Art(s) and (some) Thoughts

Art(s) et (quelques) réflexions

ReSilent – Transforming Sound Intrusiveness

Andrea Cera and Antonio Camurri

We propose an approach towards the modeling and automated analysis and enhancement of intrusive soundscapes, based on a series of timbral, temporal and contextual dimensions around which the potential intrusiveness of a sound is articulated.
Furthermore, we provide concrete examples from a series of prototypes and algorithms we are currently developing in the EU Horizon Europe ICT STARTS ReSilence *project.*

Part A: Intrusive Sounds

Intrusive sounds are events which emerge from a background in such a way that involuntary reactions are elicited: the activation of certain brain regions; heartbeat slowing down; muscular rigidity, etc.[1]

This does not mean that intrusive sounds are always a negative experience: almost every kind of music has intrusive elements, and we probably like the sensations created by this intrusiveness. We like heavy metal because we probably like the sensation of activation, energy, force created by its abrasive sounds.

Intrusiveness becomes a problem when it is experienced in a passive way, without decision: being exposed to an aggressive blinker sound while driving; receiving the sound of a ringtone too rich and loud; being in a hospital while medical equipment sends out urgent alerts; being in a park while a modified scooter passes by.

That is why sound designers, who create such types of sound, must pay particular attention to intrusiveness.

Intrusiveness as relation between background and foreground

Intrusiveness emerges as a relation between background and foreground.

There are many ways in which a sound can emerge from a background. The very definition of what constitutes a "background" is still problematic, and it can be safer to give just an operational definition of a salient sound as a sound who captures attention in a bottom-up direction, in a given context.[2] We could characterize a "background soundscape" in terms of a number of audio and context descriptors at different temporal scales, characterized by a limited statistical variability (i.e., small variations not causing salient events raising attention in the listener). And a salient sound would be a change in one or more of these descriptors big and fast enough to activate attention.

In the past, several fields of study have investigated the mechanisms in our auditory system which allow for this phenomenon to occur, without finding yet a definitive conclusion.

Since the sixties, the fascinating field of auditory attention[3] research investigated how the emergence of sound from a background elicits drastic changes in brain structures connected with arousal and vigilance. We are constantly listening to the sound of the world around us, and any change in it can awaken our attention.

In the 1990s, the field of auditory scene analysis, defined by Albert Bregman in his seminal book,[4] analyzed how our auditory system is able to deconstruct complex auditory scenes into individual elements. How are we able to understand the voice of the person speaking to us in a party, while many other voices arrive at the same time to our auditory system.

More recently, another important field of research took inspiration from the visual domain to investigate auditory salience,[5] and understand why every element emerging from a background does not capture our attention in the same way.

Research about this topic attempts at understanding how much a particular sound must be different from a background in order to

be salient and capture our auditory attention, which are the privileged dimensions, how different temporal scales are in play, etc.
Many studies have already found high correlations between the audition of this type of intrusive sounds and an increase in arousal, changes in skin conductance and in heart-rate, etc. independently from the fact that in a subsequent appraisal one can like the sound.[6]

Intrusiveness and timbre

Several timbral dimensions contribute to the contrast between background and foreground and are responsible of intrusiveness effects. We will focus on four of them: auditory roughness, spectral centroid, spectral sharpness, spectral skewness. In the following, we describe these audio descriptors.

