



Sketching step in sound design: the sound designers' point of view

Patrick BOUSSARD¹; Clément DENDIEVEL¹; Hélène LACHAMBRE¹

GENESIS, Domaine du Petit Arbois, 13545 Aix-en-Provence, France

ABSTRACT

Sound design consists in creating or modifying the sound of an object (car engine, Human Machine Interfaces, ...) in order to add relevant information, improve the feeling, the perceived quality, or reduce noise annoyance. The European project SkAT-VG aims at understanding the process of creating a sound ex-nihilo, with a focus on the sketching step (which may be done by the designer alone or during a discussion with the client). It also aims at creating new tools and methods to facilitate the designers' activity. To understand the sound design process and sound designers' needs, interviews have been conducted with a focus on the two points addressed by the project: the sketching step and the use of vocalizations and gestures. First, designers are asked to present themselves and talk about their experience and their creation process in sound design. Then they describe the tools they would like to have. They are also asked how relevant they would find the use of vocalizations and gestures and how they would imagine this use (for sketching a sound in an interactive process, or in a discussion with the customer). Finally, some prototype tools are presented to get preliminary feedback on their usefulness and usability.

Keywords: Sound design, Sketches, Vocalizations and Gestures.

1. INTRODUCTION

Sound design has been existing for very long (at least since the beginning of cinema, probably since the invention of theater). Since the last decades, the available electronic and numeric tools (computers, analogic and numeric synthesizers) changed a lot the creation process. This changes being quite recent, the creation process is not well known. Recent researches have been conducted to explore the whole design process (1), the definition of what a customer may expect from a sound design (2), the usefulness of sketches (1) and the use of vocalizations and gesture to communicate about sounds (3). The SkAT-VG project (Sketching Audio Technologies using Vocalizations and Gestures) (4) aims at studying several points:

- How are people conveying sound ideas using vocalizations and gestures? Recent researches have proven that non-verbal vocalizations are more efficient than words to describe a sound.
- What is the sound design process? Does it involve a sketching step? Should it involve a sketching step?
- Considering the two previous points, the project aims at developing new tools addressing the sketching step, using vocalizations and gestures.

In the project, Genesis (Aix-en-Provence, France) is the industrial partner, and is in charge of the contact with the sound designers. In a first task, Genesis is in charge of understanding how sound designers work, and what are their needs, particularly for the sketching. In a second task, Genesis is in charge of integrating technologies developed by the other project partners. The last task is to test the tools developed in the project by making them tested by real sound designers.

The first and last tasks involve a strong contact with sound designers. The first task has been addressed by interviewing professional sound designers about their creation process. The last task has been addressed by organizing an artistic event: the "48h of sound design". During this event, five sound designers were asked to design sounds to accompany existing pieces of art using tools developed to facilitate the sketching step using vocalization and gestures.

The rest of this article is organized as follows: in section 2, the interview method is presented. Then, in section 3, we summarize the answers of the sound designers to the diverse points addressed during

¹ pboussard@optis-world.com, hlachambre@optis-world.com, cdendievel@optis-world.com

the interviews: sound designers background, sound design practice, use of vocalizations and gestures. Finally, in section 4, the artistic event “48h of sound design” is presented, along sound designers feedbacks about the preliminary sketching tools developed in the project.

2. INTERVIEW CONTEXT AND METHOD

2.1 Interviews context

Interviews have been conducted by four engineers from Genesis, Aix-en-Provence, France. Seven interviews were made face-to-face, in diverse places (Venice, Genesis office, Tokyo, ...).

Four designers were interviewed in English: three British designers, therefore native English speakers, and one Japanese designer. Genesis being a French company, the seven other interviews were conducted in French, with seven French native speakers and one Italian designer (French fluent speaker).

At the end, eleven interviews have been conducted, with professional sound designers. All interviews have been recorded (audio or video), in order to get material to be studied later.

