Izvorni znanstveni rad
Rukopis primljen 23. 1. 2023.
Prihvaćen za tisak 21. 7. 2023.
https://doi.org/10.22210/govor.2023.40.01

Renáta Gregová, Dmytro Hrytsu
renata.gregova@upjs.sk, dmytro.hrytsu@upjs.sk
Pavol Jozef Šafárik University in Košice
Slovakia

An acoustic analysis of a sample of English and Ukrainian onomatopoeias: A pilot study

Summary

Much has been written about the position of onomatopoeias in languages nevertheless, sources dealing with the phonetic nature of these sound-echoing words are very rare. Some notes about the phonetic substance of onomatopoeias can be found in the work of Bredin and also Tsur who offers a more detailed phonetic investigation of sound-imitating expressions in relation to real sounds. On the basis of acoustic analysis of the onomatopoeic expression /ku:ku:/ and of the real sound the cuckoo produces, Tsur comes to the conclusion that such speech sounds are used as equivalents of natural sounds/noises whose phonetic properties best fit the imitated natural sounds. To evaluate this assumption, this pilot study includes acoustic analysis of onomatopoeias from English and Ukrainian. The selected English and Ukrainian verbal imitations of natural sounds (a cat's meowing, an owl's hooting, hitting two metal objects together and knocking on a door) were acoustically analyzed in terms of the structure of their sound spectra for comparison with the sound spectra of the natural sounds they imitate.

Keywords: onomatopoeia, sound spectrum, acoustic analysis, English, Ukrainian

1. INTRODUCTION

Starting from Saussure's (1959) theory about the arbitrariness of the linguistic sign, onomatopoeia is usually defined as the verbal imitation of a sound from the extra-linguistic reality and is considered the representation of a non-arbitrary relationship between linguistic form (sounds) and meaning (see, e.g., Perniss, Thompson, & Vigliocco, 2010; Schmidtke, Conrad, & Jacobs, 2014).

In recent decades, much has been written about these expressions, and onomatopoeias have been addressed from various angles: from the viewpoint of sound symbolism, through the focus on the meaning of onomatopoeias and their grouping in terms of lexical categories and word classes, up to the attention to the position of onomatopoeias in communication, etc. (for further details, see, e.g., Bredin, 1996; Gregová, 2021; Körtvélyessy, 2020; Sasamoto & Jackson, 2016). Nevertheless, very little has been written about the phonetic, that is, articulatory and acoustic structure of onomatopoeias in relation to the acoustic structure of the sounds they are said to mimic.

Theoretically, humans can try to imitate any sound they hear. But the degree of the similarity between the real sound and its human imitation depends on various factors, such as the number of sounds/phonemes in the given language, limitations imposed by the anatomy and physiology of human speech organs, as well as the limitations resulting from the human perception of sounds (for further details, see Bredin, 1996). It follows from this that the degree of the similarity of an onomatopoeia as the expression echoing the sound from the extra-linguistic reality with the natural sound it imitates is different in various languages. Consequently, the acoustic-iconic sign is not a copy of the external acoustic phenomenon; on the contrary, it is often only its vague or even rough imitation, causing intralingual variability (several onomatopoeic expressions for one and the same sound; consider, for example, English *scream – screech* or *crack – crake – crash*) as well as extralingual variability (compare, for example, English *bow-wow* and Slovak *hav-hav* 'bow-wow') (Krupa, 1992: 27).

"The English expression for the sound a rooster makes (*cock-a-doodle-doo*) is quite different from the German expression (*kikiriki*), which is different again from the French expression (*cocorico*), and all of which are arguably quite different from the actual sound which emerges from a rooster" (Perniss et al., 2010: 2). The importance of the frequently mentioned "natural resemblance" is overestimated[1], and onomatopoeias are also determined by convention (Bredin, 1996: 561), that is, they are conventionally shared by the speakers of the given language. As Bredin adds, an onomatopoeia "imitates, echoes, reflects, resembles, corresponds to, sounds like, expresses, reinforces" or "has a natural or direct relation with" (Bredin, 1996: 555). The word which aptly depicts the nature of onomatopoeia rests on the selective "linguistic interpretation of an objective acoustic phenomenon" (Krupa, 1992: 27) and depends on the phoneme inventory of a language. Thus, why are some onomatopoeias almost

[1] This view questions Saussure's conception of onomatopoeia (see Körtvélyessy, 2020 for more detailed discussion about the given issue).

identical in languages (*kikiriki* in German and Slovak) and some entirely different (*cock-a-doodle-doo* in English vs. *kikiriki* in German and Slovak)? What is the "leading power" governing the choice of linguistic sounds (phonemes) when imitating natural sounds? To answer these questions, a deep acoustic analysis of natural sounds and their language "versions" is necessary.

However, as already mentioned, much has been written about onomatopoeias from various perspectives, but very little about their phonetic nature. It seems that the only work dedicated directly to a phonetic analysis of onomatopoeias is Tsur's paper *Onomatopoeia: Cuckoo-language and tick-tocking. The constraints of semiotic systems* (Tsur, 2001). In this work, Tsur indicates that the selection of linguistic material used for the imitation of natural sounds depends on the similarity of the acoustic properties of linguistic sounds to those of the sounds of nature. To prove or disprove the validity of this claim, the authors decided to carry out comparative research focused on the acoustic analysis of onomatopoeias from a sample of Indo-European languages. However, prior to analyzing a large sample of data, a pilot study was performed on a sample of onomatopoeias from two languages – English and Ukrainian[2] – to verify the appropriateness of the research procedure.

In this article, Tsur's theory will first be explained, then the selection of the material and the method of analysis will be introduced. The following part will focus on the research itself. Finally, the results of the analysis and the implications for further research will be summarized in the conclusion.

2. SOUNDS OF NATURE AND THEIR VERBAL IMITATIONS

It is well-known, acoustically, that vowels are characterized by formants (see, e.g., Johnson, 2012). These are the resonance frequency peaks of the vocal tract usually referred to as F_1, F_2, etc. F_1 reflects the resonance of the front cavity and depends on the height of the tongue. F_2 is the resonance of the back cavity and is based on the frontness/backness of the tongue. F_0 refers to the fundamental frequency associated with the pitch. Individual vowels are characterized by their distinct formant pattern (Hayward, 2013), as illustrated in Figure 1, which captures the formant structure (indicated as the dark strips in the spectrogram) of the English long monophthongs [iː], [ɑː], and [uː].

[2] At this stage of the research, the selection of the languages was quite random, depending on the availability of the data.

