



International Journal of  
**MULTILINGUAL EDUCATION**

**ISSN: (Print)** ISSN 1987-9601

**(Online)** E ISSN 1512-3146

**Journal homepage:** <http://multilingualeducation.org/>

---

## **Investigating Speech Characteristics of Georgian Native and Non-Native Speakers: A Forensic Phonetics Study**

**Mariam Navadze**  
Ivane Javakhishvili Tbilisi State  
University, Georgia  
Email: [mariam.navadze@tsu.ge](mailto:mariam.navadze@tsu.ge)

To cite this article: Mariam Navadze (2022): Investigating Speech Characteristics of Georgian Native and Non-Native Speakers: A Forensic Phonetics Study: International Journal of Multilingual Education, #21;  
DOI: 10.22333/ijme.2022.21000; pp. 19-30.

To link to this article: <https://doi.org/10.22333/ijme.2022.21003>

**Mariam Navadze**

*Ivane Javakhishvili Tbilisi State University, Georgia*

## **Investigating Speech Characteristics of Georgian Native and Non-Native Speakers: A Forensic Phonetics Study<sup>1</sup>**

### **ABSTRACT**

Forensic Phonetics aims to identify speakers through various speech characteristics that may differentiate them from others. This paper discusses the importance of selecting appropriate parameters that are independent and have high inter-speaker and low intra-speaker variation. Speech rate and fundamental frequency are analyzed to collect statistical information on the Georgian language for both native and non-native speakers. The study recorded oral speeches of 60 Georgian native speakers from three different age categories, and 20 high-competence Azerbaijani-speaking Tbilisi State University students. Results show the normal articulation tempo of the Georgian language to be between 5.1 to 6.3 syllables/second. Additionally, fundamental frequency differences were observed in the non-native speaker group compared to the native speaker group.

**Keywords:** *Fundamental frequency, Articulation rate, Speech analysis, Forensic phonetics, Speaker identification.*

### **1. Introduction**

Forensic Phonetics is a cross-disciplinary field and is one of the branches of applied linguistics. The field aims to enhance the quality of audio recordings used in the investigation process, to decode the content of hard-to-understand or damaged sections, verify the authenticity of audio recordings and to identify speakers. Identifying speakers is a key issue in forensic phonetics. As is typical in the speaker identification process, two or more audio recordings are compared. One of them is the questioned recording in which the investigation attempts to identify the voice of the suspect and the

---

<sup>1</sup> Research was funded by Shota Rustaveli National Science Foundation of Georgia (Project ID PHDF-18-405)

other is a known voice sample that is obtained from suspect and the speaker's identity is known for certain.

In the process of examination, voice samples are compared on the basis of many speech characteristics through which speakers may differ from one another. These characteristics should be chosen on a case-by-case basis, considering the unique features of the search samples. While there are some criteria that should guide the expert in the selection of search parameters. Parameters must:

- Have a high degree of inter-speaker and a low degree of intra-speaker variations;
- Not be subject to attempts of imitation or disguise;
- Be simple and accurate to report;
- Be relatively easy to measure (Nolan, 2009).

Philip Rose (Rose, 2002) puts an additional condition in the process of selecting parameters - each parameter should be as independent as possible from the other parameters, so that the research is strong in terms of the speakers' discriminant power.

In the research process, it is very important to determine whether the similarity (or difference) between the compared samples according to a certain parameter is speaker specific or is a common in the relevant population. The more uncommon a characteristic is among speakers, the more effective it will be in differentiating between them. Consequently, it is important for each language to have statistical data on the prevalence of certain speech parameters. The research presented in our current article aims at collecting this statistical information on the Georgian language.

Research parameters can be categorized as segmental and supra-segmental units of speech, voice and manner of speaking. Some characteristics can be analyzed using both auditory and acoustic methods. The most common characteristics are: language, dialect, accent, speech rate, fundamental frequency, voice quality, formant frequencies, hesitation, speech pathologies, and etc. (ENFSI, 2022). In the current paper, we discuss the speech rate and fundamental frequency for the Georgian-speaking population.

## **2. Research data description**

To collect data on the Georgian native speaker population, we selected three age categories - [18-30], [31-40], [41-50] - for the study of Georgian native speakers. In each age category 10 women's and 10 men's oral speeches was recorded (60 speakers in total). There selected different sections from a speaker recording and measured their fundamental frequency and articulation rate based on the number of syllables uttered in a ten-second interval excluding pauses, hesitation elements and disfluencies.