2.2 Questions

The interviews are run as a discussion. The questions asked are very open in their potential answers. The interviewer should be careful not to influence the designers' answers. However, if an important point is not developed by the designer, the interviewer may guide the discussion to this direction. In order not to influence the answers, the questionnaire is not sent to the sound designers before the interview. The interview is made of 5 main stages:

1. Quick description of the SkAT-VG project. As said before, this introduction is made so as not to influence the designer's answers.
2. Sound designers' presentation. The sound designer is asked to present himself: his name and profession, along with his previous experience as sound designer, his areas of work.
3. Sound design practice. The designer is asked to describe his general approach of sound design. If he does not spontaneously speak about a sketching step, this point is specifically asked by the interviewer.
4. Uses of vocalizations and gestures in sound design. The designer is asked whether he already used his voice or gesture at any stage of the process of sound design: early discussion with the customer, sketches, production of intermediate or final sounds.
5. Expectations. The designer describes what would be an ideal tool. This question is specifically oriented to the usefulness, advantages and limits of the potential use of voice and gesture, and to the usefulness of having specific tools for the sketching step.

3. INTERVIEWS SUMMARY

3.1 Sound designers' presentation

Andrea Cera (AC) is a musician, composer and industrial sound designer. He graduated in piano and music composition at the Conservatorio Cesare Pollini in Padova. He has worked with choreographers, has composed music for ballets, has worked on research projects (IRCAM, NOTAM, ...) and has designed the sounds of the Renault electric car Zoé.

Andy Farnell (AF) was trained as a computer scientist with a degree from University College London. He is now a researcher in procedural audio and is involved in both teaching and consultancy at Queen Mary University London.

Allister Sinclair (ASi) is a sound designer mostly working for artistic project. He is interested in random or pseudo-random sound creation and uses exclusively Pure Data to compose sounds. He studied at les Beaux Arts, Cergy, France.

Adam Stark (ASt) is a sound designer working for the C4DM (Centre for Digital Music) at Queen Mary London. He received his PHD and masters at Queen Mary in signal processing. He is also a computer scientist and writes a lot of software. He has worked on the sound design of two short films and has also played in a lot of bands where he did electronic sound generation.

Christian Heinrichs (CH) is a PhD student at Queen Mary University of London in the C4DM. He followed classical musical studies before turning to music technology, particularly the procedural audio. He practices sound design independently, mostly to the video game, and is participating in the development of an application of collaborative music creation.

Isabelle Ballet (IB) is a sound engineer. She worked for radio, before joining Ubisoft, where she is now audio director. She mainly worked on video games sound tracks, such as "Splinter Cell" or "Red Steel".

Julien Bayle (JB) is a sound and visual artist, who creates programmed installations and audiovisual live performances. His creation is almost exclusively artistic. He also provides lessons on the use of Max/MSP, Ableton Live and Max for Live.

Mathieu Pellerin (MP) is a freelance sound designer graduated from the Higher Institute of Sound Techniques ESRA. He has accumulated experiences in various music recording studios before devoting to animated films sound design in the production company Ramses II. He recently signed on sound editing and sound design of the animated film "The Suicide Shop" by Patrice Leconte.

Norio Kubo (NB) is a Japanese sound designer. He has been the CEO of "Yokohama institute of acoustics" since 2005. His activities are industrial sound design, mainly for automotive industry, but also for alarm or everyday goods.

Simon Cacheux (SC) is a sound designer graduated from the French sound engineer school ENS Louis-Lumière. He started by doing a little cinema and then worked a lot on sound scenography (spaces' sound design). He is now working in the automobile industrial field, and teaches sound design in Paris Condé school and in Limoges Limousin IUT.

Xavier Collet (XC) is a freelance composer and sound designer. His creation activities are as varied as the composition of movie sound track, or video games and MMI sound design. He also teaches sound design, sound technique, music computer science, and is the author of video tutorials.

3.2 Sound design practice

3.2.1 Steps in sound design process

Overall creation process

Few designers follow a precise and recurrent procedure in their different sound designs. JB is the designer with the clearer and most defined process of creation. The first step is the "*writing and research step*". This is done almost without creating any sounds, but by using a visual description of sounds: words, shapes, drawing. In a second step, he creates sound sketches. The final sound creation is only done in a third step. CH explains that for traditional sound design, he begins the creative process by starting from an image. He watches a video to determine the mood and starts to try out ideas and goes through this process until reaching a final result. CH feels that instead of starting with the most basic sounds, it is important to work with the psychological effects of the sound and "*looking at the big picture*" first and then focus in on smaller details. He believes that you need to understand how a sound is generated in order to replicate it and you need to combine the feeling and physics of the sound to make it successful. MP essentially works with sounds from banks, due to very strong time constraints which do not let him much time to purely explore sounds.

