https://doi.org/10.31820/f.35.1.6

*Iva Bašić, Gordana Varošanec-Škarić*

# THE NORMALIZATION OF VOWEL FORMANTS – A CASE STUDY OF CROATIAN AND SERBIAN

*dr. sc. Iva Bašić, Sveučilište u Zagrebu, Filozofski fakultet*
*ibasic@ffzg.unizg.hr orcid.org/0000-0002-2171-9300*
*dr. sc. Gordana Varošanec-Škarić, Sveučilište u Zagrebu, Filozofski fakultet*
*gvarosan@ffzg.unizg.hr orcid.org/0000-0002-0864-4252*

*izvorni znanstveni članak*
UDK 811.163.42'342.41
811.163.41'342.41
rukopis primljen: 18. srpnja 2022; prihvaćen za tisak: 12. lipnja 2023.

*The aim of this study was to compare the results of a traditional formant analysis of vowels with the results of normalization systems on the example of Croatian and Serbian speech. Male native speakers of Croatian and Serbian were used for this study (N=92). Traditional results of formant analyses express differences among analysed groups of speakers caused by linguistic, sociolinguistic, but also physiological factors. Considering that the values of formant vowels are influenced by many factors, including idiosyncratic physiological characteristics of the vocal tract, normalization approaches remove those variables among speakers that are caused by mutual physiological differences. Therefore, the dialectal, inter-linguistic and/or sociolinguistic differences among speakers whose speech is being analysed are isolated in a scientifically more objective way. The results of this study have shown that formant values are more grouped together and centralized (especially in vowels [a] and [i]), than in non-normalized results within each language individually. This contrastive analysis has shown that in Croatian [i], [o] and [u] are more closed and frontal, the vowel [a] is more closed and back, and the vowel [e] is more open and front, in relation to the vowels in Serbian. This study exemplifies the advantage of normalization systems in the interpretation of acoustic results.*

**Keywords:** *vowels; formants; normalization; Croatian; Serbian*

## 1. Introduction

Vowel analysis is the subject of many sociolinguistic and socio-phonetic studies. Phonetic, phonological and other features of vowels are analyzed among speakers of different languages, dialects, genders, age, ancestry, level of education, etc. Except for individual analyses of the vowel systems of Croatian and Serbian (for Croatian: Bašić 2018, Bašić and Biočina 2020, Jurišić and Bašić 2022, Varošanec-Škarić and Bašić 2015, Bakran 1990, Bakran and Stamenković 1990, Bakran 1996, Škarić 1991, Varošanec-Škarić 2010, Varošanec-Škarić 2005, Pletikos 2003, 2005, Kišiček 2012, and for Serbian: Ivić and Lehiste 1963, 1965, 1967, 1996, Simić and Simić 1989, Rađenović and Jovičić 1997, Jovičić 1999, Jovičić and Kašić 2009, Jovanović and Jovičić 2011, Jovičić et al. 2015, Gudurić 2004, Marković and Bjelaković 2006, Marković 2007, 2012, Prica and Ilić 2010, Sudimac 2016a, 2016b), the last ten or so years have seen the conduction of several contrastive analyses of Croatian and Serbian. The first study dealing with this subject was published by Marković and Bjelaković (2008), who compared the vowel systems of Croatian and Serbian speakers of both sexes from two different corpora. A review of this study by Marković and Bjelaković (2008) was presented in Bašić (2018). The conclusion of the conducted analysis point to the differences in the articulation of the vowels [e] and [o], which are more open in Croatian (when compared to Serbian long vowels, but more closed than short vowels). Furthermore, the authors indicate that the front vowel [i], the back vowel [u] and the central vowel [a] are more peripheral in Serbian.

One of the most frequently used acoustic methods for analyzing vowel production is formant analysis, that is, the analysis of vowel formant (frequency) values. The values of formant frequencies and the level of their variability are influenced by many factors. If we define formants as resonant frequencies of the vocal tract (Johnson 2012, Harrison 2013), spectral peaks of the acoustic spectrum (Fant 1960, Fry 1979, Crystal 2010) or the consequence of vocal passage resonance (Clark and Yallop 1995) it becomes clear that the very length of the vocal tract, along with the circumference, shape and size of resonant cavities, results in different formant values. For example, due to a change in the mentioned categories of the articulation tract, the same speaker will realize different vowels that are, according to Harrison (2013) realized in a wide specter of formant patterns. Although the relationship between formant values and the articulatory behavior of

the vocal tract during articulation is very complex, its illustration is usually simplified through the correlation between different tongue positions (with regard to height) in the dimensions front/back and open/closed (that is high/low) and the first two formants. The first formant (F1) correlates with the height of the tongue, i.e. with the openness or closeness of the vowel, while the second formant (F2) correlates with the front or back positions of the tongue. If the vowel is more open (or lower), then the value of its F1 will be higher, and if the vowel is more front, it will have a higher second formant (F2) value.

Apart from more expressive variations in formant values resulting from the multiple reorganization of the vocal tract (such as among the formant values of vowels [a], [i], [u] etc.), each speaker also exhibits speech variability manifested as a change of speech style, different coarticulatory influences, speech health, etc. For example, formant values for the vowel [a] will be different depending on whether we have analyzed the vowel in speech while reading, spontaneous speech, or in isolated words (disfluent speech). Furthermore, formant values of vowels differ according to their phonemic environment, and not only in the transient area, but also in the central stable part of the vowel (Bašić 2018). For example, the vowel [i] in the Croatian word *sito* will have different formant frequencies than in the words *kipar* or *milo*. Also, many authors have investigated and shown that the consecutive multiple articulation of an isolated vowel by the same speaker will result in different formant values, despite the removal of influences from other factors.

