the generation of grounded design hypotheses, especially at the very start of the conceptual phase, when the level of discussion is essentially verbal.

Sonic concepts can be represented by means of experience prototyping approaches, that is any kind of representation, in any medium, used to understand, explore or communicate the engagement with a product, space or system design. Early demonstrations and physical interactive sketches can exploit the Wizard-of-Oz and Foley approaches to simulate the sound-driven interaction with the concept. Typically, an invisible performer plays the role of the computational system, and manipulates sounding objects to provide real-time interaction to the user/designer [34]. This is an economical way to explore diverse strategies in sonic feedback, that can be eventually refined in functional mock-ups.

Sonic overlay is a form of video prototyping in which a filmed interaction is subsequently enriched with added or replaced sonic elements[8]. The sonic overlay affords a quick way to communicate scenarios, compare different solutions by simply replacing, muting or enabling the sound tracks, and provides a chronological history

[7] An interesting example is the Circuit Scribe rollerball pen, which is provided with a non-toxic, conductive silver ink, for quickly drawn sketches of electric circuits, `http://electroninks.com/`.

[8] `https://vimeo.com/12549217`.

of the sonic sketch, in both terms of lateral and vertical transformations (i.e., exploration of new or alternative ideas, and refinement of the selected ones) [25]. The video sketch can even originate from Wizard-of-Oz sessions, added sounds can be synthetic, produced by playing physical objects, or even generated through vocal imitations.

Indeed, vocal mimicry is a natural way by which humans communicate complex concepts and events by means of non-verbal descriptions. In vocal imitation, creative non-verbal utterances are intended to be acoustically similar to a given sound, or to the sound manifestation of a referent thing. In a set of recognition experiments, Lemaitre and colleagues showed that vocal imitations are as effective as verbal descriptions in communicating identifiable sounds, while outperforming when the referent sounds are non-identifiable [23].

Recently, this innate ability has brought the attention of the SID community, as intuitive and immediately available means to produce sonic sketches. The phonatory apparatus ideally represents the drawing tool par excellence, available to the sound designer to act out sound design ideas. Vocal sketching does not require particular skills to enact meaningful vocalizations, and is highly performative, especially when sketches are made for tightly coupled interactions. Practices and exercises for education purposes have been devised and explored in workshop setting, providing evidence of the beneficial capacity of vocal sketching in capturing, and manipulating the temporal aspects of the sound design inquiry [35, 36]. The pros of immediacy of use and absence of external technological constraints are counterbalanced by the cons of the inherent ephemerality of vocal sketches and natural limitations of the human voice. Not all the acoustic characteristics of sounds are reproducible by vocal imitations, although it is possible to communicate sounds by mimicking the salient perceptually-relevant features for their identification (e.g., a bell, or a polyphonic sound).

The ultimate goal is to facilitate what Goldschmidt defines as *dialectics of (sonic) sketching* [26]. This describes a peculiar pattern in the circular conversation between (sonic) sketch and human: Interpretive acts alternate with representational acts in a continuous feedback loop of new knowledge creation (*hearing as*) and form production (*hearing that*), as depicted in Figure 1.

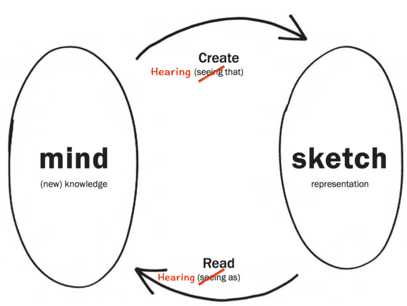

Figure 1: Sonic sketch and human in conversation, adaptation from Buxton [5, p. 114].

In other words, hearing while sketching shifts across two modalities, a reading phase in which new connections among the actual sonic elements are revealed (e.g., the configuration of the salient components of an envelope), and a creative phase in which the new internal image is actualised on the external representation, the sonic sketch.