Auditory roughness

Auditory roughness is the effect caused by fast modulations in amplitude.[7]
Auditory roughness can happen in many ways: humans and animals are able to create roughness when they scream, or cry. We could think that this intrusive effect must have some evolutionary function:[8] create a sensation of alert, of urgency, or simply assuring that a call for help cannot be ignored for too long and can be perceived easily even from afar. Vocal effects in singing related to roughness are part of the vocal technique and related to motor activities of specific discrete vibratory pattern of supraglottic structures, to obtain and control distortion, growl, rattle, grunt.
In many musical genres, good singers master the art of controlling these effects, evoking rage, high energy, excitation and possibly infecting their listeners with the same primitive emotional qualities.
Orchestral composers also control roughness in music by managing dissonance (not only harmonic but also timbral): they organize complexes of tones with fundamentals or partials so close between them that the periodical phase cancellations/reinforcements which occur create this particular modulation, slightly and deliciously uncomfortable. Listeners' ears are somehow titillated by subtle modulation of these roughness effects, which come and go, in dialog with smooth, fluid and consonant timbral agglomerations.
In the universe of sounds created by machines, roughness is often present: car and motorbike engines present varying levels of roughness, since these sounds are very fast successions of explosions (at the beginning of every cycle of every cylinder in endotermic engines), i.e. very fast and very ample modulations of amplitude. This is often connected with the sensation of sportiness and competition, because the more clearly the sound of these fast succession of shots is audible, and not concealed by mufflers, resonances, damping devices, the more it evokes explosions, fuel, fire, and therefore emotions related to danger, aggressiveness, power.
In other cases, mostly with electric appliances, roughness is considered as a defect.
When we passively listen to sounds with high roughness (for instance, when immersed in urban traffic), we could unknowingly experience emotional states which enhance our nervousness, energy levels, arousal.

Spectral centroid

The spectral centroid indicates in which frequency zone the center of gravity of a sound is located.[9] An electric bass with a dark timbre, playing a low E has a low centroid. If the same note is played with a brighter timbre, like in a slap note, the centroid will slightly rise.
Research about intrusive sounds has often found that the higher the centroid, the more annoying and intrusive the sound is. Several reasons explain this phenomenon: in particular, our auditory system is much more sensitive to higher frequencies (i.e. in the range of human voice) than to lower frequencies.
This feature of our auditory system has been studied since the 1930s, with the first examples of the equal-loudness contour curves by Fletcher and Munson:[10] these diagrams explain how different frequencies have to be played louder or softer in order to generate the same sensation of loudness: the lower we descend in frequency, the less we are sensitive. The peak of our sensitivity is around 3KHz: this means that sounds with a centroid around 3KHz are the ones that we hear as the loudest, the

clearest, the most evident. After this threshold, our auditory system tends to become less sensitive: at the same time, very few and rare sounds have their center of gravity this high.
If sounds have a clear fundamental and a harmonic partials structure, the centroid tends to overlap with the fundamental, but not always: in any case, fundamentals which evoke the pitch of a voice screaming or talking with an altered tone can also be intrusive and elicit primitive emotional states of high energy and possibly negative valence (which is what happens when we hear a child crying).

Spectral sharpness

Sharpness indicates the amount of high-frequency components in a sound.[11] As mentioned above, the human auditory system has an increasing sensitivity to frequencies in the range of about 0.5-5KHz, with the highest peak at around 3KHz.
One possible evolutionary explanation for this phenomenon is that in human language the timbral difference between vowels is focused around this zone, and our auditory system became hyper-sensitive to this zone in order to perfectly distinguish between different vowels.[12]
In any case, it is a fact that a sound with high sharpness has a higher impact on the auditory system than a sound with low sharpness.

Spectral skewness

Skewness relates to the distribution of frequencies around the centroid, the center of gravity of a sound.[13] The centroid too is also connected to intrusiveness: the correlation between skewness and centroid is an example of the complexity of the problem of modeling intrusiveness.
A sound with skewness higher than zero will have a greater amount of energy in the lower frequency regions, with respect to the higher regions. A sound with zero skewness will be well balanced around its centroid. A sound with negative skewness will have more energy above the centroid.
A personal hypothesis from the authors is that we tend to give more attention to sounds whose skewness is high, or negative: high or negative skewness may contribute to raise intrusiveness.
Environmental sounds, background sounds found in the wild (e.g., mixtures of sounds of leaves moved by breeze in distant trees, water, wind, rain, etc.) tend to have a balanced skewness, neither positive nor negative. Sometimes, even urban soundscapes have a similar quality, mostly if captured in places where roads are distant – and in this sense, they can paradoxically be perceived as "natural".
On the contrary, sounds generated by a human artifact, like a musical instrument, or a piece of machinery, often show positive or negative skewness, intentionally or not. This unbalance makes sounds unnatural, strange, non-usual, and therefore more likely to attract attention, and to cause intrusion.

