PROSODIC CHARACTERISTICS OF ACOUSTIC SIGNAL IN SERBIAN AND EMOTION RECOGNITION

UDC 811.163.41'342.1:159.942.

Nikolina Kuruzović, Valentina Đorić, Ivana Novakov

University of Educons, Faculty of Sport and Psychology, Novi Sad, Serbia

Abstract. *The study focuses on the process of emotional recognition based on prosodic characteristics of acoustic signal in Serbian. Its main objective is to answer the question of the success of prosodic recognition of several emotions in Serbian. Furthermore, it explores whether the success in recognizing emotions in Serbian varies with regards to the following variables: the type of emotion, gender and the type of sentence. A total of 25 participants of average age around 23 participated in the study. The data was created by using Praat software for speech analysis. The results show general success in recognizing emotions on the level of M=12.43 (SD= 6.11). The participants were the most successful in recognizing the emotion of happiness, and the least successful in recognizing sadness. Post hoc analysis found a statistically significant difference in recognizing happiness and emotions of the negative spectrum (sadness, boredom, anger and fear). The results are discussed from the point of social functions of emotions, as well as the characteristics of the affective prosody in Serbian, with recognition of inevitable methodological limitations.*

Key words: *prosody, intonation, pitch, emotion recognition, Serbian, Praat program*

1. INTRODUCTION

The life of man as an inherently social being entails constant functional managing of often very complex social situations. Competent social behavior requires integration of a great deal of information about the interaction partners such as emotions, standpoints and values, intentions, needs and goals, wishes and many other (Morningstar et al. 2018). It is the emotions of other people that contain special social importance and based on them, we form impressions about others and their behavior. This significantly determines the direction of present and future social interactions in the direction of more or less cooperation, affiliation, mutual liking. The ability to observe, recognize, interpret and use information correctly about other people's emotions (so-called emotional/affective cognition) is related to, better social

orientation, social interaction and relationships both in childhood and adulthood, mental health, professional and academic success (Mayer et al. 2008), and individual's general wellbeing (Halberstadt et al. 2001). Emotional cognition begins with observing verbal and non-verbal emotional manifestations. So-called paralinguistic (non-linguistic) communication segments, like facial expression and prosodic characteristics, are very important and are also marked as not less significant than linguistic segments in the process of emotional cognition (Leena 2012). Although earlier research tradition was focused on researching facial expression as the most important modality of emotional expression, a number of studies of acoustic modalities have emphasized that acoustic components such as speech prosody have an important function in affective communication.

The objective of the study was to determine the success of emotional recognition based on prosodic characteristics of acoustic signal in Serbian, as well as to assess eventual variations in success of prosodic emotion recognition according to the type of emotion, listener's gender, and the type of sentences. Another objective was to study the adequacy of implemented methodological procedure, as well as to determine whether the type of sentence content influenced the participant's success in emotion recognition.

1.1. Recognition of emotions based on prosodic characteristics

There is a strong focus on studying emotional prosody in clinical population, as pathological changes in the process of expression and recognition of emotions through prosodic characteristics of speech are a significant diagnostic indication of numerous psychiatric and neurological states. They include schizophrenia, different types of dementia, behavior disorders, and developmental disorders such as ADHD (Bąk 2016). By combining research results of clinical and non-clinical population, a heterogeneous picture of emotional prosody is created, with researchers' consensus on important points. First of all, the authors agree about the fact that basic emotions can be recognized based on the prosodic characteristics of speech, as well as based on the facial expression, but with a mild influence of speaker's and listener's gender, the type of emotion and a manner of expression (Bąk 2016), as well as cultural background (Laukka, Elfenbein 2021).

For some researchers, the equal use of facial expression and prosody in emotion recognition tasks displays clear evidence of the universality of basic emotions (Ekman 1992). Hence, it is not only specific face configurations, but also prosodic markers that enable the recognition of emotions. Listeners are thus able to correctly recognize emotions based on acoustic characteristics, even if the speech is missing certain lexical-semantic content, or it is presented in a foreign language (Pell et al. 2009). There is also indication that the convergence of vocal and facial information is necessary for optimal emotional cognition (Morningstar et al. 2018). Although certain cases show a preference for one of these two modalities, it is only through their combination that a possibility of realistic understanding of emotional information is achieved. Preference for a modality in expression and recognition of individual emotions can be a reflection of biological functions and evolutionary pressures, if we observe vocal signals from the perspective of functional adaptation (Bryant 2020).

Studies of relationship of congruence and incongruence of semantics and prosody show that in situations of congruent semantics and prosody, "prosodic characteristics facilitate expressive and illocutionary power of what has been said" (Bąk 2016: 81). In situations of incongruent semantics and prosody, the listener relies more strongly on prosodic characteristics, which increase the accuracy of emotion recognition (Ekman 2003).

In the situations where the listener relies exclusively on prosody (without semantics), the ambiguity of emotional information increases (Bąk 2016). There will often be a confusion of emotions, especially those of similar intensity. It is possible to confuse the emotions even when they have opposite valence (positive and negative). This indicates that the intensity of the emotion is of great importance when observing and recognizing emotions. There are vocally expressed emotional categories which are not determined by valence, but by unique acoustic characteristics of each individual category (Jaywant, Pell 2010). Although valence is not a completely reliable discriminating factor, it is possible that it arises as significant due to the methodology of studies based on forced choice questions as in artificial situation, which is different from the usual social interactions.