The so far mentioned factors of variability and formant frequencies are unified in forensic phonetics under the term intra-variability, or the variability of a single speaker. On the other hand, the differences and variability of formant frequencies among different speakers are covered by the term inter-variability, or the variability among speakers. In this way we can, for example, analyze the differences in formant values of a single speaker or among different speakers, depending on the posed research questions. Variability among speakers is also caused by the differences in the length, shape and circumference of the vocal tract (especially between groups of men, women and children), which was highlighted by Peterson and Barney (1952). Differences on formant values in identical twins show that variability is also possible between two identical physiological bases (an identical vocal tract), due to an acquired idiosyncratic articulatory behavior (Nolan and Oh 1996, Loakes 2006, 2008).

### 1.1 Vowel normalization

Vowel normalization is a statistical procedure with which we can justifiably compare acoustic data, such as formant values among different groups of speakers. For example, among speakers of different sex, language, dialects, etc. The idiosyncratic physiological characteristics of speakers (head size, vocal tract length and shape) are eliminated as unnecessary acoustic data variability. In other words, if the corpus we are analyzing contains both male and female speakers, it is not justifiable to compare formant values of female speakers with those of male speakers in the interpretation of results, because formant values in male speakers are lower in relation to female speakers due their longer and larger vocal tracts. Vowel normalization eliminates the physiological differences among speakers, while not interfering with their mutual dialectal, intra-linguistic and sociolinguistic differences. Non-normalized results of formant frequencies do not offer information on how much the differences between analyzed groups of speakers are influenced by linguistic and sociolinguistic factors, and how much by physiological ones. The official NORM website, which offers some of the better known normalization packages, highlights that "vowel normalization is crucial when comparing vowel realizations among different speakers in a linguistically and sociolinguistically sensible way" (Thomas and Kendall 2007).

### 1.2 Types of normalization methods

Normalization methods are basically divided into intrinsic and extrinsic ones. When all information for normalization is found in a single token, we are dealing with the vowel intrinsic normalization method, which combines formant values (F1, F2, F3 and sometimes F4), fundamental frequency values, and often the formant bandwidth values as well. On the other hand, the vowel extrinsic normalization method compares formant values of different vowels between different speakers. The parameters pertaining to the speakers, vowels and formants can be changed in both methods, which accordingly leads to multiple combinations: intrinsic/extrinsic in terms of speakers, intrinsic/extrinsic in terms of vowels, and intrinsic/extrinsic in terms of formants. If we use information pertaining to only one speaker, it is advisable to use the speaker intrinsic method. In case we are dealing with more speakers, it is best to use extrinsic normalization (for example, normalization according to Labov, Ash and Boberg (2005)).

Adank (2003) has examined 12 normalization methods on a large sample of vowels from standard Dutch. Her results have shown that sociolinguistic information (on the regional varieties of Dutch) is most successfully preserved by methods that are intrinsic in terms if formants and speakers, while extrinsic ones are more suitable for vowels. The author especially highlights the advantages of normalization methods by Lobanov (1971) and Nearey (1978). The same results were obtained in a study by Flynn (2011), in which the author examined the efficiency of neutralizing the variability of formant data with consideration to the physiological and anatomical differences between speakers, by using 20 different methods of vowel formant normalization.

Many authors emphasize that the aim of the normalization procedure should be questioned prior to vowel normalization, and that it should be decided which method is best suited for the corpus at our disposal. A research question should determine whether we even need vowel normalization, and if we do, which type. A random choice of a normalization package leads to the manipulation of data, and finally the results, without any clear awareness of which acoustic data has been eliminated. Ferrari Disner (1980) and Thomas (2008) list four general goals of normalization:

a) to eliminate variability caused by physiological differences between various speakers,
b) to preserve sociolinguistic, dialectal and intra-linguistic differences in vowel quality,
c) to preserve phonological differences among vowels, and
d) to shape cognitive processes that enable us to normalize vowels articulated by different speakers.

In phonetic research, the goal of normalization differs from the goals of sociolinguistic and dialectal studies, and is mostly concerned with the shaping of cognitive processes that aid us, as listeners, to normalize the vowels of our interlocutors. On the other hand, sociolinguistic and dialectal studies stress the latter as the least important, while the first two goals (a) and (b) are of prime importance.

## 2. Data, materials and methods

A corpus of 46 native Croatian speakers and 46 native Serbian speakers was used for the present analysis (N=92). This sample was collected as

part of a larger corpus[1] comprising the research by Bašić (2018). Speakers were chosen according to the following criteria: speech status, place of birth and longer residence, origin of parents, level of education, and age. The first three criteria were chosen in order to mitigate as much as possible the local characteristics of speech. The fourth criterion was chosen based on the assumption that the group of medium and highly educated speakers would probably contain more speakers with a general pronunciation. According to the final criterion related to age, relatively young speakers were chosen, forming a very homogenous group of the average age of 24 for Croatian (min=18 y., max=29 y., median=23), and the average age of 23 for Serbian (min=19 y., max=35 y., median=21).

Those speakers whose speech was judged to be non-regional by a verification process were chosen from the corpus. Based on the speech material composed of 50 shorter sentences, target words containing the analyzed vowel in the initial syllable were prepared. Each Croatian vowel ([i], [e], [a], [o] and [u]) was presented through 10 words[2], which also proved to be methodologically justified in the preliminary study by Varošanec-Škarić and Bašić (2015). Serbian speakers were verified by phonetic experts, native Serbian speakers, while the verification of Croatian speakers was performed by the authors of this study, also phonetic experts, and native Croatian speakers. Formant values were measured in the stable part of the accentuated vowel. All speakers were recorded with high quality acoustic equipment in studio conditions or in silent rooms with reduced noise levels.

