Association of psychoacoustic metrics with Italian words describing perceptual sound attributes

Raffaele Mariconte[a]* | Giovanni Brambilla[b] | Diego Annesi[c] | Maurizio Diano[d] | Fabio Lo Castro[e] | Claudia Giliberti[a]

[a] INAIL, DIT,
Via Roberto Ferruzzi, 38-40, 00143 Roma
[b] DISAT, University of Milano-Bicocca,
Piazza della Scienza, 1, 20126 Milano
[c] INAIL, DiMEILA
Via di Fontana Candida, 1, 00040
Monte Porzio Catone (Roma)
[d] INAIL, Regional Directorates Calabria,
Via Vittorio Veneto, 60, 88100 Catanzaro
[e] CNR-INM, Section of Acoustics
and Sensors O.M. Corbino,
Via del Fosso del Cavaliere, 100,
00133 Roma
* Corresponding author:
r.mariconte@inail.it

Ricevuto: 6/2/2024
Accettato: 20/5/2024

DOI: 10.3280/ria1-2024oa17236
ISSNe: 2385-2615

Psychoacoustic parameters, being closely related to sound perception, are usually applied in product sound quality and, recently, also in environmental soundscape analysis or at workplace, to investigate its potential in describing acoustic comfort.
Lexicons of descriptive words of perceptual sound attributes are available in literature, but the language is often a crucial issue, being the translation not always easy to keep the original meaning.
This paper describes two different preliminary experiments dealing with such words in Italian and the evaluation of their association with psychoacoustic parameters. For these experiments, 12 sounds recorded in three different environments (at workplace, in nature and in the community) were selected and processed to determine some psychoacoustic parameters. These sounds were randomly played in a quiet room at the same equivalent level L_{eq} (dB) by headphone in the two experiments with the participation of two different groups of listeners, each formed by 24 subjects. Multivariate statistical analysis and correlation have been applied to compare their responses with some acoustic and psychoacoustic descriptors.
Keywords: psychoacoustics, sound quality, environmental noise, sound perception, descriptive words

Associazione di parametri psicoacustici con parole italiane descrittive di attributi sonori percepiti
I parametri psicoacustici, essendo correlati con la percezione del suono, sono di solito applicati alla qualità sonora del prodotto e, recentemente, anche nell'analisi del paesaggio sonoro ambientale o nell'esposizione sonora lavorativa, con l'intento di valutarne il loro potenziale nella descrizione del comfort acustico.
Lessici di parole descrittive della percezione di attributi sonori sono disponibili in letteratura, ma la lingua utilizzata è un fattore rilevante in quanto la traduzione da altre lingue, la più frequente è l'Inglese, non sempre è tale da mantenere il significato originario.
Questo articolo descrive due esperimenti preliminari riguardanti l'uso di queste parole nella lingua italiana e la loro associazione con i parametri psicoacustici. A tale scopo, 12 suoni registrati in tre diversi ambienti (in ambiente di lavoro, in natura e nella comunità) sono stati selezionati ed elaborati per determinarne alcuni parametri psicoacustici. Questi suoni sono stati riprodotti allo stesso livello equivalente L_{eq} (dB) in cuffia con modalità casuale in una stanza quieta in entrambi i due esperimenti con la partecipazione di due differenti gruppi, ciascuno composto da 24 ascoltatori. Analisi statistiche sui descrittori acustici e i responsi soggettivi sono state applicate per valutare la loro associazione.
Parole chiave: psicoacustica, qualità sonora, rumore ambientale, percezione sonora, attributi sonori

1 | Introduction

Descriptive words of sound perception are largely context-dependent, and most of them have no clear relationship to properties that acousticians know how to measure. Furthermore, the variability in sound expertise among individuals is typically large, from sound experts (acousticians, sound engineers and sound designers) to non-experts (consumers, naive people). The application field involved is also different, from the assessment of sound reproduction quality to product sound quality and soundscape analysis. Another important issue is the language, since the translation from other languages does not always keep the same concept and meaning. Indeed, the words must not only be translated well linguistically, but also must be adapted culturally to maintain the content validity at a conceptual level across different cultures [1]. For instance, dealing with soundscape the "Soundscape Attributes Translation Project" (SATP) [2] has been aimed at validating the first 15 translations of the soundscape attributes as reported in the ISO/TS 12913-2:2018 [3,4].

Consensus vocabularies are extensively applied in the field of sensory science [5].

 Articolo scientifico/Scientific paper

As stated in [6], “Although language is not a very precise tool for characterizing sounds there are many words for describing sounds. Their meaning may not be precisely the same from person to person and there may be unambiguous relations between words and attributes. We may or may not have words for all attributes and often the words we have (descriptors or labels) are not one dimensional. Anyway, it may be worthwhile to create a lexicon of sound describing words”.

The benefits and desired characteristics of consensus vocabularies to define attributes or descriptors used by assessors to characterize the sound perceptual differences are described in [7,8].

The frameworks for the description of everyday sounds available in the literature differ greatly in terms of their methodology and complexity [9].

Lexicons of descriptive words of perceptual sound attributes are available in literature (e.g., [10]). In particular, a systematic literal survey has been carried out for the characterization of everyday sounds developed in several research fields, including auditory cognition, soundscape research, artificial hearing, sound design and so forth [11]. At the time of running the present experiments, only one lexicon was retrieved for Italian words, selected from Italian Web Corpus 2016®, and concerning the perception of sounds in areas surrounding ports [12].

The extension to further environmental contexts than those in [12] was deemed interesting towards the development of an Italian lexicon of perceptual sound attributes.