There are indications that gender influences perception of emotions, i.e., that men and women perceive emotions differently, as women rely more on prosodic characteristics of speech in tasks of recognizing emotions. It has been found that women (to a lesser extent) recognize emotions on the basis of prosody more quickly and accurately (Schirmer, Kotz 2003, Goerlich et al. 2011). There is also a tendency of gender stereotyping in the process of interpreting emotions, where the gender of the person showing emotion is a significant factor in the process of attributing emotional states (Schirmer, Kotz 2003). Some studies of Eastern cultures show the opposite situation, in which males of a population achieve better results in recognizing emotions, which is explained by the differences in gender roles assigned by the Eastern culture (Abdel-Hamid 2020, Meftan et al. 2015).

One can also notice the influence of research methodology on the accuracy of recognizing emotions. During a task of recognizing emotions, the participant is most often exposed to an artificial acoustic stimulus (emotions are fake), which is recognized better than a real one in real social situations (Jaywant, Pell 2010).

Hypothesis on universality of emotions in expression and recognition is supplemented by rare cross-cultural studies. The conclusion of the studies is that both universal and culturally-specific processes influence the recognition of emotions based on prosody (Elfenbein, Ambadi 2002, Paulmann, Uskol 2014), as recognition is somewhat more successful if the vocal stimulus is vocalized by a speaker of the corresponding language and not a speaker of a foreign language (Laukka, Elfenbein 2021). At the same time, we need to consider the length of the exposure to vocal stimulus as a significant factor which influences cross-cultural emotion recognition. The effect of the cultural factor is lost with longer exposure (Paulmann, Uskol 2014).

2. METHOD

2.1. Problem and research objectives

The research objective is to examine the main conclusions of studies (Scherer et al. 2001, Bąk 2016, Gendron et al., 2014, Paulmann, Uskol 2014, Laukka, Elfenbein 2021, Sauter et al. 2010, Huang et al. 2019, Abdel-Hamid 2020, Meftan et al. 2015, Elfenbein, Ambadi 2002) which have addressed emotional recognition based on the prosodic characteristics in different languages, and therefore contribute to the understanding of this phenomenon. Considering the relatively small number of studies covering the topic of prosodic emotional recognition in Serbian, we have tried to determine the characteristics of possible cultural distinctiveness of the emotional cognition phenomenon in our population. Results of studies and meta-studies of emotional prosody in different languages, with a smaller number of cross-cultural studies, do

not allow us to make unambiguous conclusions due to variations in success of emotion recognition according to the category of emotions. Although numerous studies point to the greater success of acoustic recognition of negative emotions (Scherer et al. 2001, Bąk 2016, Paulmann, Uskol 2014, Laukka, Elfenbein 2021), there are studies which claim the opposite (Gendron et al., 2014) or define equal success in recognizing both positive and negative emotions in certain cultures (Sauter et al. 2010, Huang et al. 2019). The question of the importance of gender for prosodic recognition of the language has also not been answered clearly, although it is claimed that women are more successful in recognizing emotions based on acoustic characteristics. this advantage of the female population is very small, and sometimes it is not even fully confirmed.

Besides determining the success of emotional recognition based on prosodic characteristics of acoustic signal in Serbian, the goal of this study was to assess subsequent variations in success of prosodic emotion recognition according to the type of emotion, listener's gender, and the type of sentences (stimulus). Another goal was to examine the adequacy of the methodological procedure implemented. Therefore, we tested the differences in success of emotion recognition considering the type of sentence (stimulus), in order to explore whether the sentence content influenced the participant's success in emotion recognition.

2.2. Research sample

A total of 25 participants took part in the study, 16 females and 9 males. The average age of participants was 23.42 (SD=2.23). The participants were students of the IV year of the Faculty of Philosophy at the University of Novi Sad. The participation in the study was voluntary and anonymous.

2.2.1. Research procedure

The research was conducted in the laboratory for phonology of the Faculty of Philosophy in Novi Sad, in two phases: a) preparatory phase and b) main research. The preparatory phase included the design of the stimuli and the preparation of the participants for the main research. After creating the stimuli base, using Praat program, the selection of the participants and their preparation for the main research started. During the lecture on the subject Phonology, we described the research to the students, its purpose and its realization and we asked the students if they were interested to apply. Twenty five volunteers applied. Before the research, the participants were separately interviewed in order to screen for possible hearing or cognitive impairments, using the technique of assessment of affective prosody (Ocić 2012). The Interview was completed by a psychologist. We explained to the participants the steps in the main research in order to arrange for adequate understanding of the procedure. We also clarified that the research had to be conducted individually, without any possible distractions.

We started the main research in the laboratory, which was a quiet space. The participants used headphones while they were listening to the audio. In this way, we enabled the control of the research in terms of interference of other sound signals from the environment. The main part of the research was conducted by using the online program Google Forms. This program was used both for the display of the stimuli as well as for responding. First, by using this program, the acoustic stimuli were played to the participants in the following way: each participant was presented with 30 randomly chosen stimuli – 5 sentences read in 6 different emotional prosodic characteristics, and the participants' task

was to recognize correctly the emotional tone of each sentence read, and to choose 1 of 6 offered answers related to the type of emotion. The average completion time was 8 minutes and 52 seconds.