We selected a group of 20 speakers (10 men & 10 women) from the age category of 18-30 to research F0 and articulation tempo of Georgian as a second language. The speakers were selected

from the groups with high linguistic competence of the Preparatory Educational Program at Tbilisi State University.

### 3. Speech rate

As mentioned above, in the process of comparing speakers, one of the most frequently used comparative characters is the speech rate. Speech rate is a quantitative characteristic that can be measured by counting the number of speech units uttered within a specific time interval. Speech unit may be a word or a syllable. As the length of words can vary significantly both across different languages and within the same language, it is recommended to use the number of syllables as the unit of measurement. There are two methods for measuring speech rate: speaking rate and articulation rate. To measure speech rate, the number of speech units uttered within a continuous period of time is calculated, including pauses, disfluencies, and hesitation elements. Articulation rate is measured by considering only the time required to utter speech units as a unit of time, excluding pauses, hesitation elements, and disfluencies (ENFSI, 2022).

Speech rate can also be described qualitatively using auditory analysis as "slow," "normal," or "fast," but this type of description can be very arbitrary and may represent an illusion of the listener.

The aim of this study is to investigate the articulation tempo of both native and non-native Georgian speakers. As mentioned above, for the group of nonnative speakers we chose high language competence Tbilisi State University students whose native language is Azerbaijani.

#### 3.1 Inter and intra-speaker variations in articulation rate

The comparison of the rate of articulation between gender groups gave us the following picture:

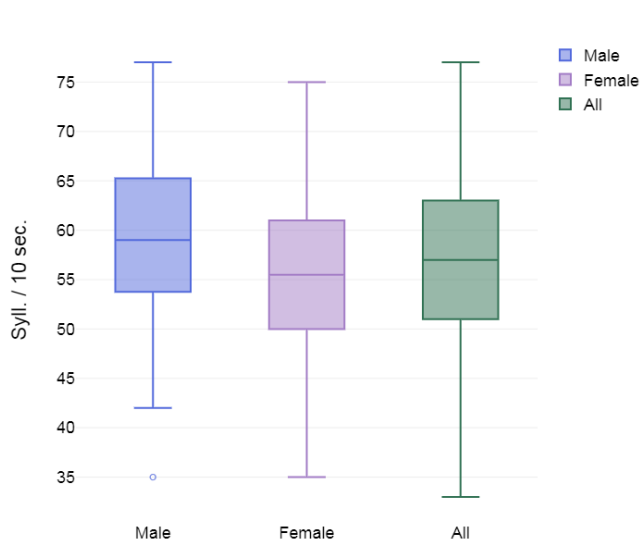


Figure 1. Number of uttered syllables in 10 second interval by native speakers of Georgian language.

Groups:	men	women	all
Sample size (n):	90	90	180
Minimum:	35	35	33
Q1:	54	50	51
Median:	59	55.5	57.5
Q3:	65	61	63
Maximum:	77	75	77
Mean ( $\bar{x}$ ):	58.9	55.4	57.1
Outliers:	35, 33		

Table 1. Statistical data of uttered syllables in 10 seconds by native speakers of Georgian language.

Data analysis shows that the minimum number of syllables spoken in a ten-second interval is 33, while the maximum is 77. The mean and median values of syllables are both approximately 57 (i.e., 5.7 syllables/second). The first quartile of the total data  $Q1 = 51$ , and the  $Q3 = 63$ . As approximately 50% of the data falls between the first and third quartiles, it is possible to determine the 'normal' or most common tempo in Georgian as 5.1 to 6.3 syllables/second. Articulation rates that fall outside this range can be considered 'slow' or 'fast'.

An analysis of the articulation rate across age and gender groups revealed the following findings:

