The use of onomatopoeias to describe environmental sounds

Susana M. Capitão Silva •; Luis M. T. Jesus ♦; Mário A. L. Alves *

• Escola Superior de Saúde da Universidade de Aveiro (ESSUA) e Direcção Regional da Educação do Norte, Ministério da Educação, Portugal; e-mail: susanacapitao@ua.pt

♦ Escola Superior de Saúde da Universidade de Aveiro (ESSUA) e Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA), Portugal; e-mail: lmtj@ua.pt

* Centro Hospitalar de Vale do Ave, Guimarães, Portugal; e-mail: marioandrealves@gmail.com

1. Introduction

In everyday life people listen to different kinds of sounds. Speech and music have been widely investigated, while environmental sounds have only recently been a focus of major studies (Gygi, 2001; Silva, 2007). Environmental sounds provide important information about the events of daily living, and their study has several applications including medicine, artificial intelligence, noise control and design of virtual auditory environments.

In musical listening we pay attention to the features of the sound itself and to the emotional sensations conveyed by the sound, while in everyday listening we concentrate on the events and not sounds, listening to what generates them (Gaver,1993). Gygi (2001) also emphasized that when we listen to environmental sounds the purpose is to identify the source, while in the perception of speech and music there is a greater concern with the sound semantics and expressive value.

Gaver (1993) argued that environmental sounds acoustic features are related with the materials and interactions that occur between the objects that produce sound. In Gaver's (1993) classification of environmental sounds three types of interacting materials are addressed: *solids, gases* and *liquids*. In the first type, the interactions that can produce sound are *impact, deformation, scraping* and *rolling*. Possible interactions between gases are *explosions, gusts* and *wind.* The sounds produced by liquids, are classified in four interactions: *drip, pour, splash* and *ripple.*

Several studies have addressed the sounds produced by a single type of object or event, such as breaking and bouncing bottles (Warren and Verbrugge, 1984), clapping (Repp, 1987) and steps (Xiaofeng *et al.*, 1991). In all the studies, acoustic features in the temporal and spectral domain are analyzed. Animal sounds are analyzed in several studies concerning behaviour and communication (Ikeda *et al.*, 1997; Lengagne, 2001; Riede *et al.*, 2001). In most of the studies, acoustic features are determined at a preliminary stage of the analysis.

In a number of previous studies, the relationships between the acoustic properties of environmental sounds and onomatopoeic features have been analysed (Tanaka *et al.*,

***Textos Seleccionados. XXIII Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa, APL, 2008, pp. 413-419**

1997; Iwamiya and Nakagawa, 2000; Tanaka *et al.*, 2006). Only one study (Tanaka *et al.*, 2006) has addressed the relationship between the auditory impressions of sounds and onomatopoeic representations with a variety of sounds that included different sources (e.g., machinery, animals and liquids). Psychoacoustic experiments were carried out to study the possibility of using onomatopoeic representations to estimate the impressions associated with several environmental sounds (Tanaka *et al.*, 2006). A free description experiment using onomatopoeic representations was performed. The sounds were grouped in terms of the onomatopoeia's phonetic parameters, but it was difficult to relate sounds of the same groups (Tanaka *et al.*, 2006).

In the present study, acoustic analysis and psychoacoustic experiments were conducted. The results were related with Gaver's (1993) classification of environmental sounds to test the hypothesis that the acoustic features of sounds and their perception is influenced by the materials and interactions involved in sound production.

2. Method

In the first authors' professional practice, as a Speech and Language Therapist, environmental sounds are used as a tool for auditory training. Although the number of environmental sounds can be virtually infinite, 37 stimuli were selected from the following commercially available CDs: "Digo o que faço, Faço o que Digo", Areal Editores, Portugal; "Ginásio dos Sons", Edições Convite à Música, Portugal. These included sounds produced by solids (impacts, deformations and scrapping), liquids (dropping, splash and ripple), gases (explosion, gust and wind) and animals (mammals, birds and insects), as shown in Table 1. The data from the original CDs were digitally transferred to .wav mono files at 16 bits, with a sampling frequency of 44 kHz.