3. RESULTS

Results of the study are presented below, starting from acoustic characteristics of sentences (Figure 1 to Figure 6) in Serbian, which expresses individual emotions. The stimuli have a comparable intonation structure to the pitch in Serbian proposed by Jovičić (Jovičić et al., 2006). Figure 1 shows the following acoustic characteristics in case of the emotion of happiness: rising intonation, high F0, as well as a wide scope of F0. Articulation is normal. The intensity is high while the tempo is faster. The duration is higher because the length of the syllables is above average.
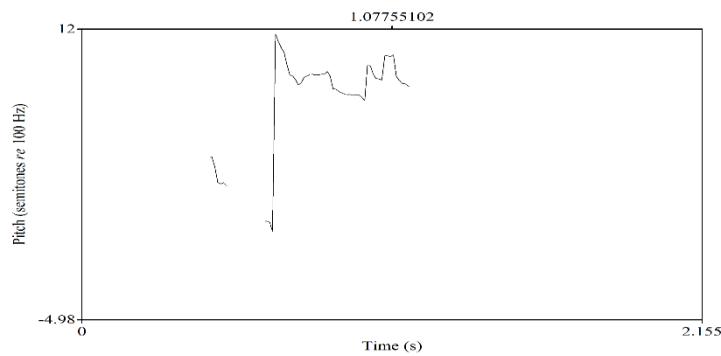

Fig. 1 F0 of sentences expressing the emotion of happiness

Next is the representation of acoustic characteristics of neutral sentences. Figure 2 shows fall-rise intonation. F0 average is lower as well as the range of F0. The intensity is normal, tempo is slower and the duration is shorter.
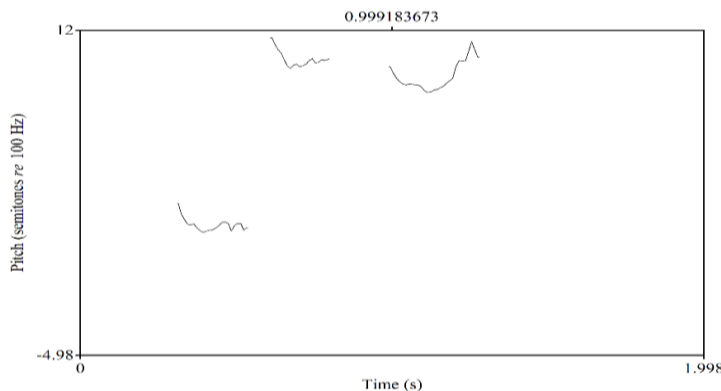

Fig. 2 F0 of neutral sentences

Figure 3 shows characteristics of sentences expressing the emotion of boredom. Intonation for the emotion of boredom is falling. F0 average is lower as well as the range of F0. The intensity is lower, tempo is slower and the duration is shorter.
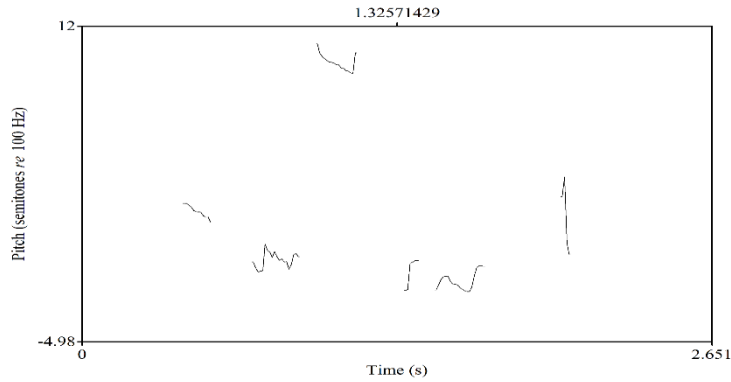

Fig. 3 F0 of sentences expressing the emotion of boredom

Figure 4 represents the characteristics of sentences expressing the emotion of anger. One of the acoustic characteristics of this emotion is falling intonation, similar to every emotionally colored speech. The intensity of speech is higher, F0 is very high, and the tempo is faster, while the quality of speech is tenser. F0 range is wide and duration is higher.
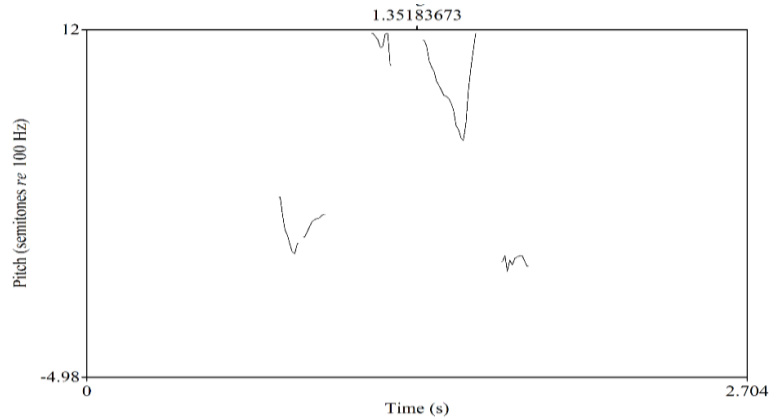