In this framework, the present work describes an experimental study to explore:

- Italian descriptive words of perceptual attributes of sounds;
- for some attributes determine their perceived extent (into three categorical intervals);
- association of the selected attributes with acoustic and psychoacoustic metrics.

Two laboratory listening experiments have been carried out using 12 sounds, recorded in three different environments. The subjective responses collected in the listening tests have been compared with some acoustic and psychoacoustic metrics by statistical analyses (e.g., multivariate and correlation analyses) implemented in the “R” software [13].

2 | Materials and methods

2.1 | Sound samples

A set of 12 sound stimuli were used in the two laboratory listening experiments, each lasting 10 s. This duration was short primarily to prevent hearing fatigue during the listening session. However, during the test the participant was allowed to listen in loop the sound until the submitted questionnaire was filled in. The sounds were recorded in three different environments, namely in a working environment (W), in nature (N) and in the community (C) (Tab. 1). Some of the recordings were binaural (the community sounds mainly), others stereo or monoaural, as detailed in Table 1, being the last without spatial cues. However, this aspect was deemed not so crucial for the objectives of the listening tests, since none of the attributes proposed dealt with spatial sound features. Thus, for homogeneity purpose, the sound stimuli were processed into monaural tracks (binaural tracks were previously appropriately equalized to compensate the effects of ear canal and pinna) and normalized at the same sound level (–30 dB rms) by the Audacity® audio editor. Afterwards, they were imported in ArtemiS Suite® v14.1, setting 0.1 s time resolution and Fast time weighting, to determine various acoustic descriptors and six psychoacoustic metrics [14], namely loudness N [15], sharpness S [16], fluctuation strength F [17], roughness R [18], tonality T [19] and impulsiveness I [20]. The last three metrics (R, T, I) were computed according to the Hearing Sottek Model HSM, implementing a large number of highly overlapped critical-bandwidth filters [21]. The centre of gravity G of the 1/3 octave spectrum was determined in the frequency range 16-20,000 Hz by:

$$G = \frac{\sum_{k=bl}^{bh} f_k \cdot 10^{(L_k/10)}}{\sum_{k=bl}^{bh} 10^{(L_k/10)}} \quad [\text{Hz}] \tag{1}$$

where *bl* and *bh* are the lowest and the highest frequency bands delimiting the spectrum, respectively, f_k is the central frequency of the kth band and L_k is the corresponding band level.

Tab. 1 – Sound stimuli, type of recording and acoustic descriptors
Stimoli sonori, tecnica di registrazione e descrittori acustici

Environment	*Sound*	*Descriptors*
Work (W)	W1 Keyboard typing (MR)	L_{eq} [dB] L_{Aeq} [dB(A)] Dev. st. sL_A [dB(A)] N5 [sone GF] S average [acum] R* average [asper] F average [Vacil] T* average [tu] I* average [iu] 1/3 octave spectrum and its centre of gravity G [Hz]
	W2 Weaving loom (MR)	
	W3 Diesel engine (MR)	
	W4 Fan (MR)	
Nature (N)	N1 Seagulls (MR)	
	N2 River (MR)	
	N3 Rain (MR)	
	N4 Sea waves (SR)	
Community (C)	C1 Outdoor market (BR)	
	C2 Indoor metro (SR)	
	C3 Urban square (BR)	
	C4 Urban street (BR)	

*Hearing Sottek Model HMS [20]
MR = monoaural recordings, SR = stereo recordings, BR = binaural recordings

Listening tests in laboratory quiet room are time consuming, especially when the people sample is not too small. To get a reasonable number of subjects in short time and to facilitate their participation, instead of a specific room a quiet ordinary one, at participants’ convenience, was used to reduce transferring times to the laboratory. This approach required to setup a lightweight and portable sound reproduction system. The implemented system was formed by a digital audio player Creative Zen, connected to a semi-closed circumaural headphone AKG K 44. The system was calibrated in an anechoic chamber by a head and torso simulator (Fig. 1), and during the sound playback frequency equalization was applied to take into account the headphone frequency response.

Fig. 1 – Calibration of the sound reproduction system used for the listening tests
Calibrazione del sistema di riproduzione sonora utilizzato nelle prove di ascolto

2.2 | Listening tests

The 12 sounds were presented diotically by binaural headphone in a random order to minimize the bias on responses due to the presentation order [22]. The test room was quiet, without any significant sound interfering with the listening. The listening sound level was fixed and previously calibrated (L_{eq10s} = 73 dB). At the beginning of the listening test, after the introduction and the instructions given by the experimenter, each sound was one shot played for 10 s and, on participant request, loop listening was available until the questionnaire was filled in (Fig. 2).

In the selection of participants, those with high education were preferred, taking into account the task to be performed, requiring a deep knowledge of the language suitable to run the session without too much assistance of the experimenter.

Fig. 2 – Experimental setup of the listening test
Configurazione sperimentale delle prove di ascolto

2.2.1 | Experiment 1

In the questionnaire the participant was asked to select in a list of 22 Italian words (Tab. 2) those considered most appropriate to describe her/his perception of the sound just heard, without assigning any rank order among the selected ones. In Tab. 2 a possible English translation, within (), is reported only in order to be consistent with the language used to write this this paper, but not shown to the listeners. The words in Italian were presented in alphabetical order. The option to indicate other words not present in the list was also available. In choosing the 22 descriptive words of sound perception, the outcome of the study in [12] was taken into account.

