The SpeaK methodology for sound design

Patrick Susini, Olivier Houix, Nicolas Misdariis, Romain Barthélémy

The SpeaK methodology for sound design

Patrick Susini[1]
STMS (Ircam-CNRS-SU)
1 place Igor Stravinsky, 75004 Paris

Olivier Houix[2]
STMS (Ircam-CNRS-SU)
1 place Igor Stravinsky, 75004 Paris

Nicolas Misdariis[3]
STMS (Ircam-CNRS-SU)
1 place Igor Stravinsky, 75004 Paris

Romain Barthélémy[4]
Ircam Amplify
7 rue de Turbigo, 75001 Paris

ABSTRACT

The first meetings of a sound design project aim to present the request which may concern the identity of a brand or the sound of a new product. It is then crucial to elaborate recommendations for the sound designer. However, the major difficulty encountered is to specify the request in terms of sound features. The Speak methodology is based on an efficient co-design workshop to help participants to specify their request with words related to sound features. This methodology is based on a lexicon composed of 35 words often used by professionals to described sound features. Each word is related to a sound feature explained by a definition and highlighted by a corpus of sound examples, and is associated with a card used by the participants during the brief meeting. Using this methodology, participants are actively involved in the process to develop sound recommendations.

1. INTRODUCTION

Product sounds are not only pleasant or unpleasant. They serve many other purposes: they contribute to the brand image and the coherence of a product, elicit emotional reactions in users, and even have functional aspects in terms of information. As such, product designers not only want to diagnose the quality of a product sound, they also want to design its timbral and temporal characteristics to address different interdependent aspects, such as pleasure, identity, and functionality, as well as taking into account the environment in which it will be heard. As an example, most people in France associate the jingle played before any vocal announcement in French railway stations with the French national railway company (SNCF).

---

[1] Patrick.Suini@ircam.fr

[2] Olivier.Houix@ircam.fr

[3] Nicolas.Misdariis@ircam.fr

[4] barthelemyds@gmail.com

The timbral features and temporal properties of the jingle have been specifically designed to attract the attention of users and to communicate the values of the company. In addition, this sound has been designed to be enjoyable in the complex sonic environment of railway stations. Another example is the sound designed for the ZOE electric car - produced by the French carmaker Renault - to inform pedestrians of its movement on the street; its timbre is now emblematic of the car's identity and is nicely integrated into the urban sound environment [1]. The news sounds created in the two previous examples are referred to as intentional sounds. They are the result of a sound design approach implemented to create new sounds in order to make intentions audible in a given context of use [2].

But what happens when the product does not exist yet, or when the device was silent but to which it has become necessary to add sounds? Typically, an electric car is quieter than a car with an internal combustion engine. However, it is widely accepted that it is necessary to add intentional sounds to alert pedestrians and also to inform the driver about the car's state of functioning (e.g., its speed). New sounds must be imagined and created that satisfy functional constraints (e.g., detectability) as well as constraints in terms of pleasantness, identity, and ecology. The fruitful approach providing relations between users' preferences and timbre attributes based on an analysis of a collection of existing sounds (i.e., the sound quality methodology) is not useful in this case. Fortunately, practice in sound design is led by a strong creative process based on different sources of inspiration; in addition to their technical skills, sound designers are characterized by creative abilities to make sound sketches composed of different timbres, which can make all the difference in producing a successful articulation between functionality, pleasantness, and identity of a new product sound with respect to the sound environment. As it has been done for science fiction movies, sound designers have to imagine and create new intentional sounds for our everyday environments.

However, brands have expectations or values that constrain the creative process, and it is not obvious for sound designers, at the first step of the process, to understand and to translate those expectations by sound features as recommendations for the creative process. For example, during co-design sessions with partners from the French railway company (involving marketing professionals, sound designers, ergonomists, etc.), three values – benevolent, simple, and efficient – were presented as the main brand values. But which timbre features and temporal morphologies could be associated with those values?