---

[1] Ministry of Science and Education project *Forensic phonetics: auditory recognition and acoustic analysis of speech (no. 130-0000000-0786)* and two short-term grants provided by the University of Zagreb: *Forensic phonetics$_2$: the standardization of acoustic procedures* and *Forensic phonetics$_3$: measures of the fundamental frequency (f0) of male speakers*, under the supervision of Prof. Gordana Varošanec-Škarić, Phd.

[2] Each vowel of Croatian and Serbian language was represented by 10 target words:
- /patka/, /mati/, /faktor/, /tako/, /dabar/, /žaba/, /vata/, /lakat/, /džabe/, /gadno/ for vowel [a]
- /pekar/, /metar/, /leti/, /neto/, /deka/, /šetnja/, /redom/, /četa/, /sedmi/, /keper/ for vowel [e]
- /pita/, /vika/, /rikne/, /bitno/, /čipka/, /nikad/, /sito/, /tigar/, /dika/, /šiba/ for vowel [i]
- /poklon/, /kopar/, /noga/, /fokus/, /soba/, /šogor/, /roba/, /voda/, /joga/, /čopor/ for vowel [o] and
- /buka/, /tupo/, /kupka/, /šupa/, /zubni/, /sutra/, /čudo/, /ruka/, /jutro/, /luka/ for vowel [u].

### 2.1 Acoustic analysis – measuring formant frequencies

The acoustic analysis of the recorded spoken material included sound editing and measuring formant frequencies. At the beginning of the editing process, speaker introductions were isolated, along with pauses and similar material that was not informative for this study. This was followed by prepared target words placed at the ends of sentences. Three points for which formant frequencies were measured were determined in the central part of the target stressed vowel. Frequencies were measured (in Hz) for the first three formants (F1-F3) with the Praat program for the acoustic analysis of sound (Boersma and Weenink 2022). Based on the conducted acoustic processing, a total of 41.400 tokens was acquired, which means that 450 tokens were analyzed per speaker. Although long and short vowels are not differentiated in Croatian and Serbian, some researchers showed differences between formant frequencies of vowels in Croatian and Serbian, according to vowel duration (vowels under short or long stress). However, despite duration being one of the main prosodic dimensions in speech organization, formant contours do not correlate with pitch accents in Croatian (Pletikos 2003). According to that, in this paper primary research aim was to analyze and describe differences in the acoustic characteristic of vowels in Croatian and Serbian. Normalization procedures used in this paper and other normalization procedures are usual and mandatory methods when comparing different speakers.

### 2.2 The normalization of vowel formants (NORM)

After the acoustic processing and speech material analysis, acquired data was adapted for processing in the normalization packages available through the link to the NORM website: The Vowel Normalization and Plotting Suite (Thomas and Kendall 2007). In this study, three normalization methods were used for data analysis (of formant frequencies): a modified Watt and Fabricius method, along with Lobanov's and Nearey's methods. All three methods belong to the same classification group (vowel extrinsic and formant and speaker intrinsic), which according to Adank (2003) and Flynn (2011) scores the best results when it comes to balancing and adjusting/aligning the vowel systems of different speakers. A non-normalization representation of formant frequencies, whose aim was to compare the representation results to those produced by the normalization packages, was also used for the present study. The conduction of normalization methods required several tokens to be categorized into vowel classes ([i], [e], [a], [o]

and [u]), speakers, and target words. Finally, median values were calculated for each vowel category of every individual speaker.

## 3. Results and discussion

### 3.1 Non-normalized vs. normalized representations of results – individual and group values

Based on the numerical data obtained through the acoustic formant analysis of Croatian and Serbian vowels, which were adapted for use in different normalization packages (NORM), non-normalized (left) and normalized (right) average formant values of each speaker and each vowel are shown individually (*Figure 1*), in accordance with the modified Watt and Fabricius method (Watt and Fabricius 2002). This study used the modified Watt and Fabricius method because Thomas and Kendall (2007) warn that the original method proposed by the authors could skew, i.e. distort the lower part of the vowel area. All normalization results were subsequently scaled, as the use of the Watt and Fabricius method, and Nearey's and Lobanov's methods calls for the scaling of results due to the products of the specified methods being vowel areas that are neither in hertz, nor in barks. The scaling factor is mirrored in the modification of an individual's vowel area in a value similar to hertz. All other normalization methods stress that result scaling should be avoided (Kendall and Thomas 2010).

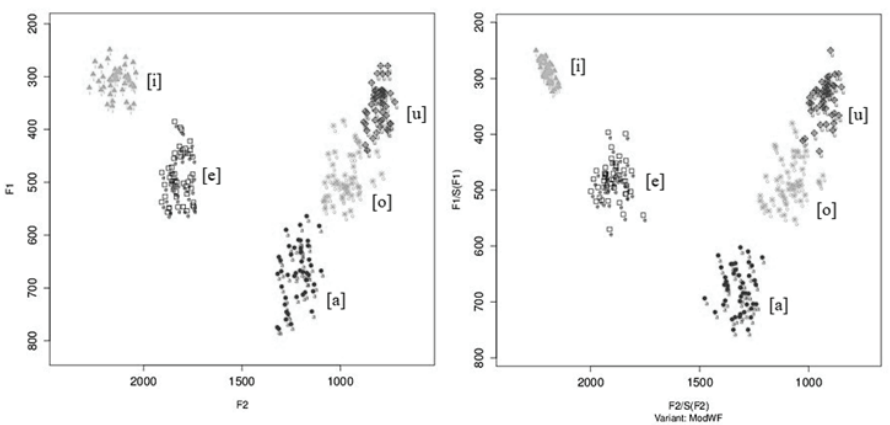


***Figure 1*** *Non-normalized (left) and normalized (right) formant values of Croatian vowels*

In the *Figure 1* the average values of all speakers (N=46) has been marked by different color for each vowel ([a], [e], [i], [o], [u]). Based on this representation, we can see that the dispersion of values in all vowels is reduced in the normalized representation of results, that is, the values are more grouped together and centralized (especially in the vowels [i] and [a]). In this way, the vowels in the vowel area F1 x F2 are mutually more separated and their mutual overlaps are mitigated simultaneously. The same results are also proposed by Lobanov (1971), and Watt and Fabricius (2002).