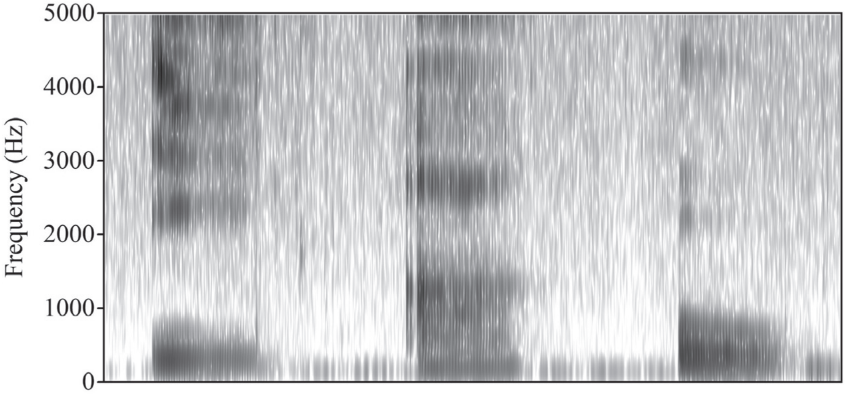

Figure 1. The formant structure of the English vowels [i:], [ɑː], and [uː]

Slika 1. Formantska struktura vokala [i:], [ɑː] i [uː] u engleskome jeziku

Consonants are also characterized by their recognizable acoustic structure, but the situation is slightly more complicated in case of this category of sounds since "information relevant to individual consonants overlaps with preceding and following speech sounds" (Hayward, 2013: 337). In addition, the "acoustic classification of consonants must bring together a number of acoustic dimensions of different types, including not only frequency but also intensity and timing" (Hayward, 2013: 337). Following phonetic criteria for consonant description, two of them are mutually dependent. Specifically, acoustic impression depends on the manner of articulation. Thus, sounds created by complete closure (stops) are called plosives acoustically. If there is a narrowing during the production of a consonant (constrictive, spirant), a sound is acoustically labelled as fricative. In the case of the approximation of two articulators, a sound is called a frictionless continuant. All these 'acoustic consequences' are observable on the sound spectrum of consonants. For example, the audible friction of fricatives is represented by a "strong band of noisy energy" (Hayward, 2013: 337), as illustrated in Figure 2, which shows the sound spectrum of the English consonant [s]. The function of vocal cords plays an important role when describing consonants. If a consonant is voiced, vocal cords vibrate, which is represented as a dark strip at the bottom of a spectrogram (Figure 3), and if not, they do not (for details, see, e.g., Hayward, 2013: 334–394; Sabol & Zimmermann, 2014: 91–98)[3].

[3] Of course, we are aware of the fact that all those basic articulatory and acoustic characteristics of vowels and consonants are familiar to all phoneticians, but they are crucial for our further analysis and that is why we decided to summarize them in one brief paragraph.

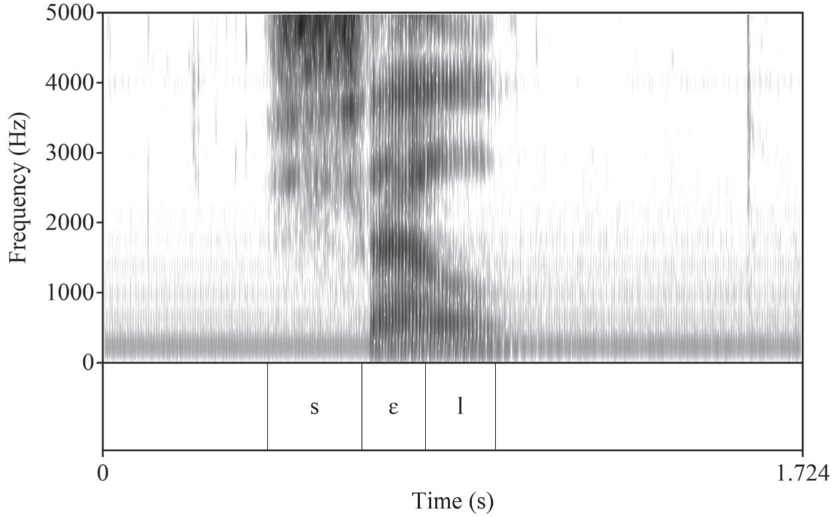

Figure 2. The sound spectrum of the English consonant [s] pronounced in the word *sell* [sɛl]

Slika 2. Spektralni prikaz konsonanta [s] u engleskome jeziku izgovorenom u riječi *sell* [sɛl]

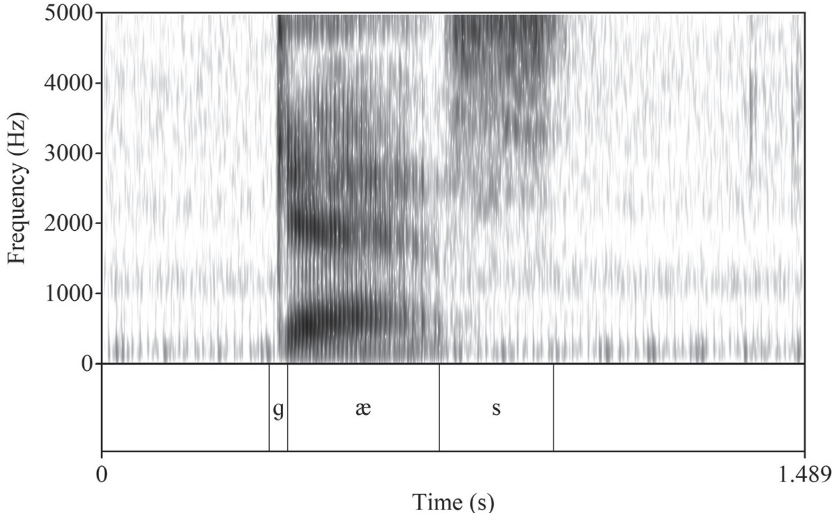

Figure 3. The sound spectrum of the English consonant [g] pronounced in the word *gas* [gæs]

Slika 3. Spektralni prikaz konsonanta [g] u engleskome jeziku izgovorenom u riječi *gas* [gæs]

In Tsur's (2001) view, the human imitation of sounds from the extra-linguistic reality is like a translation from one semiotic system to another: it is limited by the options of the target system. In other words, the quality of translation depends on the target system. As to onomatopoeias, they imitate sounds from the world around us. However, there is an infinite number of noises in nature but only approximately 50 (up to a maximum of 100) speech sounds in languages, which are, moreover, represented by only 20-something letters in an alphabet (Tsur, 2001: 9). Tsur uses the cuckoo and the sound it produces as an argument to prove this claim. In English, the bird's sound is *cuck-coo* /ku:ku:/. But the bird produces "neither the speech sound [k] nor [u:]; it uses no speech sounds at all" (Tsur, 2001: 9). The author adds that imitation is not the translation of a sound but the translation of the perceived sound qualities from reality (a semiotic system) to another semiotic system (a language). The quality of this translation depends on how fine-grained the sign units of the target system are (Tsur, 2001: 9). That is, the quality of the imitation of a sound depends on the sounds of a target language. The representation of the cuckoo's sound in a particular language can therefore be considered adequate if the language uses "those speech sounds that are most similar in their effect to the cuckoo's call" (Tsur, 2001: 9). Tsur tested this hypothesis on the English word [ku:ku:], the results of which are captured in Figure 4.