The above proposed audio descriptors are among the most cited in the scientific literature on sound annoyance and intrusiveness and are currently investigated in our work. It is worth noting that they are not independent (e.g. sharpness is related to spectral centroid; roughness to amplitude); further, other audio descriptors can be considered contributing to intrusiveness.

Intrusiveness and temporal dimensions

A fundamental component in modeling intrusiveness concerns the number of different temporal scales involved in audio and context descriptors.
Several studies have explored how many temporal dimensions contribute to the intrusiveness of sounds.[14]
The most widely accepted finding is that duration is positively correlated with intrusiveness. This is why a short shot of klaxon can be perceived as less aggressive that a prolonged and sustained one.
At the same time, short sounds might induce a sensation of hurry, of urgency, so perhaps there are other dimensions that should be considered.
An important dimension is the morphology of attack and decay. A sound can induce less urgency if its end is slowly fading, maybe in conjunction with a transformation in timbre.

And the attack of a sound is less intrusive if it grows in intensity in a very gradual and slow way. It is important to note that this slow shape has to be really slow: a sound with a fade-in up to 2 seconds is not yet out of the territory of intrusiveness and can still evoke stirring and arousing primitive emotional states.
From the authors' experience in previous work and in scientific experiments, in order to avoid any sensation of aggression and intrusion, a sound with a significant timbral difference from its background should take at least 6 seconds to emerge and to fade-out.
These numbers seem to belong to a particular temporal dimension: a low-intrusiveness temporal dimension.
The following image resumes a detail of an analysis of the temporal dimensions we elaborated in the *DanzArTe* project,[15] which structure human experience in an interactive sonification of individual as well as joint actions in a dyad.

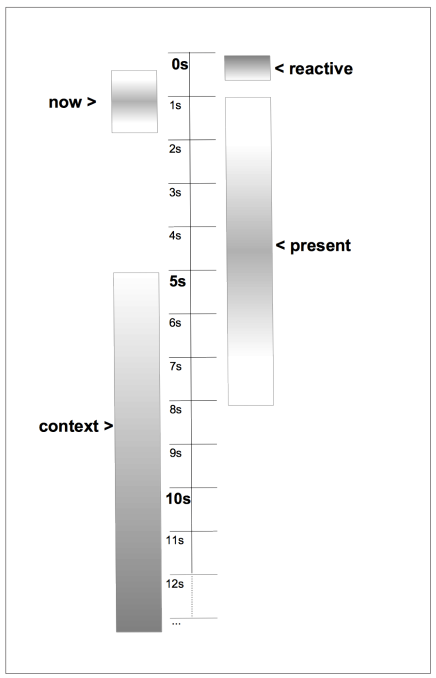


Fig. 1. The temporal dimensions elaborated in the *DanzArTe* project.

The temporal-scales we identified in the DanzArTe project are the following:

- reactive [0,1s–1s]: this layer contains events which fluctuate between appearing fused together and as separated; which start to evoke a sensation of pulsation, and a clear sense of urgency.[16] Events in this layer oscillate between being too fast to be used as synchronization mark for clapping, and allowing a comfortable clapping rate.[17]
- now [0,5s–2s]: the lower part of this layer contains events which progressively diminish the evoking of a sensation of urgency, and enter the zone of preferred BPM, the rate of healthy walking, and the easiness of clapping synchronization.[18] The upper part of this layer contains pulsations which make synchronized clapping difficult and correspond to the slower rates of traditional metronomes.
- present [1s–8s]: this layer represents what French psychologist Paul Fraisse called the "perceptual present",[19] a temporal dimension where events can be grasped as coherent and homogeneous, and where it is easier to encode, memorize and recognize musical patterns.[20] According to Fraisse, the center of this layer is around 2/3 seconds.
- context [5s–12s and more]: this layer represents the exit from the "perceptual present" and characterizes events whose length exceeds what can no more be consciously perceived as one coherent event.
In this layer, saliency phenomena starts to appear: for instance, Botteldooren and De Coensel suggest that requests for bottom-up auditory attention be modeled as timing out in 10 seconds.[21]