Tab. 2 – List of Italian words proposed to the participant to describe the sound attribute perception; English words within () not shown to the listeners
Elenco delle parole in italiano proposte al partecipante per descrivere gli attributi sonori percepiti; parole in inglese tra () non mostrate agli ascoltatori

Chiaro (clear)	Familiare (familiar)	Invadente (intrusive)	Squillante (shrill)
Confuso (blur)	Fastidioso (annoying)	Naturale (natural)	Stabile (steady)
Conosciuto (known)	Forte (loud)	Noioso (boring)	Variabile (fluctuating)
Cupo (dull)	Gradevole (nice)	Piacevole (pleasant)	Vivace (lively)
Debole (soft)	Innaturale (unnatural)	Sibilante (hissing)	Altro (other)
Discreto (fair)	Insolito (unusual)	Sgradevole (ugly)	

Twenty four subjects (average age 40 ± 14 years) participated to this experiment, 50% male and 54% with degree or PhD education. They self-reported an average noise sensitivity of 6.5 (± 1.8) on a scale from 1 (not at all) to 10 (very much). The average duration of the listening sessions was 9 minutes and 30 s, with an average response time of 30 s for the choice of each attribute. Notwithstanding the possibility to indicate other attributes not present in the proposed list, this option was chosen only by 8 (33.3%) participants for a total of 9 different descriptive words.

2.2.2 | Experiment 2

Among the various psychometric scaling [23], such those applied in [24,25], in this experiment the semantic differential bipolar scale [26] was used as tool to collect the perceived extent of antonymic sound attributes located at the opposite ends of the scale. Respondents were asked to choose one of the seven equal intervals on the bipolar scale indicating the perceived extent of the selected attribute. The central interval corresponds to the neutral "Neither/Nor" response and the answer coding was not shown to the participants but used only for rating processing (Fig. 3).

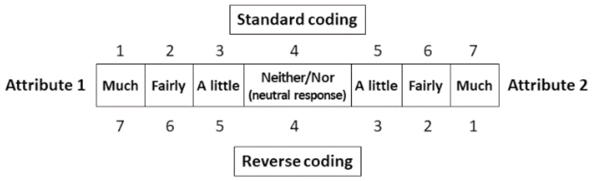

Fig. 3 – Semantic bipolar scale used in experiment 2
Scala semantica bipolare impiegata nell'esperimento 2

The proposed seven scales with perceptual attributes, in Italian words, are reported in Tab. 3 and they were chosen taking into account the outcome of experiment 1, with translation of the attributes in English words, reported within (), herewith given only in order to be consistent with the language used to write this paper, but not shown to the subjects.

Tab. 3 – Bipolar semantic scales in Italian used for the ratings; English words within () not shown to the subjects
Scale semantiche bipolari in italiano utilizzate per le valutazioni; parole in inglese tra () non mostrate ai partecipanti

Scale	*Attribute 1*	*Attribute 2*	*Coding*
S1	Sconosciuto (unknown)	Conosciuto (unknown)	Standard
S2	Cupo (dull)	Sibilante (hissing)	Standard
S3	Sgradevole (ugly)	Gradevole (nice)	Standard
S4	Stabile (steady)	Variabile (fluctuating)	Standard
S5	Noioso (boring)	Vivace (lively)	Standard
S6	Piacevole (pleasant)	Fastidioso (annoying)	Reverse
S7	Confuso (blur)	Chiaro (clear)	Standard

A group of 24 subjects (average age 35 ± 15 years) participated to this experiment, 50% male and 46% with degree or PhD education. None of these participants took part in the experiment 1. They self-reported an average noise sensitivity of 7.0 (± 2.1) on a scale from 1 (not at all) to 10 (very much).

The average duration of the listening sessions was 9 minutes and 10 s with an average response time of 46 s for each sound.

3 | Results and discussion

3.1 | Sound stimuli

Figure 4 shows the range of the L_{Aeq} levels in dB(A) of the sound stimuli versus their centre of gravity G, Eq. (1), of the 1/3 octave band spectrum, with the coloured scale reporting the standard deviation s of the $L_{A,0.1s}$ levels. A reasonable separation among sound stimuli and the recording environments is observed.

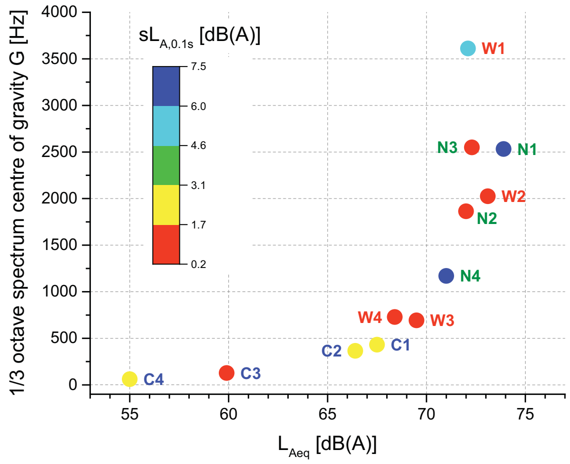

Fig. 4 – Range of sound L_{Aeq} levels in dB(A) versus the centre of gravity G of the 1/3 octave band spectrum
Variabilità dei livelli continui equivalenti L_{Aeq} in dB(A) con il centro di gravità G dello spettro a 1/3 di ottava

Further outcomes are provided by the Pearson's correlation matrix of the nine sound descriptors (Fig. 5), ordered by hierarchical clustering (agglomeration algorithm Ward D2 [27]) into four groups. The group at the top-left of the correlogram includes the continuous equivalent level L_{Aeq} and the 5th percentile of loudness N5, both describing the energy content of the sound. The frequency features of the sound are described by the group including sharpness S and spectrum centre of gravity G, as well as the group formed by roughness R and tonality T. The sound temporal features are represented by the group at the bottom-right of the correlogram, including impulsiveness I, standard deviation s of $L_{A,0.1s}$ sound levels and fluctuation strength F. This outcome confirms that at least three perceptive dimensions, namely the noise intensity, its temporal variations and its spectrum, emerge when dealing with sound environment assessments [28].