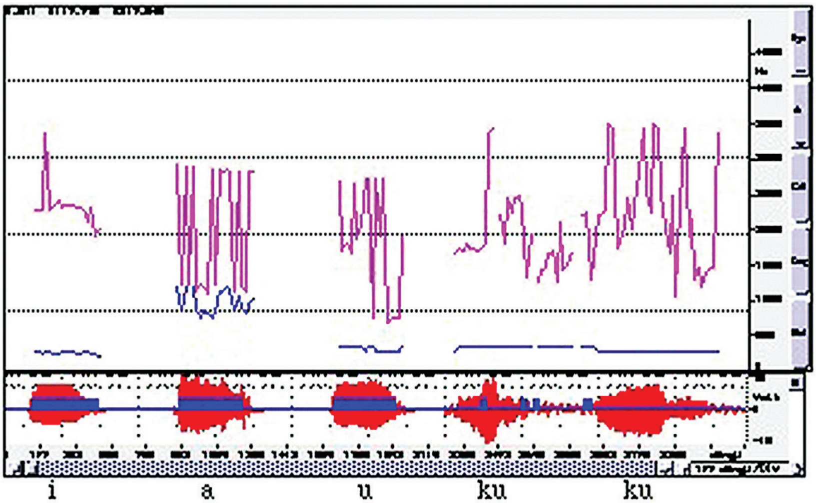

Figure 4. The English onomatopoeic word [ku:ku:] (Tsur, 2001: 11)

Slika 4. Onomatopejska riječ [ku:ku:] na engleskome jeziku (Tsur, 2001: 11)

Figure 4 demonstrates that the first formant of the sound of the bird (on the right-hand side) has a similar shape and frequency range to the human sounds [i] and [u] (on the left). It is evident that "the overtone structure of the cuckoo's song displays greater resemblance to the [u] than to the other two cardinal vowels" (Tsur, 2001: 9). Nevertheless, there is nothing like [k] in the cuckoo's song. To solve the issue, Tsur assumes that each sound system chooses those consonant sounds whose articulatory and acoustic properties that have their reflection in distinctive features "are nearest to the target sound" (2001: 9). Thus, in the case of the cuckoo's sound, the speech sound [k] is characterized by a similarly abrupt pitch onset as the bird's sound (Figure 5). In human language, at least three consonants are characterized by abrupt articulation – [p], [t] and [k] – but only [k] is a compact consonant, that is, more abrupt. The other two consonants are scattered (Tsur, 2001: 9). Tsur's finding thus indicates the possibility of making generalizations as to which bundles of distinctive features are the closest to imitating natural sounds.

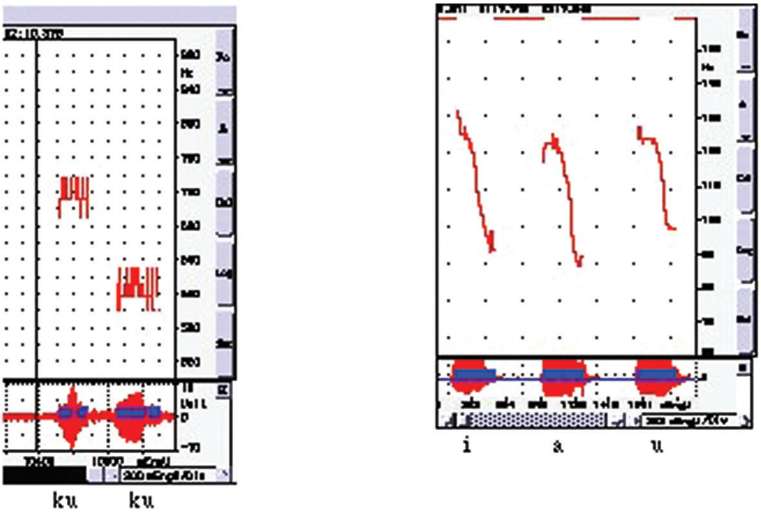

Figure 5. The pitch contours of the cuckoo's song and of the phonetic vowels i-a-u (Tsur, 2001: 12)

Slika 5. Kretanje visine tona u kukavičjoj pjesmi i u vokalima i-a-u (Tsur, 2001: 12)

3. METHOD OF ACOUSTIC ANALYSIS

3.1 Selection of sample material

Four onomatopoeic words were chosen for the pilot analysis: two from the "animal sounds" category and two from the "inanimate objects" category. The animal onomatopoeias are a cat's meowing, represented as /miːaʊ/ in English and /mʲaʊ/ in Ukrainian, and an owl's hooting, which is /huːt/ in English and /pʊɦu/ in Ukrainian. The first inanimate object onomatopoeia is the sound made as a result of hitting metal, which is /dɪŋ/ in English and /dzʲinʲ/ in Ukrainian, and the second is the sound made when knocking on a door, which sounds like /nɒk/ in English and /tʊk/ in Ukrainian (Table 1).

These onomatopoeias are lexical and can be found in the dictionaries of the respective languages. The idea was to pick two pairs of onomatopoeias that are similar across the two languages and two pairs that are dissimilar in order to make generalizations about the choices of sounds employed by either language[4].

Table 1. English and Ukrainian onomatopoeias selected for analysis
Tablica 1. Onomatopejske riječi iz engleskoga i ukrajinskoga korištene u analizi

	Cat's meowing / Mijaukanje	Owl's hooting / Hukanje	Sound made as a result of hitting metal / Udaranje metala	Sound made as a result of knocking on a door / Kucanje na vrata
English / Engleski	/miːaʊ/	/huːt/	/dɪŋ/	/nɒk/
Ukrainian / Ukrajinski	/mʲaʊ/	/pʊɦu/	/dzʲinʲ/	/tʊk/

3.2 Acoustic processing of the data

The natural sounds were downloaded from the internet (mixkit.co) and run through the software Praat, version 6.2.12. The onomatopoeia corresponding to each natural sound was recorded using Praat at a 44.1 kHz sampling rate. The words were pronounced in isolation by one male native speaker per language. Then, the natural sound and its corresponding onomatopoeia in English and Ukrainian were

[4] The purpose of this comparative/contrastive analysis is (1) to prove/disprove Tsur's assumption about the relationship between sounds of nature and their human imitations and (2) to make some generalizations on the basis of the data from two langauges with different sound systems.

put together for analysis. Variables such as formant pattern, pitch, and intensity were taken into consideration and compared. Where needed, an LPC analysis[5] was carried out to extract particular formant frequencies.

3.3 Results of the analysis

3.3.1 The acoustic analysis of animal sounds

3.3.1.1 A cat's meowing versus English and Ukrainian onomatopoeia

The spectrogram of a cat's meowing (Figure 6) can be divided into four phases – a short beginning with a rapid pitch change, a gradual increase in frequency, steady frequency and pitch, then the gradual lowering of formant frequency and pitch. The onomatopoeias of a cat's meowing – /mʲaʊ/ in Ukrainian (Figure 7) and /miːaʊ/ in English (Figure 8) – can also be divided into four phases – consonant closure, frequency increase, steady vowel, and gradual formant lowering.

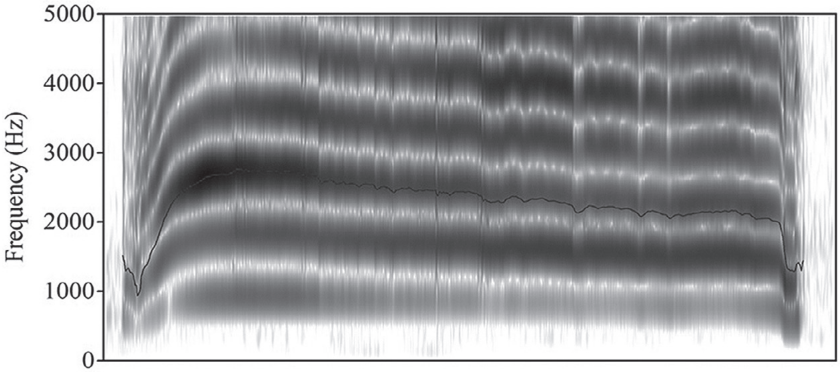

Figure 6. Spectrogram of a cat meowing (line indicates pitch)
Slika 6. Spektrogram mijaukanja mačke (linija prikazuje tonsku visinu)

The first phase of the natural sound is accompanied by a rapid pitch change. Onomatopoeias incorporate nasal segments at the beginning of the utterance in both Ukrainian (Figure 7) and English (Figure 8). This is made clear by the fainter energy bands that indicate anti-formants. Anti-formants are created when the nasal cavity "absorbs" the resonant frequencies of the sound (Johnson, 2012: 155). Since the sound

[5] LPC (linear predictive coding) is used to represent spectra in a more compressed form so that the formant frequencies are easier to measure (see, e.g., Johnson, 2012).

wave has two outputs (mouth cavity and nasal cavity), it passes through the oral cavity while resonating back to the nasal cavity at the same time. Because of this, the formant frequencies cancel each other out, creating anti-resonance (Johnson, 2012: 155).