In section 2, we discuss the different strategies to speak about sounds. In section 3, we focus on the lack of a shared vocabulary to speak about sound features by introducing the SpeaK project. In section 4, a proposed lexicon to speak of sound features – words4sounds – is described. In section 5, the SpeaK methodology based on this lexicon associated with cards used in co-design sessions is presented in order to overcome the problem of a common vocabulary to communicate about sound features during a sound design process that involves different participants

2.   DIFFERENT STRATEGIES TO SPEAK ABOUT SOUNDS

Unfortunately, nonexperts in sound are not used to speak about sound features, or to describe a sound for itself. The most common strategy is to describe the source of the sounds ("this is the sound of a hairdryer," "it is a vacuum cleaner," "this is a trumpet") or the action that produced them ("someone is hitting a glass," "this is the sound of a string being pinched," "she is pushing a switch"). This is the causal strategy, which is the most intuitive way to speak about sounds for nonexperts. Descriptions are sometimes solely related to a specific meaning in a specific context or location: alarm sounds in intensive care units have a specific meaning only for the staff. This is the contextual strategy: Verbal descriptions are not specific to a sound's features but are more

context-dependent. Finally, it seems that descriptions are seldom based on the sound itself in terms of acoustic characteristics and timbre features. This is the reduced listening strategy: Descriptions are directly related to the features of a sound independently of the meaning, the process that produced the sound, or its location. This distinction between the different strategies was highlighted by Pierre Schaeffer [3], and later by Michel Chion [4]. Then, Gaver [5] opposed two listening strategies: everyday listening (similar to the causal strategy) vs musical listening (similar to the reduced listening strategy).

Although daily life listening is usually focused on sound as a reference to a source or a meaning, it happens that one is led to describe the characteristics of a sound, and there again, one will find biases to give its opinion in term of a general judgment "It is an annoying sound", or still by imitating the sound characteristics with voice. For example, we produce the sound "Pam! Pam!" for describing an impact on wood, and "Tinnng! Tinnng!" for an impact on glass or metal. We can have also recourse to onomatopoeias: "Toc! Toc!" in French or "Knock! Knock!" in English to describe a sound made by someone knocking on a door.

In the sound design process, especially when it involves participants with diverse levels of expertise, this diversity of strategies can be a serious obstacle for communication. Sound designers usually need information in terms of acoustic characteristics related to timbre features or temporal properties, but initial intentions are often expressed by the client with terms related to a meaning, a function or an emotion. For example, the intention for an alarm sound in the context of a hospital could be described as "alerting but kind" rather than sound features such as long, smooth, continuous, highpitched, loud enough, which are directly informative for the sound designers. Unfortunately, there is no common practice for speaking about sounds, and in particular about timbre features; sound designers often complain about the lack of tools to communicate about sounds in a sound design process.

3.   THE SPEAK PROJECT

Thus, speaking about the characteristics of a sound is difficult because there is **a lack of words to promote communication on sounds**, or to express the sensory experience with sounds. Using a sound lexicon is a first step to overcome this lack. Since the work by Pierre Schaeffer and Michel Chion, several authors have proposed different extend taxonomies tacking into account emotional dimensions [6] or kinaesthetic aspects of sounds [7]. In the last 20 years, several studies have investigated the use of verbal descriptions or semantic correlates of timbre for different musical instruments such as the organ [8], the guitar [9], the piano [10, 11], or the violon [12], among others. More recently, vocabulary employed by sound professionals was explored [13, 14, 15].

However, the meaning of words used to describe sounds can vary from person to person, even for sound professionals ranging from sound engineers to musicians [16, 17], **revealing a lack of consensual definitions for the words relevant to describe sounds.**

In addition, words, and their definitions, are sometimes not sufficient to describe a hearing experience associated with a sound characteristic**. Sound examples are a good way to understand the relationship between words and sounds**, and to experience this relationship.

Therefore, the SpeaK project is a collaborative web platform for organizing, presenting, and sharing sound lexicons combining words, definitions and sound examples. The SpeaK project allows contributors to create and share lexicons related to a specific project. For example, it is an

opportunity to account for the specific vocabulary shared by musicians to speak about their sound experience with their musical instrument, or for the specific vocabulary used by composers for electronic music, or by sound designers for soundscapes, etc.