Apart from Croatian, values of formant frequencies were also processed by non-normalization and normalization methods for Serbian. *Figure 2* clearly shows the differences between the representations of vowel areas of the same group of speakers, depending on whether the results of the formant analysis were normalized or not. The normalized representation of vowel systems (right) is uniform and centralized, which resulted in a minor dispersion of results and minor overlaps of vowel presentations of individual vowels in the representation. Furthermore, normalized data offer more precise and exact results, based on which we can make stronger conclusions, especially when comparing normalized results between two groups of speakers (of different sexes, age, language, dialect, etc.).

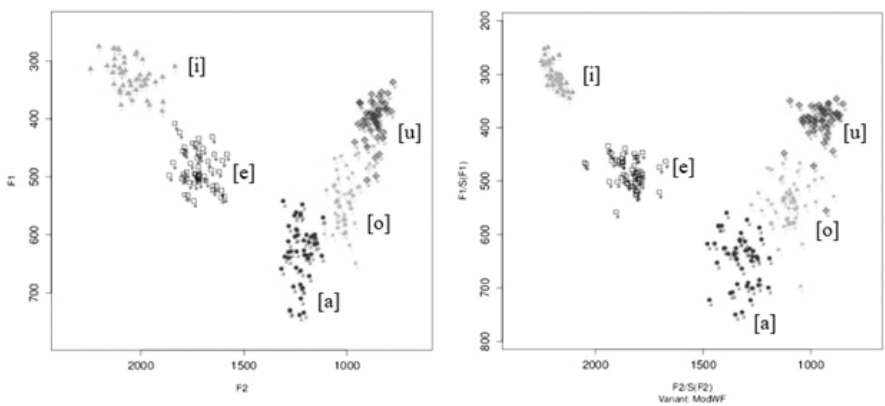


***Figure 2*** *Non-normalized (left) and normalized (right) formant values of Serbian vowels*

Apart from individual values, this study also shows group values in both normalized and non-normalized representations. These representations are frequently used when we wish to highlight the average values of groups of speakers being compared. Differences in these representations are less expressed, precisely because the averaging of non-normalized results included a kind of normalization as well. In other words, group averages of normalized results were doubly normalized (through the averaging and normalization method), so the vowel position resulting from the group average will not be overly centralized and uniformed (see *Figure 3* for an example of Croatian).

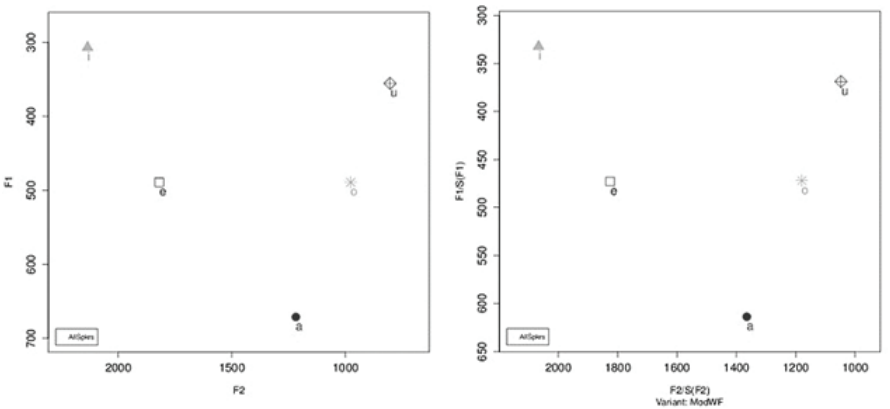


***Figure 3*** *Non-normalized (left) and normalized (right) averaged formant values of Croatian vowels*

### 3.2 Non-normalized vs. normalized results representation – linguistic differences

The previously described results representations of formant analyses on the level of a single language (normalized and non-normalized representation, single and group average values representation) enable a more precise description of the speech of an individual group of speakers. Apart from this, normalization methods offer a representation of dialectal, sociolinguistic and contrastive language differences among various groups of speakers. In the following section, we will show the linguistic differences between the analyzed groups of Croatian and Serbian (male) speakers,

along with the differences in the normalized and non-normalized results representation. *Figure 4* shows the differences between the two types of representations and the linguistic differences between speakers of Croatian (bright grey color) and Serbian (dark grey color).