When looking at the spectrogram of a cat meowing (Figure 6), one can notice clear, distinct formants. The lack of anti-formants indicates that the cat's meowing does not incorporate the nasal cavity at all. Moreover, the spectrogram shows that there is no constriction to the sound wave whatsoever, which indicates vowel-like articulation. Interestingly, there is a rapid change in pitch at the beginning of the recording. This pitch variation is replaced by consonants and vowels in the human "version" of this sound. It is not really an issue of high or low pitch but rather the pattern. Cats' vocal tracts can be of various lengths, producing higher or lower fundamental frequencies. Still, big or small vocal tract, we associate the sound produced by a cat with a cat[6].

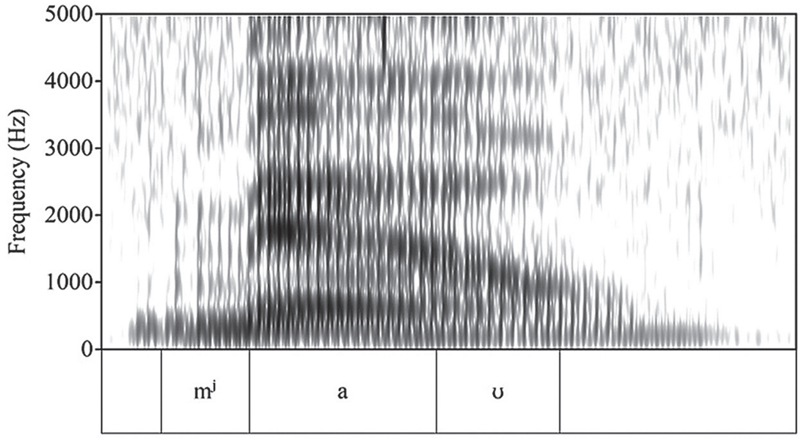

Figure 7. Spectrogram of Ukrainian onomatopoeia /mʲaʊ/
Slika 7. Spektrogram ukrajinske onomatopejske riječi /mʲaʊ/

All in all, the beginning of a cat's meowing does not contain any constriction of the vocal tract. However, the rapid change in pitch causes one to perceive it as occlusion or disruption in the airflow since such a rapid pitch change is normally a result of consonant articulation. A nasal stop seems to be the best option to imitate the beginning of a cat's meowing due to its sonority (to imitate a vowel-like pattern) and its occlusiveness (to imitate an abrupt pitch change).

[6] In the same way, using an example from the human world, the "Happy birthday" melody can start with a higher or lower note, but it still follows the same pitch interval/pattern.

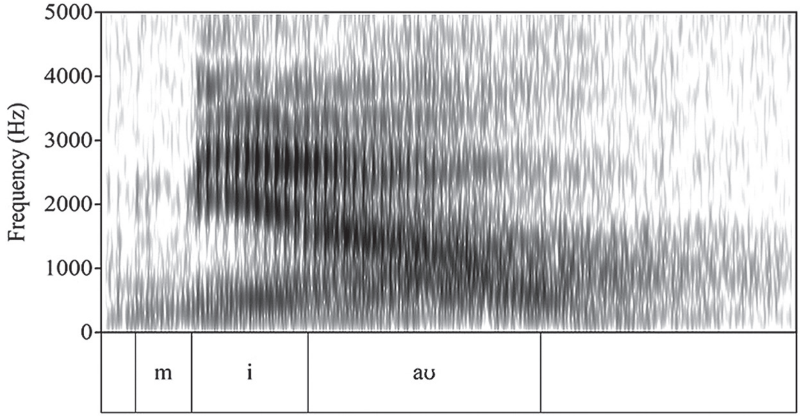

Figure 8. Spectrogram of English onomatopoeia /miːaʊ/

Slika 8. Spektrogram engleske onomatopejske riječi /miːaʊ/

The second phase of the cat's meowing incorporates an increase in formant values. The English onomatopoeia (Figure 8) reflects the formant increase by the high close long vowel /i:/. Ukrainian, on the other hand, allows for consonants to be half-palatalized. Palatalization occurs during the process of secondary articulation, where the body of the tongue reaches the palatal region in the vicinity where /i/ is produced. In this way, one segment can have two articulations at the same time. However, in Figures 7 and 8, the distinction between /m/ + /i/ and /mʲ/ is almost invisible. Hence, the imitation of an increase in formant values is rooted within the palatalization of the nasal consonant.

The third phase of the cat's meowing is a steady sound. The first two frequency peaks of sound waves in a cat's meowing (Figure 6) were calculated using LPC measurements: F_1 = 798 Hz, and F_2 = 1593 Hz. If put on a vowel chart, the formant frequencies are somewhere in the region of /a/ (Figure 9). Both Ukrainian and English use open round vowels in a similar frequency range to imitate the sound – in English, it is the beginning of the diphthong /aɪ/, and in Ukrainian, it is /a/.

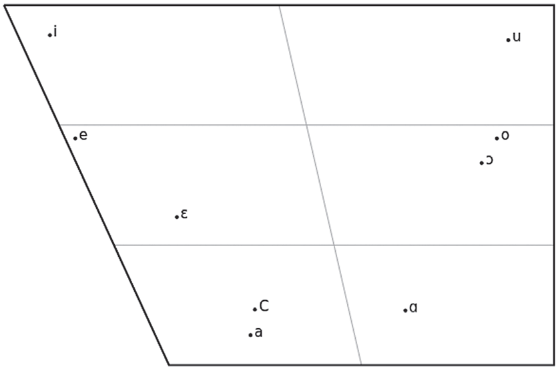

Figure 9. Vowel chart representing F_1 and F_2 in a cat's meowing. The "cat vowel" is marked as [C]

Slika 9. Vokalski dijagram koji predstavlja F_1 i F_2 u mijaukanju. "Mačji vokal" u dijagramu je označen [C]

Finally, the fourth phase is gradual formant lowering. In the cat's meowing, the formant structure gradually descends from its peak at the beginning throughout the whole recording and reaches its minimum right at the end with an abrupt fall in pitch. This corresponds to the glide from /a/ to /ʊ/ in the diphthong in English and the change from /a/ to /ʊ/ in Ukrainian. Since there is a lowering of the pitch in a cat's meowing, the corresponding onomatopoeia uses back vowels due to being low-pitched.

3.3.1.2 An owl's hooting versus English and Ukrainian onomatopoeia

The sound of an owl is visualized in Figure 10, and Figures 11 and 12 represent the Ukrainian /pʊɦu/ and English /huːt/ onomatopoeias for the given sound.