Fig. 4 F0 of sentences expressing the emotion of anger

The tempo in case of the emotion of sadness is slower, the F0 is lower and range is lower, while the quality of speech is unintelligible. Intonation is falling and the intensity is lower. The duration for sadness is the longest and the tempo is slower. Acoustic qualities are shown in Figure 5.
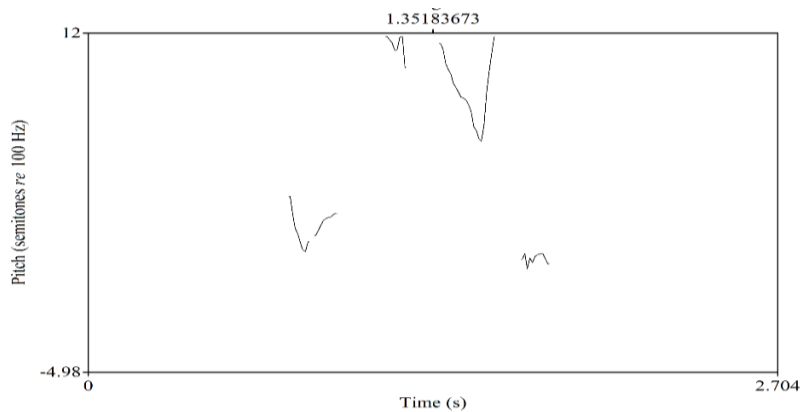
Fig. 5 F0 of sentences expressing the emotion of sadness

The last figure shows the emotion of fear. On Figure 6 we notice that with the emotion of fear the F0 is high and the range of F0 is wide. Intonation is rising. The intensity is normal; duration is low and the articulation itself is precise. The tempo is fast.
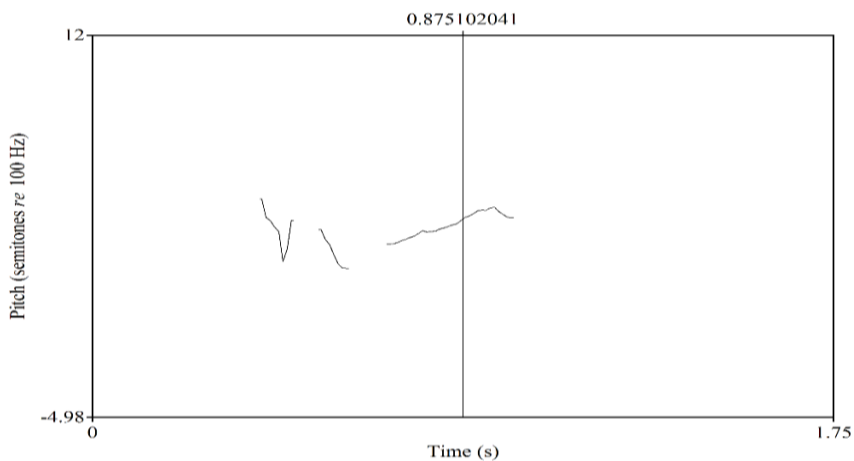

**Fig. 6** F0 of sentences expressing the emotion of fear

After presenting acoustic characteristics of sentences conveying particular emotions in Table 1 the results of success in recognition of individual emotions are presented. Generally speaking, at the level of the entire task, the lowest score was 3 correct answers, and the maximum score was 25 correct answers out of 30 possible correct answers. The average number of correct answers was 12.43 (SD = 6.11).

As indicated in Table 1, the participants could recognize happiness with the greatest success, followed by the neutral emotion. The same table indicates that the participants have demonstrated the least success in recognizing sadness.

Table 1 Percentage of correct answers according to sentences and emotions

Emotion	% of correct answers per sentence				
	They bought a new car.	Milica has gone to the hairdresser.	Ana is travelling by plane.	It is almost 9 o'clock.	The lamp is not on the desk.
Neutral emotion	80%	36%	44%	72%	68%
Happiness	76%	88%	100%	96%	68%
Sadness	24%	44%	32%	32%	16%
Boredom	48%	36%	44%	60%	32%
Anger	36%	36%	16%	68%	60%
Fear	44%	20%	72%	12%	32%

