DAVIDE CHIARELLA,
PAOLA CUTUGNO,
MELISSA FERRETTI
ILC (Istituto di Linguistica Computazionale) sede secondaria di Genova, CNR (Consiglio Nazionale delle Ricerche) Via De Marini 6, Genova ITALY.
e-mail: davide.chiarella@ilc.cnr.it, paola.cutugno@ilc.cnr.it, ferrissa.mf@gmail.com

A linguistic approach of sound characterisation and polarization: first steps

Introduction

The activities of the "TRIPLO: TRasporti e collegamenti Innovativi e sostenibili tra Porti e piattaforme LOgistiche" project, funded with funds from the Interregional Operational Programme Italy-France Maritime 2014-2020, have as their particular goal to increase the sustainability of commercial ports and associated logistic platforms, helping to lessen noise pollution [1][2]. In some project activities, the acoustic impact on the people exposed to noise from back port activities is evaluated in connection to how each person perceives the noise.
Only technical investigations, which cannot ensure a phenomena's universality in terms of perception, can objectively describe a phenomenon in environmental surveys [3]; A sound can be viewed as both a physical reality that can be measured using objective criteria and a sound perception phenomenon that is of a subjective character and related to the subject's psycho-physical-emotional state. Because these two traits are inextricably linked, it is not enough to just look at them independently.
Driven by these motivations, we created questionnaires concerning the perception of sounds, the structure and first results of which can be consulted in [4] [5] [6]. In this article, in the first part we present a methodology to identify adjectives characterising each sound via TF-IDF (term frequency - inverse document frequency) [7][8][9][10]; in the second part we analyse the positive or negative emotions described by the adjectives given for each sounds with TexBlob, a sentiment analysis classifier, and subsequently we compare the results obtained with the ones shown in [6].

Characterising sounds by means of adjectives: TF-IDF

In analysing the results of the questionnaires, we wanted to check whether it was possible to identify the most characterising adjectives for each of the sounds using the TF-IDF (term frequency – inverse document frequency). The TF-IDF algorithm allows us to measure the importance of a term within a collection of documents: in our case the term is an adjective and the documents are the set of adjectives referring to the sound we wanted to characterise. This metric, in fact, allows us to associate a word with a higher weight when it is more frequent in the document and when the same word is more absent in the other documents (thus identifying the most relevant for a given document).

In detail, the TF-IDF function is defined by two distinct parts:

TF = (# occurrences of the term t in the document) / (# terms in documents)

IDF = log(# documents / # documents containing the term t)

The result of the function gives rise to 21 vectors, one for each sound, consisting of the 8 adjectives weighted with the TF-IDF function. All vectors are represented in Table 1.

Sounds	**First 8 adjectives from TD-IDF**							
Attracco Docking	fastidioso	rumoroso	ripetitivo	forte	rimbombante	cupo	metallico	ridondante
Camion Truck	fastidioso	rumoroso	metallico	caotico	elettrico	rilassante	basso	profondo
Campane Bells	festoso	domenicale	fastidioso	risuonante	gioioso	forte	rumoroso	rindondante
Cigolio Plastico Plastic squeaking	fastidioso	stridente	stridulo	cigolante	strano	acuto	gommoso	rilassante
Container	fastidioso	rumoroso	veloce	sordo	cittadino	meccanico	pesante	normale
Corde Ropes	scricchiolante	fastidioso	teso	strano	legnoso	rilassante	antico	vecchio
Gabbiani Seagulls	naturale	rilassante	acuto	fastidioso	marittimo	piacevole	stridulo	ripetitivo
Grilli Crickets	rilassante	naturale	tranquillo	sereno	acuto	piacevole	pacifico	fastidioso
Gru Crane	fastidioso	ripetitivo	allarmante	intermittente	acuto	crescente	sopportabile	ritmico

