How do you reproduce a synthetic sound of an ellipse in water ? A phenomenological investigation of pre-reflexive contents of consciousness.

Jean Vion-Dury, Marie Degrandi, Gaëlle Mougin, Thomas Bordonné, Sølvi Ystad, Richard Kronland-Martinet, Mitsuko Aramaki

How do you reproduce a synthetic sound of an ellipse in water ? A phenomenological investigation of pre-reflexive contents of consciousness.

Vion-Dury J., Degrandi M., Mougin G., Bordonné T., Ystad S., Kronland-Martinet R., and Aramaki M.

Aix Marseille Univ , CNRS, PRISM , Marseille, France

jvion-dury@ap-hm.fr

Abstract. This article describes a listening experiment based on elicitation interviews that aims at describing the conscious experience of a subject submitted to a perceptual stimulation. As opposed to traditional listening experiments in which subjects are generally influenced by closed or suggestive questions and limited to predefined, forced choices, elicitation interviews make it possible to get a deeper insight into the listener's perception, in particular to the pre-reflexive content of the conscious experiences. Inspired by previous elicitation interviews during which subjects passively listened to sounds, this experience is based on an active task during which the subjects were asked to reproduce a sound with a stylus on a graphic tablet that controlled a synthesis model. The reproduction was followed by an elicitation interview. The trace of the graphic gesture as well as the answers recorded during the interview were then analyzed. Results revealed that the subjects varied their focus towards both the evoked sound source, and intrinsic sound properties and also described their sensations induced by the experience.

Keywords: phenomenology, elicitation interview, auditory perception, sound synthesis, graphic gestures, audio-motor loop.

1 Introduction

When preforming perceptual evaluations of sounds, it is important to be aware of the fact that listeners may focus on different aspects. Gaver [6] distinguished everyday listening from analytical listening. In the case of everyday listening of a simple source, the listener pays attention to the sound producing object, such as its size [12] and the material of which it is composed [[1 ,9], In the case of more complex situations reflecting for instance interactions between sound sources, the listener perceives properties related to the event as a whole. Warren and Verbrugge [31] showed that objects that bounce and break can be distinguished by listeners with a high degree of accuracy, while Repp [19] revealed that subjects were able to recognize their own recorded clapping and the hand position from recordings when someone else is clapping. More recently, Thoret et al. [26, 27] showed that subjects were able to recognize biological motions and certain shapes from friction sounds produced when a person is drawing on a paper.
To favor analytical listening where the listeners focus on intrinsic sound properties linked, for instance, to loudness, pitch, and timbre other approaches have been used. Merer [15] used acousmatic sounds for which the source could not be easily recognized to reveal sound structures responsible for the evocation of movement categories. Other approaches such as sensory analysis during which a group of subjects identify sensory descriptors such as onomatopoeias have been used, for instance to characterize the formantic transition from "ON" '' (pronounced [ɔ̃]) to "AN" (pronounced [ɑ̃]) that characterizes sounds from car engines [20,23].
Other approaches, such as vocal imitations, that do not specifically focus on everyday or analytical listening have been used to extract relevant features of kitchen sounds [13], and more recently to reveal invariant structures responsible for the evocation of movements and materials [3, 4]. Psycholinguistic analyses have been used to characterize sounds from musical instruments through spontaneous verbalizations. One such study that investigated violinists' preference judgements during a playing task, led to a model that linked auditory and haptic sensations to the timbre, quality, and playability of the instrument [20,21]. Sound perception is a conscious experience that can be described not only in so-called "third person" protocols (from the point of view of the experimenter within a given paradigm, e.g. a psycho-physical paradigm), but also by protocols aiming at describing the experience from the subjects' perspective (subjective methods) mainly based on the Husserlian phenomenology. Most of the time, spontaneous descriptions of experiences and cognitive processes are poor [18] because the experience does not guarantee immediate access to its background contents [28]. Several kinds of information usually remain undisclosed, masked or "pre-reflexive" as they are called in phenomenological language [18]. Various methods allow to accurately describe the conscious experience in its reflexive and mostly pre-reflexive part. Among them, the elicitation interview (EI) [14, 29] is a disciplined introspection method conceptually based both on neurolinguistic programming (NLP) and Husserlian phenomenology

[10]. EI makes it possible to return to the non-reflexive part of the conscious experience of a subject, hereby limiting influences from closed or suggestive questions.
Whereas the qualitative research methods used in sociology, such as Glaser and Strauss' anchored theory (see [20]) or the "repertory grid" method use textual corpora of reflexive descriptions of experiences to extract emerging themes and their variations, EI is essentially interested in the non-reflexive component of the experience. For this reason, whereas in the qualitative methods, the subjects use their autobiographical memory, in the EI, the subjects must relive their experience and activate their "integral memory", in particular corporeal.
We previously described pre-reflexive conscious experiences in passive listening of sounds [16]. In the current work we analyze pre-reflexive content of conscious experiences in an active task consisting in reproducing a sound by drawing on a graphic tablet.