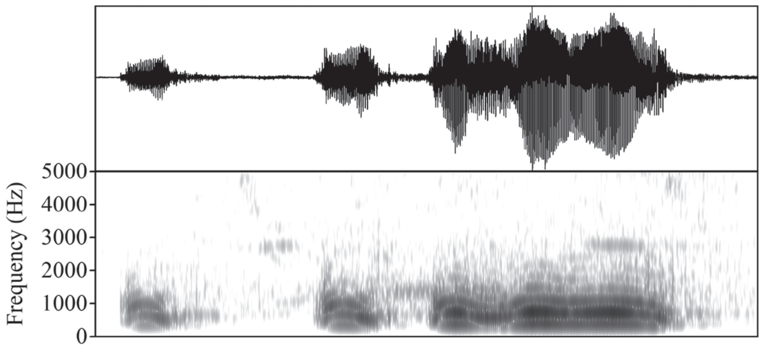

Figure 10. Oscillogram and spectrogram of an owl's hooting

Slika 10. Oscilogram i spektrogram sovina hukanja

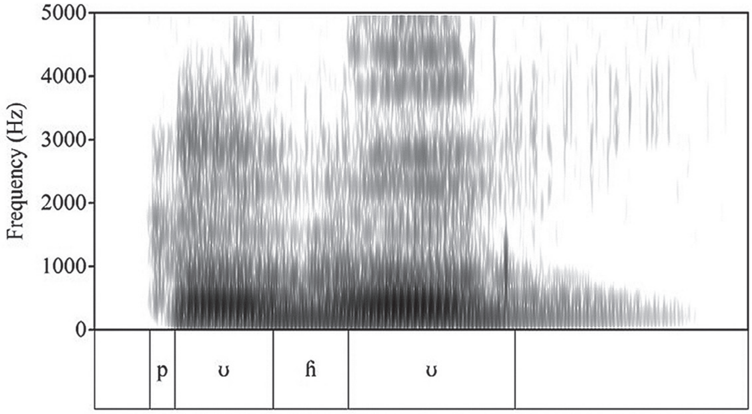

Figure 11. Spectrogram of the Ukrainian onomatopoeia /pʊɦu/

Slika 11. Spektrogram ukrajinske onomatopejske riječi /pʊɦu/

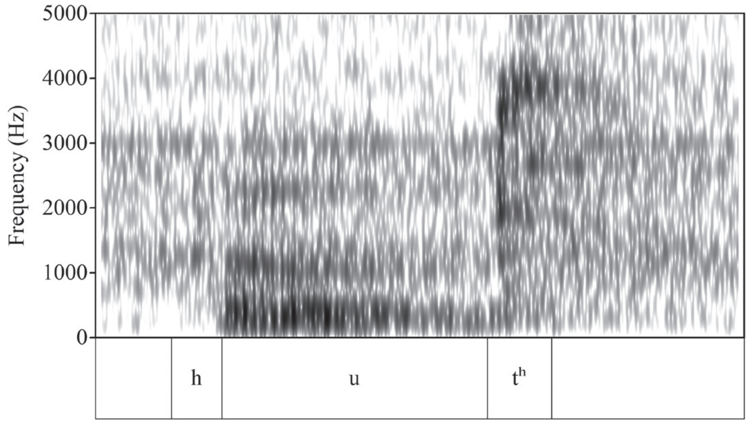

Figure 12. Spectrogram of the English onomatopoeia /huːt/

Slika 12. Spektrogram engleske onomatopejske riječi /huːt/

The sound the owl makes is produced at lower frequencies, as shown by the bottom of the spectrogram having the biggest concentration of energy. Higher frequencies are non-existent; therefore, the sound can be described as having low tonality. These two characteristics are typical of the distinctive features [grave] and [acute][7] in Jakobson's

[7] The sound is acute when it is characterized by a high pitch, that is, the upper side of the spectrum dominates. When the lower side of the spectrum dominates, the sound is characterized by the feature [grave] (Jakobson, 1978).

theory of distinctive features (Jakobson & Halle, 1956). The sound an owl makes has the concentration of energy at the bottom of the spectrogram, which indicates [+grave]. At the same time, there is no dominance in the high region of the spectrogram, which indicates [-acute].

In both Ukrainian and English close back vowels with [+grave] and [-acute] specifications in their verbal imitations of an owl's call are used. However, the application of consonants in the two languages differs. In Ukrainian, the bilabial plosive /p/ and the voiced glottal fricative /ɦ/ imitate natural sounds. In English, the glottal fricative /h/ and the alveolar plosive /t/ are employed. The owl's call follows a certain pattern that cannot be directly transformed into human speech. Instead, both languages use plosives to account for the interruptions in the owl's call. However, it is necessary to mention that the interruptions are not quick or compact but spread or diffused over time. That is why diffuse consonants are preferred here over compact ones[8].

Both languages use glottal fricatives to solve the prolonged nature of an owl's call, the main difference being voicing. The owl's call spectrogram (Figure 10) features dark bands of energy at the bottom. This connotes that the sound produced by an owl is voiced (cf. Figure 2). The same can be said about the Ukrainian sound /ɦ/ (Figure 11); the bands are slightly less faint than the surrounding vowels, but the sound is still voiced. The bands, however, are the faintest during /h/, implying lower vocal fold activity. Consequently, the voiced fricative is, phonetically, closer to the actual sound the owl makes. In this case, due to the absence of a voiced glottal fricative in English, the language chooses the place of articulation to be of greater importance than the [+voiced] feature and settles for /h/ over any other voiced consonant. The reason for this is that voiceless/voiced distinction creates less impact on the overall utterance than the change in the place of articulation. The language chooses the place of articulation to produce the sound whose acoustic quality is the closest to the sound an owl makes.

3.3.1.3 Animal sounds – summary

Regarding a cat's meowing, both languages analyzed use their sound systems perfectly to imitate the sound. Nasals appear to echo an abrupt change in pitch. The high close vowel in English and palatalization in Ukrainian are used to mimic the rise in formant values. Open round vowels are the closest formant values to the steady sound

[8] The sound (phoneme) is diffuse if no formant prevails. The feature [compact] indicates that there is one centrally located formant region in the sound spectrum (Jakobson & Halle, 1956).

in the middle of a cat's meowing. In both languages, the back close vowels function as representations of the lowering in formant values and pitch.

As for the owl's sound, the Ukrainian onomatopoeia contains phonetic information that is similar to the real owl's call due to the selection of the following linguistic material:

1. diffuse plosive accounts for prolonged interruptions in the owl's call,
2. the [-acute], [+grave] vowel imitates the owl's sound, and
3. the voiced glottal fricative imitates the voiced glottal modification present in an owl's call.

English onomatopoeia, on the other hand, seems to be phonologically more restricted in imitating the owl's call due to the absence of a voiced glottal fricative. Also, the distinctive features of the compactness and abruptness of /t/ (see note 9) do not correspond with the generally diffused quality of an owl's call. Thus, the only element that has imitational value in English onomatopoeia is the close back rounded vowel /u/ due to its [-acute] and [+grave] features (see note 8).

3.3.2 The acoustic analysis of the sounds of inanimate objects

3.3.2.1 Hitting metal versus English and Ukrainian onomatopoeia

The graphic representation of the sound resulting from two metal objects being hit together is captured in Figure 13. The Ukrainian onomatopoeia of this sound – /dz^{j}in^j/ – is illustrated in Figure 14, and the English onomatopoeic expression echoing this sound – /dɪŋ/ – is visualized in Figure 15.