<table>
<tr><th colspan="2">Category</th><th>Stimuli</th></tr>
<tr><td rowspan="3">Solid</td><td>Impact</td><td>bell, to snap one's fingers, walking, cutlery hitting plate, axe, small bell, clapping, walking down a staircase, drum, digging with a shovel, closing door, maracas, tambourine</td></tr>
<tr><td>Deformation</td><td>clapping, walking down a staircase, drum</td></tr>
<tr><td>Scrapping</td><td>digging with a shovel, closing door, maracas, tambourine, rubbing the hands, sawing, sweeping</td></tr>
<tr><td rowspan="3">Liquid</td><td>Drip</td><td>filling a glass of water, rain, waterfall, shower</td></tr>
<tr><td>Splash</td><td>rain, waterfall, shower, sea, boiling water</td></tr>
<tr><td>Ripple</td><td>river</td></tr>
<tr><td rowspan="3">Gases</td><td>Explosion</td><td>thunder, boiling water</td></tr>
<tr><td>Gust</td><td>filling up a tire, plain</td></tr>
<tr><td>Wind</td><td>whistle, wind</td></tr>
<tr><td rowspan="3">Animals</td><td>Mammals</td><td>cow, dog, pig, cat</td></tr>
<tr><td>Birds</td><td>duck, bird, rooster</td></tr>
<tr><td>Insects</td><td>cricket</td></tr>
</table>

Table 1: Environmental sounds stimuli classification.

The environmental sound features were analyzed using *Praat* to extract parameters from the acoustic signal waveforms, spectrograms and spectra (see Figure 1), such as duration and fundamental frequency, and frequency of major resonances.

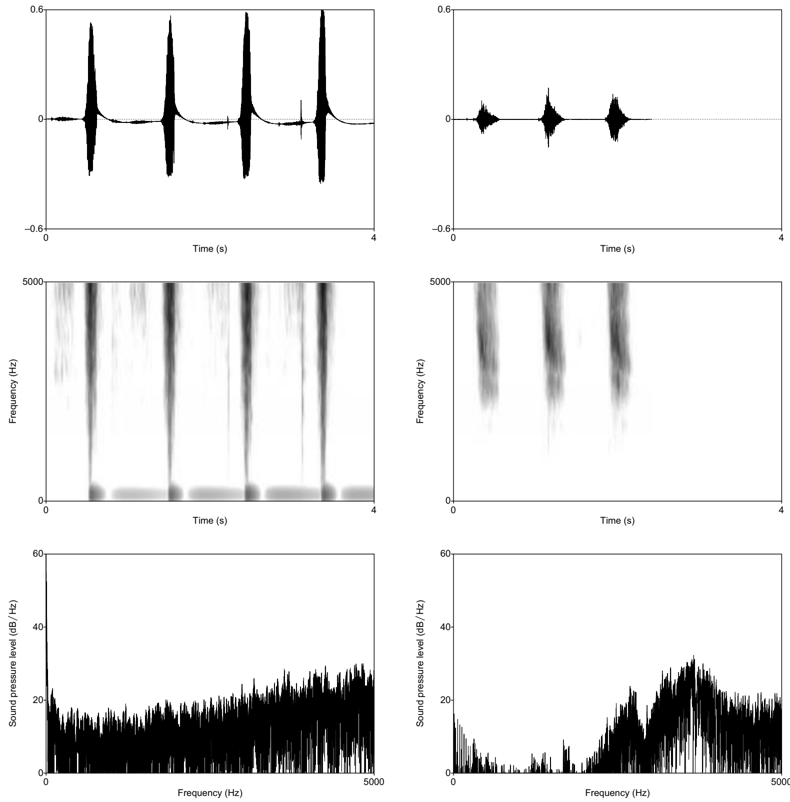

Figure 1: Waveform, spectrogram and spectra: on the left of the environmental sound *filling up a tire*, on the right an example of onomatopoeia used to describe it [St St St].

A perception experiment with 8 normal hearing European Portuguese listeners (4 male and 4 female) was also carried out. The subjects were asked to describe the environmental sound stimuli using onomatopoeias. The stimuli were presented to each participant through headphones (Senheiser eH 1430) attached to a computer. The stimuli presentation order was randomized in terms of the groups of sounds. The subjects could listen to the sound stimuli as many times as they felt necessary. At the same time, the onomatopoeias, produced orally by the subjects, were recorded using a Philips SBC ME 400 unidirectional condenser microphone and *Adobe Audition*. The sampling frequency was also 44 kHz, to allow for comparison with the original stimuli.

Each onomatopoeia waveform and spectrogram was manually analyzed to detect the beginning and end of productions. The onomatopoeias were classified using 15 phonetic parameters consisting of 7 places of articulation (bilabial, labio-dental, alveolar, palato-alveolar/palatal, velar and uvular), 4 manners of articulation (plosive, fricative, liquid and nasal), 2 voicing categories (voiced and voiceless) and 3 vowel groups (Group 1 – /6, a, E, e/, Group 2 – /i/, Group 3 – /u, o, O/). These groups were proposed previously by Tanaka *et al.* (2006), whose results showed they comprised the most important parameters when characterizing the onomatopoeic representations used by Japanese subjects. The onomatopoeias were then associated to the type of environmental sound they represented, as shown in Figure 2.