2 Material and Methods

In this section, we describe the interactive device used by the participants, the experimental protocol and the elicitation interview.

2.1 Equipment: The "tablet-Synthesizer" Device.

Sound synthesis is a powerful tool to create any kind of sounds that either imitate real or virtual situations. Current synthesis models enable high quality re-synthesis of natural sounds that can be generated in real-time. One challenging aspect linked to sound synthesis is the control of the synthesis parameters that is not always intuitive. To meet this challenging control issue, we have developed a synthesizer based on perceptual features linked to the evocation of actions and objects [1, 2]. This device is based on the ecological approach to perception proposed by Gibson [8] which considers that actions and objects are recognized through invariant structures. The sound synthesizer makes it possible to create sounds from verbal labels that describe the action (e.g. hitting, scraping, rolling) and the object (e.g. material, size, shape) associated with the sound. Any combination between actions and objects can hereby be simulated, such as scratching a small metallic bell or hitting a big wooden bar [4]. Unrealistic situations can also be simulated this way, such as rubbing the wind or scratching a wave.

In the present study we decided to use a sound texture that evoked a movement in water, since the timbre of liquid sounds vary strongly with the dynamic action. To create the reference sound that the subjects were asked to reproduce, the synthesized sound was combined with an elliptic movement recorded by the experimenter who drew on a WACOM INTUOS PRO graphic tablet. The experimenter freely chose the eccentricity and the orientation of the ellipse that he/she was asked to draw ten times. To induce a periodic movement, we used a 60 bpm metronome while the experimenter was drawing to help him/her maintain a regular speed. Among the ten repetitions, the three most regular ellipses were selected. The position of the stylus was recorded by a Max/MSP interface at a sampling rate of 129 Hz. We then derived the position to get the velocity profile. The scalar product that quantifies the difference between the two profiles (the reference profile and the profile performed by the subject) was calculated. If the two profiles are completely different, the scalar product will have a low value, whereas if they are identical, it will have a maximum value (around 1).

2.2 Experimental Protocol

The subjects were first asked to listen to the reference sound which nature and origin they ignored. They were then asked to reproduce this reference sound on the WACOM INTUOS PRO graphic tablet with the gesture that best imitated the reference sound. The subjects produced the sound in real time while they performed the gesture on the graphic tablet.

Participants Ten subjects, 7 women and 3 men (aged from 26 to 70 years) were included in this experiment. Five subjects were experienced musicians practicing an instrument on a regular basis and the the remaining 5 participants were not musicians. The ten subjects were right handed. Subjects did not have any hearing or neurological problems, such as memory-related problems or attention difficulties. The interview was conducted by one of the three medical doctors involved in the study: MD, GM, JVD. An

audiogram was performed for each subject before the beginning of the experiment to make sure that none of the subjects had hearing impairments.

The Elicitation Interview In a second step (just after the reproduction of the sound), the subjects were asked to review their experience while listening to and reproducing the sound by means of an elicitation interview, by answering the question "how did you perceive and reproduce the sound?". The EI was conducted by three experienced researchers in phenomenology and EI.

The EI requires a certain number of methodological specificities:

a) The first key of the interview is to lead the subjects to describe their experience, that is to tell what they experienced and not what they thought, believed or imagined their experience had been like [18].

b) The interviewer should lead the subjects to discuss their past experiences by helping them to find the sensory and emotional dimensions.