Low-intrusiveness events are at the border between the last two scales (present/context), while the first two scales identify events which tend to create an urgent and immediate-response dimension.
In the *DanzArTe* project, which focused on low-intrusiveness sounds to nudge participants of an interactive social experience (typically older people at risk of fragility, possibly together with caregivers) for physical and cognitive treatment, to move slowly and with

fluidity, we excluded audio events in the first two time-scales, and designed the interactive sonification mainly on events temporally organized at the border of "present" and "context" dimensions. The positive results of the validation of this project show that the focus of these slower temporal scales is beneficial in order to nudge towards a joint physical activity on certain movement quality without creating intrusion.[22]

This reflection on temporal scales needs further work, e.g., by taking into consideration other (sub)scales (e.g. 5-8 seconds, which we didn't use in the *DanzArTe* project). We should also investigate if the notion of "background" could be associated with a particular kind of event connected to a certain temporal scale, or if this notion should not be considered as an "event" among others but as a "process" characterized by a duration at a slow temporal scale.

Intrusiveness and the most obvious dimension– amplitude

After this list of possible causes of intrusiveness, a parameter seems to be missing: amplitude.
Amplitude usually is the first factor taken in consideration when trying to correct an intrusive sound: this sound is too loud, take down the volume!
In fact, amplitude is a fundamental factor, the first dimension studied to understand the negative effects of sound towards the auditory system.
In this article we take for granted that over certain amplitude levels any kind of sound will be not only intrusive but even dangerous for the listeners' health, but we want to study more subtle effects, which can arise from sounds at low sound pressure.
Just measuring the amplitude of a signal can be extremely complex. Several methods have been adopted, pre-filtering in several ways the signal to be measured, to simulate the response of the ear and even taking into account different ways to average the amount of sonic energy delivered by a particular situation.
For years researchers concentrated on studying matters of sound pressure: our current measures of sound pollution are still based on measures of sound pressure level.

While this dimension is still extremely important and crucial, the studies we mentioned before showed how many different timbral features are critical to understand why sounds can become intrusive, even if their amplitude is not very different from the amplitude of their background, even if they are "soft".
Further, the subjective perception of amplitude, i.e. loudness, as well as auditory masking and spectral relations between the background and a sound event should be considered.

A strategy to lower intrusiveness

In a previous text for *LINKs* co-written with Nicolas Misdariis (IRCAM) one of the authors proposed principles for a low-intrusiveness form of sound design: a guideline to control the intrusiveness of new sounds to be introduced in a given situation.
In the *ReSilence* European project the authors explore a strategy to lower the intrusiveness of sounds over which we have only partial control (for instance, the presence of vehicles in a nearby road).
Since intrusiveness is the product of the contrast between a foreground sound and its background, in the case where the foreground can be controlled only in small part, the solution consists in changing the background.
The solution can be split in two:
- solution 1: slightly alter what is possible in the foreground sound in order to make them similar to the background;
- solution 2: change the nature of the background, making it timbrally similar to the foreground, using very slow temporal scales.

At the moment of the writing of this paper, the experimentation is moving along two directions:
- direction A) the creation of a database of intrusive sounds and a low-intrusiveness version of the same sounds, on the basis of a static background, to be used for evaluation and validation of the two versions of sounds in scientific experiments;
- direction B) a smartphone app that allows to apply these strategies to reduce intrusiveness in realtime on the audio stream captured by

the device's microphone, and to diffuse on headsets the transformed background.