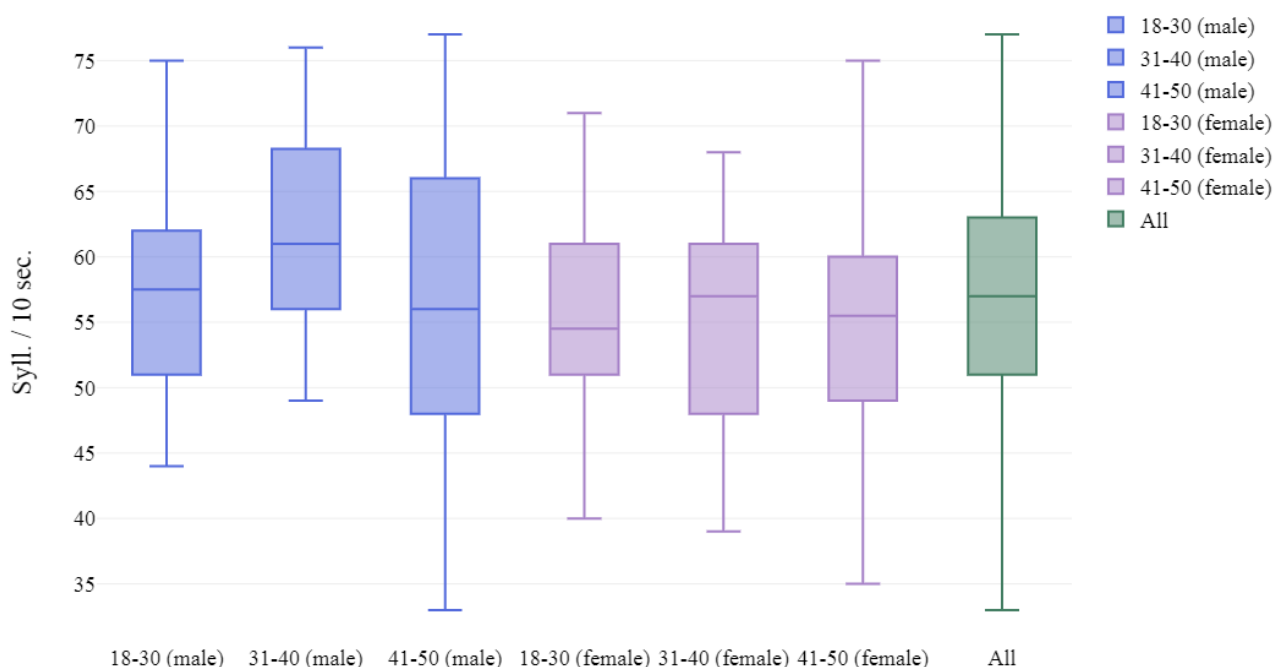


Figure 2. Box-plot of uttered syllables in 10 second interval by native speakers of Georgian language in different age groups.

<b>Groups:</b>	<b>18-30 (male)</b>	<b>31-40 (male)</b>	<b>41-50 (male)</b>	<b>18-30 (female)</b>	<b>31-40 (female)</b>	<b>41-50 (female)</b>	<b>All</b>
Sample size (n):	30	33	30	30	30	30	180
Minimum:	44	49	33	40	39	35	33
Q1:	51	56	48	51	48	49	51
Median:	57.5	61	56	54.5	57	55.5	57

Q3:	62	68	66	61	61	60	63
Maximum:	75	76	77	71	68	75	77
Mean ( $\bar{x}$ ):	57.27	61.94	56.57	56.27	54.57	55.33	57

Table 2. Statistical data of uttered syllables in 10 seconds by native speakers of Georgian language in different age groups.

The data analysis shows that in the 41-50 age group, the degree of data dispersion is relatively high for both sexes. The articulation rate of women is slightly lower than that of men, and for both genders, middle-aged speakers have a greater articulation tempo compared to other groups. No other correlation trend was found between speech rate and age groups. There's no significant difference in articulation rate between age and gender groups, so it seems that this parameter does not have discriminating power in terms of determining a speaker's age and gender group.

Since there were no significant differences according to gender and age groups with this parameter, we only compared the general data of the native and non-native speaker groups. The

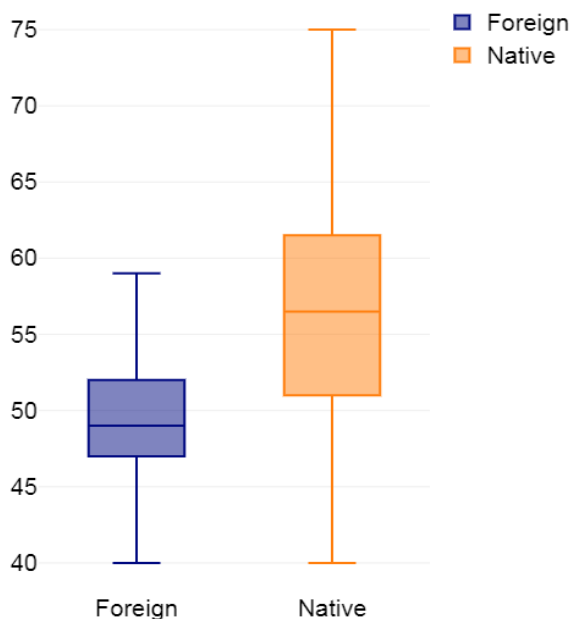


Figure 3. Box-plot of uttered syllables in 10 second interval by native and non-native speakers from 18-30 age category.