3. Results

3.1. Acoustic Analysis

The results were grouped according to the environmental sound category (solid, liquid, gas or animal) and the type of interaction.

Sounds produced by solids were the shortest (mean=1.352s, sd=1.129s), especially the ones caused by a deformation. Sounds produced by liquids and gases had longer durations corresponding to continuous sounds. Relating these during results with those from the onomatopoeia produced (fricatives and vowels mostly), we noticed that fricative is the manner of articulation with a longer duration, and also that vowels are the sounds of speech that have larger relative duration values (Stevens, 1998). The mean duration of liquids (mean=8.284s, sd=2.499s) was superior to the one of gases (mean=4.852s, sd=4.368s). The animal stimuli had a mean duration of 1.752s (sd=0.684s).

The stimuli waveform, spectrogram and power spectral density were compared to the ones of the onomatopoeias produced to describe the corresponding environmental sounds. Sounds in the group solids-impact or solids-deformation had a high amplitude at the beginning that decreased in a short period of time, which was also observed in some of the phones in onomatopoeias used by subjects to describe these sounds ([tu~tu~], [pa~pa~] and [tSiktSktSi]). The sounds classified as solids-scrapping and its onomatopoeias showed a continuous acoustic signal, with low energy in the frequencies below 500 Hz.

All the sounds generated by liquids showed a continuous signal with constant amplitude, also observed in some of the phones used in the onomatopoeias ([tuSt tuSt tuSt], [tS:], [ua:] and [fliflifli]). In liquids-drip and liquids-ripple category low energy was observed below 500 Hz.

Sounds generated by gases were continuous noise signals. Most of the sounds had more concentration of energy below 1000 Hz.

Animal sounds showed a periodic signal, except for insect sounds that were aperiodic. In mammal sounds, periodic and aperiodic characteristics were observed as well as more concentration of energy below 3000 Hz. In bird sounds, most of the energy was above 2000 Hz and in insect sounds between 2500 and 3500 Hz.

3.2. Onomatopoeia Analysis

The productions used by the subjects to describe the environmental sounds had similarities between categories, as shown in Figure 2.

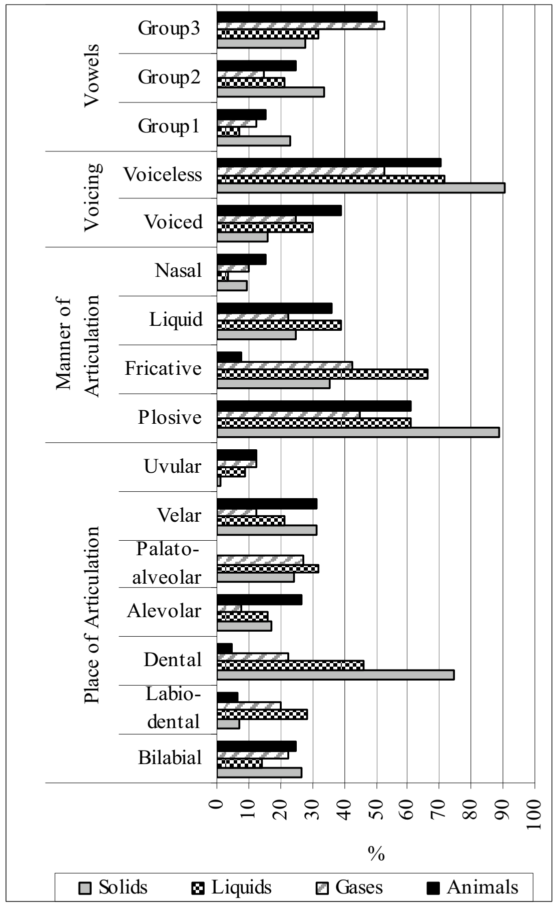

Figure 2: Results of the onomatopoeias phonetic parameters by category of environmental sound.

Sounds produced by solids-impact were described with [t]. In the case of solids--deformation subjects used [p] and vowels from group 1 and group 3, which are characterized by a low F2 (Stevens, 1998). Both plosives and solids-impact sound production involve a brief contact between two structures. In plosives there is a reduced energy production (Stevens, 1998) as observed in the environmental sounds they were used to describe. Solids-scrapping were described mainly with [t], [d], [S] and the vowel [i], as shown in Figure 2. Fricatives have longer durations than plosives (Stevens, 1998), which justifies their use to describe solids-scrapping.