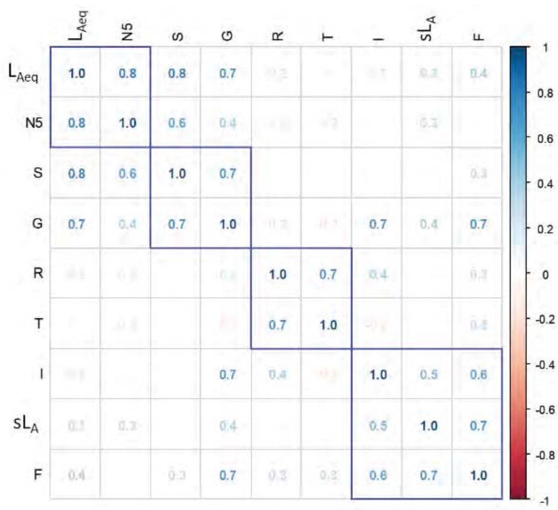

Fig. 5 – Pearson's correlation matrix of the 9 sound descriptors, ordered by Ward D2 hierarchical clustering into four groups
Matrice di correlazione di Pearson dei 9 descrittori acustici, aggregati nei quattro gruppi ottenuti con l'analisi gerarchica con algoritmo Ward D2

3.2 | Experiment 1

Dealing with the sound attributes selected by the participants, as shown in Fig. 6 the sounds were rather Familiar or Known to the participants (19.4% of all the 1159 responses). Thus, it may be plausible that the chosen attributes were the outcome of an aware selection rather than one by chance. The sounds recorded in the working environment were mainly described as Intrusive (82.4% of all the 51 selections of this attribute), Ugly (73.7% of all the 57 selections of this attribute) and Annoying (70.4% of all the 71 selections of this attribute). Natural sounds were described as Nice (73.8% of all the 65 selections of this attribute) and Pleasant (71.9% of all the 64 selections of this attribute). Community sounds were described mainly Blur (56.7% of all the 67 selections of this attribute) and Fluctuating (54.0% of all the 50 selections of this attribute), as well as more Lively (16.1% of all the 31 selections of this attribute) than Annoying (14.1%).

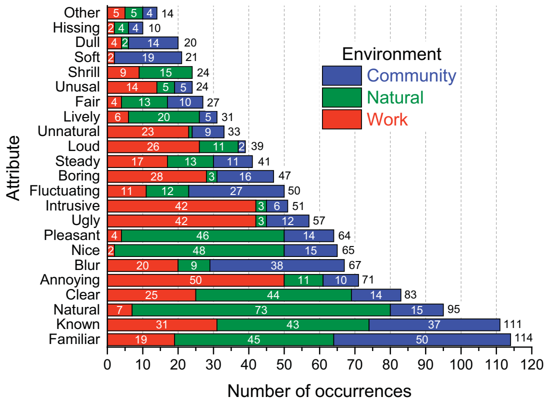

Fig. 6 – Number of occurrences of the selected attributes
Occorrenze degli attributi prescelti per i suoni

The attribute occurrences were also processed by the Principal Component Analysis [29], an unsupervised learning algorithm technique used to examine the interrelations among a set of variables. For this purpose the package FactoMineR was used [30]. For each attribute, the occurrence values observed for the twelve sounds were normalized by subtracting their mean and dividing by their standard deviation. The outcome of the first two PCA dimensions, explaining 62.7% of the dataset variability, is reported in Figure 7.

Attributes distant from the circle centre are well represented on the factor map. Those giving similar information are grouped together (like Nice/Pleasant and Intrusive/Annoying), and if negatively correlated they are positioned on opposite sides of the circle centre (like Pleasant/Annoying), in diagonally opposed quadrants (like Clear/Bur).

Table 4 reports for each recording environment the four most occurred attributes with their percentage referred to the total responses on each attribute.

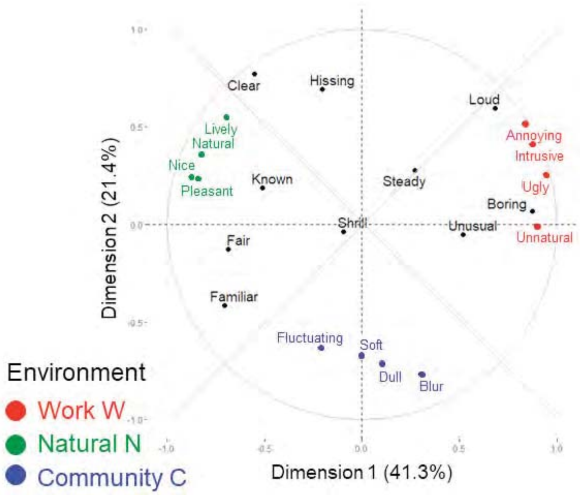

Fig. 7 – First two PCA components of the attributes based on their occurrences
Prime due componenti dell'analisi PCA sugli attributi in base alle loro occorrenze

Tab. 4 – Four most occurred attributes (percentage referred to the total responses on each attribute) for each recording environment
Quattro attributi più ricorrenti (percentuale riferita al totale delle risposte per ciascun attributo) per ciascun ambiente dei suoni

Environment	*First four most occurred attributes (percentage referred to the total responses on each attribute)*
Work (W)	Intrusive (82.4%)
	Ugly (73.7%)
	Annoying (70.4%)
	Unnatural (69.7%)
Nature (N)	Natural (76.8%)
	Nice (73.8%)
	Pleasant (71.9%)
	Lively (64.5%)
Community (C)	Soft (90.6%)
	Dull (70.0%)
	Blur (56.7%)
	Fluctuating (54.0%)

The four most occurred attributes of sounds in the working environment are in the top-right quadrant, opposite to that top-left where the four most occurred attributes of natural sounds are located, whereas the four most occurred attributes of community sounds are in-between in the bottom quadrants.