Other sound designers do not have precise procedure. For SC, the design process "*mainly depends on the time available to create a sound*". IB insists on the importance of "*kludge*" in sound research. ASi often "*reinvents the wheel*" and has "*no well-defined steps in [his] creations*". For instance, instead of using an existing beat, he prefers trying to understand how this beat was made. Then he reproduces it from scratch and, doing so, new ideas come to result in a well created module. His ideas come from logical and mathematical concept. AF even claims that "*there is not one, direct approach one must take. The beauty of sound design is that you can start however you want and use any and all tools at your disposal, whether they are physical artifacts one day such as bicycle wheels or high end tools the next*". He explains that "*the general approach to sound design is to approach the problem from many different angles to try to get a fix on what you are trying to achieve. There is no real first step in this process, you may already have the sound in a digital library or sound cloud, or you may have to understand what the sound would be and create it. A designer can take a technical approach focusing on the structural and procedural method or a more abstract approach focusing on the phenomenological or psychological effect of the sound*".

However, most designers agree on the fact that chance plays a role in creation. For IB, sound design is "*a domain of serendipity*", as "*the most interesting sounds are often created by accident*". SC and AF also agree on that as they believe in "*happy accidental creations*". ASi really feels that "*trial and error is the best and that while a lot has been written about the correct way to get good mapping, all you are doing is playing with stuff until you find what you think feels good*". SC and AF believes in "*happy accidental creations*".

Use of sketches

About the usefulness of having a sketching step, opinions are very different from a designer to another, depending how they consider its use. JB is the only designer who spontaneously speaks about sketches. For him, this is the second step of creation. Sketches are a tool he uses for himself during creation. They should represent *“one or several key aspects of the sound: its type, texture, evolution over time...”* Other sound designers do sketches for their self, to explore sounds.

About the use of sketches as a way to discuss preliminary ideas with the customer, most designers are very skeptical. For SC, *“for customers, sketches are always tricky, either because it's difficult for them to extrapolate and imagine the situation, or because they'll hear and focus on something that is unimportant. If we want sketches to work, you must push very far the sketch”*. AC highlights that *“only musicians can extrapolate a sketch to the final sound”*. XC reports that most of the time, the customer opinion is *“it works”* or *“it does not work”*. Therefore, the designer should present a sound as close to the final one as possible, to avoid being refused because of a wrong extrapolation of a non-definitive sketch.

3.3 Uses of vocalizations and gestures in sound design

3.3.1 Uses of vocalizations

Vocalizations are firstly used as raw material (AC, ASt, IB, JB, SC, XC). Recordings of voice could be used either in sound design for movies (character voices, impact noises), or in industrial cases. For instance, SC proposes turn signal sounds only made with voice (and then modify with some effects). AC also use recordings of his voice with other materials in the engine sound that he designed. Vocal material is also used as database for granular synthesis (JB, ASt).

Vocalizations are sometimes used during the exchanges with customers (AF, ASt, CH, MP). It could be more efficient than sitting and re-synthesizing something several times. Because of the lack of vocabulary of some customers, the vocalization is sometimes more efficient than describing sounds with words.

Vocalizations are used as sound parameters' control (ASi, JB, XC). They are used to trigger sounds, to control synthesizer parameter (with voice envelope and pitch), or as a MIDI controller to play melodies.

Sound designers mention some limits of vocalizations. They may not be so efficient when vocalizations are presented to non-musician (“non-expert”) listeners. Then for a personal use, vocalizations are useless for some sound designers when they already have the sound in their head. It could also be easy to recognize the vocal dimension of the sound if not well deforming (maybe because of formants). Imitations are not so common in some culture. The use of voice for choosing a sound, or as a mean of control would be more difficult for a Japanese person for instance, even though lots of onomatopoeias are used to describe sounds, such as “boko-boko” to describe “bubbles jumping in water”.