measurement of the tempo of articulation and comparing it to the articulation tempo of Georgian language native speakers from the relevant age category shows the following picture:

Groups:	Foreign	Native
Sample size (n):	60	60
Minimum:	40	40
Q1:	47	51
Median:	49	56.5
Q3:	52	61.5
Maximum:	58	75
Mean ( $\bar{x}$ ):	49	56.76

Table 3. Statistical data of uttered syllables in 10 second interval by native and non-native speakers from 18-30 age category.

As can be seen from the diagram, if the average tempo for native speakers is 5.6 syll./s, the articulation tempo for the group of speakers of Georgian as a second language is significantly lower

and represents 4.9 syll./s. It should also be taken into account that the data of this group is less dispersed, and the range is 18 syllables (compared to the range of 35 syllables in the group of native speakers). Basically, the tempo of articulation ranges from 4.7-5.2 syll./s.

As stated above, the expert must determine whether the questioned and known samples came from the same or different sources after comparing the linguistic profiles of the two speakers.

As stated above, the expert must determine whether the questioned and known samples came from the same or different sources after comparing the linguistic profiles of the two speakers. The more parameters we match in the process of comparison, the more categorical the conclusion can be. How prevalent a particular trait is in the relevant population is also crucial in the decision-making process. The rarer the matching characteristic, the stronger the hypothesis that we are dealing with the same speaker. Consequently, while evaluating speech profiles, matching with a rare feature has higher caliber than matching with a common feature. Thus, we prefer to break these classes into subclasses and rate them on a 5-point scale rather than a 3-point scale because different forms of data were included in the same class ("slow," "rapid") in terms of frequency of repetition (see histogram). Statistical data can also be used for articulation rate verbal characterization.

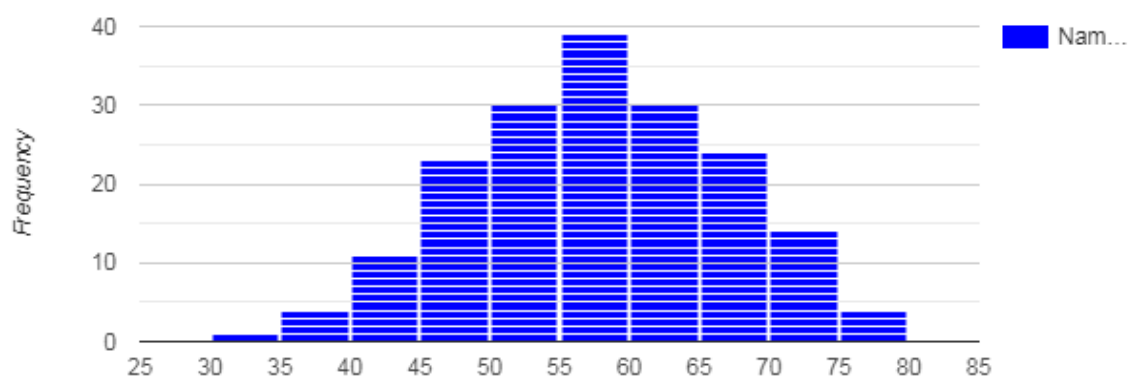


Figure 4. Histogram of uttered syllables in 10 second by native speakers of Georgian language.

- 0 to 40 syllables ( $\approx 3\%$  of total data) – "very slow";
- 40 to 51 syllables ( $\approx 19\%$  of total data) – "slow";
- 51 to 63 syllables ( $\approx 54\%$  of total data) – "normal";
- 63 to 75 syllables ( $\approx 21\%$  of total data) – "fast";
- Above 75 syllables ( $\approx 3\%$  of total data) – "very fast"

We evaluated the number of syllables spoken by the speakers in a 10-second section according to the 5 point-scale and got the following picture:

Mean syllables per 10 sec.	Verbal description	Number of native speakers	Number of non-native speakers
0 to 40 syllables	very slow	1	0
40 to 51 syllables	slow	10	14
51 to 63 syllables	normal	39	6
63 to 75 syllables	fast	9	0
Above 75 syllables	very fast	1	0

Table 4. Evaluation of the average articulation rate of speakers on the 5-point scale.