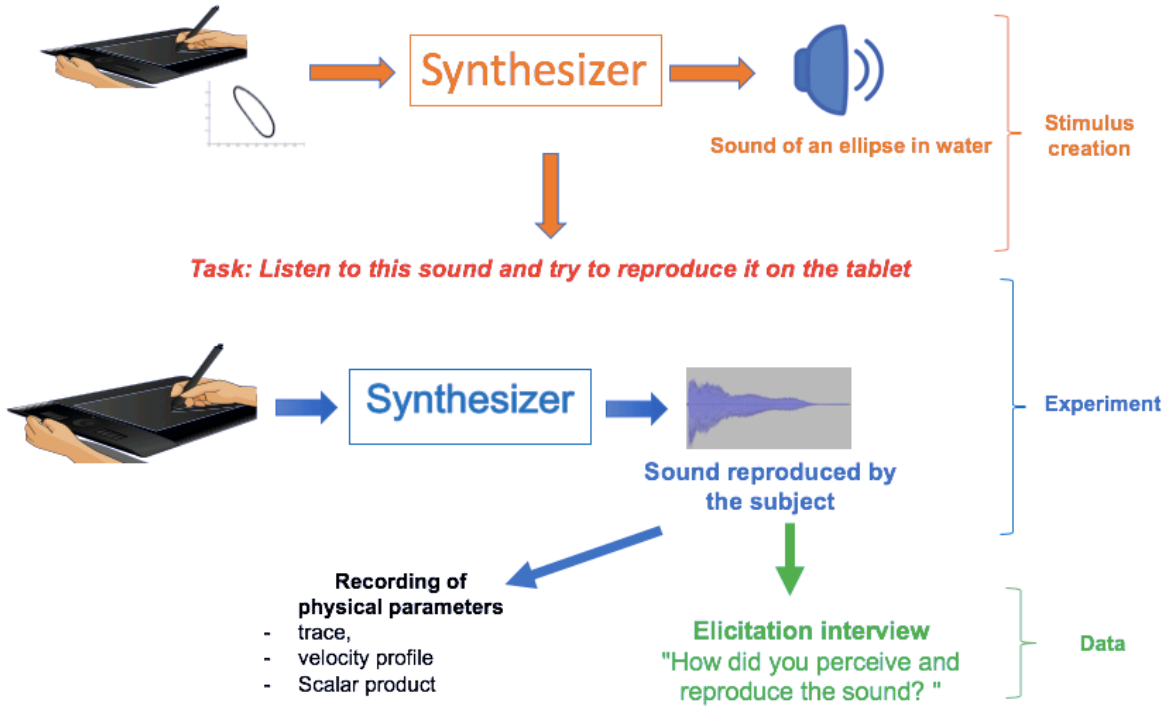


Fig. 1. Experimental protocol

c) The interview consists in helping the subjects redirect their attention from the content of their experience (the "what"), to its diachronic and synchronic structure oriented towards the experiential (non-causal) "how". The diachronic structure of the experience corresponds to the stages of its deployment over time. The synchronic structure of the experiment corresponds to the configuration at a given moment of the sensory registers used, the type of mobilized attention... etc. The aim is to make the subjects relive their experience rather than to remember it.

d) To collect such a description, the interviewer's questions should be "empty of content", non-inductive and "point" to the structure of the experiment without providing any content. Questions are, for example: "From what did you start? What did you feel ? How did it appear to you?", etc. This mode of questioning emphasizes the "how" of the conscious experience and excludes the "why".

e) The structure of an interview is iterative while guiding the attention of the subject towards a diachronic or synchronic mesh which progressively becomes more detailed each time. The average duration of an

interview is about an hour to describe a few seconds of experience (as Stern puts it, “there is a world in a grain of sand” [25]). The interviewer must remain totally neutral. A good harmonization of affects (motor and prosodic affective tuning [25]) is a critical condition for the quality of the interview.

Data collection and analysis. All the EIs were recorded, with the subjects’ agreement. The physical data (pen movement, speed, pressure etc ...) were collected from the computer connected to the graphic tablet. The records of EIs were entirely transcribed. The analysis of verbatim was carried out to extract the descriptive categories (saliencies) from each interview. The choice of descriptive categories for each interview was validated by 7 people in an inter-judge session.

Data collection and analysis. All the EIs were recorded, with the subjects’ agreement. The physical data (pen movement, speed, pressure etc ...) were collected from the computer connected to the graphic tablet. The records of EIs were entirely transcribed. The analysis of verbatim was carried out to extract the descriptive categories (saliencies) from each interview. The choice of descriptive categories for each interview was validated by 7 people in an inter-judge session.

3 Results

The physical data from the tablet were analysed together with the EIs. Only the data from the EIs, as well as the drawings recorded on the tablet are presented in Tables 1 and 2.

Types of sound listening.

The EI enabled to collect the synchronic and diachronic structure from the listening experience of each subject. These data respond to both the ”what” of their experience but also to ”how”, to the proper way of perceiving and reproducing this sound. They give a fine and precise description of an experience that lasted for a few seconds by allowing an awareness of the different processes. Each of the 3 types of listening can be analyzed from a) the main sensory modalities used, b) the attentional disposition of the subject, c) the position of the subject with respect to the sound and d) the moment this type of listening occurs. Each subject has a preferred type of listening (in this experiment), but this does not mean that he or she does not use other types of listening in a less marked way. This part of the analysis is presented in Table 1.

Three descriptive categories (attractors) which are common to all the subjects can be identified concerning sound listening. These categories are related to the way the subjects hear a sound while they prepare its reproduction. The first three descriptive categories which correspond to three types of sound listening are in line with categories identified in our previous work [17]. They are: 1) the direction of listening, 2) the sensory listening modalities, 3) the attentional disposition.