The SpeaK web page is reachable here https://speak.ircam.fr/en

The benefit of a sound lexicon is to improve and facilitate communication on sounds, because sounds are sensations related to words.

As introduced above, in the SpeaK project, a lexicon is defined by:
 – a list of words used to describe sound characteristics
 – a definition to explain the meaning of each word of the list
 – a corpus of sound examples to foster the perception of the characteristics

Words can be classified in different typologies; for example, to differentiate temporal and spectral characteristics. Sound examples can be organized in different categories; for example, to compare artificial and natural sounds.

4.   THE SPEAK LEXICONS

4.1.   words4sounds.speak

The lexicon words4sounds.speak is the first sound lexicon available on the SpeaK web platform (figure 1). It was elaborated by the Sound Perception and Design group (Ircam STMS Lab) on the basis of Maxime Carron's PhD [19]. It is based an academic review of a large number of works dealing with verbal descriptions of timbre for different kinds of sounds, from abstract to everyday sounds. Then, this review was combined with interviews with French speaking sound professionals from different fields (e.g., composers, sound designers, sound engineers, etc.).

Then, a lexicon of 35 relevant words, frequently used by professionals, to describe the perceived characteristics of a sound was proposed (e.g., tonal/noisy, low/high, dry/resonant, dull/bright, rough, warm, round, nasal, rich, strident, dynamic, crescendo/decrescendo, ascending/descending, fast/slow attack …) as an extension of Schaeffer's work. The lexicon is proposed in English and French based on a collaboration with the Timbre Semantics group within the Actor project (https://www.actorproject.org/). The list of words is structured in three classes of general aspects (e.g., high/low, short/long, etc.), temporal morphology (e.g., crescendo/decrescendo, continuous/discontinuous, etc.), and timbre attributes (e.g., dull/bright, nasal, warm, etc.).

The lexicon aims at enhancing and supporting communication in sound design collaborative sessions with definitions and sound samples. This lexicon is part of the SpeaK methodology presented in section 5.
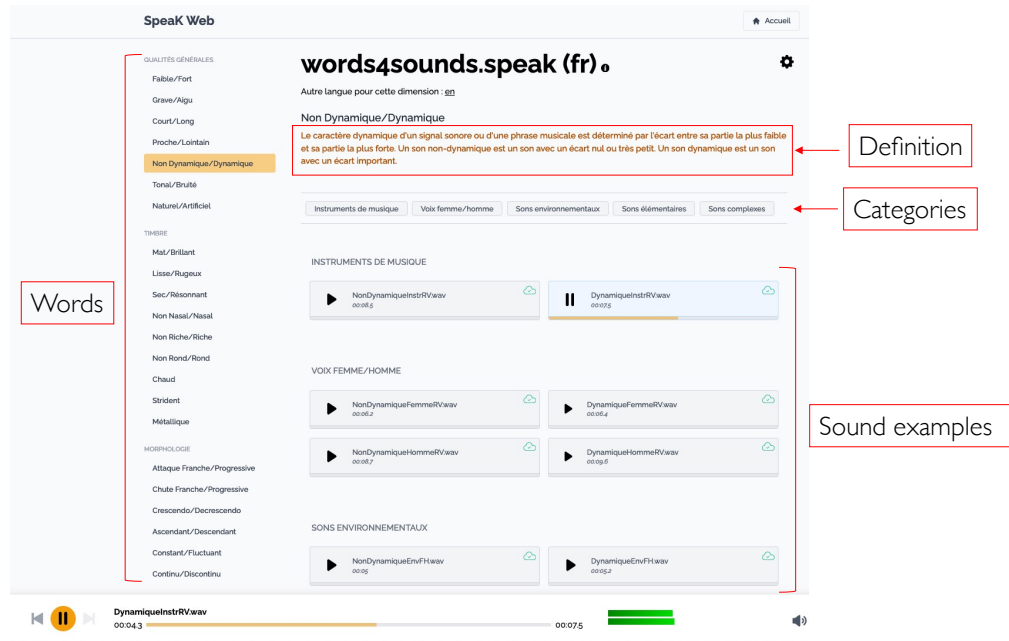
Figure 1: The SpeaK Web interface for the words4sounds lexicon (https://speak.ircam.fr)