Direction A: database

The creation of the intrusive side of this database follows a well established practice,[23] consisting in augmenting a realistic background with tightly controlled foreground sounds.

The background is composed using fragments of the "Soundscapes of the World"[24] project: a collection of urban soundscapes, recorded and calibrated in several cities, sometimes in zones far from traffic lines, ideal to create a constant, uneventful background against which to oppose foreground sounds. The background is a loop whose length is impossible to be detected by a listener.

The foreground intrusive sounds are made from a collection of urban sounds, that one of the authors made during the years with a Zoom H4. It consists of cars, motorbikes, trucks, construction-work noises, klaxons, loud smartphone ringtones, ventilation devices, skateboards, and three simulations of music coming off windows, stores, etc.

A typical presentation of this database is a montage of different intrusive events, one after the other (but not repetitive, not periodic), sometimes slightly overlapping, sometimes with pauses where only the background is audible. We chose a density of 12 events per minute, which seems the more annoying, according to a study by Kazmarel and Preis.[25]

The low intrusiveness versions of these sounds are prepared according to the following techniques.

Solution 1 (slightly alter the intrusive sounds)

1) change the spectral centroid and sharpness (in our case, diminish) of the foreground sound by filtering it in order to bring these dimensions closer to the corresponding ones of the background; this type of filtering (usually a low-pass filter, but in certain cases more complex mixtures of low-pass and peak/notch) could be done also in real-life, with physical objects, for instance sound-absorbing panels designed to work as low-pass filters.

2) reduce the roughness of the foreground sound by a) analyzing its fluctuations and even out them using delays to add energy in the zones where the signal has low amplitude, and b) adding reverbs to generate sound with a similar function (masking the low-energy portion of the fluctuation with sound which prolong the high-energy part).

This operation would be more difficult to perform in real-life, with physical objects, but could be approximated by resonators.

3) change spectral skewness by filtering out possible imbalances. This operation could not be performed in real life, because it is too much dependent on the morphology of each single sound.

Solution 2 (alter the background)

This is an important task for the *ReSilence* project. The basic concept is to create a smooth, uneventful version of the foreground sounds, to become part of the background sound, in order that the emergence of the intrusive sound creates less contrast.

According to our observations on the temporal scales, this "shadow" of the intrusive sound must precede and follow the intrusive sound, in order to envelope it in a slower temporal scale.

For our database, this "shadow" takes 6 seconds to emerge and 6 seconds to disappear. At the center of this very slow event we add the intrusive sound, slightly transformed as described in the previous paragraph.

To create this "shadow" we use a granulator in MaxMSP,[26] whose head is placed in the center of the intrusive sound, setting the grain size and a quantity of random small deviations on the head position to obtain a prolonged version of the intrusive sound. On this basis, several transformations from direction A are performed, until the resulting sound has a level of smoothness and equilibrium similar to the background, while maintaining a reasonable resemblance to the original intrusive sound.

The "shadow" is finally mixed with one to three sound-files which evoke natural sounds. These sound-files are: 1) a simulation of wind made with the Sound Design Toolkit;[27] 2) a recording of wind passing through trees with a distant road noise; 3) a mix of recordings of crickets and cicadas, real and synthesized.

These three sound-files differ in timbral content. Sound 1 is rich in low frequencies and has an artificial sense of stability and evenness; Sound 2 is less skewed than the other two sounds (it is the only recording non mixed with synthesized sounds) but has a faint presence of artificial elements; Sound 3 is very sharp and bright and quite evocative of open-air spaces.

For each shadow we decide to add some mixture of these three sounds, to create an agglomerate which bears some familiarity with the original intrusive sound, but at the same time (because of this last augmentation) has a better coexistence with the background. For instance, if a shadow is too bright (because it comes from a sound with a high sharpness), we will add to the shadow some presence of Sound 1 and 2 in order to lower the sharpness of the agglomerate and make it easier to mix with the background.