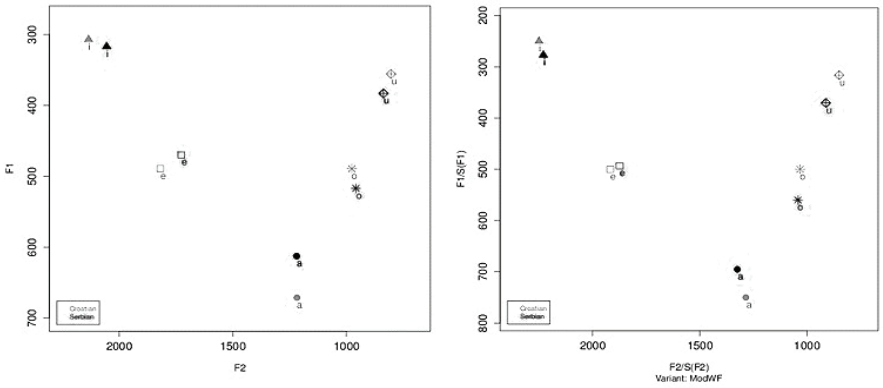


***Figure 4*** *Non-normalized (left) and normalized (right) formant group values of Croatian and Serbian vowels*

The specified figure clearly shows larger differences between Croatian and Serbian speakers in the articulation of the central vowel [a] and front vowels [i] and [e] in a non-normalized representation. A normalized representation (*Figure 4* right), on the other hand, shows greater differences in the articulation of back vowels [u] and [o]. Since normalization eliminates the differences in the length of the articulation passage, and the size and shape of resonance cavities among speakers, and highlights their mutual sociolinguistic differences, we can conclude that the linguistic differences among speakers of the analyzed languages are more significant in the articulation of back vowels and the central vowel [a], while they are insignificant in the articulation of front vowels, i.e. they are caused by physiological differences among speakers.

First three formants (F1-F3) are usually interpreted in terms of vowel height (height of tongue during the vowel pronunciation) to the degree of backness/frontness of a vowel and lip rounding/protrusion. First formant (F1) is related to vowel height (open vowels have higher F1 values, and closed vowels have low F1 values). On the other hand, second formant (F2)

is related to the frontness/backness of a vowel (fronted vowels have high F2 values and back vowels have low F2 values). Third formant is associated with the lip protrusion and rounding (rounded vowel with protruded pronunciation have lower F3 values, since vocal tract is longer due to lengthening and lip protrusion).

The right-hand representation of normalized results in *Figure 4* shows that in relation to Serbian vowels in Croatian are as follows:

- [i] is slightly more closed and front (lower F1 values and higher F2 values),
- [e] is somewhat more open and front (higher F1 and F2 values),
- [a] is more closed and back (lower F1 values and lower F2 values)
- [o] is more closed and front (lower F1 values and higher F2 values)
- [u] is more closed and front (lower F1 values and higher F2 values).

As could be seen, four vowels are more closed, while [e] is more open; four vowels are more front, while [a] is more back.

It should also be mentioned that the distance from group values for a single vowel points to the size of the difference in the area F1 x F2. Consequently, there is a bigger difference in the articulation of back vowels and the central vowel between Croatian and Serbian speakers, than in the articulation of front vowels.

For a better comparison of the different types of result representations, Table 1 offers a list of articulation differences of Croatian vowels in relation to analyzed Serbian vowels, considering the first (F1) and second formants (F2), along with the type of results representation (normalized and non-normalized results). The Table clearly shows that the normalization of results led to a change in the direction of the difference in the articulation of vowels [a] and [o] (shown in red), of which the vowel [a] in the non-normalized representation is determined (only) as more closed, while the normalized representation also shows it as more back. Furthermore, the vowel [o] is determined as more closed and back in the non-normalized representation, while the normalized representation shows it as more closed, but also as more front.

***Table 1*** *Linguistic differences between the vowel systems of Croatian and Serbian in normalized and non-normalized formant results*

| vowel (Croatian) | non-normalized results | formant | normalized results |
|---|---|---|---|
| [i] | more closed | F1 | more closed |
| | more front | F2 | more front |
| [e] | more open | F1 | more open |
| | more front | F2 | more front |
| [a] | more closed | F1 | more closed |
| | – | F2 | **more back** |
| [o] | more closed | F1 | more closed |
| | more back | F2 | **more front** |
| [u] | more closed | F1 | more closed |
| | more front | F2 | more front |

It is evident from what has been described so far that normalization results of acoustic analyses (especially of formants) lead to more precise data that can serve as a basis for conclusions about linguistic differences among analyzed speakers. This data is void of physiological speaker variability that could greatly impact research results.

Although this study strives to draw attention to the necessity of normalizing results of acoustic analyses, primarily of formants, here exemplified through the contrastive linguistic comparison of Croatian and Serbian, it is also necessary to compare it to the results of previous studies as the results differ. Differences are confirmed only for front vowels, while the articulation of other vowels is different.

In the present study, Croatian vowels are sometimes located on the more peripheral parts of the trapeze (the more closed and front vowel [u]) in relation to Serbian vowels, while other times the vowels are more peripheral according to one criterion, and less peripheral according to another. For example, the vowel [i] in Croatian is more closed and front than in Serbian, and this trait makes it more peripheral. On the other hand, because it is less back, it is also less peripheral. Considering that two basic characteristics of vowel articulation (open/closed and front/back) can be read from the trapeze, and that peripherality in articulation does not always have to be determined in both characteristics, for a better understanding of the

conclusions and differences between the compared articulations it would be useful to explain what is understood by peripherality in articulation, and to which characteristics it refers.

Varošanec-Škarić, Bašić and Kišiček (2016, 2017) compare Croatian vowels with Serbian and Slovenian vowels (Varošanec-Škarić, Bašić and Kišiček 2017). The results of the conducted studies, in which the basis of the analyses included a formant analysis of a methodology similar to the one used in the present study (except normalization methods), have indicated a more open articulation of the vowel [a], a more front articulation of the front vowels [i] and [e], and a more back articulation of the vowel [u], in relation to Serbian vowel articulation.

Bašić (2018) presents the results of a contrastive analysis, which differ for the vowels [e] (more closed) and [u] (more back) in relation to the present study.