Mare calmo Calm sea	rilassante	calmo	naturale	piacevole	fresco	bello	tranquillo	estivo
Motosega Chainsaw	fastidioso	assordante	metallico	tagliente	penetrante	acuto	stridulo	rumoroso
Muletto Forklift	fastidioso	rumoroso	acuto	assordante	stridulo	metallico	cupo	spiacevole
Orologio Watch	ritmico	fastidioso	ripetitivo	ritmato	preciso	regolare	noioso	elettronico
Passaggio a livello Level crossing	fastidioso	ripetitivo	squillante	acuto	allarmante	forte	pericoloso	metallico
Pioggia Rain	rilassante	bagnato	fastidioso	continuo	piacevole	piovoso	scrosciante	naturale
Rubinetto aperto Tap open	rilassante	fastidioso	bagnato	tranquillo	continuo	piacevole	veloce	morbido
Rubinetto che perde Tap leaking	ripetitivo	fastidioso	rilassante	gocciolante	monotono	snervante	ritmico	bagnato
Sirena nave Ship siren	fastidioso	forte	rumoroso	assordante	marittimo	navale	intenso	lontano
Traffico Traffic	fastidioso	rumoroso	forte	urbano	cittadino	rilassante	ventoso	metallico
Treno Train	fastidioso	rumoroso	stridente	metallico	forte	acuto	ripetitivo	assordante
Vento tra le foglie Wind in the leaves	naturale	rilassante	piacevole	gradevole	libero	confuso	fastidioso	acuto

Table 1

Intersecting each set with all the others, we have considered, as characterising each sound, the adjectives that appear only in one vector. This result can be expressed by means of the difference between the previously identified adjectives for each sound ($A_{i,8}$) and the intersection of all the sets expressing each sound ($A_{j,8}$ $j \neq i$).

Intersecando ciascun insieme con tutti gli altri abbiamo considerato come caratterizzanti ciascun suono gli aggettivi unici rispetto a tutti gli altri. Dal punto di vista insiemistico questo risultato può esser espresso mediante la differenza tra gli aggettivi identificati precedentemente per ciascun suono ($A_{i,8}$) e l'intersezione fra tutti gli insiemi che esprimono ciascun suono ($A_{j,8}$ $j \neq i$). That is, we considered the complementary of each intersection to identify the characterising adjectives:

$$B_i = A_{i,8} - (A_{i,8} \cap A_{j,8\ j \neq i}) \text{ con } i = \{1, 2, 3 \ldots 21\} \quad j = \{1, 2, 3 \ldots 21\} \quad i,j \in N$$

The result of the operation performed produces the vectors shown in Table 2.

Sounds	$A_{i,8} - (A_{i,8} \cap A_{j,8\ j \neq i})$							
Attracco Docking	rimbombante							
Camion Truck	caotico	elettrico	basso	profondo				
Campane Bells	festoso	domenicale	risuonante	gioioso				
Cigolio Plastico Plastic squeaking	cigolante	gommoso						
Container	sordo	meccanico	pesante	normale				
Corde Ropes	scricchiolante	teso	legnoso	antico	vecchio			
Gabbiani Seagulls	naturale	rilassante	acuto	fastidioso	marittimo	piacevole	stridulo	ripetitivo
Grilli Crickets	sereno	pacifico						
Gru Crane	intermittente	crescente	sopportabile					

Mare calmo Calm sea	calmo	fresco	bello	estivo				
Motosega Chainsaw	tagliente	penetrante						
Muletto Forklift	spiacevole							
Orologio Watch	ritmato	preciso	regolare	noioso	elettronico			
Passaggio a livello Level crossing	squillante	pericoloso						
Pioggia Rain	piovoso	scrosciante						
Rubinetto aperto Tap open	morbido							
Rubinetto che perde Tap leaking	gocciolante	monotono	snervante					
Sirena nave Ship siren	navale	intenso	lontano					
Traffico Traffic	urbano	ventoso						
Treno Train	fastidioso	rumoroso	stridente	metallico	forte	acuto	ripetitivo	assordante
Vento tra le foglie Wind in the leaves	gradevole	libero	confuso					

Table 2

Analysing the results obtained, we can state that where the source of the noise is recognised and not linked by a personal memory, the adjectives chosen, in this case by the students, were for the most part linked to the characteristics of the source that emitted them: two examples for all can be represented by the “motosega” (i.e. chainsaw) vector (Ita: [tagliente, penetrante] Eng: [sharp, penetrating]) and the “Sirena nave” (i.e. ship siren) vector (Ita: [navale, intense, lontano] Eng: [naval, intense, distant]).