A further analysis dealt with the Spearman rank correlation matrix on all the variables, that is attributes and noise descriptors (Fig. 8), where the variables are displayed according to their hierarchical clustering order (Ward D2 agglomeration method) considering two groups, highlighted by the two rectangles with blue borders. For instance, a sound described as Natural is positively correlated with Pleasant ($r = 0.9$) and Lively ($r = 0.8$), whereas r is negative for Ugly (–0.8) and Boring (–0.7). The correlation points out also the correspondence among synonyms (e.g., Nice and Pleasant with $r = 0.9$).

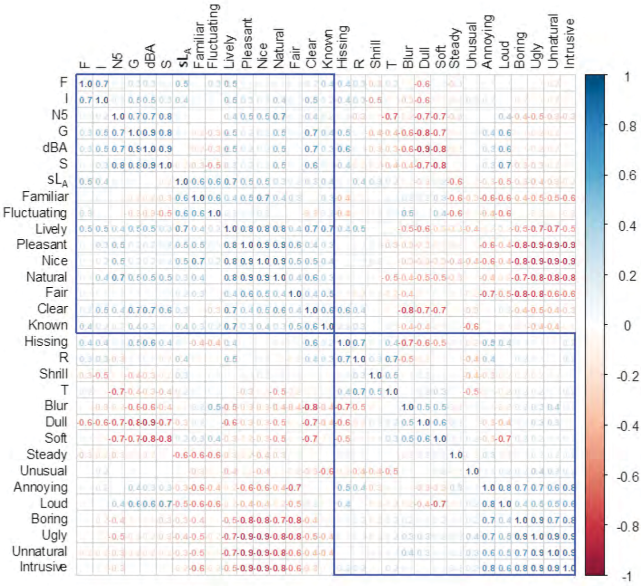

Fig. 8 – Spearman rank correlation matrix on all the variables (attributes and acoustic descriptors)

Correlogramma di Spearman tra tutte le variabili (attributi e descrittori acustici)

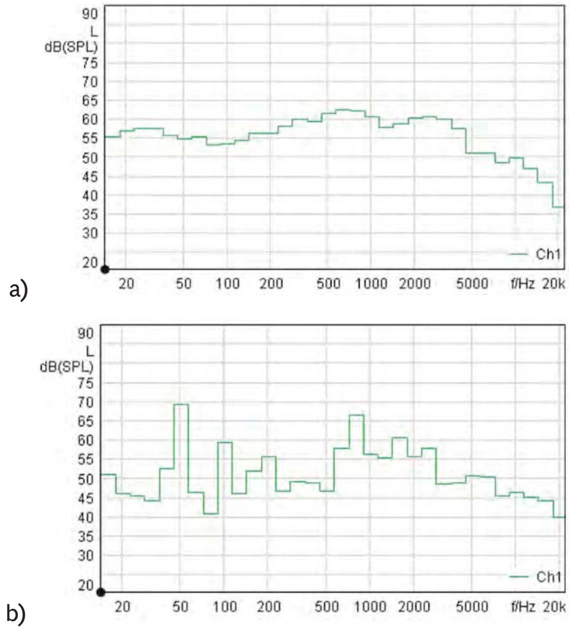

Fig. 10 – 1/3 octave band spectrum of L_{eq}: Sea Waves (a) and Diesel Engine (b)

Spettri L_{eq} a 1/3 di ottava: Onde Marine (a) e Motore Diesel (b)

On the whole, the Sea Waves (N4) was the sound rated mostly Nice and Pleasant (23.1% and 21.9% of all the 65 and 64 selections of these attributes, respectively). On the contrary, the Diesel Engine (W3) was the sound rated mostly Intrusive and Ugly (29.4% and 28.1% of all the 15 and 16 selections of these attributes, respectively). Regarding these two sounds, Fig. 9 and 10 report the FFT spectrograms and the 1/3 octave band spectrum of L_{eq}, respectively.

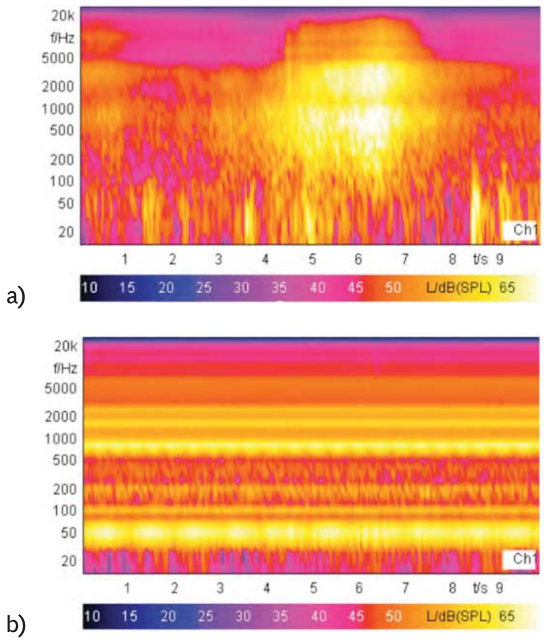

Fig. 9 – The FFT spectrograms of sounds: Sea Waves (a) and Diesel Engine (b)

Spettrogrammi FFT dei suoni: Onde marine (a) e Motore Diesel (b)

The spectrum of diesel engine shows tonal components not present in the one of sea waves and their presence in the former confirms the enhanced annoyance they produce, in addition to the type of source and its semantic content.