Formant values could be associated with articulation features (open/closed, back/front, (un)rounded vowel), but also with the perception of vowel brightness and darkness. For example, values of second formant (F2) could be interpreted as the feature of brighter/darker vowels. Although perception of vowels has not been analyzed in this paper, normalized values of F2 could be interpreted in that way. Therefore, we could conclude that front ([i], [e]) and back ([o], [u]) vowels in Croatian could be perceived as brighter in Croatian, compared to Serbian (since F2 values are higher than in Serbian), while central vowel [a] could be perceived as darker in Croatian (lower F2 than in Serbian).

## 4. Conclusions

Vowel normalization contributes to the strength of the interpretation of linguistic, sociolinguistic, and dialectal differences among speakers, eliminating in the process their individual physiological differences. The basic division of normalization methods includes intrinsic and extrinsic methods, and has enabled multiple combinations of three key parameters in the working tools: speakers, vowels, and formants. During the selection process of the vowel normalization method, research questions should be considered and adjusted in line with the characteristics of the corpus. It is necessary to understand that vowel normalization is not a one-dimensional procedure, but a complex network of interrelations among more factors. The differences in variability between speakers have been determined in

the analyzed groups, and in the (articulatory) linguistic differences in the characteristics open/closed and front/back, in relation to the plan F1 x F2. The results of this study have shown that formant values are more grouped together and centralized (especially in vowels [a] and [i]), than in non-normalized results within each language individually. This contrastive analysis of Croatian and Serbian speakers has shown that in Croatian [i], [o] and [u] are more closed and frontal, the vowel [a] is more closed and back, and the vowel [e] is more open and front, in relation to the vowels in Serbian. This study exemplifies the advantage of normalization systems in the interpretation of acoustic results to the results performed on vowel formant analyses.

**Acknowledgements**

We wish to thank all speakers of Croatian and Serbian who have participated in this study and who have agreed to be recorded. We thank Janja Čulig Suknaić, PhD (Faculty of Humanities and Social Sciences, University in Zagreb, Department of Englis) for proof reading this article.

**Funding**

This research has used the corpus collected in the framework of the Ministry of Science and Education project *Forensic phonetics: auditory recognition and acoustic analysis of speech (no. 130-0000000-0786)* and two short-term grants provided by the University of Zagreb: *Forensic phonetics$_2$: the standardization of acoustic procedures* and *Forensic phonetics$_3$: measures of the fundamental frequency (f0) of male speakers*, under the supervision of Prof. Gordana Varošanec-Škarić, Phd.

## References

Adank, Patricia Martine (2003) *Vowel normalization: A perceptual-acoustic study of Dutch vowels*, University of Nijmegen, The Radboud University. https://pure.manchester.ac.uk/ws/portalfiles/portal/33081529/FULL_TEXT.PDF (accessed January 21, 2022).

Bakran, Juraj (1990) "Djelovanje naglaska i dužine na frekvencije formanata vokala", *Govor*, 6, 2, Zagreb, str. 1–12, dostupno na adresi https://hrcak.srce.hr/file/260380, posjet 6. lipnja 2023.

Bakran, Juraj (1996) *Zvučna slika hrvatskoga govora*, IBIS grafika, Zagreb.

Bakran, Juraj; Stamenković, Milan (1990) "Formanti prirodnih i sintetiziranih vokala hrvatskoga standardnoga govora", *Govor*, 7, 2, Zagreb, str. 119–137, dostupno na adresi https://hrcak.srce.hr/file/260006 posjet 6. lipnja 2023.

Bašić, Iva (2018) *Akustička analiza općeprihvaćenoga hrvatskoga i srpskog govora – formantska analiza i mjere fundamentalne frekvencije*, doktorska disertacija, Sveučilište u Zagrebu, Filozofski fakultet.

Bašić, Iva; Biočina, Zdravka (2020) "Metodologija akustičkih snimanja", *Zbornik radova Filozofskog fakulteta u Splitu*, 13, Split, str. 19–40, dostupno na adresi https://hrcak.srce.hr/file/361009, posjet 6. lipnja 2023.

Boersma, Paul; Weenink, David (2022) "Praat: Doing Phonetics by Computer" dostupno na adresi https://www.fon.hum.uva.nl/praat/, posjet 24. kolovoza 2022,

Clark, John W.; Yallop, Collin (1995) *An introduction to phonetics and phonology*, Blackwell, Oxford.

Crystal, David (2010) *The Cambridge encyclopedia of language*, Cambridge University Press, Cambridge.

Fant, Gunnar (1960) *Acoustic theory of speech production*, Mouton, The Hague.

Ferrari Disner, Sandra (1980) "Evaluation of vowel normalization procedures", *Journal of the Acoustical Society of America*, 67, Melville, str. 253–261.

Flynn, Nicholas (2011) *Comparing vowel formant normalisation procedures*, dostupno na adresi https://www.york.ac.uk/language/ypl/ypl2/11/YPL2-11-01-Flynn.pdf, posjet 6. lipnja 2023.

Fry, Dennis Butler (1979) *The Physics of Speech*, Cambridge University Press, Cambridge, dostupno na adresi https://doi.org/10.1017/CBO9781139165747, posjet 6. lipnja 2023.

Gudurić, Snežana (2004) *O prirodi glasova*, Zavod za udžbenike i nastavna sredstva, Beograd.

Harrison, Philip Thomas (2013) *Making accurate formant measurements: an empirical investigation of the influence of the measurement tool, analysis settings and speaker on formant measurements*, University of York.

Ivić, Pavle; Lehiste, Ilse (1963) "Prilozi ispitivanju fonetske i fonološke prirode akcenata u savremenom srpskohrvatskom jeziku", *Zbornik za filologiju i lingvistiku,* 6, Novi Sad, str. 31–73.

Ivić, Pavle; Lehiste, Ilse (1965) "Prilozi ispitivanju fonetske i fonološke prirode akcenata u savremenom srpskohrvatskom jeziku", *Zbornik za filologiju i lingvistiku,* 8, Novi Sad, str. 75–119.