3.3.2 Uses of gestures

Having often had a musical practice, sound designers are familiar with gestures as sound control means. They use gestures in lot of fields and always as sound control inputs. AC already used the whole body movement, through the position of sensors on dancers' body or on his own body, to control sound generation. The movement may be used live, or recorded for a later use. ASi used the gestures of a flamenco dancer to control the volume of the sound. He used a security camera to catch the movements, when movements were abrupt, the volume was reduced to zero. He also used the movement of a robot (Robothespian) to control sound parameters in live. ASt made a project involving musical gloves used to control anything from lighting to synthesizers. These gloves have bend sensors in the fingers and orientation sensors on the wrist that record the movements of the hand and send the information wirelessly to a computer. The gloves use qualifiers (directions), movements (posture) and events (karate chop). JB uses gesture indirectly during his sketching step, for the drawing of sounds. During the lessons he has been giving in art school, he and his students were using gesture to communicate about sound characteristics: amplitude and frequency.

Once again, NK highlights the fact that if gestures are common in communication in some culture (south Europe for instance), they are not used a lot in the Japanese culture.

4. SOUND DESIGNERS' FEEDBACK ON PRELIMINARY TOOLS

4.1 SkAT-VG tools

Two tools have been presented and/or tested by the sound designers during the interviews and during the 48h of sound design (see next section). These tools are named **SkAT Studio** and **Mimes**. They are described below.

4.1.1 SkAT Studio

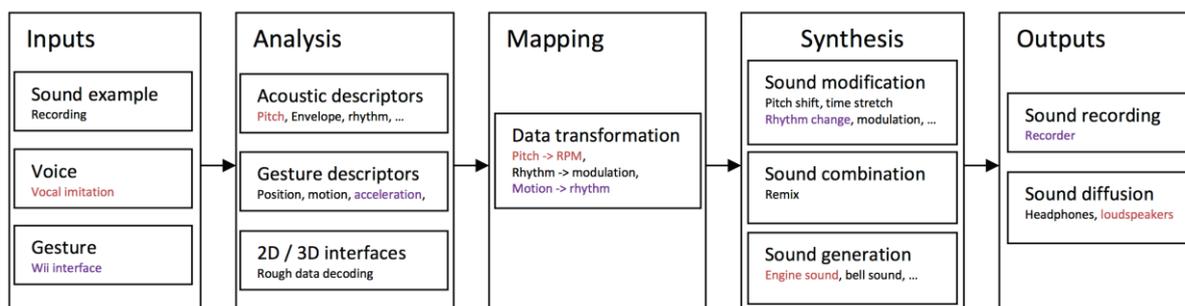


Figure 1: SkAT Studio 5-stages process. Two scenarios are highlighted in red and in purple. Red scenario: control of a engine sound using voice. Purple scenario: change of the rhythm of a sound using hand motion.

SkAT Studio is a Cycling 74 Max modular framework, conceived to host diverse technologies developed by the other (academic) partners of the project. The general workflow is composed of 5 stages, as described on Figure 1:

- Input: vocal and gestural signals acquisition.
- Analysis: audio and gestural feature extraction.
- Mapping: transformation of the extracted features in control values for sound production.
- Synthesis and processing: sound production.
- Output: sound rendering and recording.

In the main GUI, shown on Figure 2 it is possible to load a collection of modules, according to the general workflow. Each group, representing a stage of the workflow, can host multiple modules simultaneously (i.e., multiple feature extractors, control maps, and sound synthesizers). Signal and control data routing throughout the SkAT Studio framework can be programmed at will by means of patchbays. Modules (analysis, mapping, sound synthesis and processing) can be easily implemented as Max patches, according to a template that is provided with the software release.

A configuration of modules in groups defines a “control scenario”, that can be stored locally and recalled as a global preset. Predefined control scenarios are provided with the software release. They are based on synthesizers developed by the IUAV partner and derived from their SDT² (Sound Design Toolkit) toolbox. Each synthesizer is integrated as SkAT Studio module, and a mapping module has been carefully built to construct relevant control scenarios.

² <http://soundobject.org/SDT/>



Figure 2: SkAT Studio GUI. The scenario is the control of a blowing model using voice.

4.1.2 Mimes

Mimes is a tool developed by the IMTR team of IRCAM, which is also a partner of the SkAT-VG project. This tool allows to control a concatenative sound synthesis by using either vocalizations or gestures through the use of interactive objects (see Figure 3). Its use is described below.