3.3 | Experiment 2

Figure 11 reports the percentages of occurrences observed for each bipolar scale pooling all the responses to all sounds. Most of the stimuli were much (60.1%) and fairly (19.1%) Known for the participants. It may be plausible that their ratings could not be influenced by hearing unknown sounds. Moreover, the neutral response Neither/Nor was reported only 12.7% of the 2014 collected ratings, instead of the potential 2016 due to two missing responses, indicating that the participants were often

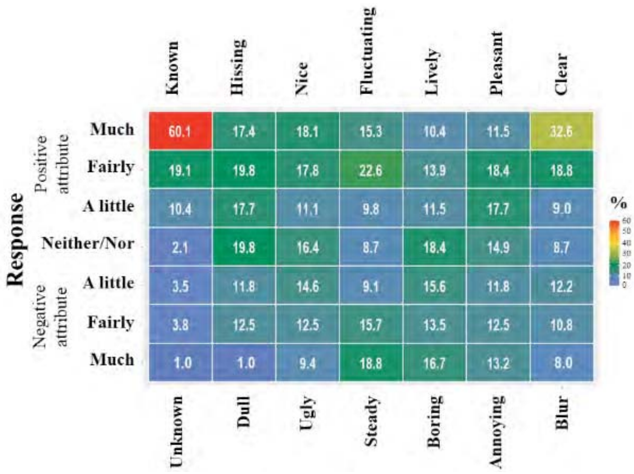

Fig. 11 – Percentages of occurrences for each bipolar scale pooling all the responses to all sounds

Percentuali di occorrenze per ciascuna scala bipolare di tutte le risposte all'insieme dei suoni

able to select the magnitude of the proposed attributes on the semantic scales. This could be likely also because a large part of respondents (43.1%) found the sounds fairly (18.8%) or much Clear (32.6%), rather than fairly (10.8%) or much (8.0%) Blur.

The highest percentage of neutral response Neither/Nor (19.8%) was observed for the scale Dull/Hissing. The 12 sounds were, overall, more often Annoying (45.8%) than Pleasant (35.8%).

More hints may be drawn considering the mean scores on the seven scales, versus the three different sound recording environment (Fig. 12). It is clear that the perception of natural sounds was rated more positively (37.0%) than that in the working environment, whereas the sounds in the community is rated in-between. The less clear sounds were those recorded in the community environment but, still, close to the neutral response ($\overline{S7}$ = 3.9), whereas those in the working environment are the most Annoying ($\overline{S6}$ = 2.9) and Ugly ($\overline{S3}$ = 3.1).

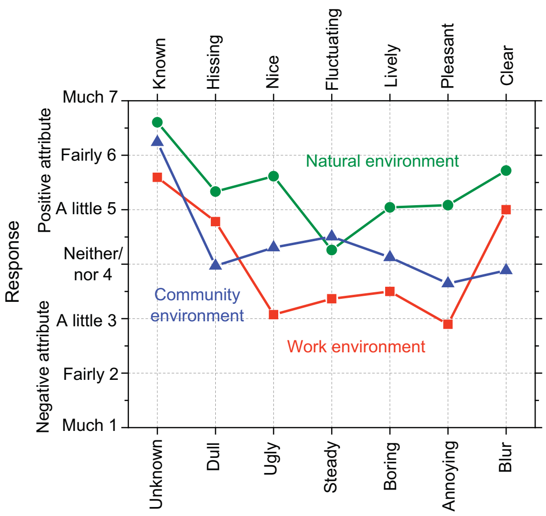

Fig. 12 – Mean score for each bipolar scale versus the sound recording environment

Valutazione media per ciascuna scala bipolare di tutte le risposte diversificata per l'ambiente dei suoni

A more detailed analysis has been performed for the sounds recorded in the same environment, as shown in Fig. 13. Regarding the work environment (Fig. 13a), the diesel engine noise (W3) was rated the most Annoying, Ugly and Steady, whereas the fan noise (W4) was the most Unknown, Dull and rated as neutral on the scale Blur/Clear ($\overline{S7}$ = 4.0). The natural sounds were rated positively almost on all the scales (Fig. 13b). It is confirmed that people enjoy the perception of natural sounds. The highest overall mean scores were observed for the seagull sound (N1), namely for Hissing ($\overline{S2}$ = 6.3), Fluctuating ($\overline{S1}$ = 5.0), Lively ($\overline{S5}$ = 6.1), and Clear ($\overline{S7}$ = 6.2) attributes, whereas the rain sound (N3) was rated slightly Steady ($\overline{S1}$ = 3.8), Lively ($\overline{S5}$ = 4.2), Pleasant ($\overline{S6}$ = 4.5) and Clear ($\overline{S7}$ = 5.1). The outcomes observed for community noises are in-between the previous ones (Fig. 13c). The urban square sound (C3) was the most Nice ($\overline{S3}$ = 5.4) and the outdoor market sound (C1) was rated the most Lively ($\overline{S5}$ = 4.2).