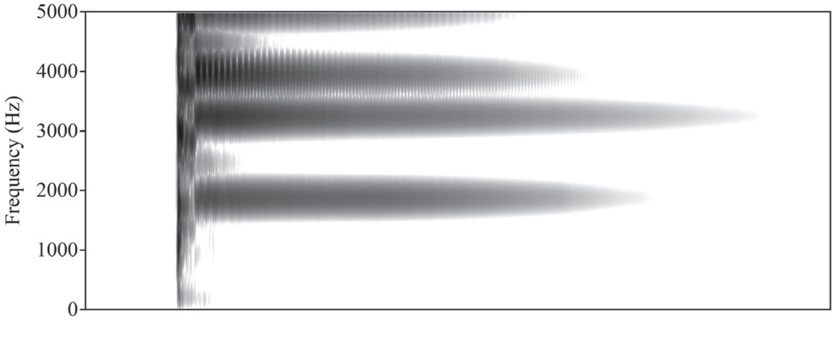

Figure 13. Spectrogram of the sound made by hitting two metal objects together
Slika 13. Spektrogram zvuka koji nastaje udaranjem dvaju metalnih predmeta

In his analysis of the relationship between a natural sound and its linguistic imitation, Tsur (2001) mentions that there is a connection between a sound perceived by the human ear as "metallic" and the diffuse/compact distinction of phonemes (see note 9). The sounds made by vibrating metal decay slowly, with high frequencies showing less damping than low frequencies. In contrast, the sounds made by vibrating wood decay more quickly and show more high-frequency damping. In other words, in the former, the sound has well-defined frequency peaks, while in the latter, the sound has frequency peaks smeared all over the spectrogram (Tsur, 2001: 13–14).

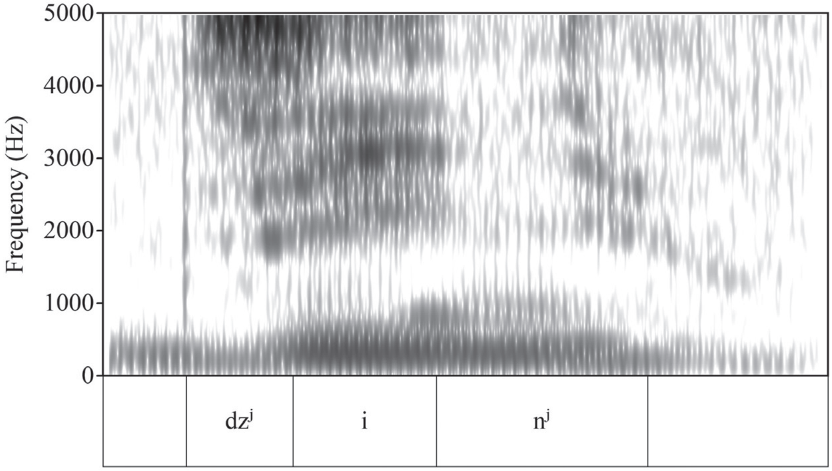

Figure 14. Spectrogram of the Ukrainian onomatopoeia /dz^{j}in^j/ imitating the sound of hitting two metal objects together

Slika 14. Spektrogram ukrajinske onomatopejske riječi /dz^{j}in^j/ koja oponaša zvuk udaranja dvaju metalnih predmeta

The spectrogram of a natural sound in Figure 13 supports Tsur's claims since one can trace well-defined frequency peaks for the first three formants. Combined with the damping of the lower frequencies, it is possible to generalize that the sound has [+compact] and [+acute] features. Consequently, the usage of sounds that are [+compact] and [+acute] will reach the maximum closeness when imitating the natural sound.

In Ukrainian onomatopoeia, the alveolar affricate /dz^j/ stands for the moment of striking. The formant shape of /dz^j/ resembles the formant shape of the beginning of the natural sound. The fricative component in affricate articulation results in the darkening of the higher region of the spectrogram. This indicates the [+acute] feature of the sound. Palatalization plays an important role as well. As mentioned above (3.3.1.1), palatalization is secondary articulation where the tongue reaches the hard palate. Due to this shift in the place of articulation, the palatalized sound will always have a slight concentration of energy in the central region of the spectrogram. This causes the sound to become [+compact][9]. Thus, /dz^j/ is both [+acute] and [+compact]. The high close vowel /i/ echoes a prolonged resonance after the contact of two metal surfaces. The behaviour of the formants during the articulation of /i/ and the resonance in natural sound is almost identical. F_1 is low, while F_2 and F_3 are drawn together, indicating the high-pitched sound. Finally, Ukrainian contains the palatalized alveolar nasal /n^j/ to imitate decaying resonance at the end of the natural sound. Nasals look like weak vowels on the spectrogram (Ladefoged & Johnson, 2010: 200–201). This means that, generally speaking, nasals have less intensity during the articulation, which results in fainter energy bands on the spectrogram. Due to this, it is possible to connect the decaying resonance of metallic sound with weak intensity and fainter formants during the production of nasals. Ukrainian also palatalizes the nasal at the end to make it [+compact] so that the phoneme resembles the [+compact] feature of the natural sound.

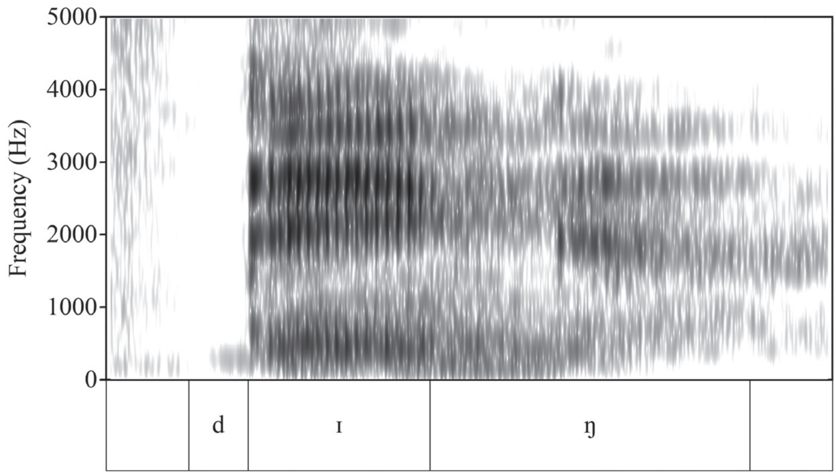

Figure 15. Spectrogram of the English onomatopoeia /dɪŋ/ imitating the sound of hitting two metal objects together

Slika 15. Spektrogram engleske onomatopejske riječi /dɪŋ/ koja oponaša zvuk udaranja dvaju metalnih predmeta

[9] In Ukrainian, only [-compact] consonants can be fully palatalized. Feature-wise, palatalization can be best described in opposition, such as /d/ and /d^j/, where two phonemes differ only in their [+/-compact] opposition.

The English onomatopoeia /dɪŋ/ (Figure 15) includes the alveolar plosive /d/ for the moment that two metal objects are hit together. The darkening at the bottom of the spectrogram indicates that the sound is voiced, which corresponds to the voicing at the beginning of the natural sound. The obstruction to the airflow during the articulation of the plosive results in complete silence on the spectrogram. While the spectrogram of the natural sound clearly resembles a restriction of the acoustic wave, the plosive is not as precise in imitating the beginning of the natural sound as the affricate in Ukrainian. The darkening of the higher region of the spectrogram, caused by the frication in the articulation of affricates, indicates that the latter mimics the beginning of the natural sound with more precision. Feature-wise, the alveolar plosive /d/ is [-compact], which does not correspond with the natural sound's overall [+compact] character.