The first type of listening is turned to the source of the sound and involves attention directed to the origin of the sound with an active search for familiar scenes associated with the source. In this type of causal listening the imagination is very active. The subject is thus projected into an imaginary scene evoked by the sound heard which is integrated into the scene, and a given context in the visual modality. This listening structure appears spontaneously and early in the diachronic description of the experience. This type of listening, characterized as everyday listening by Gaver [6], represents the main listening mode of three subjects but is, for 8 out of 10 subjects, associated with the other types of listening.

The second type of listening, characterized as analytic listening by Gaver, is directed to the characteristics of the sound. This way of perceiving sounds appears when subjects focus on the reproduction task. This time the sound is brought back to its different components (rhythm, pitch, timbre, intensity), and the subjects focus on the sound itself and not on the causality. This is the main listening type for four subjects, but 8 out of 10 subjects used it in the experiment.

The third level of listening is a particular listening modality that is usually not spontaneously described in our daily lives and rather evoked at the end of the diachronic description of the listening experience. It focuses on the effect of the sound, specifically the dynamics, the movement it induces relative to the whole body. It is an ”internal” or “embodied” listening modality in which the boundaries between the sound and

the corporal space become porous. Subjects adopt a more passive position related to the sound, in a way they are" impregnated" by the sound. This is the main listening modality for three out of ten subjects, but 8 out of 10 subjects used it in the experiment. Finally, we did not find any difference between musicians and non-musicians with respect to the type of listening.

Table 1. Three types of listening experiences

Types of listening experiences						
Listening focus (LF)	Main LF in subjects	Number of subjects presenting this LF	Main sensory modalities involved	Attentional disposition	Sound/ auditor position	Moment of appearance
Origin of sound	2, 4 and 5	8/10	Scenes (sea, beach…) perceived by the auditory and visual modalities	Directed attention towards the source. Active search remembrance, familiar scenes evoking the source.	Location of the subject in relation to the scene.	Appears spontaneously first while listening
Acoustic characteristics of sound	1, 7, 8 and 10	8/10	Timbre, intensity, rhythm, height perceived by the auditory modality but may be associated with other modalities (rhythm with kinesthetic sensitivity)	Attention directed towards the different parts of the sound Active position of the subject in relation to the sound.	Accurate location of the sound, external to the subject....	Appears when subjects focus on the task of reproduction
Effect of the sound	3, 6 and 9	8/10	Dynamics of the sound perceived by kinesthetic sensitivity, mainly	Attention less focused, more global Position of the subject rather passive compared to the sound	Blurred boundary between body space and sound Effect of sound throughout the body.	Particular listening modality, generally not described spontaneously rather evoked at the end of EI.

Reproductive strategies.

The drawings made by the subjects can be classified into 5 shapes: ellipses (n = 2), lemniscates (n = 1), sinusoids (n = 2), lines (back and forth) (n = 3), complex shapes (n = 2) (Table 2).
The velocity profiles of the movements used to draw the different shapes are perceived as dynamic variations in the sounds and serve as auditory cues to recognize different shapes [27].
The velocity profiles are broadly similar, regardless of the drawn shapes performed by the subjects on the tablet. These velocity profiles, like the drawings on the tablet, show that the subjects are actually trying to reproduce an alternate motion. The scalar products comparing the reference velocity profile with the velocity profile produced by each subject are rather high (> 0.87) reflecting a correct reproduction of the reference velocity profile.

Several unexpected results are observed in the reproduction task which can be analysed according to three criteria: a) how the subject conceptually represents the sound, i.e. what the sound looks like (Table 3, column 3), b) the imaginary visual content (an imaginary scene) associated with the sound (Table 3, column 4), and c) how the subject actually drew the sound on the tablet (Table 3, column 6).

The representation of the sound contents to reproduce the sound heard depends on each subject's main listening mode (Table 3, comparison of columns 2 and. 3). For all the subjects with predominant listening based on the origin of the sound (i.e. everyday listening), the sound heard resembles more or less typical waves. For subjects presenting predominant listening based on the acoustic characteristics of the sound (i.e. analytical listening) the reproduction strategy (4 subjects out of 4) is based on the physical dynamics of the sound (rhythm, oscillations, intensity). For subjects whose predominant listening is based on the sound effect, the reproduction strategy is based more on the feeling of oscillations and pulsations (2 subjects out of 3).

The visual scenes associated with the task consisting of reproducing movements in a liquid were summarized in column 4 of Table 3. These visual scenes are consistent with the subjects' representation of sounds (column 5) in 7 subjects. The coherence between the produced shape and the imaginary scene is

good for half of the subjects (column 7). The coherence between the produced shape and the sound representation is good for 6 subjects, (column 8), who are mainly the same as those with a good coherency in the other representation.

Table 2. Recorded traces, velocity profiles and scalar product related to the reproduction task of each subject.