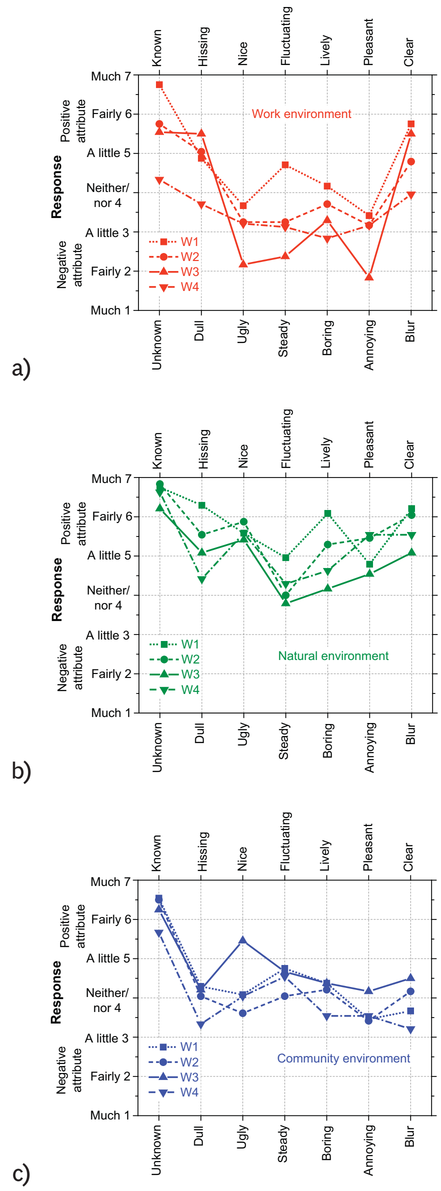

Fig. 13 – Mean score for each bipolar scale S for each sound in the three recording environments

Valutazione media per ciascuna scala bipolare S per ciascun suono nei tre diversi ambienti

Moreover, the Spearman rank correlation matrix (Fig. 14) was computed considering the mean of coded scores obtained on the bipolar semantic scales for each sound and the corresponding acoustic parameters listed in Table 1. The variables are displayed according to their hierarchical clustering order (Ward's D2 agglomeration method) considering two groups, highlighted by the two rectangles with blue borders.

In the top-left group, only positive correlations are observed between two scales (S2 and S7) and four acoustic descriptors, namely loudness N5, sharpness S, L_{Aeq} and spectrum centre of gravity G. In particular, scale S2 (Dull/Hissing) has a good correlation (r = 0.85) with S7 (Blur/Clear), L_{Aeq} (r = 0.81), S (r = 0.70), G (r = 0.69) and N5 (r = 0.50). It is worth to note that sharpness S and perceived hissing feature of sound, rated on scale S2, are positively correlated (r = 0.70).

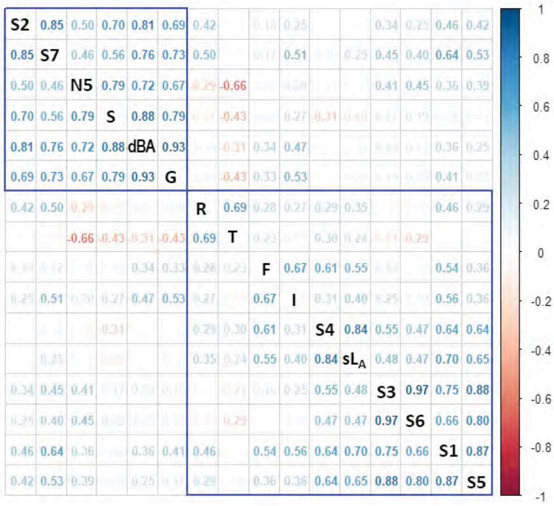

Fig. 14 – Spearman rank correlation matrix between mean scores on semantic scales and acoustic descriptors

Correlogramma di Spearman tra le valutazioni medie sulle scale semantiche e i descrittori acustici

The remaining five scales are in the bottom-right group of the correlation matrix, including also roughness R, tonality T, fluctuation strength F, impulsiveness I and standard deviation s of sound level $L_{A,01s}$. As expected, scales S3 (Ugly/Nice) and S6 (Annoying/Pleasant) are well correlated (r = 0.97), as well as with S5 (Boring/Lively) with r = 0.88 (S3/S5) and r = 0.80 (S6/S5). Thus, a sound perceived as Lively is likely to be rated as Pleasant and Nice. The standard deviation of sound level $sL_{A,0.1s}$ is positively correlated (r = 0.84) with scale S4 (Steady/Fluctuating).

The above results show that the set of acoustic descriptors considered can be reasonably associated to the descriptive words of sound perception as reported by participants in these two experiments, and they can quantify the various perceptual dimensions (e.g., time and frequency patterns, semantic content).

4 | Conclusions

The two experiments above described are a preliminary approach towards the development of a lexicon of Italian descriptive words of the perception of environmental sounds.

Even though the results cannot be generalized, due to the small number of sound samples and limited to three different contexts only, they show that the set of acoustic descriptors considered correspond satisfactorily to the perceptual sound attributes selected and quantified by the participants to describe the various perceptual dimensions, namely the noise intensity, its temporal variations and its spectrum.

Further investigations are planned to enlarge the sample of sounds under test and the listening panel, to improve the perceptual descriptive words set and to apply further experimental protocols and statistical analyses.

Conclusioni

I due esperimenti descritti costituiscono la fase preliminare di uno studio volto a formulare un lessico di parole italiane descrittive della percezione di suoni ambientali.

Allo stato attuale i risultati ottenuti, pur se non generalizzabili per il numero ridotto dei suoni e dei contesti di registrazione esaminati, sono interessanti. L'insieme dei descrittori acustici considerati, infatti, appare abbastanza associabile agli attributi prescelti e quantificati dai partecipanti per descrivere la loro percezione sonora in termini di intensità, andamento nel tempo e in frequenza.

Si intende pianificare ulteriori esperimenti per ampliare sia l'insieme dei suoni da analizzare che il gruppo di ascoltatori al fine di approfondire il lessico da sviluppare, utilizzando anche ulteriori protocolli sperimentali e analisi statistiche.

References

[1] D.E. Beaton, C. Bombardier, F. Guillemin, M.B. Ferraz, Guidelines for the process of cross-cultural adaptation of self-report measures, Spine 2000, 25, pp 3186–91, DOI: 10.1097/00007632-200012150-00014.

