TRADITION AND MODERNITY IN VETERINARY MEDICINE, 2020, vol. 5, No 2(9): 65–69

EVIDENCE FOR ACCURACY OF APPLIED METHOD IN DECODING SOUND SIGNALS

Pavlina Spiridonova, Violeta Alexandrova

University of Forestry, Faculty of Veterinary Medicine, Sofia, Bulgaria
E-mail: vgerasimova@ltu.bg

ABSTRACT

The current study aimed to provide experimental results for proving that computer analysis is an accurate and valuable method in decoding sound signals. We used biomonitoring as well as specific computer tool for machine analysis of one word with stress on the first and on the second syllable. The same methodology we applied with recordings of dog's barking, occurred in one concrete situation, to see if we would accomplish promising results. We established a difference in both examinations.

Key words: audiogram, computer analysis.

Introduction

The larynx was originally developed as a device to protect the lower airways against the entrance of food and liquid, which is its main purpose. The other function of the larynx is in phonation – the production of voice, which developed later in the evolutionary history to achieve its perfection with humans (Kent, G., R. Carr., 2001)

The sounds of human speech are more complex than those produced by other species, although no greater complexity of laryngeal structure is present. Indeed, the complex laryngeal mechanism is not indispensable to this task.

The formation of sounds is a complex nevroreflectory act. It involves organs and systems, such as respiratory, digestive, muscle, bone, etc. The adjustment of the type and volume of the sounds, songs, etc., is performed by the central nervous system.

According to *Wild, J. M.* (1997), in birds, as in humans, vocal control involves the intricate coordination of three major groups of muscles, namely, those of the vocal organ, the respiratory apparatus, and the vocal tract, including the jaw and tongue. However, songbirds and parrots, like humans, but unlike other non-songbirds, have developed a special telencephalic vocal control system for the production of learned vocalizations.

According to Fitch, W. (2000), the evolution of speech can be studied independently of the evolution of language, with the advantage that most aspects of speech acoustics, physiology and neural control are shared with animals, and thus open to empirical investigation. At least two changes were necessary prerequisites for modern human speech abilities: (1) modification of vocal tract morphology, and (2) development of vocal imitative ability.

Kohonen self-organizing neural networks, also called self-organizing maps (SOMs), have been used successfully to recognize human phonemes and in this way to aid in human speech recognition. This paper describes how SOMS also can be used to associate specific information content with animal vocalizations. (John Placer et all, 2006).

Materials and methods

To understand human and animal language we examined different audiograms using computer analysis with specific tool. Data was registered and shown as oscillogram. The oscillographic curve presents the frequency of the sound waves and their intensity. To proof the accuracy of this analyzing method words were recorded carrying stress on two different syllables – Sofia etc. The same principle was applied with different dog's sound signals in one concrete behavioral reaction.

We needed a technical solution to outline the biphasic amplitude of the sound pulses and the pause period between them. With the available resources and applied methodology we achieved the desired aim of reading the dynamic sound pulses in frequency line specter. Thus we developed a visual concept of the stressed syllable in coded signal.

Results

Analyzing the recorded audiograms of the word "Sofia", pronounced differently, we concluded that there is significant contrast in phoneme and syllable formation (Figure 1 and 2).

In the pronunciation of the word "Sofia" with stressed "o", the following characteristics of the audiogram were found – the record (Fig. 1) starts with three separate sharp peaked modules, of which the middle one has the highest intensity (87,0 dB). The curve continues with serial of lower amplitude fluctuations with increasing frequency and ends with a plateau. The "o" segment shows, moreover, the highest frequency (215 Hz) in the whole record. The audiogram ends with several subsequent waves with low amplitude. Every syllable in this recorded speech sound is clearly pronounced.

In the pronunciation of the word "Sofia" with stressed "i", were registered remarkable different sound and frequency modulation in audiogram – the record (Fig. 2) consists of continuous one-peaked module with variations of its frequency and amplitude. In the middle of this segment is registered the highest amplitude – 84,7 dB following the highest frequency (1464 Hz). The peak is followed by groups of serial low in frequency modulations. The pronunciation of "Sofia" with stressed "i" sounds more fluent and the characteristics of this audiogram proves it.

The amplitude range in both records, measured in dB, is very similar in spite of the fact that the frequency, measured in Hz, is significantly higher in the second record.

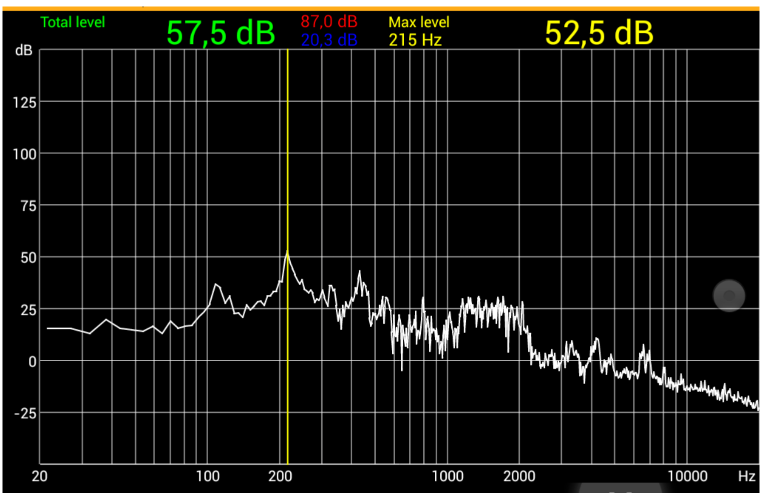

Figura 1: Record of the word "Sofia" with stressed "o".