In the following image we show some graphs representing measurements of loudness, spectral centroid, roughness, sharpness, skewness, of intrusive and low-intrusiveness sounds, compared to the background. It is possible to see the effect of the different transformations, which, for each sound, tend to force the relative timbral feature towards the values of the background (which in general are lower). In particular, the loudness scheme shows the difference between clear attacks/releases of intrusive sounds, and the slower fades in the low-intrusiveness sounds. The overall sound pressure of intrusive and low-intrusiveness database is the same (the two sound-files deliver the same RMS average energy): in the low-intrusiveness version the peak energy is diffused before and after the main sound.

Fig. 2. Plots of timbral dimensions of background sound (left column) and foreground sounds (right column) used in the ReSilence experiments.

At the end of this ongoing experiment in the *ReSilence* project a detailed list of all transformations will be available, along with the individual sounds. At the time of writing, a few informal listening sessions have been conducted, confirming the efficacy of the techniques for transforming the sounds.

A generic set of effects inspired by this work is being implemented in an Android app, called ReSilent, which can perform them in real-time using the smartphone microphone as input, and delivering the transformed sound to the output, to be listened to with headphones.

This app will allow a low-intrusiveness treatment to an intrusive soundscape – the headphones will function as a generic loudness reduction device, while the sounds coming from the app will create the "shadows" described above, with the limitation that these augmented sound structures won't precede the intrusive sounds (which, coming from the external world, cannot be predicted), but only follow them.

Conclusion and future work

This paper presents an ongoing work originated in two European projects ICT DANCE and the FET PROACTIVE EnTimeMent 28 and is part of the DanzArTe project and the Horizon Europe ICT STARTS ReSilence project[29].

Several research directions are part of the ongoing and future work.

Sound is originated by movement (see for example how Friberg and Sundberg investigated analogies between the stopping of running and the termination of a piece of music30): the cross-correspondences between audio descriptors and expressive human movement descriptors are a promising direction of investigation, for example in real-time interactive sonification of expressive (individual as well as joint) human full-body movement in predictive health and cultural welfare applications.

The multi-temporal scales proposed in this paper can be considered part of a broader investigation, started in the EnTimeMent European project, towards spatio-temporal multi-layer conceptual frameworks.

The multiple temporal scale approach here proposed focuses on one of the multiple dimensions of the phenomenon; ongoing investigation focuses on integrating saliency and context in the conceptual framework.

[1] Ken Hume & Mujthaba Ahtamad, "Physiological responses to and subjective estimates of soundscape elements," *Applied Acoustics* 74(2), 2013, p. 275-281; Robert Hughes & Dylan M. Jones, "The intrusiveness of sound: Laboratory findings and their implications for noise abatement," *Noise and Health* 4(13), 2001, p. 51-70; Mercede Erfanian, Andrew J. Mitchell, Jian Kang and Francesco Aletta, "The Psychophysiological Implications of Soundscape: A Systematic Review of Empirical Literature and a Research Agenda," *International Journal of Environmental Research and Public Health* 16(19), 2019, p. 3533.

[2] Baptiste Bouvier, *Saillance auditive : de la caractérisation psychoacoustique à la perception de l'environnement sonore*, PhD dissertation, Acoustique [physics.class-ph], Sorbonne Université, 2024, https://theses.hal.science/tel-04562041.

[3] Jonathan B. Fritz, Mounya Elhilali, Stephen V. David and Shihab A. Shamma, "Auditory attention - focusing the searchlight on sound," *Current Opinion in Neurobiology* 17(4), 2007, p. 1-19.

[4] Albert S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*, Cambridge, MA, The MIT Press, 1990.

[5] Nicholas Huang & Mounya Elhilali, "Auditory salience using natural soundscapes," *The Journal of the Acoustical Society of America* 141(3), 2017, p. 2163-2176.