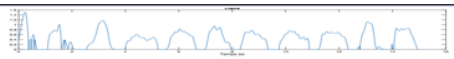
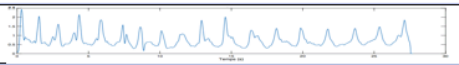
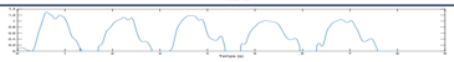
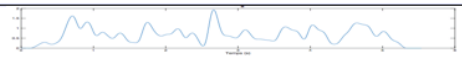
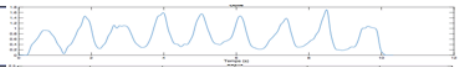
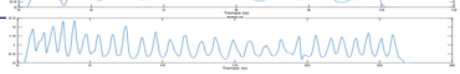
Subject n°	Recorded trace	Velocity profiles	Scalar product
2			0,89
4			0,93
5			0,89
1			0,87
7			0,90
8			0,95
10			0,92
3			0,98
6			0,99
9			0,91

Table 3. Results of the elicitation interviews related to type of sound listening and reproduction task. The coherency between the representation of the sound and the imaginary content or the reproduction gesture is given by following symbols: (+) = good coherence, (+/-) = medium coherence, (-) = poor coherence.

1	2	3	4	5	6	7	8
Subject n°	**Listening focus**	**The sound is like a… (= representation) (Sound listening)**	**Associated imaginary visual content (Sound reproduction)**	**3 vs 4**	**Recorded trace (movement)**	**6 vs 4**	**6 vs 3**
2	Origin of sound	Wave	Scene of waves. I'm riding the waves.	+		+	+
4		Waves and bubbles	Scene of close waves with bubbles bursting, like scum.	+		-	-
5		Waves and bubbles	Scene of wave with splashing droplets. Slow "glou glou"	+		-	-
1	Acoustic parameters of sound	Something perfectly rounded with rythm	Ellipse on a table, or a cardiac motion, "wroum wroum"	+		+	+
7		Dynamics, a sinusoid, a wave	I wanted to make sinusoid shapes	+		-	-
8		A growing flow	It is like a refrain.	-		-	-
10		Waves	Scene of waves, with"vrrr" and "boum"	+		+	+
3	Effect of the sound	Oscillation, oval shape (hourglass)	I am washing a sweater by hand. I sing the rhythm, I dance	+/-		+	+
6		Dynamics. A zigzag shape, an ellipse	I danse on the sound. I'm rowing on the water.	+/-		-	+
9		Pulsation, a spindle	I wanted to do a spindle, or a wave, with internal movements	+		+	+

4 Discussion and conclusion

The phenomenological analysis of the pre-reflexive contents of the consciousness in a reproduction task of a sound using a sound-based graphic tablet makes it possible to confirm the main types of listening previously described by Gaver [5] or Petitmengin et al. [17]. The fact of having a reproduction task to be accomplished modifies, with respect to an isolated passive listening, the diachronic and synchronic content of this experience (the moment of appearance of the experiential content, in particular).

In this preliminary work involving a small population of subjects, we did not find differences in listening and sound reproduction based on age, gender, or musical experience. It would be interesting to increase the number of subjects to assess whether differences appear according to these factors. However, we can not perform EIs on large populations because of the considerable time required for data processing. We (GM, JVD) are currently testing faster and more efficient data processing methods to increase the number of subjects involved in this type of study.

When comparing our current and previous studies [17], several differences must be reported. The initial study focused on describing the modalities of listening to the sound, as such, and without any task required at the end of the listening. The study aimed to highlight the descriptive categories of the non-reflexive part of the sound listening experiences and to define the general structure of such an experience. For this reason, various sounds were used (sounds from nature, sounds from everyday life, abstract sounds). Some individual differences linked to the way subjects listened to sounds were observed, but the constitution of subgroups of subjects did not appear. In our current study, only one sound is proposed with an associated reproduction task. If the same types of non-reflexive experiences can be observed, the task to be done changes the type of intentionality [8] and attentional focus.

Perceptual modalities also change between listening and reproducing tasks. During a listening task, it is the auditory system that is mainly solicited, whereas in a reproduction task, perception is multimodal with a solicitation of auditory, visual and proprioceptive activities in addition to motor skills.
In this study we did not focus on the way in which the heard sound is transformed (or not) in the consciousness at the moment of initiating the drawing on the tablet, just after hearing the sound. This probably depends on the type of listening that the subject adopts and probably on the time it takes before he/she starts moving the pen on the tablet. In phenomenological terms, this amounts to asking the question of how retention and protention [11] are respectively organized in this pivotal moment (this thick present) between listening and reproduction.
This study also made it possible to highlight the fact that even if each subject possesses a preferential type (focus) of listening, other types of listening are also mobilized to find the resources for carrying out the reproduction task. In reality, perceptual processes are rarely unimodal. The so-called primary receptor areas of the cerebral cortex that were thought to be specific to a sensory modality are in fact multimodal [7]. Multimodality in perception seems to be the rule and the degree of synaesthesia varies from one individual to another [24]. This, for example, could explain why we find an entanglement of different types of listening in all the subject responses, along with the presence of a preferential type. This entanglement of available perceptual dispositions reveals the complexity of the processes involved in the reproduction task and opens a new field of research on the co-presence of pre-reflexive perceptual processes.