3.1. Sketching audio technologies using vocalisations and gestures

The EU project SkAT-VG[9] (Sketching Audio Technologies using Vocalizations and Gestures, 2014 - 2016) is framed in the field of multidisciplinary research on computational support to sound design practices, especially regarding the early and conceptual stages of the design process. Two main strategies for vocal imitations have been highlighted, imitations that describe the event generating the sound (e.g., a squeaking door) and imitations that describe the spectro-temporal contours of the sound itself. Integrated sketching tools will allow to exploit the potential of vocal imitations and manual gestures to select and drive the sound synthesis, by controlling either the mechanical properties of physics-based sound models [37], or the parameters of signal-based synthesisers [22]. Indeed, voice-converted sound models, that is dynamic configurations of synthesis algorithms, can be further refined, easily communicated, and potentially exchanged among designers and stakeholders. For this purpose the basic research agenda of SkAT-VG is built around three main objectives:

1. **Understand**, that is extending existing knowledge in perception and production of vocal imitations and expressive gestures;
2. **Classify**, that is to develop automatic classifiers of vocal and gestural imitations, base on what is imitated, by integrating signal analysis with the physio-mechanics of vocal production;
3. **Design**, that is to explore the effectiveness of vocal and gestural sketching in sound design, by exploiting automatic classification for the selection and parametrisation of sound synthesis models.

In the following section we report about the research-through-design (RtD) activities that we carried out in several workshops on vocal sketching. Starting from a general workshop where the focus was on vocal imitations and product design [24], two more targeted workshops on vehicle sounds have been held. Automotive applications, especially concerning HEV (Hybrid and Electric Vehicles), represent a powerful design scenario, which is currently under exploration in the context of SkAT-VG research. The workshops were aimed at collecting information and investigating underlying concepts for the communication of vocal imitations.

4. TOWARDS A PROPAEDEUTICS ON VOCAL SKETCHING

The workshop on vocal sketching is structured in RtD activities aimed at a three-folded objective:

1. Ground the design exercises and their evaluation in phonetics (i.e., elicitation and articulation) and auditory perception of vocal imitations;
2. Study the sound designer's behaviour: Explorations in unconstrained setting are aimed at collecting emerging use strategies of voice and gestures for design purposes;

[9] `http://www.skatvg.eu/`.

3. Develop a propaedeutics on vocal sketching: Collect and evaluate exercises, and progressively refine and organise them in the form of a manual, or a sketchbook, build a structured format of workshop on sketching practices for sound design.

The current framework rests on a reduced version of the SID workshop[10] format, described in [36], in which 10-20% of frontal teaching is balanced with 90-80% of hands-on, learning by doing activities.

4.1. His engine's voice, a workshop on vocal sketching for motor sound design

An example of RtD activities on vocal sketching, spanning on a two-days workshop is the following:

First Part - Propaedeutics on vocal sketching
- Collectively -

- *Introduction* to the EU project SkAT-VG;
- *Ear-cleaning exercise*, everyday listening and action-sound affordances;
- *Vocalization techniques for the production of basic sound effects*.

- Teamwork -

- *The imitation game*, competitive guessing game on vocal imitations;
- *Soundomotion*, guessing game on morphological attributes of vocal sounds;
- *Acousmatic narratives*, polyphonic vocal sketches of fictional machines;
- *Product sound design*, sketching the sonic interaction with physical props;
- *Vocal mimicry of motor sounds*, improvisation session around idling and running motor sound;
- *Typo-morphological exploration of motor sounds*.

Second Part - Design session, vocal sketching and video prototyping
- Collectively -

- *The acoustics of motor sounds and sound synthesis approaches*.

- Teamwork -

- *Sound design session*, video prototyping of vocal sketches of motor sounds, according to a given brief.

The workshop is paced in two parts. The first part is introductory to everyday listening, to foster reflection on action-sound affordances. Some vocal techniques to produce basic sound effects [38] are introduced, within the framework of voice production attributes, according to sound source types and initiation mechanisms in speech, proposed by Helgason in [39]. This part is aimed at fostering the exploration of one's own vocal abilities, and sensitising elicitation strategies in vocal imitation tasks and possible use of accompanying gestures, while breaking the ice in one's own body and voice exposure. In the second part of the workshop, a general overview on motor sounds and synthesis introduces the teams to a proper design session. The assignment is to sketch the sound for a silent vehicle, according to a set of well-defined constraints. Some exercises are carried out collectively, some others as teamwork. So far, the workshops were attended by students with a background in computer music and technologies of sound, in a range of twelve-twenty participants, split in teams not exceeding the number of six members.

In the next subsection we cover the description of the most relevant exercises, by discussing their rationale and objectives in the context of SkAT-VG project researches. For this purpose, we refer the discussion of established exercises, such as ear-cleaning and vocal sketching of sonic interaction with physical props, to previously published articles [35, 20].