[6] E. A. Bjork, "Startle, Annoyance and Psychophysiological Responses to Repeated Sound Bursts," *ACUSTICA - acta Acustica* 85, 1999, p. 575-578.

[7] Pantelis N. Vassilakis, "Amplitude modulation depth versus degree of amplitude fluctuation: implementation error, adjustment and implications," *The Journal of the Acoustical Society of America* 08(5/2), 2000, p. 2597.

[8] Marine Taffou, Clara Suied & Isabelle Viaud-Delmon, "Auditory roughness elicits defense reactions," *Nature Scientific Reports* 11(1), 2021, p. 956.

[9] Goeffroy Peeters, "A large set of audio feature for sound description (similarity and classification) in the CUIDADO project," *Ircam, Analysis/Synthesis Team*, 2004.

[10] Harvey Fletcher & W. A. Munson, "Loudness, its definition, measurement and calculation," *The Journal of the Acoustical Society of America* 5, 1933, p. 82-108.

[11] Geoffroy Peeters, "A large set of audio feature for sound description (similarity and classification) in the CUIDADO project," *op. cit.*

[12] Alexander Stoessel, Romain David, Matthias Bornitz et al., "Auditory thresholds compatible with optimal speech reception likely evolved before the human-chimpanzee split," *Nature Scientific Reports* 13(1), 2023, p. 20732.

[13] Geoffroy Peeters, "A large set of audio feature for sound description (similarity and classification) in the CUIDADO project", *op.cit.*

[14] Kozo Hiramatsu, Koichi Takagi, Takeo Yamamoto & Jun Ikeno, "The effect of sound duration on annoyance," *Journal of Sound and Vibration* 59(4), 1978, p. 511-520; E. A. Bjork, "Effects of Inter-Stimulus Interval and Duration of Sound Elements on Annoyance," *ACUSTICA - acta acustica* 88, 2002, p. 104-109.

[15] Antonio Camurri, Emanuele Seminerio, Wanda Morganti et al., "Development and validation of an art-inspired multimodal interactive technology system for a multi-component intervention for older people: a pilot study," *Frontiers in Computer Science* 5, 2024, p. 1290589.

[16] Clara Suied et al., "Evaluating Warning Sound Urgency With Reaction Times," *Journal of Experimental Psychology Applied* 14(3), 2008, p. 201-212.

[17] Bruno H. Repp, "Sensorimotor synchronization: A review of the tapping literature," *Psychonomic Bulletin & Review* 12(6), 2005, p. 969-992.

[18] Daniel J. Cameron & Jessica Grahn, "Perception of Rhythm," in *The Cambridge Companion to Rhythm*, Cambridge, Cambridge University Press, 2020, p. 20-38.

[19] Paul Fraisse, "Perception and Estimation of Time," *Annual Review of Psychology* 35, 1983, p. 1-36.

[20] Stephen McAdams, "Psychological constraints on form-bearing dimensions in music," *Contemporary Music Review* 4, 1989, p. 181-198.

[21] Dick Botteldooren & Bert De Coensel, "Informational masking and attention focussing on environmental sound," *NAG/DAGA Proceedings*, 2009, p. 399-402.

[22] Antonio Camurri, Emanuele Seminerio, Wanda Morganti et al, "Development and validation of an art-inspired multimodal interactive technology system for a multi-component intervention for older people: a pilot study," *op. cit.*

[23] Kenneth Ooi, Zhen-Ting Ong, Karn N. Watcharasupat et al., "ARAUS: A Large-Scale Dataset and Baseline Models of Affective Responses to Augmented Urban Soundscapes," *IEEE Transactions on Affective Computing* 15(01), 2024, p. 105-120.

[24] https://urban-soundscapes.org

[25] Tomasz Kaczmarek & Anna Preis, "Annoyance of Time-Varying Road-Traffic Noise," *Archives of Acoustics* 35(3), 2010.