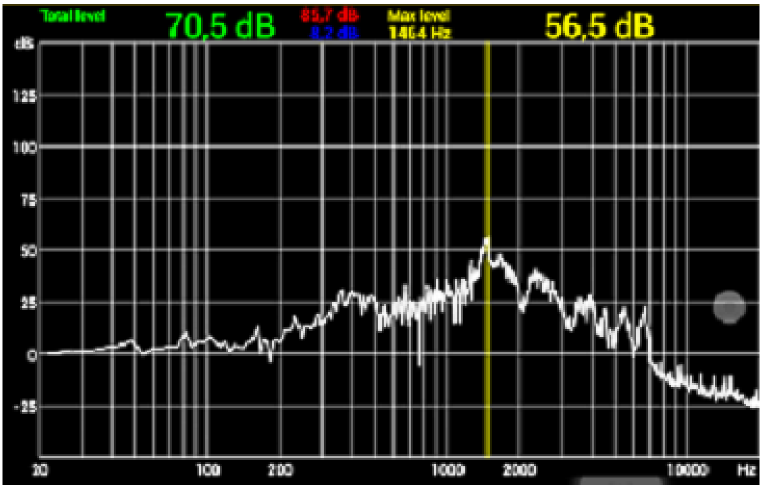

Figura 2: Record of the word "Sofia" with stressed "i".

In animal language there is no such clear difference in phoneme and syllable formation as it is in human speech. Records of dog making whining sounds to go outside and play (Fig. 3 and Fig. 4), graphically presented as audiograms, has very similar characteristics of the forming sound module. Both records consist of contiunous module without rapid changes in frequency and amplitude. Modules start with several (3 and 4) higher amplidute peaks, smoothly followed by higher frequency peaks which form a plateau. After that the curve continues with low in frequency and amplitude sound pulses.

Amplitude and frequency diapason in both records (90,4 dB and 73,9 dB; 2051 and 1680 Hz) are approximately equal. This outcome confirms, one more time, the identity of dog's sound signals and the accuracy of applied method.

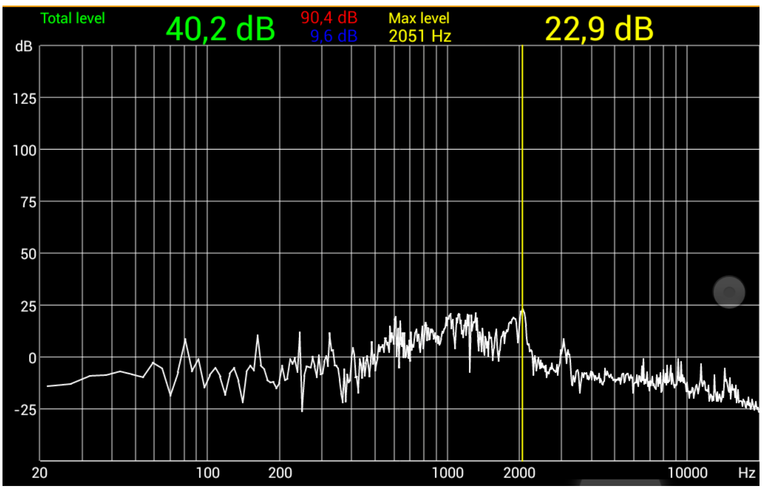

Figura 3: Record of dog (Sayren) – making crying noises to go outside and run.

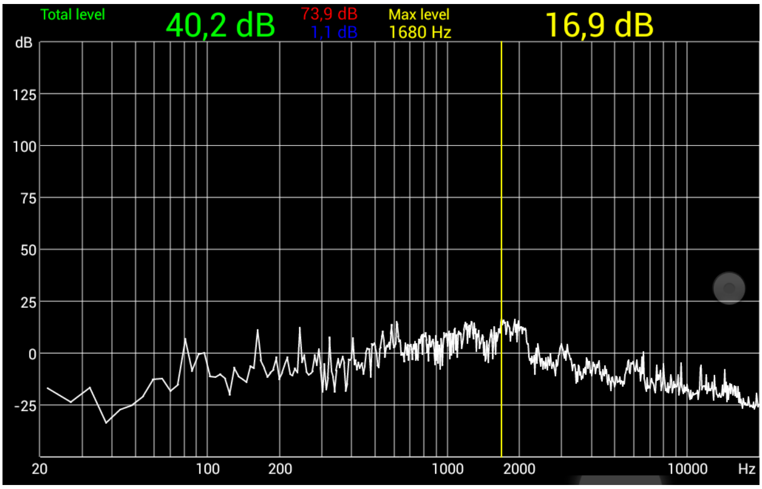

Figura 4: Record of dog (Sayren) – making crying noises to go outside and run.