We were interested to investigate whether the differences in the success of recognition of prosodic characteristics of different emotions are statistically significant. Differences in task success regarding the type of emotional tone were analyzed using a one-way ANOVA, where the independent variable was the type of emotion, and the dependent variable the number of correct answers for the given emotion (considering all presented sentences). We wanted to see whether the participants would be more successful in recognizing particular emotions compared to others. Levene's test of equality of variances has shown that the group variances did not show statistically significant differences, $F_{(5,24)} = 1.508$, $p = 0.225$. Values of skewness and kurtosis for number of correct answers in a prosodic task according to type of emotion are within the limits of adequate values (Sk = 0.413, Ku = -0.766), which indicate absence of more significant deviations from normal distribution. One-way ANOVA has shown there were statistically significant differences in success of recognition of different emotional tones, $F_{(5, 24)} = 7.040$, $p < 0.001$. As seen in Figure 7, the participants were most successful in recognizing happiness (M=21.400, SD=3.361), followed by neutral emotion (M=15.000, SD=4.743). The Participants have shown less success in recognizing boredom (M=11.000, SD=2.739) and anger (M=10.800, SD=5.215), and the least success in recognizing fear (M=9.000, SD=5.874) and sadness (M=7.400, SD=2.608). Post-hoc Scheffe test has shown there were statistically significant differences in success of recognizing happiness and sadness, (p=0.002), boredom (p=0.032), anger (p=0.028) and fear (p=0.007).
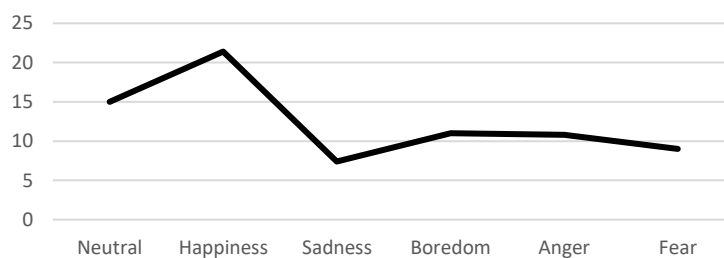
Fig 7. Average number of correct answers according to type of emotional tone

Furthermore, we examined the statistically significant differences in success of recognition of emotional tone according to the type of sentence presented. More precisely, we wanted to check whether the participants would achieve better results in recognizing emotions in case of

particular sentences compared to some other sentences. This question was also answered by using a one-way ANOVA, where the independent variable consisted of five different sentences presented to the participants, and the number of correct answers for each sentence was the dependent variable. Levene's test of equality of variances has shown that group variances did not show statistically significant differences, $F(4, 25) = 0.379$, $p = 0.822$. One-way ANOVA has shown no statistically significant differences of success in the task of recognition of emotional tones depending on the presented sentence, $F(4, 25) = 0.243$, $p = 0.911$. This finding shows that sentence content did not influence the participants' success in emotion recognition based on prosodic characteristics, as shown in Figure 8.
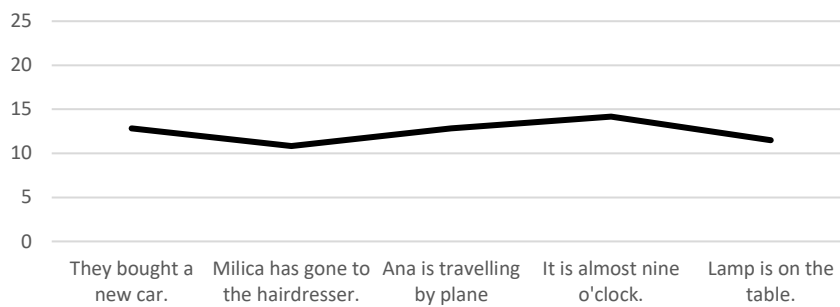

Fig 8. Average number of correct answers based on sentences

Gender differences in general success of emotion recognition based on prosodic characteristics were tested with the T-test for independent samples, where the independent variable was gender, and dependent variable the total number of correct answers in the task, which was 0-30. As the sample size of the two groups was different, we first tested the hypothesis of homogeneity of variants, where Levene's test of equality of variances has shown no statistically significant differences in variants of the two groups, $F = 0.023$, $p = 0.880$. Values of skewness and kurtosis parameters for the results achieved in the task were within the limits of adequate ($Sk = -0.236$, $Ku = 0.114$) and do not show a significant deviation from normal distribution. T-test results have shown no statistically significant differences between the participants of female and male gender in general success in the task of recognition of emotional tone $t(23) = 2.003$. $p = 0.057$, although it can be observed that male participants have achieved slightly greater success in the task ($M = 16.444$, $SD = 2.877$) as compared to the female participants ($M = 13.875$, $SD = 3.181$), as shown in Figure 9.
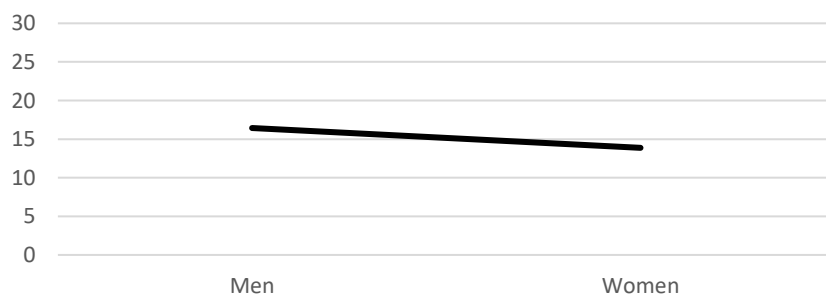

Fig 9. Success in the task according to gender

4. DISCUSSION

Observing the whole sample, the general accuracy rate of emotional recognition based on prosodic characteristics is a little over 40%. This result is generally lower than certain studies (Abdel-Hamid 2020, Paulmann, Uskul 2014, Gendron et al. 2014, Huang et al. 2019, Sauter et al. 2010, Laukka, Elfenbein 2021), although it is difficult to make any comparisons and offer final results. Namely, not only is there no unified measurement for expressing a general level of success of prosodic emotion recognition, but there are no clear cultural standards of comparison (Scherer et al. 2001). The conclusion in most of the aforementioned studies, partly confirmed by this investigation, is that prosodic characteristics of acoustic signals provide participants with enough information in the process of recognizing certain emotions.