In the evaluation process, it should be taken into account that the articulation rate is more susceptible to attempts at disguise in terms of deceleration than acceleration. To be more specific, the speaker can deliberately slow down their articulation rate, while they are limited by their physical-morphological condition to increase the articulation rate beyond a certain level, even if they wish to do so.

We compare measurements taken from various parts of the same speaker's recording in order to evaluate intra-speaker changes in articulation tempo. The minimal and greatest measurement differences between identical speakers are 1 and 31 syllables, respectively. The average range is 11 syllables. The standard deviation between each speaker's measurements ranges from 0 to 17 syllables. The average standard deviation of the measures is 6. As a result, in speaker comparison, the upper limit of the standard deviation of intra-speaker variation in the Georgian-speaking population can be considered as 17 syllables.

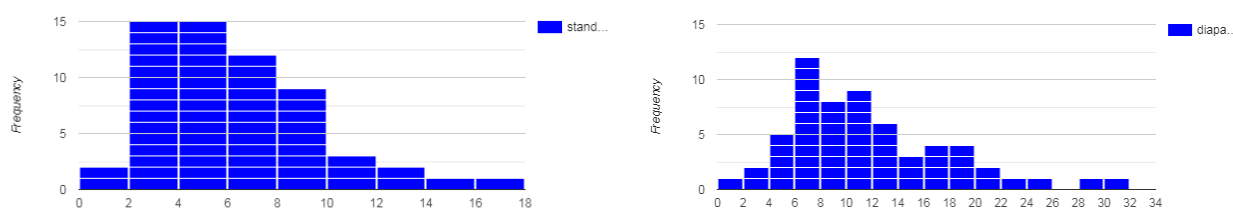


Figure 5. Histograms of standard deviation and diapason of intra-speaker variations of articulation rate.

#### 4. Fundamental Frequency (F0)

The fundamental frequency (F0), also known as the first harmonic is the number of repetitions of a (quasi-)periodic wave (of any type actually) during the time unit of one second (Hollien, 1990,



20). It plays a crucial role in the process of sound production and perception. Many individuals exhibit habituated pitch patterns that can aid the listener in identifying them by voice (Hollien, 1990, 198). It is the lowest frequency in the sound wave and is perceived as the "tone quality" of the sound. The fundamental frequency is usually measured in Hertz (Hz). It can vary significantly between speakers or even between different utterances of the same speaker. The fundamental frequency can be affected by factors such as gender, age, physical or emotional state, and more. Variations in fundamental frequency may also indicate cultural, regional, or individual identity.

As mentioned above, fundamental frequency is measured in Hertz and as a rule is represented as the arithmetic mean, although it can also be reported as the median, maximum and minimum value or standard deviation (Drygajlo, et al., 2015).

Fundamental frequency changes between various genders as a result of physiological differences between men and women, particularly related to the length of the vocal cords. Because their vocal chords are considerably shorter than men's, women's fundamental frequencies are higher than men's. "typical ranges of interspeaker variation for F0 mean value are 80-200 Hz for men and 150-400 Hz for women who normally have higher F0 due to shorter vocal fold length" (ENFSI, 2022, p. 8).

The correlation between fundamental frequency and age may be individual and population dependent and requires further research. However, in general, this parameter tends to decrease with age.

#### 4.1 Inter and intra-speaker variations in F0

We measured the average fundamental frequency of the voices of the speakers we recorded and presented the data as a box-plot.