As proof of this, the vector representing “campane” (i.e. bells) shows that the source that issued it was recognised and evokes in the student personal memories of festivity; indeed, the adjectives used are [festoso, domenicale, risuonante, gioioso] (Eng: [festive, Sunday, resounding, joyful]). Similarly, for the vector representing “mare calmo” (i.e. calm sea), in addition to having recognised the source, the student visualised the calm and refreshing characteristics of the sea by associating it with beauty and summer (Ita: [calmo, fresco, bello, estivo] Eng:[calm, refreshing, beautiful, summer]).

Conversely, where the source that produced the sound played was not recognised, the characteristics referring to the quality of the sound were highlighted as for “Gru” (i.e. crane). The sound named “Gru” referred to the sound of the buzzer, placed above the crane itself, which emitted different kind of sounds during its movement. In this case, the vector referring to the crane contains adjectives referring to the characteristics of the sound produced (Ita: [intermittente, crescente] Eng: [intermittent, increasing]). In the crane vector we also find the adjective “sopportabile” (i.e. bearable), at the third position of the vector: in this case we can say that it expresses a sensation of the personal sphere, i.e. the ability to be tolerated by the respondent.

We found an anomaly on two vectors, “gabbiani” and “treno” (i.e. seagulls and train), as they could not be characterised via TF-IDF vectors intersection since they both had all adjectives in common with the other sounds. Both sounds were recognised, however, the train vector has all adjectives connoting the characteristics of the sound made to be heard, whereas the vector representing seagulls includes both adjectives describing the sound and adjectives qualifying its perception.

Sentiment analysis on adjectives: a comparison between classifier

Sentiment Analysis is a technique used to analyse the positive or negative emotions contained in a text: we made the choice to employ it in order to evaluate the perspectives stated by interviewees by connecting adjectives to the sounds heard. In this regard, two different libraries have been used: VADER (Valence Aware Dictionary and sEntiment Reasoner) and TextBlob. The results obtained with VADER can be consulted in [6], while this work presents the results obtained with TextBlob. VADER and Textblob return a score, respectively “compound score” and “polarity score”, for each word or sentence within the range [-1.0,1.0] where -1 represents most extreme negative sentiment and +1 most extreme positive sentiment. It must be said that both software libraries have the advantage of providing a predefined model for sentiment analysis, but both only work on English words and both are general purpose, resulting in a lack of precision with regard to adjectives referring to sounds: by way of example, we may mention that some adjectives referring to sounds are identified as neutral while in the spoken language they have a markedly negative or positive valence. All of the above, a summary of the sentiment perceived for a limited set of sounds (the same of [6]) by respondents can be seen in the chart below.