The category of sounds liquids-drip were represented by onomatopoeias that had [t, d] and [S] combined with the vowel [i], liquids-splash onomatopoeias consisted mainly of labio-dental and alveolar fricatives and [t], and liquids-ripple were described using plosives, liquids and voiceless fricatives. The fact that liquids are the stimuli with larger mean duration is related to the use of speech sounds that also have longer durations, namely fricatives. Moreover, it was noticed that solids and liquid sounds were both described with voiceless phones, as shown in Figure 2, which may be related to the absence of a periodic signal in those categories.

In the case of gases-explosion, the use of bilabial plosives and [R], with vowels from the third group, was observed (see Figure 2). In gases-gust and gases-wind, [f] and group 3 vowels were used. Just like in the liquids category, fricatives were used to describe gases, which also have long durations. The use of the vowels from the third group (/u, o, O/), mainly the vowel [u] that has low values of F1 and F2 (Stevens, 1998), may be related to the fact that the sounds produced by gases have more concentration of energy below 1000 Hz.

The onomatopoeias of the animal sounds had more voiced phones than any of the previous categories, as shown in Figure 2, which can be explained by the perception of a periodic signal in most of the stimuli from this group. The sounds of animal-mammals were described with [m] and [R], and vowels from the first and third group. Sounds of animal-birds were described using [p, k] and [r], with the vowel [i]. The use of plosives may be related to the fact that these stimuli were a sequence of consecutive transient sounds, and the liquids may be used to describe the speed of the transitions. The sounds in the animal-insects category were described with [g], [r] and the vowel [i], that has F1 and F2 values within the frequency range of the corresponding environmental sound (Stevens, 1998).

Conclusions

In the present study, preliminary work on the acoustics of environmental sounds has been described. An attempt to relate this type of sounds with the speech sounds was also made. When relating the environmental sounds with the onomatopoeic representation used to describe them, waveforms, spectrograms and spectra of both stimuli were used to show similarities in the temporal pattern and in the frequency domain between the environmental sounds and onomatopoeias. These characteristics are also similar for sounds produced by the same material.

These results indicate that onomatopoeias offer a valid method for identifying the acoustic properties and auditory impressions associated with sounds. Onomatopoeias can, therefore, be used as a measure of human impressions. There may be shared features between environmental sounds and speech sounds. If so, environmental sounds may be used to improve auditory abilities without necessarily involving linguistic competencies like those of perceiving speech. Especially when the speech and language therapists are treating young hearing-impaired children, the use of carefully selected environmental sounds may enhance the stimulation of specific audition abilities concerning frequency, duration or intensity, and facilitate the perception of speech and environmental sounds, and improve the auditory training results.

5. Acknowledgments

This work was developed as part of the MSc in Speech and Hearing Sciences at the Universidade de Aveiro, Portugal.

References

Gaver, W. (1993) What in the world do we hear?: An Ecological approach to Auditory Event Perception. *Ecological Psychology* 5, pp. 1-29.

Gygi, B. (2001) *Factors in the Identification of Environmental Sounds*. PhD Thesis. Department of Psychology, Indiana University, USA.

Ikeda, Y., Jhans, G., Nishizu, T., Sato, K. & Morio, Y. (2003) Individual identification of dairy cows by their voice. *Precision Livestock Farming*, pp. 81-86.

Iwamiya, S. & Nakagawa, M. (2000) Classification of audio signals using onomatopoeia. *Soundscape* 2, pp. 23-30.

Lengagne, T. (2001). Temporal stability in the individual features in the calls of eagle owls (bubo bubo). *Behaviour* 138, pp. 1407-1419.

Repp, B. (1987). The sound of two hands clapping: an exploratory study. *JASA* 81, pp. 1100-1109.

Riede, T., Brunnberg, L., Tembrock, G., Herzel, H. & Hammerschmidt, K. (2001) The harmonics-to-noise ratio applied to dog barks. *JASA* 110, pp. 2191-2197.

Silva, S. (2007) *Traços Acústicos e Perceptivos dos Sons Não Verbais e da Fala*. MSc Thesis, Universidade de Aveiro, Portugal.

Stevens, K. H. (1998) *Acoustic Phonetics*. MIT Press, Cambridge.

Tanaka, K., Matsubara, K. & Sato, T. (1997). Onomatopoeia expression for strange noise of machines. *JASJ* 53, pp. 477-482.

Tanaka, K., Takada, M. & Iwamiya, S. (2006). Relationships between auditory impressions and onomatopoeic features for environmental sounds. *Acoustic Science Technology* 27, pp. 67-79.

Warren, W. &Verbrugge, R. (1984). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. *JEP* 10, pp. 704-712.

Xiaofeng, L., Logan, J. and Pastore, R. (1991). Perception of acoustic source characteristics: Walking sounds. *JASA* 90, pp. 3036-3049.