4.2. The exercises

Each exercise is ideally devised in order to exploit time constraints as means to foster iteration and improvisation in the accomplishment of the assignment. It has been demonstrated that rapid iterations in rough prototyping generates more valuable insights than allocating the same time for a single iteration [40]. In addition, the creative collaboration and spontaneity enabled by improvisational practices not only support the design work, but open space for the emergence of meaningful behaviours in vocal sketching practices.

4.2.1. Vocalization techniques for the production of basic sound effects

In this exercise, a few basic vocal techniques for producing several sound effects are introduced, by listening and mimicking target sounds, previously analysed and discussed in the ear-cleaning session. These are common techniques used by prominent imitators and artists such as Fred Newman[11], author of the well-known book Mouthsounds [38], and Michael Winslow[12], master of vocal gymnastics.

The Mouthsounds book is a valuable practical resource to start with. The table of contents is organised according to two main principles, the increasing difficulty of the techniques, and a pragmatic taxonomy of vocal imitations which reflects both the perceptually-relevant physical attributes of the referent sound (i.e., interacting materials and temporal development) and the correspondent articulatory characteristics of vocal production. With a few basic techniques and their virtuoso combination, it is possible to cover a large variety of sound categories: frictions, impacts, aerodynamics, machines, animal sounds, and so forth. For instance, the *Glottal Fry* technique is used to produce low frequency vibrations of the vocal folds, in order to obtain sparse patterns of clicking sounds. When combined with falsetto and glottal stops, this basic effects can be used to mimic several kinds of squeaking and creaking sounds (e.g., doors, wheels, a Geiger counter, etc.). Similarly, the *Palate Grind* technique is functional for the vocal production of many sounds with significant aerodynamic components. It is basically a *guh* sound used to mimic guns and explosions, air flows and whooshes, scraping, grinders, hand saw, etc.

Several other techniques are introduced in order to cover the taxonomy of source types and voice production attributes, recently proposed by Helgason [39] to support the annotation of non-linguistic vocal productions and transitions. Figure 2 shows

[10] https://vimeo.com/16655747.

[11] http://www.mouthsounds.info/home.htm.

[12] http://michaelwinslow.net/bio/.

only the first level of the taxonomy, without the further ramifications. Within this framework, the *Palate Grind* can be described

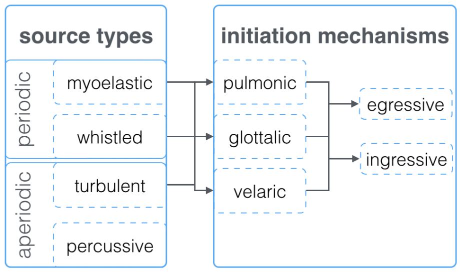

Figure 2: Vocal imitations are described according to the periodicity of the vocal signal, the articulatory characteristics of sound types, in conjunction with the airstream initiation mechanisms.

as an aperiodic, turbulent sound, initiated by means of the lungs, by pushing the airflow either out or in, through the mouth or nose (i.e., egressive / ingressive pulmonic initiation).

Integrating the Mouthsounds techniques within the system of voice production attributes proved to be effective in stimulating analytical skills, and foster the participants' iterative exploration and control of the initiation mechanisms. However, the warming-up goal of this exercise requires a careful pacing in order to keep the improvisational and playful approach in the foreground, without overloading the immediacy of the performance with the search of a compulsory coherence with the taxonomy.

4.2.2. *The (vocal) imitation game*

This exercise is structured as a guessing game[13]. Guessing games are largely used to reinforce concepts in (children) education and foster experiential learning. In addition, packs of cards are often used as tools to support learning and decision-making in design education and practice[14].

We devised a collaborative game based on a card pack of verbal description of sounding objects, the goal is to guess the sound-producing action or object mimicked with the voice. Each player/performer at turn chooses a card from the pack of verbal descriptions of objects/interactions, mimics the sound of that object, and has to lead the listeners to guess it. The game develops analytical skills through listening, and guessing constitutes an implicit evaluation of the performer's vocal imitation, thus enabling iteration and virtuoso exploration of the vocal capabilities.

The design research is focused on collecting evidence of emerging strategies in vocal imitations produced outside the laboratory. The performer's gestures are excluded from the listeners' sight in order to keep the vocal practice in the foreground and avoid a guessing mainly focused on *illustrators*, namely those iconic gestures closely related to speech [41]. Indeed, the set of referent sounds used in the pack of cards is characterised by a relatively high degree of action-sound causality.