The sound examples were created, recorded, and mastered under the direction of composer Roque Rivas at Ircam (figure 2), except the environmental sounds which were proposed and recorded by François Hamon as part of his internship at DNSEP Design Sonore, ESAD TALM Le Mans. They were created to highlight the sound features related to the words for different categories of sounds.
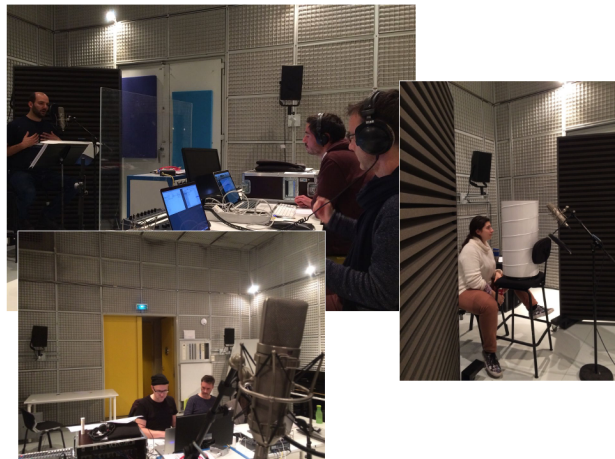

Figure 2: Voice recording sessions at Ircam for the lexicon with Sylvain Cadars and Jeremy Bourgogne as sound engineers, and Roque Rivas as musical director. Singers are Nicolas Certenais and Marina Ruiz

The definitions are based on analysis of interviews with French speaking sound professionals [14, 16]. Result for Dull(Mat)/Bright(Brillant) can be seen on figure 3.

Figure 3: Result of an analysis of interviews with sound professionals for Bright [16].

The definition obtained is : "The word Bright (brillant) is often used to describe a timbre characteristic of different sounds; the opposite word used is dull (mat). The words dull and bright refer to the amount of high-frequency energy perceived within a sound. A dull sound has a low amount of high-frequency components. The term muffled is also used. A bright sound contains a substantial amount of high-frequency components. The term sharp is also used." Good musical examples for bright are the glockenspiel and the trumpet.

4.2. Other lexicons
– **words4RoomAcousticalQuality.speak**
It is a lexicon of words to describe Room Acoustical Quality by professional acousticians. It is based mainly on the work published by [19].
– **words4EV.speak**
It is a lexicon of words to describe sounds of Electric Vehicle by professional acousticians. It is based mainly on the work published by [20].

5. THE SPEAK METHODOLOGY

5.1. The brief in a sound design project
The first briefs in a sound design project are crucial to elaborate recommendations for the sound designer. However, in the first briefs which bring together different non-expert sound practitioners ranging from the project manager to the communication manager, the major difficulty encountered is **to specify the request with words related to sound features**. Thus, in several collective co-design sessions for different projects with industrial [21] or institutional partners, the words4sounds.speak lexicon was used in order:

– to have a shared and unique list of words to describe the expected sound features
– to involve all the participants in the collective co-design sessions
– to foster the contact with the sound designer

5.2. Different steps of the methodology
Different steps can be proposed to the participants during a project using the lexicon, depending on the time available, the number of possible sessions and the duration of the project.

1/ A training session in two phases (figure 4)

– Phase #1 Lexicon inventory: collective listening > individual listening > debriefing
– Phase #2 Evaluation: level#1 > level#2 > final check of a shared representation
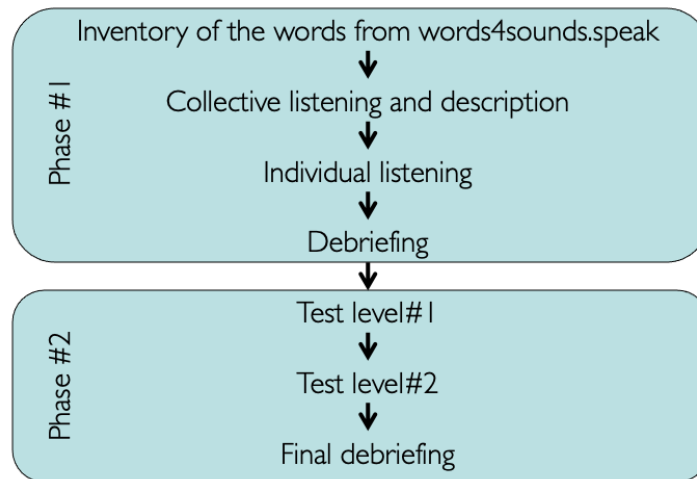
Figure 4: Training session in two phases