The above-mentioned studies, have determined differences in accuracy of prosodic recognition of individual emotions, which is also explored in our study. The results indicate a high rate of success in prosodic recognition of emotion of happiness. We can notice that the participants recognized happiness very successfully based on acoustic characteristics, while their recognition of the negative emotion spectrum (sadness, fear, boredom, anger) had much lower and statistically significant success. Observing the results of other studies encompassing several different world languages (English, Japanese, Taiwanese, Chinese, Italian, Spanish, German, Arabic, etc.) we can notice that most participants are more successful at recognizing negative emotions (Abdel-Hamid 2020, Paulmann, Uskul 2014, Gendron et al. 2014, Huang et al. 2019, Sauter et al. 2010, Laukka, Elfenbein 2021). However, a smaller number of studies, as well as studies comparing several cultures, indicate that certain cultures are either better at recognizing emotion of happiness (Gendron et al. 2014) or equally good at recognizing positive and negative emotions (Sauter et al. 2010, Huang et al. 2019). The results of this study correspond to the smaller number of studies, but differ from the studies which suggest that a greater number of participants of different cultures are more successful when faced with recognition of different negative emotions.

Perhaps it is possible to look for the explanation of the results regarding the recognition of happiness in the social function of the emotion of happiness. Happiness is recognized as an affiliate social sign which is primarily shared among the members of one's own group (culture) (Sauter et al. 2010), whose important function is to improve and facilitate social interactions (Laukka, Elfenbein 2010). Happiness as a part of behavior repertoire of approaching (Elfenbein, Ambady 2003), and contributes to creating a positive social atmosphere, facilitating more frequent social contacts and retaining membership in important social groups. If we agree that Serbian society can still be characterized as a traditional society with high appreciation for group values, i.e., group belonging and sociability within a wider community, then the fact that the participants are very accurate in recognizing sentences with the emotion of happiness is not surprising. It is possible that the participants view socialization in a traditional society, which highly values sociability and group belonging, as more sensitive to affiliative social signs of emotion of happiness. It is believed that each cultural group has its own specific emotional program acquired during socialization, which results in differences in emotional expression and recognition (Elfenbein, Ambady 2003).

We can observe a significant difference between recognition of positive (happiness) and negative emotions (sadness, boredom, fear, anger) in this study. It is possible that the participants are less sensitive to acoustic signs of negative emotions, which is a reflection of avoidance behavior, and as such can lead to impaired social relationships more often. On the other hand, this might not be a lack of participants' sensitivity to acoustic signs of

negative emotions, but minimizing negative emotions. Namely, minimizing communication signs of negative emotions can also be a sign of tolerance, especially in cultures which are collectivist and as such gravitate to the social group (Huang et al. 2019). Members of these cultures can inhibit negative emotions as a way of maintaining interpersonal harmony and avoiding interpersonal conflict or burdening others. This kind of deactivation strategy can reflect on the expression and observing of negative emotions. A smaller sample of this study prevents us from making more definite conclusions about more successful recognition of emotion of happiness. We need to inquire whether the very situation of participating in a study has contributed to this specific result. It might be less likely that the participants do not have a developed ability to recognize negative emotions, especially if we know their biological and evolutionary function of signaling danger (Öhman 2002). It is possible that the very fact the participants were in an artificial situation of emotion recognition, which totally eliminates the need for caution and focus on danger signals, has caused the reduced accuracy of negative emotion recognition. It might be that the acoustic characteristics (e.g., high intensity) of sentences expressing happiness have caused greater accuracy in their recognition. We should not ignore the fact that the participants must choose between several answers offered, more precisely one positive and four negative emotions, and one neutral. Possibly, it was easier to identify the single positive emotion, which is acoustically different. On the other hand, maybe the fact that we have four negative emotions, some of which have similar acoustic characteristics, has produced artificially a greater number of mistakes within this category of emotions. A bigger number of studies and bigger sample may give more specific answers to these questions.

The study indicates that the participants are most successful in recognizing neutral sentences, right after happiness, which is in accordance with numerous other studies (Scherer et al. 2001, Paulmann, Uskol 2014, Laukka, Elfenbein 2021). It is believed that exposure to a certain emotion in daily life can influence one's ability to recognize it. The more we are exposed to a certain emotion, the more successful we are in recognizing that emotion through different communication signals (Paulmann, Uskol, 2014). This is especially true for those emotions which have no prominent acoustic characteristics, which would make them distinctly recognizable. If we follow this line, as well as the fact that neutral quality of everyday communication is one of the predominant, we are not surprised by a relevant success in prosodic recognition of neutral tones in Serbian.

The absence of gender differences in this study is not in accordance with most studies which researched gender differences in ability to recognize emotions based on prosodic characteristics of acoustic signal in different world languages (Schirmer, Kotz 2003, Goerlich et al. 2011). Although in some of those studies the gender difference favoring women is extremely small, it still reaches the necessary level to be statistically significant (Schirmer, Kotz 2003). This study did not confirm such claims, which requires additional research due to the small size of the respondents. For now, we can only speculate whether the absence of gender differences is really particular to our culture when compared to others, or the reason may be the individual characteristics of men and women who participated in the study. The study shows the tendency of men to have somewhat greater accuracy of prosodic emotion recognition, not statistically significant, which makes us nearer some studies of Eastern cultures (Abdel-Hamid 2020, Meftan et al. 2015) trying to discover specific socialization processes related to gender roles imposed by the culture.