The last version of the pack is made of fifty cards encompassing the verbal description of *mechanical interactions* that produce sounds involving liquids, solids, and gases without focusing on the sources, and *machines* of different kinds, such as button and switches, hums, air conditioning, vehicles, food processors, and wipers. Abstract artificial sounds were excluded from this pack: The verbal description would result highly ambiguous in a game which rests on a strict economy of time to convey sound impressions at a glance. In an analogous game on the relevant auditory features of abstract sounds, players have to guess specific spectro-temporal morphologies (e.g., increasing, decreasing, impulsive, stable, etc.).

The analysis of several video-recorded game sessions highlighted a limited and diverse use of gestures by players. The clear use of gestures is mainly elicited when the imitators are standing. Iconic, metaphoric and ancillary gestures are exploited to support and fine-tune the control on the voice articulation, especially when the first imitation attempts do not work. Standing up may be useful for educational purpose, in order to foster bodily learning and performance. On the other hand, it has to be considered, for the development of future sketching tools, that in the everyday practice of design it is more likely to have working situations where professionals mostly sit around a table. In addition, it was observed how the effective elicitation of vocal imitations not only depends on the familiarity with the referent sound, but also whether the verbal description of the sounding object on the card refers to the sound-producing action (e.g., typewriting) or to the artefact itself (e.g., typewriter), producing the latter a reduced priming effect on conceptual memory [42].

4.2.3. *Video prototyping of vocal sketches of motor sounds*

This is a proper design assignment used to asses the entry level of workshop participants, prior to any kind of exposure, or as checkpoint for intermediate evaluation, after the introductory exercises. The assignment mimics a real design case (in which a sound design agency is asked to produce a study/proposal for an automotive company), yet it focuses on the creative processes of conceptual sound design by means of vocal imitations. As shown in Figure 3, the assignment is to create the sound for a silent vehicle according to an associated list of three keywords and two evocative sounds. Each team has to choose one vehicle among the available proposals, and produce two alternative sound designs for a silent video-clip of the corresponding moving vehicle[15]. Vocal sketching is polyphonic, no speech nor onomatopoeia are allowed, as well as Foley and sound processing. A video editor is provided to make audio recordings, and basic sound editing. The list of proposed vehicles is composed of concept cars with a futuristic streamlining, yet capable of evoking very diverse product sound qualities.

In the last workshop, this exercise was exploited as checkpoint to assess the value of the various basic exercises on teams which were exposed only to some sensitising activities and not others. For instance, the video prototype of the team trained on the imitations of engine sounds showed a greater attention to the sound identity and functional aspects of the chosen concept car, compared to the more expressive design outcome of the team sensitised on complex, multilayered and sonically rich vocal sketching.

However, the exercise stressed how difficult is to move on from the verbal-only level and enact the actual sound-producing stage of the conceptual phase of a design. This difficulty was reflected in a low number of iterations, compared to the ease of iteration observed in the imitation exercises. However, once the teams

[13] `http://en.wikipedia.org/wiki/Guessing_game`

[14] `http://www.methodkit.com/research-method-cards/`, a comprehensive overview of existing types of packs for design purposes.

[15] `https://vimeo.com/128886746`.

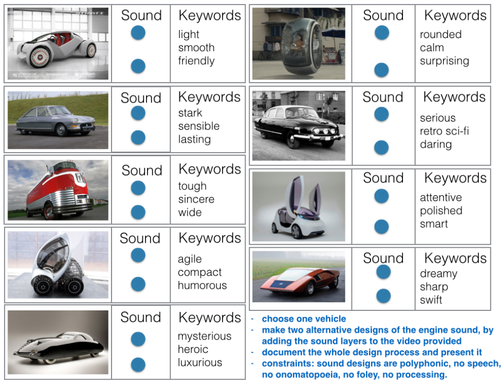

Figure 3: The list of concept cars with associated keywords for the sonic overlay exercise.

managed to engage in vocal production, they produced quite refined presentations in a very short time. The expressive use of manual gestures during the actual sketching was very limited, being the designers mainly intent on the quality of the recordings in front of the microphone. Instead, they made a creative use of several tools, pipes and boxes, to augment their vocal sounds. These clues will be taken in account for the improvement of the exercises, and the devising of an evaluation protocol.

5. CONCLUSIONS

The design hypothesis is that voice and gestures represent for the sound designer what hand and pencil are for the designer in the visual domain. Vocal imitations and drawings are natural forms of representations used by humans to give shape to notions, and focus their thoughts while reasoning. However, while the art of drawing in visual sketching has a long history, the same cannot be said for vocal mimicry, and especially within the younger field of sound design. The practice of visual pensieri has been studied, coded and refined over centuries of studies, and sketch-thinking became an inborn behaviour of visual artists and designers. On the contrary, it can be argued that sketching in the aural domain (if one rules out music composition) is still a convoluted practice.