Discussion

Human audiogram shows every syllable of the word as distinct submodule. Every submodule has its own specific form, amplitude and frequency characteristic. There is a pause period with different duration between each submodule. The submodules represent phonemes, articulation and intonation of speech, forming and perception of the latter, and the language as method of communication, regulated by certain nerve structures (zones, poles) in brain cortex (Kostov Y., V. Alexandrova. 2009; 2011). In addition to integrating all types of sensory information and coordinating voluntary movement, all cognitive functions (speech, math, learning, memory) are located here as well.

Kardong, K. (2008), define that human words are built up from carefully formed sounds called phonemes. Animal communication with sounds is mainly an emotional response to immediate circumstances, but in humans phonemes carry ideas and thoughts about past events or future actions. The combination of sounds, the relationships between them build words, which placed in order form sentences. Our speech apparatus have changed anatomically to serve speech by lengthening the pharynx, which was accomplished by the separation of soft palate and epiglottis. By such lengthening, air can be effortlessly channeled on a sustained basis through the mouth, where it is shaped into sounds.

Kostov, I. and Alexandrova, V. (2009; 2011) presume that animals with well developed speech apparatus (vocalization) probably use synthesis (impulse modulation) for sound wave modeling. These sound waves are placed in the range of frequency line. They are defined as "modules". Each module contains various amount of "submodules" and also additional pulses with different amplitude and frequency may take place in it. In this case, if specific terms has to be used, modules can be associated with syllables and submodules – with phonemes.

The duration of every module and the diversity of very unique combinations, as well as their specific arrangement in frequency line, ensures the abundance of language codes. Consequently, the language phrases are composed (by identical or very different in structure modules) and afterwards themes are formed (by combining similar or various phrases) in animal language.

This allows the conclusion that it is relevant to use the term "audiogram", which provides a better understanding of the graphical model of produced animal sounds.

Although the described above method, there is another more commonly used in decoding language sounds. It includes the so called spectrograms, which provide information only for spectral density of the signal proceeding in time (a graphical presentation of speech sound specters). The spectrogram is usually used for reading phonetic sounds; analyzing animal sounds or in other science fields (e.g. music, examination of speech communication, seismology etc.). Specific instruments, called spectrographs or sonographs, are required for spectrograms to be analyzed properly and in details.

The smallest unit in the spectrogram is the note, the leitmotif to be precise and in theoretically speaking – syllables. Repeated syllables are called tours. Strophes are composed of different tours.

By Siddhardha Balemarthy et all, all spectrograms are generated with parameters as 16-bit rate, mono source, sampling rate as 16000 Hz and are of duration 100 milliseconds.

Kohonen used a SOM to create a recognition system for human speech that could detect phonemes from a continuous speech signal (Kohonen, 1988).

Conclusion

1. The applied methodology allows detailed reading of the dynamic sound pulses in frequency line specter.
2. Therefore, an extremely accurate visual comprehension about the modulation of syllables in coded signal is obtained.
3. By using this method, it is possible to decode even the tiniest difference in physical characteristics of the sound signal (amplitude and frequency), presented as modulated syllable in audiogram.
4. The made comparison between identical words in human and animal sounds in one concrete situation, shown as audiograms, confirms one more time the huge opportunities of the human speech apparatus for articulate communication.

References

1. Y. Kostov, V Alexandrova. (2011). *Methods for decoding and demodulating of the animals language*. Zhivotnovadni Nauki, 2, XLVIII, 38–42.
2. Kardong, K. (2008). *Vertebrates: Comparative Anatomy, Function*, Evolution. Fifth Edition, 779.
3. Kent, G., R. Carr. (2001). *Comparative Anatomy of the Vertebrates*. 9th ed. New York: McGraw-Hill.
4. Kohonen, T. (1988). *The 'neural' phonetic typewriter*. Computer 21, 11–22.
5. Kostov, Y., V. Alexandrova. (2009). *How to recognize (decode) animal language*. Bulgarian Journal of Agricultural Science, 15, N5, 475-477.
6. Wild, J. M. (1997). *Functional anatomy of neural pathways contributing to the control of song production in birds*. European Journal of Morphology 35:303-325.
7. Fitch, W. (2000). *The evolution of speech: a comparative review*. Tends in cognitive sciences, Vol. 4, Iss. 7, Pages 258–267.
8. Siddhardha Balemarthy, Atul Sajjanhar, Xi Zheng. (2018). *Our practice of using machine learning to recognize species by voice*.
9. John Placer, C. N. Slobodchikoff, Jason Burns, Jeffrey Placer, Ryan Middleton. Northern Arizona University. (2006). *Using Self-Organizing Maps to Recognize Acoustic Units Associated with Information Content in Animal Vocalizations*. Animal Studies Repository.