The user first vocalizes a sound idea and thus produces implicitly some dynamical variations of sound features (for example roughness, pitch or energy). An automatic acoustical analysis of the vocalization (5) produces profiles of acoustical descriptors that are mapped to trajectories in different sound spaces of concatenate sounds also organized in 2D spaces of sound descriptors (6). The different sound spaces correspond to different aesthetic qualities (abstract sounds, concrete sounds of different materials, ...). The user could listen to the different sound proposals corresponding to each sound corpus and after selecting one or weighting the different corpora, the user associates a gesture with the sound proposal. The associate gesture is produced by manipulating an interactive object (7), embedded with sensors in order to analyze the different gestures, and the system maps dynamically the gesture with the sound proposal in order to play it dynamically.



Figure 3: Mimes control objects

4.2 "48h of sound design" in Château la Coste

The "48h of sound design" is a workshop organized by Genesis in Château la Coste, a vineyard, acquired in 2004 by Patrick McKillen to establish a center of both oenology and art.

In 48 hours, 5 sound designers (ASi, AC, MP, SC, XC) tackled one artwork and proposed their aural interpretation of the selected piece, by creating a sound representing it. The workshop was the occasion to introduce the sound design tools and methods being developed within the SkAT-VG project. Sound designers had the opportunity to experiment and incorporate them in their work.

The event started on April 28 and ended on May 1 2016 and had 3 main stages.

4.2.1 Introduction and training

SkAT-VG team introduced the overall workshop organization and SkAT-VG tools and methods to sound designers. The training session was about making games and concrete exercises on vocalizations, gestures, with and without SkAT-VG tools. At the end of the day, each sound designer selected an art piece for which they created a sound in the next 48h.

4.2.2 Sketching stage

On April 29 morning, each sound designers met un-official customers on the morning to discuss the sound concept of the selected art piece, and better define the details of their project by adding an art-expert point of view. They then produce sketches all day long, by using exclusively SkAT-VG tools (basic editing operations were allowed, e.g. tracks editing and alignment). By late afternoon, they discussed again with their customer about the sketches they produced in 24h, to guide the final sound composition.

4.2.3 Refining stage and public presentation

On April 30, sound designers refined and expanded their sketches, by creating and / or adding new sounds. They were allowed to use either SkAT-VG tools or their own tools. More than 100 visitors had the opportunity to listen to sound designers creation, on site in the Domain.

4.3 Sound designers' feedback on SkAT-VG tools

The feedbacks presented in this section were obtained during the interviews and at the end of the 48h of sound design event.

4.3.1 SkAT-VG tools' usefulness

During the 48h of sound design, the designers found the tools useful for exploring sound models capabilities, and for finding new ideas. For example, ASi said that even if he did not use sounds produced during the sketching step, "*[he] would not have had this idea without the tools*".

Another key point of the sketching tools is to create non predictable sounds. With this last idea, existing models could be interpolated/extrapolated to understand new gestures and create new sounds. In addition, the fluidity of vocalizations and gestures adds more expressivity and spontaneity to the sound controls. Sound designers also expected a tool that got natural and effective way to play with it.

According to sound designers, the sketching tools should not only be dedicated to create sound, but also to navigate through a sound bank. Several sound designers express their wish to have a tool that can search for specific sounds, from vocalizations they did.

4.3.2 SkAT-VG tools' usability

Even though all sound designers tend to use similar tools, it is really important to notice that each sound designer has a specific way to create sounds, with a specific workflow operating. To them (AC, AF, ASi, SC, JB), the most important point is that the sketching tools can be easily integrated in their workflow. As SC says, "*this is a point that makes you STILL use things*". One solution is to export well formatted files. Another solution could be to integrate new tools in existing tools, for example as VST plug-in (SC, ASi).

According to sound designers (AF, ASt, JB, SC, XC), the mapping is a key component of satisfactory interaction. This mapping should be well presetting and/or easy to create. ASt highlights that "*the controlled parameters are more or less adapted to the movement which operates. Thus the allocation of the pitch on the rotation of the wrist is a weak interaction while the mix dry/wet reverb works well*". It would be ideal for some sound designer to perform this mapping also with gesture and voice, to get away from too many numbers.

During the interview, SC thinks that from scratch it could be complicated to do something for non-trained users. Therefore, the tools should propose presets to create material that are directly usable. There should also be different levels of use/complexity of the tools: "*you can have main elementary parameters that you can easily and quickly modify, then you can go deeper by arranging more finely*

some modules hidden until now, and finally if you really want to go into details you can access to a scripting tool”.