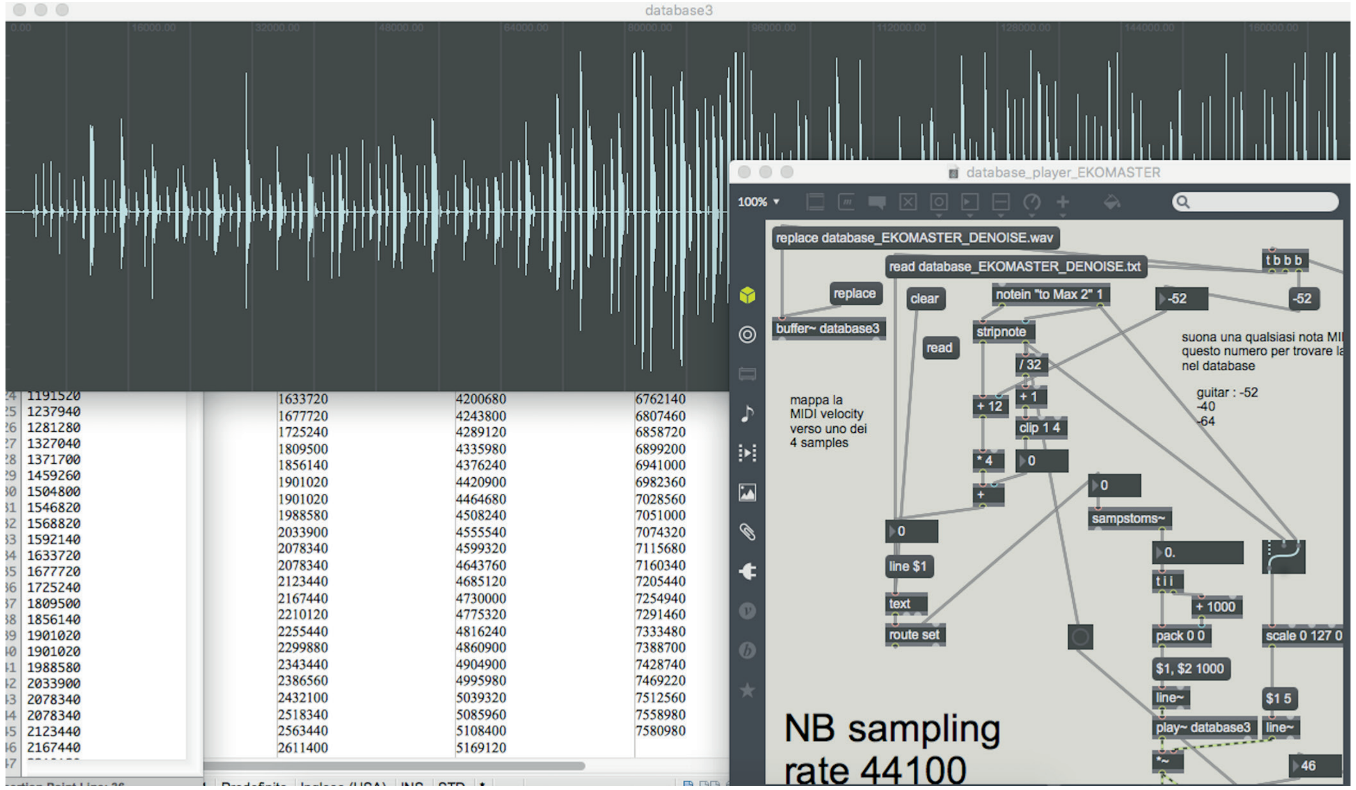
[26] https://cycling74.com/

[27] https://github.com/SkAT-VG/SDT

[28] https://entimement.dibris.unige.it

[29] https://resilence.eu

[30] Anders Friberg and Johan Sundberg, "Does music performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners", J. Acoust. Soc. Am. 105, p.1469–1484, 1999.



MaxMSP