The results of our study have confirmed the adequacy of applied methodological procedure in the sense that no statistically significant differences were detected in success

of emotion recognition relating to content of sentences, which shows the uniformity of stimuli presented. On the other hand, making final conclusions on the basis of this study must be seen in the context of several methodological limitations. Limitations of *Praat* software lie in the difficulty of use of certain functions in phonological analysis, as well as the fact that some other types of software are better for some types of phonological analyses. Generally speaking, this is a program that offers a wide spectrum of possibilities and a great data accuracy.

5. CONCLUSION

Basic motivation for realization of this study was that the topic of emotional cognition is contained in a smaller number of studies in Serbia. To this, we add the significance of emotional prosody for social interaction, as a main reason for the study. Besides determining basic prosodic characteristics of acoustic signal in Serbian language, the main goal of the study is to estimate the success of emotion recognition based on prosodic characteristics. Results point to several main conclusions.

First of all, the general rate of success of emotion recognition based on prosodic characteristics in Serbian language is a little over 40%. Although we have a somewhat lower accuracy rate than the mentioned studies in other languages, we believe that the prosodic characteristics of acoustic signal in Serbian language provide enough information during the process of emotional cognition. Especially, as we know that the process of emotional cognition never stops at one modality of information, but strives towards integration of all available information channels.

The next conclusion discusses the presence of differences in the accuracy of prosodic recognition of individual emotions in Serbian language. As in many other languages, recognition of some emotions is acoustically more successful in Serbian language as well. Thus, the participants were most successful in recognizing the emotion of happiness, followed by the neutral tones, while negative emotions such as sadness, fear, boredom, anger were not so successfully recognized Furthermore, the possible explanation can be predominantly traditional quality of our society, which highly values the ideal of social group and group belonging. In this context, minor success rate at recognizing negative emotions can be a reflection of deactivation strategy, whose end goal is to avoid negative social interaction and preserve social relationships.

When generalizing results, we also need to consider methodological limitations. Although we have confirmed the adequacy of this aspect of methodological procedure by examining differences in accuracy of emotion recognition according to the sentence content, other methodological limitations, such as smaller sample size, *Praat* software deficiencies, forced choice of emotions and generally artificial character of research situation, etc., call for repeated studies and additional examination of the results.

In addition to the academic importance of the results for subject fields that study relations of language, emotions and acoustics, like cognitive psychology, linguistics, social psychology, etc., the study could also have practical implications. This refers especially to the area of neuropsychology, which studies aprosodia, as well as speech therapy. We also see wider applications in the areas of communication and social relations.

REFERENCES

Abdel-Hamid, Lamiaa. 2020. "Egyptian Arabic speech emotion recognition using prosodic, spectral and wavelet features." *Speech Communication* 122: 19-30. https://doi.org/10.1016/j.specom.2020.04-005

Bąk, Halszka. 2016. *Emotional prosody processing for non-native English speakers*. Cham: Springer. https://doi.org/10.1007/978-3-319-44042-2_8

Bryant, Gregory A. 2020. "The evolution of human vocal emotion." *Emotion Review* 13(1): 25-33. https://doi.org/10.1177%2F1754073920930791

Douglas-Cowie, Ellen, Nick Campbell, Roddy Cowie & Peter Roach. 2003. Emotional speech: Towards a new generation of database. Speech Communication, 40, 33-60.

Ekman, Paul. 1992. "An argument for basic emotions." *Cognition and Emotion* 6(3/4): 69-200. https://doi.org/10.1080/02699939208411068.

Ekman, Paul. 2003. "Sixteen enjoyable emotions." *Emotion Researcher* 18: 6-7. https://1ammce38pkj41n8xkp1iocwe-wpengine.netdna-ssl.com/wp-content/uploads/2013/07/Sixteen-Enjoyable-Emotions-Paul-Ekman.pdf

Elfenbein, Hillary A. and Nalini Ambady. 2002. "On the universality and cultural specificity of emotion recognition: A meta-analysis." *Psychological Bulletin* 128: 203-235. https://doi.apa.org/doi/10.1037/0033-2909.128.2.203

Gendron, Maria, Debi Roberson, Jacoba Marieta van der Vyver and Lisa Feldman Barrett. 2014. "Cultural relativity in perceiving emotion from vocalizations." *Psychological Science* 25(4): 911-920. https://doi.org/10.1177%2F0956797613517239

Goerlich, Katharina S., Jurriaana Witteman, Andre Aleman and Sander Martens. 2011. "Hearing feelings: Affective categorization of music and speech in alexithymia, an ERP study." *PLoS ONE* 6(5): 1-11. https://doi.org/10.1371/journal.pone.0019501

Halberstadt, Amy.G., Susanne A. Denham and Julie C. Dunsmore. 2001. "Affective social competence." *Social Development* 10(1): 79-119. https://doi.org/10.1111/1467-9507.00150

Huang Yu-Lien, Sue-Huei Chen and Huai-HsuanTseng. 2019. "Attachment avoidance and fearful prosodic emotion recognition predict depression maintenance." *Psychiatry Research* 272: 649-654. https://psycnet.apa.org/doi/10.1016/j.psychres.2018.12.119

Jawant, Abhishek and Marc D. Pell. 2010. "Listener impressions of speakers with Parkinson`s disease." *Journal of the International Neuropsychological Society* 16(1): 49-57. https://doi.org/10.1017/S1355617709990919

Jovičić, Slobodan, Zorka Kašić i Miodrag Đorđević. 2006. *Paralingvističke i ekstralingvističke informacije u govornim tehnologijama*. Zbornik radova 50. Konferencije za ETRAN, Beograd, 2, 443-449.