Certainly, a major bottleneck is technological, and is represented by the lack of appropriate tools, “paper and pencils” to enact the emergence of sonic pensieri, and eventually their collection in sonic sketchbooks. The SkAT-VG project is addressing the sketching of integrated tools to support sonic pensieri with a double strategy: The understanding and classification of vocal imitations and expressive gestures integrate in the “paper and pencil” (a physical, vocal imitation-to-synthetic sound converter) with a top-down approach; a bottom-up process is instead centred on devising and supporting the emergence of relevant sketch-thinking behaviours in sound design, and ultimately providing guidance for the use of the computational tool. From the latter standpoint, in a first set of workshops, the conception of introductory exercises was aimed at probing the effectiveness of vocal imitations in acting out sound design ideas. Forthcoming workshops will emphasise instead the focus on the sketching behaviours, by involving expert sound designers.

6. ACKNOWLEDGMENT

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.

7. REFERENCES

[1] D. A. Schon, *The reflective practitioner: How professionals think in action*. Basic Books, 1984.

[2] A. Purcell and J. S. Gero, “Drawings and the design process: A review of protocol studies in design and other disciplines and related research in cognitive psychology,” *Design studies*, vol. 19, no. 4, pp. 389–430, 1998.

[3] G. Goldschmidt, “The backtalk of self-generated sketches,” *Design Issues*, vol. 19, no. 1, pp. 72–88, 2003.

[4] B. Verplank, “Interaction design sketchbook,” 2003, unpublished paper for CCRMA course Music 250a. [Online]. Available: https://ccrma.stanford.edu/courses/250a/lectures/IDSketchbok.pdf

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

[6] B. Edwards, *The New Drawing on the Right Side of the Brain: The 1999*. Penguin, 1999.

[7] P. Desmet and P. Hekkert, “Special issue editorial: Design & emotion,” *International Journal of Design*, vol. 3, no. 2, pp. 1–6, 2009.

[8] G. Crampton Smith, “What is interaction design?” in *Designing Interactions*, B. Moggridge, Ed. Cambridge, MASS: MIT Press, 2007, pp. 2–14.

[9] L. Hallnäs, “On the foundations of interaction design aesthetics: Revisiting the notions of form and expression,” *International Journal of Design*, vol. 1, no. 5, pp. 73–84, April 2011.

[10] K. Nakakoji, “Special issue on computational approaches for early stages of design,” *Knowledge-Based Systems*, vol. 18, no. 8, pp. 381–382, 2005.

[11] W. Gaver, J. Bowers, T. Kerridge, A. Boucher, and N. Jarvis, “Anatomy of a failure: How we knew when our design went wrong, and what we learned from it,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’09. New York, NY, USA: ACM, 2009, pp. 2213–2222.

[12] J. Zimmerman, E. Stolterman, and J. Forlizzi, “An analysis and critique of research through design: Towards a formalization of a research approach,” in *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, ser. DIS ’10. New York, NY, USA: ACM, 2010, pp. 310–319.

[13] I. Horváth and R. W. Vroom, “Ubiquitous computer aided design: A broken promise or a sleeping beauty?” *Computer-Aided Design*, vol. 59, pp. 161–175, 2015.

[14] H. Ishii, E. Ben-Joseph, J. Underkoffler, L. Yeung, D. Chak, Z. Kanji, and B. Piper, “Augmented urban planning workbench: overlaying drawings, physical models and digital

simulation," in *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, ser. ISMAR '02. Washington, DC, USA: IEEE Computer Society, 2002, pp. 203–211.

[15] G. Johnson, M. D. Gross, J. Hong, and E. Yi-Luen Do, "Computational support for sketching in design: a review," *Foundations and Trends in Human-Computer Interaction*, vol. 2, no. 1, pp. 1–93, 2009.

[16] J. D. Gould, J. Conti, and T. Hovanyecz, "Composing letters with a simulated listening typewriter," *Communications of the ACM*, vol. 26, no. 4, pp. 295–308, 1983.

[17] D. Schleicher, P. Jones, and O. Kachur, "Bodystorming as embodied designing," *Interactions*, vol. 17, no. 6, pp. 47–51, 2010.

[18] C. Snyder, *Paper prototyping: The fast and easy way to design and refine user interfaces*. Newnes, 2003.