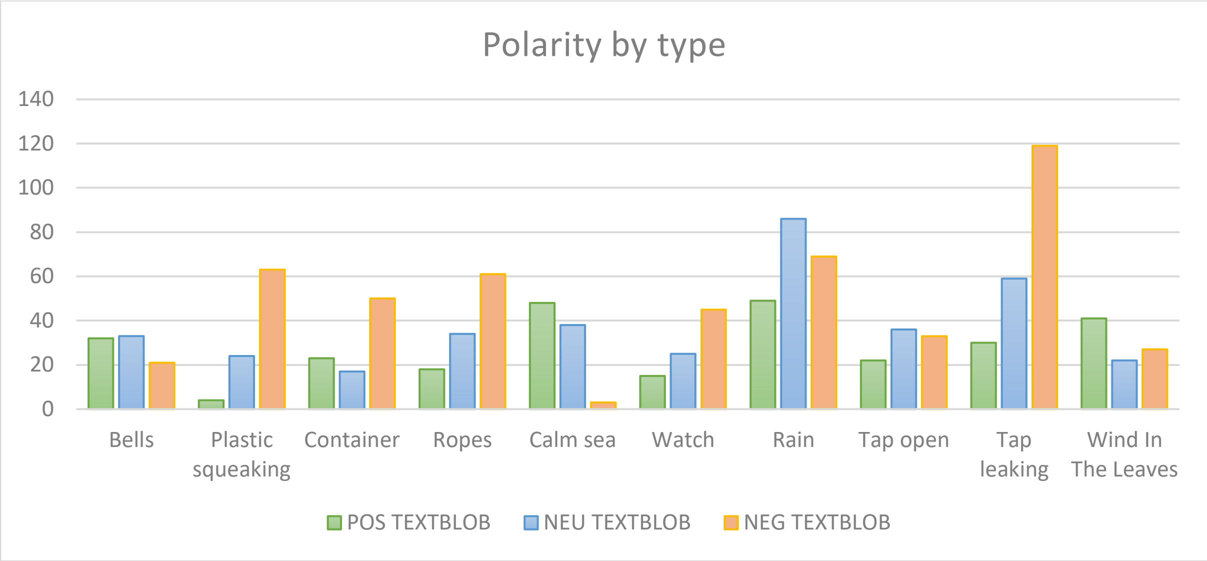

Figure 1

To get the sentiment for each audio heard the list of adjectives associated has been considered as a whole sentence and this pseudo-sentence has been given to TextBlob to analyse. It quickly catches the eye that the number of answers considered neutral is very high, and looking at the data it is evident that TextBlob is unable to correctly classify adjectives referring to sounds: for example, the list of adjectives “relaxing, quiet” referring to the sound “Rain” is categorized as neutral. This is evident when looking at the average sentiment graph where many values are close to zero (see Figure 2).

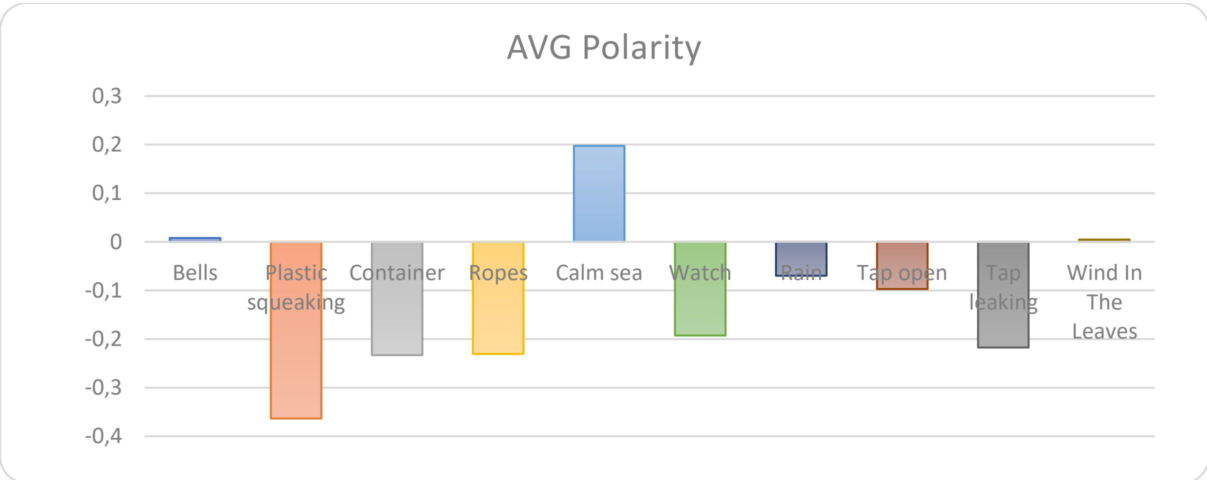

Figure 2

To get around this problem and to be able to understand the sentiment of a given sound, we can merge the two categories of neutrals and positives in only one, given the fact that a sound catalogued as neutral is not perceived as unpleasant and therefore, for a noise pollution analysis, can be catalogued as positive (no acoustic discomfort perceived).

In fact, before merging (Figure 1) we had five sounds categorized as negatives ("Plastic squeaking", "Container", “Ropes”, “Watch”, “Tap leaking”), three neutrals (“Bells”, “Rain” and “Tap open”) and two as positives (“Calm sea” and “Wind in the leaves”) while after merging (Figure 3) we have three sounds classified as positives (“Bells”, “Calm sea”, “Wind in the leaves”) and seven as negatives ("Plastic squeaking", "Container", “Ropes”, Watch”, “Rain”, “Tap open”, “Tap leaking”).