Ivić, Pavle; Lehiste, Ilse (1967) "Prilozi ispitivanju fonetske i fonološke prirode akcenata u savremenom srpskohrvatskom jeziku", *Zbornik za filologiju i lingvistiku*, 10, Novi Sad, str. 55–95.

Ivić, Pavle; Lehiste, Ilse (1996) *Prozodija reči i rečenice u srpskohrvatskom jeziku*, Izdavačka knjižarnica Zorana Stojanovića, Novi Sad.

Johnson, Keith (2012) *Acoustic and Auditory Phonetics*, Willey Blackwell, Oxford.

Jovanović, Nikola; Jovičić, Slobodan (2011) "The effect of the manner of mobile phone usage on Serbian vowel formants", *2011 19th Telecommunications Forum (TELFOR) Proceedings of Papers,* Belgrade, Serbia, 1083–1086, dostupno na adresi https://www.researchgate.net/publication/261212310_The_effect_of_the_manner_of_mobile_phone_usage_on_Serbian_vowel_formants#fullTextFileContent, posjet 6. lipnja 2023.

Jovičić, Slobodan (1999) *Govorna komunikacija - fiziologija, psihoakustika i percepcija*, Nauka, Beograd.

Jovičić, Slobodan; Jovanović, Nikola; Subotić, Miško; Grozdić, Đorđe (2015) "Impact of mobile phone usage on speech spectral features: some preliminary findings", *International Journal of Speech Language and the Law*, 22 (1), 111–126, dostupno na adresi https://www.researchgate.net/publication/281706604_Impact_of_mobile_phone_usage_on_speech_spectral_features_Some_preliminary_findings#fullTextFileContent, posjet 6. lipnja 2023.

Jovičić, Slobodan; Kašić, Zorka (2009) "Intra-spikerske varijacije u govoru: forenzičke implikacije", *International Conference on Applied Linguistics*, Novi Sad, str. 129–140.

Jurišić, Aneta; Bašić, Iva (2022) "Akustička analiza i samoprocjena govora sa zaštitnim maskama i vizirima", *Govor*, 38, 2, Zagreb, str. 181–217, dostupno na adresi https://hrcak.srce.hr/file/396964, posjet 6. lipnja 2023.

Kendall, Tyler; Thomas, Erik R. (2010) "Vowels: Vowel Manipulation, Normalization, and Plotting in R. R package, v. 1.1." dostupno na adresi http://ncslaap.lib.ncsu.edu/tools/norm/, posjet 23. kolovoza 2021.

Kišiček, Gabrijela (2012) *Forenzično profiliranje i prepoznavanje govornika urbanih varijeteta hrvatskoga jezika*, doktorska disertacija, Sveučilište u Zagrebu, Filozofski fakultet.

Labov, William; Ash, Sharon; Boberg, Charles (2005) *The Atlas of North American English: Phonetics, Phonology and Sound Change,* De Gruyter Mouton, Berlin, New York, dostupno na adresi https://doi.org/10.1515/9783110167467, pristup 6. lipnja 2023.

Loakes, Deborah (2006) "A forensic phonetic investigation into the speech patterns of identical and non-identical twins", ICPhS Proceedings, str. 691–694, dostupno na adresi https://doi.org/10.1558/ijsll.v15i1.97, posjet 6. lipnja 2023.

Loakes, Deborah (2008) "A forensic phonetic investigation into the speech patterns of identical and non-identical twins", *International Journal of Speech Language and the Law*, 15(1), Sheffield, str. 97–100, dostupno na adresi https://doi.org/10.1558/ijsll.v15i1.97, posjet 2. lipnja 2023.

Lobanov, Boris M. (1971) "Classification of Russian Vowels Spoken by Different Speakers", *The Journal of the Acoustical Society of America,* 49, Melville, str. 606–608, dostupno na adresi https://doi.org/10.1121/1.1912396, posjet 6. lipnja 2023.

Marković, Maja (2007) *Kontrastivna analiza akustičkih i artikulacionih karakteristika vokalskih sistema engleskog i srpskog jezika*, doktorska disertacija, Sveučilište u Novome Sadu, Filozofski fakultet.

Marković, Maja (2012) *Uporedna proučavanja vokala engleskog i srpskog jezika: između univerzalnog i specifičnog*, Filozofski fakultet, Novi Sad.

Marković, Maja; Bjelaković, Isidora (2006) "Neke akustičke karakteristike vokala u govoru Novog Sada", *Godišnjak Filozofskog Fakulteta u Novom Sadu,* 31, Novi Sad, str. 327–346.

Marković, Maja; Bjelaković, Isidora (2008) "Kontrastivna akustička analiza vokalskih sistema srpskog i hrvatskoga jezika", *Die Unterschiede Zwichen Dem Bosnischen/Bosniakischen, Kroatischen Und Serbischen*, ed. Branko Tošović, LIT Verlag, Münster, str. 200–215.

Nearey, Terrance M. (1978) *Phonetic Feature Systems for Vowels*, University of Alberta, Alberta, dostupno na adresi https://sites.ualberta.ca/~tnearey/Nearey1978_compressed, posjet 6. lipnja 2023.

Nolan, Francis; Oh, Tomasina (1996) "Identical twins, different voices", *International Journal of Speech Language and the Law*, 3, 1, Sheffield, str. 39–49, dostupno na adresi https://doi.org/10.1558/ijsll.v3i1.39, posjet 6. lipnja 2023.

Peterson, Gordon E.; Barney, Harold L. (1952) "Control Methods Used in a Study of the Vowels", *The Journal of the Acoustical Society of America*, 24, Melville, str. 175–184, dostupno na adresi https://doi.org/10.1121/1.1906875, posjet 6. lipnja 2023.