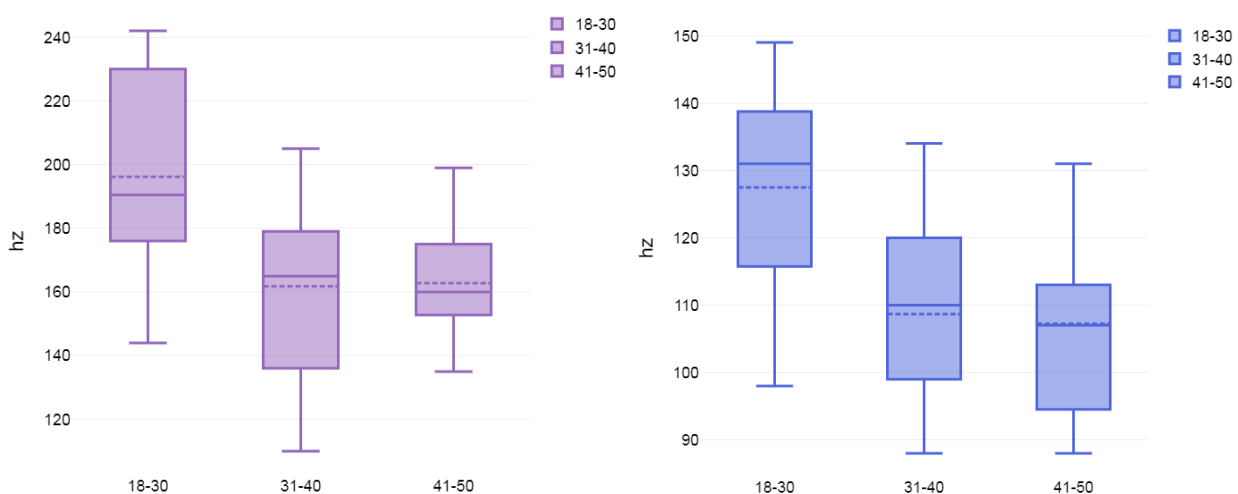


Figure 6. Box-plot of mean Fundamental frequency in different age and gender categories.

<b>Groups:</b>	<b>18-30 (Female)</b>	<b>31-40 (Female)</b>	<b>41-50 (Female)</b>	<b>18-30 (male)</b>	<b>31-40 (male)</b>	<b>41-50 (male)</b>
Sample size (n):	30	30	30	30	30	30
Minimum:	144	110	135	98	88	88
Q1:	176	136	153	116	99	94.5
Median:	190.5	165	160	131	110	107
Q3:	230	179	175	138	120	113
Maximum:	242	205	199	149	134	131
Mean ( $\bar{x}$ ):	196.17	161.77	162.76	128.37	108.67	107.25

**Table 5.** Statistics of mean Fundamental frequency in different age and gender categories.

*The diagrams show that the fundamental frequency decreases with age for both men and women. However, the gap between the age categories of 18 to 30 and 31 to 40 is substantially wider than that between 31 to 40 and 41 to 50. From a physiological perspective, the reason for this may be the still-developing body. Additionally, the fundamental frequency is less distributed, and the primary data is centered between 153-175 Hertz and the range of 135-199 Hertz in female speakers who are between the ages of 41 and 50. The data analysis leads us to the following conclusion: the following assumptions can be made when creating linguistic profiles of native speakers of the Georgian language:*

- If  $F_0 > 205$  Hz, this will be positive data in favor of the hypothesis that the voice sample belongs to a woman under 30 years old;
- If  $F_0 > 199$  Hz, this will be positive data in favor of the hypothesis that the voice sample belongs to a woman under 40 years old;
- In the case of male speakers, if  $F_0 > 134$  Hz, this is positive data in favor of the hypothesis that the voice sample in question belongs to a male person under the age of 30.

A comparison of intra-speaker variations in the fundamental frequencies of male speakers showed that the maximum difference between the speaker's measurements was 24 Hertz, and the average range value was 8 Hertz. Regarding the standard deviation, it has a range of 1 to 15, with a mean value of 2.5. As a result, when comparing the questioned and known sound samples, a fundamental frequency difference of up to 29 Hz and a standard deviation of up to 15 Hz should be

regarded as acceptable for intra-speaker differences.

The frequency range for female speakers is 1 to 42 Hz, with an average of 20 Hz. In terms of the standard deviation, it has a mean value of 10 and a range of 1 to 23. Hence, the difference in fundamental frequency up to 42 Hz and the standard deviation up to 23 Hz should be taken into account as acceptable when comparing the questioned and known sound samples.