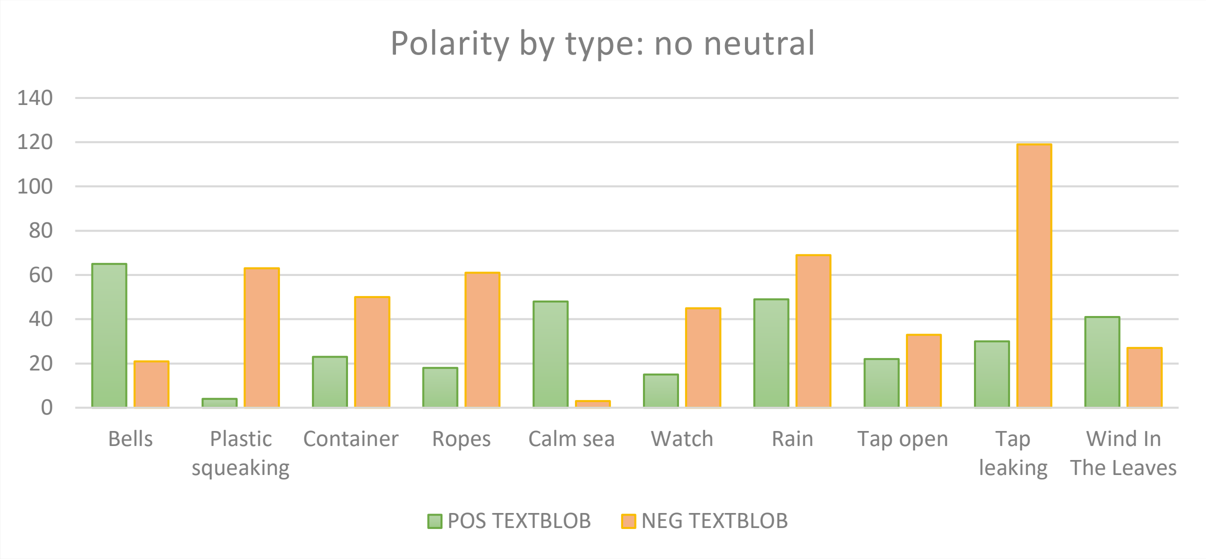

Figure 3

The same phenomenon happened with VADER to a lesser extent: a comparison of the average sentiment per sound calculated by both methods is shown in Figure 4. Here it is quite evident that regarding TextBlob the ratings categorised as neutral lower the average values and bring them, in some cases, closer to zero, again confirming that the default classifier provided by TextBlob does not succeed well in identifying sentiment in a specific domain such as sound.

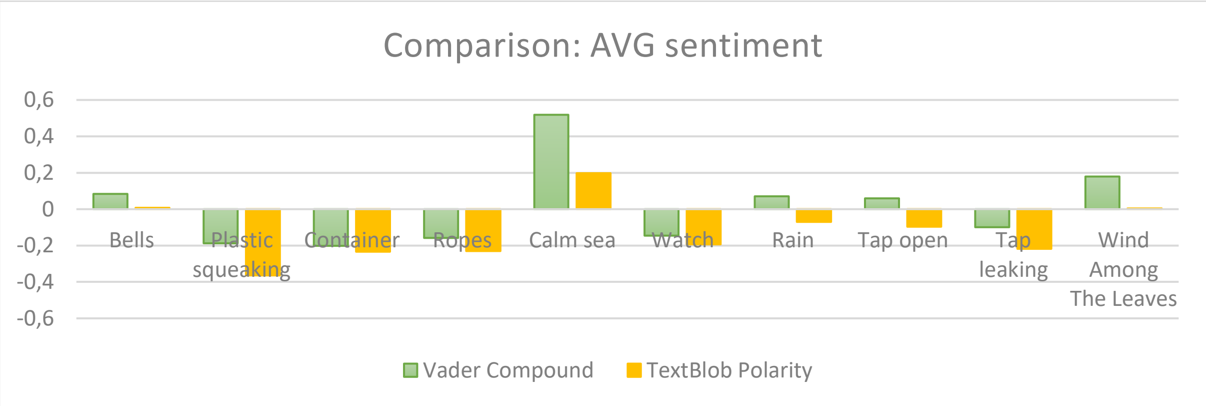

Figure 4

This deficiency is well shown in the analysis of the average polarity, where we can see that there are no well-defined sets of sounds as was the case with VADER. VADER generally classified sounds from natural/homely elements or connoted by joyful events as sounds with positive sentiment, while those from industrial or chaotic environments as negative sentiment sounds: very meaningful for TextBlob analysis it is the case of “Rain” sound which is denoted by negative sentiment. As with the analysis performed with VADER, in the one performed with TextBlob, the sound of the 'calm sea' is confirmed as having the lowest variance, demonstrating that the opinion about it is rather unanimous.