Pletikos, Elenmari (2003) *Akustički opis hrvatskih standardnih naglasaka*, doktorska disertacija, Sveučilište u Zagrebu, Filozofski fakultet.

Pletikos, Elenmari (2003) Akustički opis hrvatskih standardnih naglasaka, *Govor*, 20, 1-2, Zagreb, str. 321–345, dostupno na adresi https://hrcak.srce.hr/file/264464, posjet 6. lipnja 2023.

Prica, Biljana; Ilić, Siniša (2010) "Recognition Vowels in Continuous Speech by Using Formants", *Facta Universitatis*, 23, 3, Niš, str. 379–393, dostupno na adresi http://facta.junis.ni.ac.rs/eae/fu2k103/11prica.pdf, posjet 6. lipnja 2023.

Rađenović, Petar; Jovičić, Slobodan (1997) "Neke promene u karakteristikama govora u buci", *41. Konferencija za ETRAN*, Zlatibor, str. 582–584, dostupno na adresi https://www.etran.rs/common/archive/ETRAN_1955-2006/ET(R)AN_1955-2006/eTRAN/41.ETRAN.1997.2/Randjelovic.P.ETRAN.1997.2.pdf, posjet 6. lipnja 2023.

Simić, Radoje; Simić, Branislav (1989) *Osnovi fonologije srpskohrvatskoga književnog jezika*, Univerzitetska riječ, Nikšić.

Škarić, Ivo (1991) "Fonetika hrvatskog književnog jezika", *Povijesni pregled, glasovi i oblici hrvatskoga književnoga jezika: Nacrti za gramatiku*, HAZU, Globus, Zagreb, 17–151.

Sudimac, Nina (2016a) "Sociofonetski pogled na srpske jezične varijetete – percepcija. Akustika. Stavovi.", *Philologia Mediana*, 8, Niš, str. 555–586.

Sudimac, Nina (2016b) "Kontrastivna analiza visokih/zatvorenih vokala u produkciji izvornih govornika britanskog engleskog i srpskog jezika", Filolog, 14, Banja Luka, str. 36–55.

Thomas, Erik R. (2008) "Instrumental Phonetics", *The Handbook of Language Variation and Change*, eds. J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes, str. 168–201, Blackwell Publishing Ltd, Oxford, UK,

dostupno na adresi https://doi.org/10.1002/9780470756591.ch7, posjet 6. lipnja 2023.

Thomas, Erik R.; Kendall, Tyler (2007) *NORM: The Vowel Normalisation and Plotting Suite*, dostupno na adresi http://lingtools.uoregon.edu/norm/norm1.php, posjet 6. lipnja 2023.

Varošanec-Škarić, Gordana (2005) *Timbar*, FF press, Zagreb.

Varošanec-Škarić, Gordana (2010) *Fonetska njega glasa i izgovora*, FF press, Zagreb.

Varošanec-Škarić, Gordana; Bašić, Iva (2015) "Acoustic characteristics of Croatian cardinal vowel formants (F1, F2 and F3)", *Proceedings – Speech and Language 2015*, eds. Mirjana Sovilj and Miško Subotić, Novi Sad, str. 41–50.

Varošanec-Škarić, Gordana; Bašić, Iva; Kišiček, Gabrijela (2016) "Usporedba akustičkih vrijednosti vokalskih formanata prihvaćenoga izgovora hrvatskih i srpskih muških govornika mlađe dobne skupine", Istraživanja govora – 9. znanstveni skup s međunarodnim sudjelovanjem: knjiga sažetaka, eds. Nikolaj Lazić i Elenmari Pletikos Olof, Zagreb, str. 108–109.

Varošanec-Škarić, Gordana; Bašić, Iva; Kišiček, Gabrijela (2017) "Comparison of vowel space of male speakers of Croatian, Serbian and Slovenian language", 26th IAFPA Proceedings, eds. Gordana Varošanec-Škarić and Anita Runjić Stoilova, Split, str. 142–146.

Watt, Dominic; Fabricius, Anne (2002) "Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1–F2 plane", *Leeds Working Papers in Linguistics and Phonetics,* 9, Leeds, str. 159–173, dostupno na adresi https://www.researchgate.net/publication/239566480_Evaluation_of_a_technique_for_improving_the_mapping_of_multiple_speakers'_vowel_spaces_in_the_F1-F2_plane#fullTextFileContent, posjet 6. lipnja 2023.

SAŽETAK

Iva Bašić, Gordana Varošanec-Škarić

NORMALIZACIJA VOKALSKIH FORMANATA U HRVATSKOME I SRPSKOME JEZIKU

Svrha je ovoga istraživanja bila usporediti rezultate tradicionalne formantske analize vokala s rezultatima normalizacijskih sustava, na primjeru hrvatskoga i srpskoga govora. Za potrebe rada analizirani su muški izvorni govornici hrvatskoga i srpskoga jezika (N=92). S obzirom na to da na vrijednosti formanata vokala utječu brojni faktori, između kojih i idiosinkratičke fiziološke karakteristike govornika, normalizacijom se uklanja varijabilnost među govornicima uzrokovana njihovom fiziološkom različitošću. Normalizacijom vrijednosti formanata utvrđen je viši stupanj centralizacije svih vokala obaju jezika u usporedbi s ne-normaliziranim vrijednostima formanata, dok je kontrastivna analiza među jezicima ukazala na razlike u obilježjima prednjosti i stražnjosti te otvorenosti i zatvorenosti kod svih vokala.

**Ključne riječi:** *vokali; formanti; normalizacija; hrvatski jezik; srpski jezik*