Taking into account only the traces obtained, transcribed by the velocity profiles and the scalar products (Table 2), we can conclude from this study that the subjects reproduce the dynamic structure of an alternating motion correctly, i.e. they make a sketch or a summary of the transformational invariant underlying the evoked action. However, the two parameters (velocity profile, and scalar product) do not enable to deduce the shape drawn during the reproduction and from the drawings it is clear that the sounds do not contain the spatial cues that would lead all the subjects to draw the same figure. The phenomenological approach can give the beginning of an answer to the imagination that the sound produced among the listeners.

When the subjects are asked how they represent the sound, i.e. what it looks like (table 3, column 3), they call upon characteristics which are also related to the dynamics of a movement, according to three categories: a) oscillation: waves, oscillations, pulsations in 6 subjects, b) rhythm in 2 subjects and c) dynamics (without precision of nature) in 2 subjects. Probably this "representation" of sound participates in sound semiosis, that is to say in the sense that the subject gives to the sound. However, the variety of semiotic processes involved should be noted.

The way in which the subjects represent the sound is globally in agreement with their predominant type of listening (column 2). Subjects with a preferred listening type related to the origin of the sound (alternating movement in a liquid) hear waves and therefore focus on the structural and/or transformational invariant. Subjects who present a preferential listening oriented towards acoustic characteristics propose representations more related to signal structures and focus on the intrinsic sound properties. Subjects whose preferential listening mode is the effect of sound do not represent the sound but describe what they feel more than what they represent themselves: in this sense they have a more phenomenological approach than the other two types of subjects.

When the subjects are asked about the imaginary content (visual scenes) associated with sound reproduction (column 4), we observe a good correlation (in 7 subjects) between the representation of sound and this imaginary content (column 5). However, the imaginary content adds additional information about how the subjects were involved in the task. For example, subject 3 saw himself washing a sweater by hand. Of the 7 subjects for whom a good coherence between the sound representation and the imaginary visual scene was observed, only 5 of them generated a shape that was coherent with the representation and the imagination (comparison of columns 5 and 6).

All these results can be summarized in a simplified and hypothetic model of the processes involved (Figure 2). It turns out that the audio-motor loop leads to the reproduction of sound dynamics, i.e. the reproduction of alternating movement, while the associated shape, which cannot be predicted from the sound itself, depends on the subjects and their representation of the sound. This shape is modulated by the subject's preferred type of listening. The associated visual imaginary content, which is richer than the representation, might modulate the generation of the shape, giving it a kind of imaginary context. The type of preferential listening is also likely to modulate the imaginary content.