Onomatopoeia in English includes the high close vowel /ɪ/ for prolonged resonance. This phoneme has a different quality compared to the Ukrainian /i/, but the overall shape of the formants does not differ significantly. The nasal in English onomatopoeia indicates the decaying resonance of the natural sound. The choice of the velar nasal indicates that distinctive features are indeed significant in terms of the auditory process: the velar nasal is the only nasal that is [+compact] in the English consonant system. When the phonotactic rules of a system allow a certain consonant with an appropriate distinctive feature to occur for an imitation of a natural sound, the system will use this opportunity to reach maximum precision in imitation. In this case, English phonotactics allow the velar /ŋ/ to occur at the end of the syllables; consequently, the system is not constrained in its use of the phoneme.

3.3.2.2 Knocking on a door versus English and Ukrainian onomatopoeia

The sound spectrum of the natural sound of knocking on a door (Figure 16) can be divided into three phases: 1) an abrupt explosion on contact with the surface, 2) a short resonance, and 3) an abrupt ending. The spectrogram of this sound features three sequential knocks. Higher formants are damped and spread across the spectrogram, while the exact point of contact at the beginning has a distinct "closure-like" spike of energy. The resonance from the initial contact is short and low in frequency. Also, there is an abrupt spike of energy at the end of two knocks, indicating another "closure".

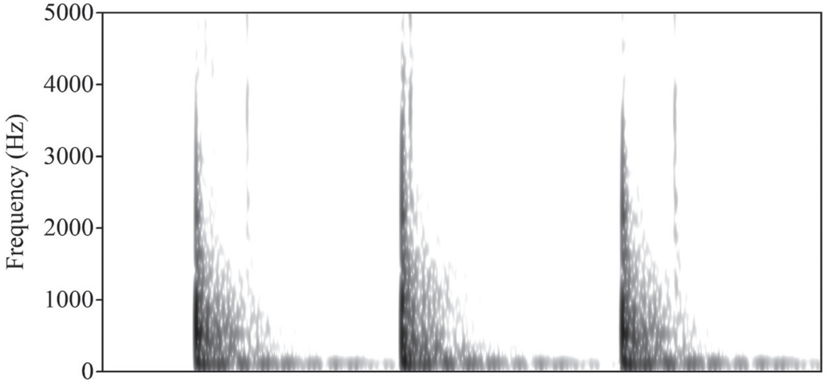

Figure 16. Spectrogram of the sound made by knocking on a door
Slika 16. Spektrogram zvuka koji nastaje kucanjem na vrata

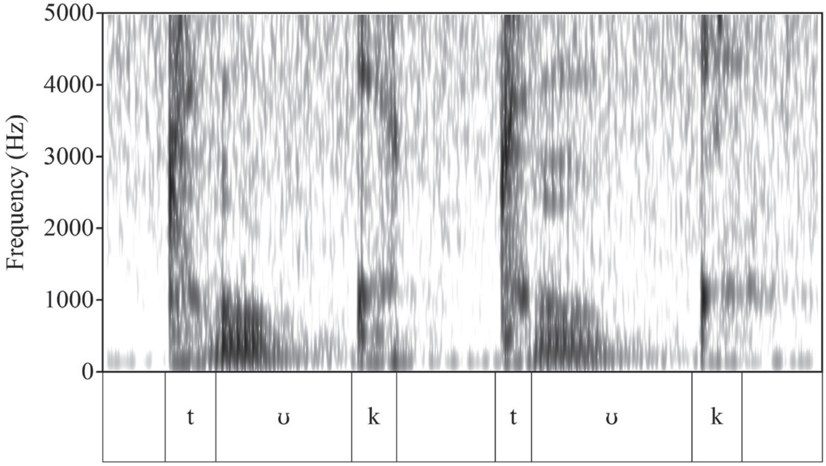

Figure 17. Spectrogram of the Ukrainian onomatopoeia /tʊk/ imitating knocking
Slika 17. Spektrogram ukrajinske onomatopejske riječi /tʊk/ koja oponaša zvuk kucanja

The Ukrainian onomatopoeia /tʊk/ (Figure 17) employs the consonant /t/ to imitate the first phase due to the plosive and diffuse nature of this consonant. The [-continuant, -sonorant][10] features rooted within the plosive correspond to an

[10] The feature [continuant] indicates that, in the process of articulation, there is no obstacle in the supra-glottal space and the air passes freely. The feature [sonorant] means that a sound is voiced, having no voiceless counterpart (see, e.g., Giegerich, 1992: 93).

abrupt explosion and a lack of resonance in the nasal cavity[11]. The feature [diffuse] is important as well since the formants at the beginning of the knocking are smeared over the spectrogram. The difference between [diffuse] and [compact] (see note 9) is neatly shown on the very same spectrogram in connection with /t/ and /k/ at the end of the recording. The frequency values are scattered over the vertical and horizontal axes during the articulation of the /t/. In contrast, the energy is concentrated in the central narrow region of the spectrogram during the articulation of the /k/.

The English onomatopoeia /nɒk/ (Figure 18) uses /n/ to represent the first phase of the natural sound of knocking. The consonant, although plosive, looks nothing like the sound at the beginning of a real knock because of its [+sonorant] feature (see note 10). Despite being occlusives, nasals still have a non-turbulent airflow through the nasal cavity, which evokes a fainter vowel on the spectrogram.

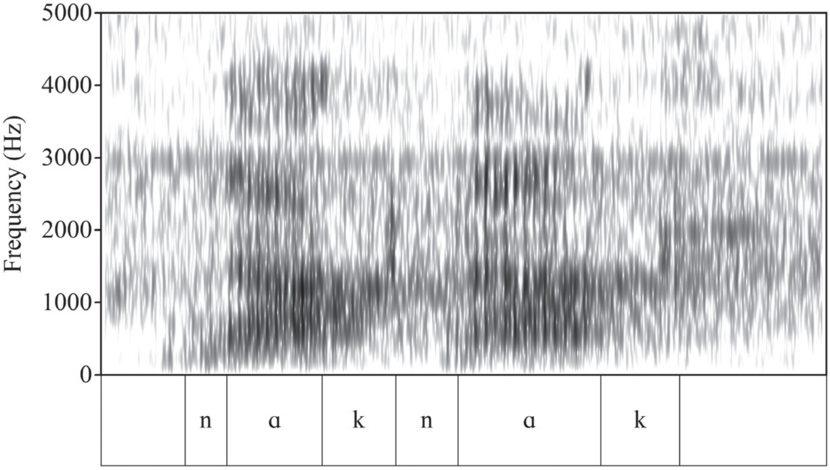

Figure 18. Spectrogram of the English onomatopoeia /nɒk/ imitating knocking

Slika 18. Spektrogram engleske onomatopejske riječi /nɒk/ koja oponaša zvuk kucanja

During the second phase, a short resonance with damped high frequencies occurs. Both languages use vowels to imitate the resonance and, importantly, these vowels are [-acute] (see note 8). The energy of acute vowels is concentrated in the higher regions of the spectrogram, so [-acute] vowels are used to represent high-frequency damping.