[2] F. Aletta et al., Soundscape assessment: towards a validated translation of perceptual attributes in different languages, Proc. 49th InterNoise, Seoul, Korea, 23-26 August 2020.

[3] ISO/TS 12913-2:2018, Acoustics-Soundscape Part 2: Data collection and reporting requirements.

[4] G.E. Puglisi, L. Shtrepi, M.C. Masoero, A. Astolfi, Evaluating soundscape in the Italian language: Validation of the translation of the English standardized perceptual attributes of the ISO/TS 12913-2:2018 and comparison with other Romance languages, Applied Acoustics 222, 11050, 2024, DOI: 10.1016/j.apacoust.2024.110050.

[5] H.T. Lawless, H. Heymann, Sensory evaluation of food – Principles and practices, 2010, Springer.

[6] T.H. Pedersen, N. Zacharov, How many psycho-acoustic attributes are needed?, Proc. Acoustics '08, Paris, France, 29 June-4 July 2008.

[7] N. Zacharov, T.H. Pedersen, C. Pike, A common lexicon for spatial sound quality assessment – latest developments, 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX), Lisbon, Portugal, 2016, DOI: 10.1109/QoMEX.2016.7498967.

[8] T.H. Pedersen, N. Zacharov, The development of a sound wheel for reproduced sound, 138th Conv. Audio Eng. Soc., 2015, Warsaw, Poland, 7-10 May.

[9] C. Guastavino, Everyday sound categorization, in Computational Analysis of Sound Scenes and Events, Eds T. Virtanen, M. D. Plumbley, D. Ellis, New York, NY: Springer International Publishing, 2018, pp. 183–213, DOI: 10.1007/978-3-319-63450-0.

[10] T.H. Pedersen, Lexicon of Sound-Describing Words – Version 1, Delta Report AV 11/05 THP, 2008.

[11] B.L. Giordano, R. de Miranda Azevedo, Y. Plasencia-Calaña, E. Formisano, M. Dumontier, What do we mean with sound semantics, exactly? A survey of taxonomies and ontologies of everyday sounds, Front. Psychol., 13, 2022, DOI: 10.3389/fpsyg.2022.964209

[12] A. Magrini, G. Di Feo, A. Cerniglia, Questionnaire analysis survey for acoustic investigation-Preliminary considerations, Proc. 48th InterNoise, Madrid, Spain, 16-19 June 2019.

[13] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2021.
[14] H. Fastl, E. Zwicker, Psychoacoustics. Facts and Models, Springer Berlin, Heidelberg, 2007.
[15] DIN 45631/A1:2010-03, Calculation of loudness level and loudness from the sound spectrum – Zwicker method – Amendment 1: Calculation of the loudness of time-variant sound.
[16] DIN 45692:2009-08, Measurement technique for the simulation of the auditory sensation of sharpness.
[17] R. Sottek, J. Becker, T. Lobato, Progress in Roughness Calculation, Proc. 49th InterNoise, Seoul, South Korea, 23-26 August 2020.
[18] H. Fastl, E. Zwicker, Fluctuation Strength, Chap. 10 in Psychoacoustics. Facts and Models, Springer Berlin, Heidelberg, 2007.
[19] ECMA-418-2, Psychoacoustic metrics for ITT equipment — Part 2 (models based on human perception), 2020.
[20] R. Sottek, T. Moll, Perception and modeling of impulsive sounds, Proc. Euronoise 2015, Maastricht, Netherlands, 31 May-3 June 2015.
[21] R. Sottek, K. Genuit, Models of signal processing in human hearing, AEU – International Journal of Electronics and Communications, 59(3), pp. 157-165, 2005.
[22] R. Zeelenberg, D. Pecher, A method for simultaneously counterbalancing condition order and assignment of stimulus materials to conditions, Behav. Res., 47 (127), pp. 127-133, 2015.
[23] A. Magrini, G. Di Feo, A. Cerniglia, Acoustic Quality of the External Environment: Indications on Questionnaire Structure for Investigating Subjective Perception, J. Otorhinolaryngol. Hear. Balance Med., 4, 4, 2023, DOI: 10.3390/ohbm4010004.
[24] U. Ayr, E. Cirillo, F. Martellotta, Disturbo da rumore. Verifica sperimentale di una scala per la valutazione soggettiva, CDA, Sezione Ricerca, Settembre 2000, pp. 925-932.
[25] D. Dal Palù, B. Lerma, L. Actis Grosso, L. Shtrepi, M. Gasparini, C. De Giorgi, A. Astolfi, Sensory evaluation of the sound of rolling office chairs: an exploratory study for sound design, Applied Acoustics 130, pp. 195-203, 2018, DOI: 10.1016/j.apacoust.2017.09.027.
[26] C.E. Osgood, The nature and measurement of meaning. Psychol. Bull. 49, pp. 197–237, 1952.
[27] J.H. Ward Jr, Hierarchical Grouping to Optimize an Objective Function, Journal of the American Statistical Association, 58, pp. 236–244, 1963, DOI: 10.1080/01621459.1963.10500845
[28] A. Can, P. Aumond, S. Michel, B. de Coensel, C. Ribeiro, D. Botteldooren, C. Lavandier, Comparison of noise indicators in an urban context, Proc. 45th InterNoise, Hamburg, Germany, 21-24 August 2016.
[29] I.T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, Phil. Trans. R. Soc. A, 374, 20150202, 2016, DOI: 10.10980/rsta.2015.202.
[30] S. Lê, J. Josse, F. Husson, FactoMineR: A Package for Multivariate Analysis. Journal of Statistical Software, 25, pp. 1-18, 2008, DOI: 10.18637/jss.v025.i01.