Laukka, Petri & Hillary A. Elfenbein. 2021. Cross-cultural emotion recognition and in-group advantage in vocal expression: A Meta-Analysis. *Emotion Review*, 13(1), 3-11. https://doi.org/10.1177%2F1754073919897295

Leena, Mary. 2012. *Extraction of prosody for automatic speaker, language, emotion and speech recognition*. Cham: Springer. Retrieved from https://www.scribd.com/document/468619789/Mary-l-Extraction-of-Prosody-for-Automatic-Speaker-Language

Mayer, John. D., Richard D. Roberts and Sigal G. Barsade. 2008. "Human Abilities: Emotional Intelligence." *The Annual Review of Psychology* 59: 507-536. https://www.researchgate.net/publication/5907081_Human_Abilities_Emotional_Intelligence

Meftah, Ali, Sid-Ahmed Selouani and Yousef A. Alotaibi. 2015. *Preliminary Arabic speech emotion classification*. In: Proceedings of the IEEE International Symposium on Signal Processing and Information Technology ISSPIT. https://doi.org/10.1109/ISSPIT.2014.7300584

Morningstar, Michele, Eric E. Nelson and Melanie A. Dirks. 2018. "Maturation of vocal emotion recognition: Insights from the developmental and neuroimaging literature." *Neuroscience Biobehavioral Review* 90: 221-230. https://doi.org/10.1016/j.neubiorev.2018.04.019

Ocić, Gordana. 2012. *Klinička neuropsihologija*. Beograd: Zavod za udžbenike i nastavna sredstva.

Öhman, Arne. 2002. "Automaticity and the amygdala: Nonconscious responses to emotional faces." *Current Directions in Psychological Science* 11: 62-66. https://doi.org/10.1111%2F1467-8721.00169

Paulmann, Silke and Ayse K. Uskul. 2014. "Cross-cultural emotional prosody recognition: Evidence from Chinese and British listeners." *Cognition and Emotion* 2: 230-244. https://doi.org/10.1080/02699931.2013.812033

Pell, Marc D., Paulmann, Silke, Chinar Dara, Areej Alasseri & Sonja A. Kotz. 2009. "Factors in recognition of vocally expressed emotions. A comparison of four languages." *Journal of Phonetics* 37: 417-435. https://doi.org/10.1016/j.wocn.2009.07.005

Sauter Disa A., Frank Eisner, Paul Ekman and Sophie K. 2010. *Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations.* Proceedings of the National Academy of Sciences of the United States of America, 107(6): (pp. 2408-2412). https://doi.org/10.1073/pnas.0908239106

Scherer, Klaus R., Rainer Banse and Wallbott Harald Wallbott. 2001. "Emotion inferences from vocal expression correlate across languages and cultures." *Journal of Cross-Cultural Psychology* 32(1): 76-92. https://doi.org/10.1177%2F0022022101032001009

Schirmer, Annett and Sonja A. Kotz. 2003. "ERP Evidence for a sex-specific Stroop effect in emotional speech." *Journal of Cognitive Neuroscience* 15(8): 1135-1148. https://psycnet.apa.org/doi/10.1162/089892903322598102

PROZODIJSKE KARAKTERISTIKE AKUSTIČKI SIGNALA U SRPSKOM JEZIKU I PREPOZNAVANJE EMOCIJA

Studija se bavi procesom prepoznavanja emocija na osnovu prozodijskih karakteristika akustičkog signala u srpskom. Osnovni cilj jeste odgovor na pitanje uspešnosti prozodijskog prepoznavanja nekoliko emocija u srpskom. Osim toga, ispituje se da li uspešnost prepoznavanja emocija u srpskom varira u zavisnosti od sledećih varijabli: vrsta emocija, pol i vrsta rečenice. U istraživanju je učestvovalo ukupno 25 ispitanika, prosečnih godina oko 23. Stimulusi za analizu govora su kreirani pomoću Praat softverskog programa. Prosečna uspešnost prepoznavanja emocija je iznosila M = 12.43 (SD = 6.11). Ispitanici su najuspešniji u prepoznavanju emocije sreća, a najmanje uspešni u prepoznavanju emocije tuga. Post hoc analiza ukazuje na statistički značajne razlike u prepoznavanju emocije sreća i emocija negativnog spektra (tuga, dosada, bes i strah). Rezultati su diskutovati u svetlu socijalnih funkcija emocija, kao i karakteristika afektivne prozodije u srpskom, uz neizbežno razmatranje metodoloških ograničenja..

Ključne reči: *prozodija, intonacija, visina tona, prepoznavanje emocija, srpski jezik, Praat program*