Conclusions

The analysis via TF-IDF proved to be quite characterising, despite the fact that two sounds were found to have an empty intersection vector: this may be due to the fact that the question asked was more about sound perception rather than sound characterisation.
Regarding the second part of our study, in contrast to the previous one, in which VADER was used, TextBlob sentiment analysis did not identify all nature-related sounds with positive sentiment, but left out the sound of 'Rain'. On the other hand, it was confirmed that sounds related to moving vehicles or human activities were perceived negatively as in the previous study. The sentiment classifier TextBlob turned out to be less optimal than VADER for the domain of acoustic perception: however, both of them confirm the conclusions of [6], namely that classifiers trained on a generalist corpus are ill-suited to more specialised domains. At the same time, it is clear that the next step for a marked improvement in results is training on a thematic corpus relating to sounds. In addition to all of this, we must say that Google Translate's semi-automatic translation from Italian to English lost the lexical richness

of the Italian language: for instance, three different adjectives in Italian were translated with only one in English and this penalised the great work of adjective selection done during the creation of the questionnaire.

Funding

The research is carried out within the framework of the TRIPLO project financed by the Interreg Cooperation Program V-A Italy France Maritime 2014 - 2020. Priority axis 3 - Improvement of the connection of the territories and of the sustainability of port activities. Specific objective 7C1 - Improve the sustainability of commercial ports and logistic platforms associated with them, in the scope to the reduction of noise pollution.

Acknowledgements

Authors want to express their deep and sincere gratitude to Lucia Marconi who over the years has guided and supported them in their research by offering advice and encouragement.

References

[1] Goines, L. and Hagler L. “Noise pollution: a modem plague.” Southern medical journal 100 3 (2007): 287-94.

[2] Stephen A Stansfeld, Mark P Matheson, Noise pollution: non-auditory effects on health, British Medical Bulletin, Volume 68, Issue 1, December 2003, Pages 243–257.

[3] Casazza, M., Boggia, F., Serafino, G., Severino, V., & Lega, M. (2018). Environmental impact assessment of an urban port: Noise pollution survey in the port area of Napoli (S Italy). J. Environ. Account. Manag, 6, 125-133.

[4] Cerniglia, A., Chiarella, D., Cutugno, P., Marconi, L., Magrini, A., Feo, G. D., & Ferretti, M., Questionnaire analysis to define the most suitable survey for port-noise investigation. Proceedings of the 26th International Congress on Sound and Vibration, ICSV26, Montreal, 7-11 July 2019.

[5] Paola Cutugno, Melissa Ferretti, Davide Chiarella, Lucia Marconi. “A Linguistic preliminary study about noise perception” in “Contribuciones a la Lingüística y a la Comunicación Social. Tributo a Vitelio Ruiz Hernández”. ISBN: 978-959-7174-40-0

[6] M. Ferretti, D. Chiarella, P. Cutugno, L. Marconi, G. Di Feo, A. Cernigli, A. Magrini. “A linguistic and psychoacoustic study for questionnaire analysis: first results”. Proceedings of the 27th International Congress on Sound and Vibration. ICSV27, Praga, 11-16/07/2021.

[7] K. Sparck Jones. "A statistical interpretation of term specificity and its application in retrieval". Journal of Documentation, 28 (1). 1972.

[8] Stephen Robertson, Karen Spärck Jones, Relevance weighting of serach terms, «Journal of the American Society for Information Science», XXVII, 1976, pp.129-146.

[9] CORNELIS JOOST van RIJSBERGEN, Information Retrieval. Second edition. London-Boston, Butterworths, 1979 (Prima edizione: 1975), in particolare le pp. 111-143.

[10] H. Wu and R. Luk and K. Wong and K. Kwok. "Interpreting TF-IDF term weights as making relevance decisions". ACM Transactions on Information Systems, 26 (3). 2008.