[19] K. Franinović and S. Serafin, *Sonic Interaction Design*. Mit Press, 2013.

[20] S. Delle Monache and D. Rocchesso, "Bauhaus legacy in research through design: The case of basic sonic interaction design," *International Journal of Design*, vol. 8, no. 3, pp. 139–154, 2014.

[21] G. Dubus and R. Bresin, "A systematic review of mapping strategies for the sonification of physical quantities," *PloS one*, vol. 8, no. 12, p. e82491, 2013.

[22] D. Rocchesso, G. Lemaitre, P. Susini, S. Ternström, and P. Boussard, "Sketching sound with voice and gesture," *interactions*, vol. 22, no. 1, pp. 38–41, 2015.

[23] G. Lemaitre, P. Susini, D. Rocchesso, C. Lambourg, and P. Boussard, "Non-verbal imitations as a sketching tool for sound design," in *Sound, Music, and Motion*, ser. Lecture Notes in Computer Science, M. Aramaki, O. Derrien, R. Kronland-Martinet, and S. Ystad, Eds. Springer International Publishing, 2014, pp. 558–574.

[24] S. Delle Monache, S. Baldan, D. A. Mauro, and D. Rocchesso, "A design exploration on the effectiveness of vocal imitations," in *Proc. of the 40th Intern. Comp. Music Conf. (ICMC) joint with the 11th Sound and Music Computing Conf. (SMC)*. San Francisco, USA: ICMA, 2014, pp. 1642–1648.

[25] 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.

[26] G. Goldschmidt, "The dialectics of sketching," *Creativity research journal*, vol. 4, no. 2, pp. 123–143, 1991.

[27] M. Suwa and B. Tversky, "What architects see in their sketches: Implications for design tools," in *Conference Companion on Human Factors in Computing Systems*. ACM, 1996, pp. 191–192.

[28] V. Goel, "Ill-structured representation for ill-structured problems," in *Proceedings of the fourteenth annual conference of the cognitive science society*, vol. 14, 1992, pp. 130–135.

[29] D. Svanæs, "Interaction design for and with the lived body: Some implications of merleau-ponty's phenomenology," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 20, no. 1, p. 8, 2013.

[30] S. Fasciani and L. Wyse, "A voice interface for sound generators: adaptive and automatic mapping of gestures to sound," in *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, G. Essl, B. Gillespie, M. Gurevich, and S. O'Modhrain, Eds. Ann Arbor, Michigan: University of Michigan, 2012, pp. 479–482.

[31] D. A. Mellis, S. Jacoby, L. Buechley, H. Perner-Wilson, and J. Qi, "Microcontrollers as material: Crafting circuits with paper, conductive ink, electronic components, and an "untoolkit"," in *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. New York, NY, USA: ACM, 2013, pp. 83–90.

[32] S. Delle Monache, D. Rocchesso, J. Qi, L. Buechley, A. De Götzen, and D. Cestaro, "Paper mechanisms for sonic interaction," in *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*. New York, NY, USA: ACM, 2012, pp. 61–68.

[33] D. Hug and N. Misdariis, "Towards a conceptual framework to integrate designerly and scientific sound design methods," in *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound*, ser. AM '11. New York, NY, USA: ACM, 2011, pp. 23–30.

[34] 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

[35] 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.

[36] S. Delle Monache and D. Rocchesso, "Experiencing sonic interaction design: Product design activities at the sid summer school 2010," in *Prossime distanze - Proceedings of XVIII CIM Colloquio di Informatica Musicale*, A. Valle and S. Bassanese, Eds., Torino - Cuneo, Italy, 2011, pp. 87–91.

[37] 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.

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

[39] P. Helgason, "Sound initiation and source types in human imitations of sounds," in *Proc. FONETIK 2014, the XXVIIth Swedish Phonetics Conference*, 2014, pp. 83–89.

[40] S. P. Dow, K. Heddleston, and S. R. Klemmer, "The efficacy of prototyping under time constraints," in *Proceedings of the Seventh ACM Conference on Creativity and Cognition*. New York, NY, USA: ACM, 2009, pp. 165–174.

[41] D. McNeill, *Hand and mind: What gestures reveal about thought*. University of Chicago Press, 1992.

[42] G. Lemaitre and L. M. Heller, "Evidence for a basic level in a taxonomy of everyday action sounds," *Experimental brain research*, vol. 226, no. 2, pp. 253–264, 2013.