Phase #1 is mandatory and is crucial to introduce the non-experts to the world of sounds – and related words. The session leader presents all the words of words4sounds.speak by playing examples using the web interface. The definition of each word is presented, and participants can ask for clarification. Then each participant discovers the lexicon individually by listening to examples using headphones. This individual exploration can last more than 30 minutes. Finally, the session leader ensures that all the words have been understood and that the examples have made it possible to identify the associated sound characteristic.

Phase #2 is optional, but crucial to ensure a good understanding of the link between a word and the associated sound characteristic, using forced-choice listening tests with only one possible correct answer. Two levels of difficulty are tested.

- Level#1: the first level, the simplest, consists for the participant to select the sound the most representative of the attribute involved. For example, 5 sounds are presented but only one is related to the word displayed, and the participant is asked to select the corresponding sound.
- Level#2: the second level is similar to the first one, but differs from it in difficulty. Indeed, the correct answer does not emerge from the corpus in an obvious way since the sound examples are constructed in order to be distinguished on smallest perceived difference along the sound feature tested. For example, 5 sounds are presented with small differences in brightness, the participant is asked to select the brightest one.

During a training session, individual and collective explorations of the lexicon are alternated with the different tests. After each test, terms are discussed collectively to ensure a common understanding. This global training ensures that participants involved in the same project have a rich, varied and shared vocabulary that is adapted to describe a large number of timbre features and temporal properties appropriate for an important variety of sounds. This procedure is an alternative to sensory evaluation often used to reveal a list of words specific to the timbre of a set of sounds in relation to consumer preferences. The sensory evaluation requires several steps of discussion, training, and testing with a panel of experts, a process which is often very long (several weeks) and specific to a set of sounds.

2/ A collective co-design session

Efficiency of a co-design session can be improved using the lexicon words4sounds with cards associated to the words. Cards (figure 5) are made available to each participant who can select or reject a word, and exchange with other participants focusing only on the sound features constrained by the cards. This codesign setup is transposed from Carron's work and was formerly inspired by specific design approaches like Kansei. Discussions are mediated by the cards and the area of exchange iss materialized with a board (figure 6), by analogy with a standard board game or role play. Supporting that, the words4sounds.speak lexicon played the role of help to which anyone can refer during the session.

Figure 5: Cards associated to the words4sounds lexicon for the SpeaK methodology

Figure 6: The SpeaK methodology applied to sound identity of Sorbonne Université by Romain Barthélémy for Ircam Amplify.

5.   FINAL COMMENTS AND CONCLUSIONS

A standardized lexicon of specific terms, such as proposed on the SpeaK web page, to describe relevant sound features is a very promising tool for the field of sound design from a practical point of view, for example, to assist in the training of the different participants involved in a sound design project to perceive and use relevant timbre features for the design process. The question of brand sound identity calls for the difficulty of communication around sound between the client and the sound designer can be solved by the SpeaK methodology. This

methodology is based on the lexicon words4sounds composed of 35 words associated with cards used to conduct co-designed sessions. This methodology has been tested during different industrial and institutional projects which allowed to confront the proposed sound lexicon with the reality of a creative process. Feedbacks reveal that briefing sessions based on the different words of the lexicon are rich discussions and exchanges. This made it possible to release quite easily a consensus for sound terms while some of the participants are unfamiliar with these terms.

The SpeaK methodology also would be useful to teach students in a sound design or post-production course who are learning to listen to timbre features and could then describe those features with a common vocabulary. From a training perspective, a set of audio tests also has been developed within the SpeaK methodology to evaluate participants' understanding of the lexicon; it is a complementary and indispensable element of applying the lexicon. The tests assess whether using the lexicon may improve listeners' perception of a specific feature as well as their ability to describe sounds with only the terms of the lexicon.