A comparison of non-native speakers' F0 from the 18-30 age group to Georgian native speakers from relevant age group showed the following picture:

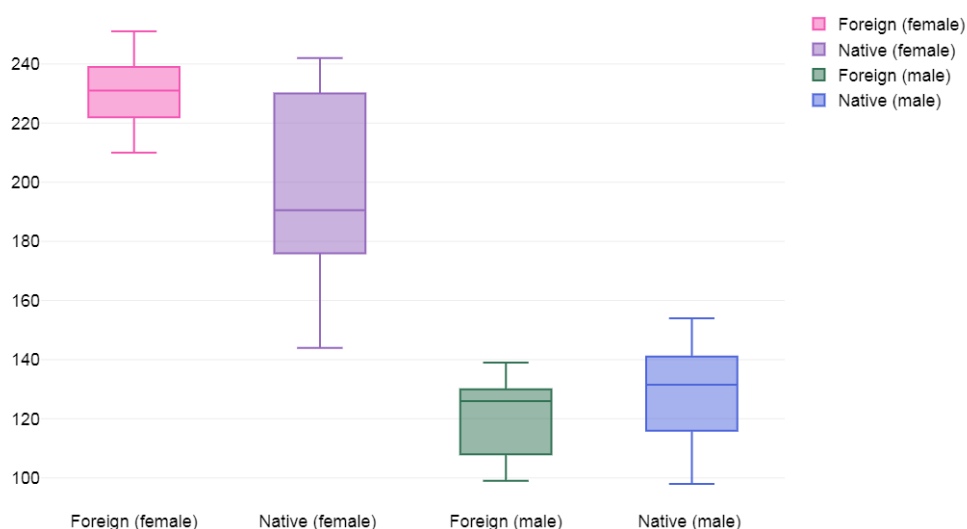


Figure 7. Box-plot of mean F0 of Georgian language native and non-native speakers from 18-30 age category.

Groups:	Foreign (female)	Native (female)	Foreign (male)	Native (male)
Sample size (n):	30	30	30	30
Minimum:	210	144	99	98
Q1:	222	176	108	116
Median:	231	190.5	126	131.5
Q3:	239	230	130	141
Maximum:	251	242	139	154
Mean ( $\bar{x}$ ):	231.38	196.17	120.53	128.37

Table 6. Statistics of mean F0 of Georgian language native and non-native speakers from 18-30 age category.

The non-native speaker women's group has an average fundamental frequency of 231 Hertz, ranging from 210 to 251 Hertz. This is 41.5 Hertz higher than the native speakers' fundamental frequency. On the other hand, the situation for men is somewhat different; in particular, the range is 99-139 Hertz (as opposed to 98-154 Hz for native speakers), and the average is 139 Hertz, which is 15 Hertz below the F0 of Georgian native speaker men. As a result, the fundamental frequency appears to be greater for women and lower for men in the Azerbaijani native speaking group when compared to the Georgian native speaking group.

## **5. Conclusion**

In conclusion, this study provides valuable statistical information on the speech rate and fundamental frequency of the Georgian language for forensic phonetics purposes. The findings suggest that speech rate and fundamental frequency can be useful parameters in differentiating between speakers, with the normal articulation tempo of the Georgian language identified to be between 5.1 to 6.3 syllables/second. These results can aid forensic phonetics experts in the identification of speakers for legal purposes. The study also highlights the importance of selecting appropriate parameters that are independent, have high inter-speaker and low intra-speaker variation, and are applicable to the specific case being investigated.

### **Acknowledgments**

We would like to express our gratitude to the Shota Rustaveli National Science Foundation of Georgia for the financial support provided under the PHDF-18-405 grant project. The author thanks Tinatin Tenterashvili for her help in data collecting process.

## References

- Drygajlo, A., Jessen, M., Gfroerer, S., Wagner, I., Vermeulen, J., & Niemi, T. (2015). *Methodological Guidelines for Best Practise in Forensic Semiautomatic and Automatic Speaker Recognition*. Frankfurt: Verlag Fur Polizeiwissenschaft.
- ENFSI, E. N. (2022). *Best Practice Manual for the Methodology of Forensic Speaker Comparison (FSA-BPM-003)*.
- Eriksson, A. (2005). Tutorial on forensic speech science. *Proc. European Conf. Speech Communication and Technology*, (pp. 1-14).
- Hollien, H. (1990). *The Acoustic of Crime* . New York: Plenum Press.
- Hollien, H. (2012). About Forensic Phonetics. *Linguistica*, 27-53.
- Nolan, F. (2009). *The phonetic bases of speaker recognition*. Cambridge: Cambridge University Press.
- Olsson, J., & Luchjenbroers, J. (2014). *Forensic Linguistics*. London: Bloomsbury.
- Rose, P. (2002). *Forensic Speaker Identification* . London: Taylor & Francis.