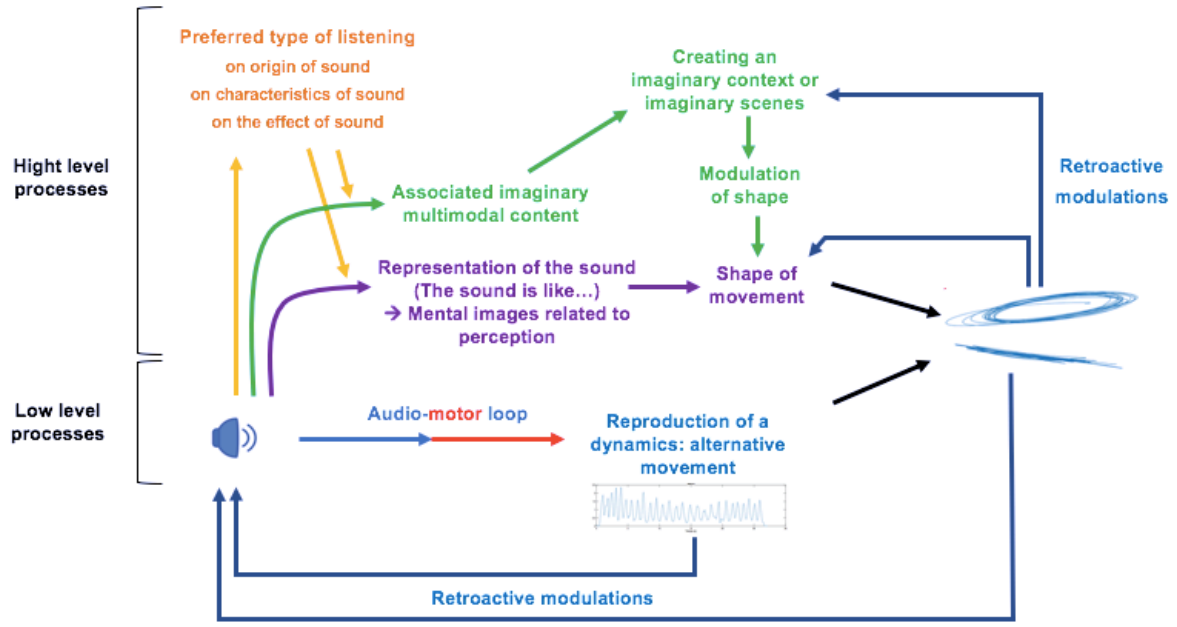


Fig. 2. Hypothetic model of the reproduction task.

One of the most important results of this study is the ability of EIs to highlight the role of imagination in the reproductive process. However, it is important to differentiate between mental imagery associated with perception and imagination. This mental imagery is either unimodal or multimodal. Imagination is that mental activity associated or not with a perception or an action. It is the imagination (fantasia) that we find liberated in dreams or hypnosis.
Thus, at the level of the representation of the sound, the subjects, describe for example sinusoids of which they might either have an image in the visual modality, or an amodal thought in the form of a concept. Recent studies show that this type of multimodal mental imagery is very often associated with perception

and is either conscious or unconscious (pre-reflective in phenomenology) [16]. It seems that this mental imagery is correlated to the activation of numerous cerebral areas, especially visual areas [32].
On the other hand, when subjects imagine how they are going to reproduce the sound, they often insert the shapes to be reproduced in a rich and systematically multimodal context. Imagination creates scenes (with motor, kinesthetic, and sometimes olfactory components) in which the subject does an action more or less related to the task, but which is not the task. This imagination creates a context, a scene or a story (washing a sweater, playing on the string of a harpsichord) related to the story and the subject's habits. It is very difficult, if not impossible, to study such rich and evanescent imaginary processes with the current means of neuroscience.

Imagination is undoubtedly the blind spot of cognitive and neurophysiological approaches. Its richness and complexity defy any reduction to simple activations or to logico-semantic or computational processes. Only introspective approaches that leave room for the subject's mental landscape, such as the Experiential Phenomenological Interview [30] (or microphenomenology) will be able to account for the richness of this imagination and the multiple backgrounds (historical, cultural, familial) from which it proceeds. In this work, the emphasis placed on the imaginary seems to open new paths in the understanding of what sound does to us in acoustics and semiotics.

At the end, this exploratory and multidisciplinary work provide an early proof of concept of the use of introspective methods in acoustics and audition in order to refine synthesis models and sound control towards an approach more and more turned towards the human experience.

Acknowledgements: We would like to thank Dr Dias-Alvez for providing us with the ellipse sound in water that made the experiment possible.

References

1. Aramaki, M., Besson, M., Kronland-Martinet, R., Ystad, S.: Controlling the perceived material in an impact sound synthesizer. IEEE Transactions on Audio, Speech, and Language Processing 19(2), 301–314 (2011)
2. Aramaki, M., Gondre, C., Kronland-Martinet, R., Voinier, T., Ystad, S.: Imagine the sounds : an intuitive control of an impact sound synthesizer. In: Ystad, Aramaki, Kronland-Martinet, Jensen (eds.) Auditory Display, Lecture Notes in Com- puter Science, vol. 5954, pp. 408–421. Springer-Verlag Berlin Heidelberg (2010)
3. Bordonné, T., Dias-Alves, M., Aramaki, M., Ystad, S., Kronland-Martinet, R.: Assessing sound perception through vocal imitations of sounds that evoke movements and materials. In: Aramaki, Davies, Kronland-Martinet, Ystad (eds.) Music Technology with Swing, Lecture Notes in Computer Science, vol. 11265, pp. 402–412. Springer Nature Switzerland (2018)
4. Bordonné, T., Kronland-Martinet, R., Ystad, S., Derrien, O., Aramaki, M., Exploring sound perception through vocal imitations, Journal of the Acoustical Society of America, 147(5), May 2020:
5. Conan, S., Thoret, E., Aramaki, M., Derrien, O., Gondre, C., Ystad, S., Kronland- Martinet, R.: An intuitive synthesizer of continuous-interaction sounds: Rubbing, scratching, and rolling. Computer Music Journal 38(4), 24–37 (2014)
6. Gaver, W.W.: What in the world do we hear? an ecological approach to auditory event perception. Ecological Psychology 5(1), 1–29 (1993)
7. Ghazanfar A.A. and Schoeder C.E. Is the cortex essentially multisensory? Trends in cognitive sciences, 10 (6) , 278-285 (2006)
8. Gibson, J.J.: The ecological approach to visual perception: classic edition. Psychology Press (2014)
9. Giordano,B.L.McAdams,S.:Materialidentificationofrealimpactsounds:Effects of size variation in steel, wood, and plexiglass plates. Journal of the Acoustical Society of America 119(2), 1171–1181 (2006)
10. Husserl, E.: Idées directrices pour une phénoménologie. Gallimard (1913 / 1985)
11. Husserl, E, Leçons pour une phénoménologie de la conscience intime du temps, Presses Universitaires de France (1905 / 1964)
12. Lakatos, S., McAdams, S., Chaigne, A.: The representation of auditory source characteristics: simple geometric form. Perception and Psychophysics 59, 1180– 1190 (1997)
13. Lemaitre, G., Dessein, A., Susini, P., Aura, K.: Vocal imitations and the identification of sound events. Ecological Psychology 4(23), 267–307 (2011)
14. Maurel, M.: The explicitation interview: example and applications. Journal of Consciousness Studies 16, 20–57 (2009)
15. Merer, A., Aramaki, M., Ystad, S., Kronland-Martinet, R.: Perceptual characterization of motion evoked by sounds for synthesis control purposes. ACM Trans. Appl. Percept. 10(1), 1–24 (2013)
16. Nanay B. Multimodal mental imagery. Cortex 105, 125-134 (2018)
17. Petitmengin, C., Bitbol, M., Nissou, J., Pachoud, B., Curallucci, H., Cermolaccce, M., Vion-Dury, J.: Listening from within. Journal of Consciousness Studies 16, 252–284 (2009)