ACKNOWLEDGEMENTS

REFERENCES
1. N. Misdariis and A. Cera. Knowledge in Sound Design – The Silent Electric Vehicle: a Relevant Case Study. In DeSForM - Sense and Sensitivity, Delft, Netherlands, October 2017.
2. P. Susini, O. Houix and N. Misdariis. Sound design: an applied, experimental framework to study the perception of everyday sounds. *The New Soundtrack*, *4*(2), 103-121, 2014
3. P. Schaeffer. Traité des objets musicaux. Editions du Seuil, 1966.
4. M. Chion. Guide des objets sonores: Pierre Schaffer et la recherche musicale, 1983.
5. W. W. Gaver. How do we hear in the world? Explorations in ecological acoustics. *Ecological psychology*, *5*(4), 285-313, 1993.
6. D. Huron. A six-component theory of auditory-evoked emotion. In Proceedings of the 7th International Conference on Music Perception and Cognition, 673-676, 2002.
7. K. Tuuri and T. Eerola. Formulating a revised taxonomy for modes of listening. Journal of new music research, 41(2), 137-152, 2012.
8. V. Rioux and D. Västfjäll. Analyses of verbal descriptions of the sound quality of a flue organ pipe. Musicae Scientiae, 5(1), 55-82, 2001.
9. C. Traube. An interdisciplinary study of the timbre of the classical guitar, Doctoral dissertation, McGill University, 2004.
10. M. Bellemare and C. Traube. Verbal description of piano timbre: Exploring performer-dependent dimensions. In Digital Proceedings of the 2nd Conference on Interdisciplinary Musicology, CIM05, 2005.
11. P. Cheminée, P. «Vous avez dit «clair»?» le lexique des pianistes, entre sens commun et terminologie. Cahiers du LCPE: Dénomination, désignation et catégories, 7, 39-54, 2006.
12. C. Saitis, C. Fritz, G. P. Scavone, C. Guastavino and D. Dubois. Perceptual evaluation of violins: A psycholinguistic analysis of preference verbal descriptions by experienced musicians. The Journal of the Acoustical Society of America, 141(4), 2746-2757, 2017.
13. T. Porcello. Speaking of sound: language and the professionalization of sound-recording engineers. Social Studies of Science, 34(5), 733-758, 2004.
14. M. Carron, T. Rotureau, F. Dubois, N. Misdariis and P. Susini. Speaking about sounds: a tool for communication on sound features. Journal of Design Research, 15(2), 85-109.
15. Z. Wallmark. A corpus analysis of timbre semantics in orchestration treatises. Psychology of Music, 47(4), 585-605, 2019.

16. V. Rosi, O. Houix, N. Misdariis and P. Susini. Investigating the shared meaning of metaphorical sound attributes: bright, warm, round, and rough. *Music Perception: An Interdisciplinary Journal*, *39*(5), 468-483, 2022.
17. V. Rosi, S. P. Arias, O. Houix, N. Misdariis and P. Susini. Shared mental representations underlie metaphorical sound concepts. *Scientific Reports*, *13*(1), 5180, 2023.
18. M. Carron. Méthodes et Outils pour Définir et Véhiculer une Identité Sonore, Doctoral dissertation, Université Paris 6 (UPMC), 2016.
19. S. Weinzierl, S. Lepa and D. Ackermann. A measuring instrument for the auditory perception of rooms: The Room Acoustical Quality Inventory (RAQI). The Journal of the Acoustical Society of America, 144(3), 1245-1257, 2018.
20. M. Duroyon, N. Misdariis, P. Susini, N. Dauchez, L. F. Pardo and E. Vialatte. How do Automotive Acousticians talk about Sounds and Comfort in Electric Vehicle?. In 10th Convention of the European Acoustics Association Forum Acusticum 2023.
21. N. Misdariis, A. Cera, R. Barthelemy, L. Worms and L.F. Pardo. Collaborative sound design for automotive applications – The sound for Electric Vehicle as an emblematic use case. In Proceedings of INTER-NOISE 2024, (to be published). Nantes, France, August 2024.