Finally, in both languages, /k/ represents the abrupt spike of energy at the end of the natural sound. As Tsur explains, the perceived abruptness is rooted within the

[11] See the shape and intensity of /t/ in /tuk/ and /nʲ/ in /dzʲɪnʲ/ in section 3.3.1. The spectrogram of /nʲ/ looks nothing like the spectrogram of /t/ due to the abovementioned anti-formants.

feature of compactness and, if possible, within a language system, the latter will be picked by users to imitate the concept (2001: 14).

3.3.2.3 Sounds of inanimate objects – summary

Onomatopoeias in both languages closely imitate the natural sound of two metal objects being hit together. In both languages, [+compact] sounds mimic the overall compact feature of the natural sound. Nasals seem to be great imitators of decaying resonance due to their anti-formants and low intensity. The only difference between the two languages is the usage of initial consonants to echo the moment of striking: in Ukrainian, it is the palatalized alveolar affricate /dzʲ/, while in English, it is the alveolar plosive /d/. In this case, the Ukrainian system appears to be more "fine-grained," as Tsur puts it (see section 2). English, on the other hand, does not use such an affricate. Instead, the system prefers the sound that is phonetically closest to the represented sound, the consonant /d/.

The analysis of the sound made by knocking on a door showed that both languages adhere to the [-acute] feature of the natural sound and use phonemes that contain the same feature specification. The difference between the two onomatopoeias lies in their imitation of the moment of striking: Ukrainian uses the alveolar plosive /t/, while English uses the alveolar nasal /n/. Formant behaviour in the /n/ sound does not correspond with the formant behaviour during the moment of striking in the natural sound. The beginning of the natural sound incorporates a rapid spike of energy, which /n/ does not have. On the other hand, /t/ in combination with the following vowel creates the impression of a rapid outburst of energy, making the imitation as precise as possible.

4. CONCLUSIONS

The phonetic properties of onomatopoeias in relation to the phonetic properties of the real sounds from the extra-linguistic reality that onomatopoeias are to imitate have still not been clearly explained in linguistic theory. Tsur (2001) suggests that speech sounds that are used by languages to mimic natural sounds/noises are those whose acoustic characteristics (resulting from articulation) best reflect the acoustic characteristics of the natural sounds.

This research focused on a detailed mutual comparison of the acoustic structure of natural sounds and their linguistic versions to evaluate Tsur's claim. The present pilot study was performed on a sample of four natural sounds (two animal sounds – a cat's meowing and an owl's hooting – and two sounds of inanimate objects – hitting two metal objects together and knocking on a door) and their language imitations in English and Ukrainian. It indicated the validity of Tsur's idea as, in our data from both languages, the speech sounds that participate in the linguistic imitations of natural sounds are those whose acoustic qualities best fit the perceived acoustic qualities of the imitated sounds[12]. In addition, it seems that this "quality correspondence" between natural sounds and their language imitations can be effectively expressed via the notion of distinctive features of phonemes, as Tsur assumes. However, this is a question for future research.

Further steps need to include more sound samples of both natural sounds and their verbal imitations, as well as the number of languages included in the analysis. Only a vast, statistically significant sample of data can provide relevant conclusions with possible typological implications.

Acknowledgments

This work was supported by the Slovak Research and Development Agency (APVV), grant no. APVV-19-0003.

REFERENCES

Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer (Version 6.2.12) [Computer program]. http://www.praat.org

Bredin, H. (1996). Onomatopoeia as a figure and a linguistic principle. *New Literary History*, *27*(3), 555–569.

Giegerich, H. J. (1992). *English phonology. An introduction.* Cambridge: Cambridge University Press.

Gregová, R. (2021). On the position of onomatopoeia in adult language. Evidence from Slovak. *Lingua Posnanienis*, *63*(2), 39–51.

Hayward, K. (2013). *Experimental Phonetics*. London/New York: Routledge.

Jakobson, R. (1978). *Six Lectures on Sound and Meaning*. Cambridge, MA/London: MIT Press.

Jakobson, R., & Halle, M. (1956). *Fundamentals of Language.* The Hague: Mouton.

[12] The nasal /n/ representing the first phase of knocking in the English onomatopoeia (see section 3.3.2.2 for further details) seem to be an exception to this preliminary conclusion. The evaluation of this exception remains open for the future more detailed analysis.

Johnson, K. (2012). *Acoustic and Auditory Phonetics.* 3rd edition. Chichester: Wiley-Blackwell.

Körtvélyessy, L. (2020). Onomatopoeia – a unique species? *Studia Linguistica*, *74*(2), 506–551.

Krupa, V. (1992). Ikonické akustické prvky v slovnej zásobe. *Jazykovedný časopis*, *43*, 26–33.

Ladefoged, P., & Johnson, K. (2010). *A Course in Phonetics.* 6th edition. Boston, MA: Wadsworth.

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, *1*, 227.

Sabol, J., & Zimmermann, J. (2014). *Akusticko-auditívna komunikácia*. Košice: UPJŠ.

Sasamoto, R, & Jackson, R. (2016). Onomatopoeia – showing-word or saying-word? Relevance theory, lexis, and the communication of impressions. *Lingua*, *175–176*, 36–53.

Saussure, F. (1959). *Course in general linguistics*. Eds. A. Sechehaye, C. Bally, & W. Baskin (Trans). New York: The Philosophical Library.

Schmidtke, D. S., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity. *Frontiers in Psychology*, *5*, 80.

Tsur, R. (2001). Onomatopoeia: Cuckoo-language and tick-tocking. The constraints of semiotic systems. https://www.tau.ac.il/~tsurxx/Cuckoo_onomatopoeia.html

Renáta Gregová, Dmytro Hrytsu
renata.gregova@upjs.sk, dmytro.hrytsu@upjs.sk
Sveučilište Pavol Jozef Šafárik, Košice
Slovačka

Akustička analiza onomatopeja na engleskom i ukrajinskom jeziku: pilot istraživanje

Sažetak

Iako postoje brojni izvori o položaju onomatopeja u jezicima, vrlo su rijetki oni koji pišu o fonetskim karakteristikama ovih riječi koje nalikuju na prirodne zvukove. Neke bilješke o fonetskim obilježjima onomatopeja dostupne su u Bredinovim te Tsurovim radovima. Tsur daje detaljniji fonetski opis oponašajućih zvučnih oblika u odnosu na stvarne zvukove. Na temelju akustičke analize onomatopejskoga izraza /ku:ku:/ i stvarnoga glasanja kukavice, Tsur dolazi do zaključka da se u onomatopejskim riječim različitih jezika koriste oni glasovi čije su fonetske osobine najsličnije prirodnim zvukovima. Kako bi se provjerila ta teza, u ovome su radu predstavljeni rezultati pilot istraživanja na uzorku engleskih i ukrajinskih onomatopejskih riječi. Akustičkom su analizom istraživane izabrane engleske i ukrajinske verbalne imitacije prirodnih zvukova (mijaukanje mačke, hukanje sove, udaranje dvaju metalnih predmeta i kucanje na vrata). Akustička obrada uključila je analizu strukture zvučnoga spektra onomatopejskih riječi i usporedbu sa spektrom prirodnih zvukova prema kojima su nastale.

Ključne riječi: onomatopeja, zvučni spektar, akustička analiza, engleski, ukrajinski