18. Petitmengin, C., Bitbol, M., Ollagnier-Beldame, M.: Vers une science de l'expérience vécue. Intellectica - Rev Assoc Pour Rech Sur Sci Cogn ARCo. 64, 53–76 (20015)
19. Repp, B.H.: The sound of two hands clapping: An exploratory study. The Journal of the Acoustical Society of America 81(4), 1100–1109 (1987)
20. Roussarie, V., Richard, F., Bezat, M.C.: Validation of auditory attributes using analysis synthesis method. In: Congr´es Francais d'Acoustique/DAGA. Strasbourg (2004)
21. Saitis, C., Fritz, C., Scavone, G., Guastavino, C., Dubois, D.: A psycholinguistic analysis of preference verbal descriptors by experienced musicians. Journal of the Acoustical Society of America 141(4), 2746–2757 (2017)
22. Saitis, C., Giordano, B., Fritz, C., Scavone, G.: Perceptual evaluation of violins. a quantitative analysis of preference judgments by experienced players. Journal of the Acoustical Society of America 132(6), 4002–4012 (2012)
23. Sciabica, J., Olivero, A., Roussarie, V., Ystad, S., Kronland-Martinet, R.: Dissimilarity test modelling by time-frequency representation applied to engine sound. In: Audio Engineering Society Conference: 45th International Conference: Applications of Time- Frequency Processing in Audio (2012)
24. Spence C. Crossmodal correspondences: a tutorial review. Attention and perception psychophysics, 73, 971-995 (2011)
25. Stern, D.: The Present Moment in Psychotherapy and Everyday Life. W. W. Nor- ton & Company (2010)
26. Thoret, E., Aramaki, M., Bringoux, L., Ystad, S., Kronland-Martinet, R.: Seeing circles and drawing ellipses: when sound biases reproduction of visual motion. PLoS ONE 11(4) (2016)
27. Thoret, E., Aramaki, M., Kronland-Martinet, R., Velay, J.L., Ystad, S.: From sound to shape: auditory perception of drawing movements. Journal of Experimental Psychology: Human Perception and Performance 40(3), 983 (2014)
28. Vermersch, P.: Conscience directe et conscience réflchie. Intellectica 31, 269–311 (2000)
29. Vermersch, P.: Describing the practice of introsopection. Journal of Consciousness Studies 16, 20–57 (2009)
30. Vion-Dury, J., Mougin, G. : L'exploration de l'expérience consciente : Archéologie d'une démarche de recherche. Vers l'entretien phénoménologique expérientiel (EPE). Chroniques Phénoménologiques, 11, 43-57 (2018).
31. Warren, W.H., Verbrugge, R.R.: Auditory perception of breaking and bouncing events: a case study in ecological acoustics. Journal of Experimental Psychology: Human Perception and Performance 10(5), 704–712 (1984)
32. Winlove C.I.P, Milton F., Ranson J., Fulford J., MacKisack M., Macpherson F., Zeman A. The neural correlates of visual imagery: a co-ordinate-based meta-analysis. Cortex 105, 4-25 (2018)