4.3.3 Limits of such tools

The first limit cited by several sound designers (AC, AF, SC, XC) is the vocal limitation: breath (for making very long sounds), tessitura, ease to use the voice. For example, AC thinks that only musicians (even singers) could use well such tools. XC points out the fact that some sounds are very difficult to imitate: *“It is easy to imitate a car or a dog, but how would you imitate the sound of ice being drop in a glass of water, or the sound of falling rocks?”.*

MP and SC point that the user will have to learn how to use the tools. From their first experience, they found the tools difficult to control: the results are very reproducible, but it is complicated to understand which modification of the sound will be produced by which modification of the input voice.

There is also a region of control and constraints, for example, sound designers do not want to use their mouth to create a guitar sound; a guitar is the best tool for that.

Finally, for AF, there is a “danger” about any numeric tool used by a designer: *“the tools can be a distraction. If the tools are too powerful, they will not allow your internal working model to have persistence. It is easy to find a sound on a synthesizer that may impress you, but is not even in the ballpark of what you meant to create, and by choosing this you are allowing the tools to drive the creative process.”*

5. CONCLUSION

In the scope of the SkAT-VG project, Genesis conducted a set of interviews of sound designers in order to understand the creation process, with a particular focus on the sketch step, and on the use of vocalization and gestures. In this article, we presented the summary and main conclusions.

For most interviewed sound designers, there is no precise creation process, it mainly depends on the project. Sketches are rarely used, only for a in-process use. Sketches are never used for communication with the customer. Vocalizations and gestures have been used by some designers: voice as raw material; voice and gesture as a control signal for sound modification.

Based on preliminary conclusions of the interviews, tools have been developed by the project's partners, for a quick production of sketches using vocalizations and gestures. A pool of sound designers accepted to test these tools during a 2-days artistic event called “48h of sound design”. The main conclusions of this event were also reported in this article.

Based on the conclusions of the most recent interviews, and on the sound designers' feedbacks from the 48h of sound design, improvements will be made on the existing tools.

ACKNOWLEDGEMENTS

The work described in this paper is part of the project SkAT-VG, which received the financial support of the Future and Emerging Technologies (FET) program within the Seventh Framework Program for Research of the European Commission under FET-Open grant number: 618067. The authors thank all the sound designers who participated the project, for the interviews and/or for the 48h of sound design event. The authors thank the sound design team from IUAV (Venice University, Italy), the Sound design and Perception team from IRCAM (Paris, France) and the Real-Time Musical Interactions from IRCAM for their tools development and their involvement in the organization of the 48h sound design event.

REFERENCES

1. Nykänen, J. Wingstedt, J. Sundhage and P. Mohlin, «Sketching sounds – Kinds of listening and their functions in designing,» *Design Studies*, vol. 39, pp. 19-47, 2015.
2. M. Carron, F. Dubois, N. Misdariis, C. Talotte and P. Susini, «Designing Sound Identity: Providing New Communication Tools for Building Brands,» in *Audio Mostly*, 2014.
3. 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. 2, n. 1135, 2014.
4. D. Rocchesso, G. Lemaitre, P. Susini, S. Ternström and P. Boussard, «Sketching Sound with Voice and Gesture,» *ACM Transactions*, 2015.
5. G. Peeter, B. L. Giordano, P. Susini, N. Misdariis and S. McAdams, «The Timbre ToolBox: Extracting audio descriptors from musical signals,» *The Journal of the Acoustical Society of America*, vol. 130, n.

- 15, pp. 2902-2916, 2011.
6. N. Schnell, A. Röbel, D. Schwarz, G. Peeters and R. Borghesi, «MuBu and Friends - Assembling Tools for Content Based Real-Time Interactive Audio Processing in Max/MSP,» in *Proceedings of the International Computer Music Conference*, 2009.
 7. O. Houix, F. Bevilacqua, N. Misdariis, P. Susini, E. Flety, J. Françoise and J. Groboz, «Objects with Multiple Sonic Affordances to Explore Gestural Interactions,» in *xCoAx conference, Computation, Communication, Aesthetics & X*, 2015.