

A Cross-Gender Statistical Analysis of Tuvinian Intonation Features in Comparison With Uzbek and Azerbaijani

D. Beziakina, E. Bulgakova

Abstract—The paper deals with cross-gender and cross-linguistic comparison of pitch characteristics for Tuvinian with two other Turkic languages - Uzbek and Azerbaijani, based on the results of statistical analysis of pitch parameter values and intonation patterns used by male and female speakers.

The main goal of our work is to obtain the ranges of pitch parameter values typical for Tuvinian speakers for the purpose of automatic language identification. We also propose a cross-gender analysis of declarative intonation in the poorly studied Tuvinian language.

The ranges of pitch parameter values were obtained by means of specially developed software that deals with the distribution of pitch values and allows us to obtain statistical language-specific pitch intervals.

Keywords—Speech analysis, Statistical analysis, Speaker recognition, Identification of person.

I. INTRODUCTION

AMONG several existing methods used in speaker identification systems (linguistic, auditory, spectral, psychological analyses and others), the comparison of pitch contours of utterances appears to be one of the most important tasks in modern speech technologies. Defining a speaker's identity requires finding a set of ranges of different acoustic parameters that are typical for a certain language, gender and age [1]-[3]. In particular, the results of statistical analysis of intonation contours and its parts help experts obtain a set of language-specific statistical ranges of pitch parameters for both male and female voices.

A number of studies were conducted on acoustic differences between female and male voices [4], [5], but not many of them focused on obtaining the typical ranges of acoustic parameters within a certain language. This work presents the results of the statistical analysis of Tuvinian intonation contours in comparison with Uzbek and Azerbaijani gender-specific intonation features. The earlier work by M. Khitrov, L. Beldiman and A. Vasiliev [6] demonstrates intonation characteristics typical for Lithuanian compared to averaged typical parameter values for Uzbek and Azerbaijani, while the present study contains the analysis of Tuvinian intonation in

comparison with Uzbek and Azerbaijani taken separately.

II. RESEARCH BASIS

Our research has been carried out on the basis of 93 recordings: 28 recordings of 10 female native Tuvinian speakers and 65 recordings of 14 male speakers. Since the set of the recordings included speech in Russian and Tuvinian, the primary statistical analysis was conducted for each language separately. The resulting data for both languages were merged for further analysis due to identical language-independent intonation structure of the utterances.

The first stage of data analysis included obtaining and correcting the pitch files for neutral declarative utterances and building the tables containing the values of pitch parameters for the structural intonation units of the utterances (prosodic phrase, head, pre-head, nuclear tone, nucleus + tail) [7], [8]. After that we performed statistical analysis of basic intonation structures used by Tuvinian speakers. Finally, pitch parameters, such as minimum, maximum and average frequency values, F0 interval measured in Hz and semitones and pitch change speed, were examined in order to obtain the ranges of speech parameter values typical for Tuvinian.

III. A CROSS-GENDER ANALYSIS OF DECLARATIVE INTONATION

The intonation system of the Tuvinian language is very poorly studied. As almost all Tuvinian speakers are bilingual (speak both Tuvinian and Russian), we can suppose that intonation patterns used in two languages are similar. Indeed, auditory analysis of Tuvinian speech does not seem to show any significant differences between the two intonation systems. At the same time, a more accurate acoustic analysis shows that Tuvinian speakers use more level tone than Russians do, the pre-nuclear part of a Tuvinian utterance usually has a more flat tone than that of a Russian intonation phrase.

2480 intonation units (prosodic phrases and nuclear tones) realized by female speakers and 5433 units performed by male speakers were examined from the point of view of pitch change direction. The results of the cross-gender analysis of Tuvinian declarative intonation are presented in Fig. 1 (the percentage of prosodic phrases with rising and falling intonation: for male speakers - on the left, for female speakers - on the right) and Fig. 2 (the percentage of rising, falling and rising-falling nuclear tones).

D. Beziakina is with the Speech Technology Center, Krasnitskogo-4, St.Petersburg, 196084, Russia (e-mail: bezyakina@speechpro.com).

E. Bulgakova is with the Department of Speech Information Systems, National Research University of Information Technologies, Mechanics and Optics (ITMO University), Saint-Petersburg, Russia (e-mail: bulgakova@speechpro.com).

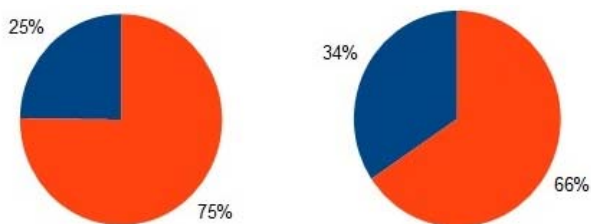


Fig. 1 The percentage of prosodic phrases with rising intonation (25% for male Tuvinian speakers, 34% for female speakers) and falling intonation (75% for men, 66% for women)

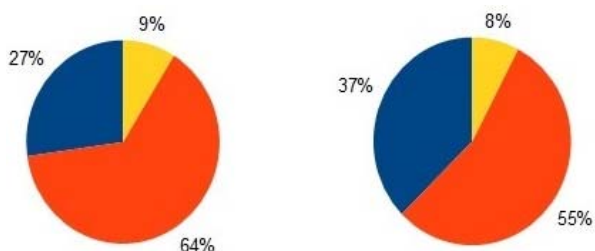


Fig. 2 The percentage of nuclear tones with rising (27% for men, 37% for women), rising-falling (9% for men, 8% for women) and falling intonation (64% for men, 55% for women)

Regarding pitch change in both prosodic phrases and nuclear tones a conclusion can be made that Tuvinian female speakers use more rising tone in declarative utterances, i.e. demonstrate more non-final intonation than male speakers (34% in prosodic phrases and 37% in nucleus in women's speech vs. 25% and 27% respectively for men's speech). However, falling tone (including level tone functioning as falling tone) remains the basic way of intonating declarative utterances.

IV. A CROSS-LINGUISTIC STATISTICAL ANALYSIS OF PITCH PARAMETER VALUES

The next stage of our work was obtaining pitch value ranges for the following parameters: maximum, minimum and average values of F0 in Hz, F0 interval in Hz and semitones and pitch change speed value measured in semitones/sec. Each parameter was presented on a diagram that demonstrated its distribution. Specially developed software was used for defining the main value distributions of pitch characteristics in order to get typical pitch value ranges for Tuvinian as well as for two other Turkic languages: Uzbek and Azerbaijani (typical value ranges were measured with 95% confidence interval).

The examples for such parameter distribution are demonstrated in Figs. 3 and 4 (for men and women respectively), showing the distribution of pitch change speed with typical interval borders specific for Tuvinian.

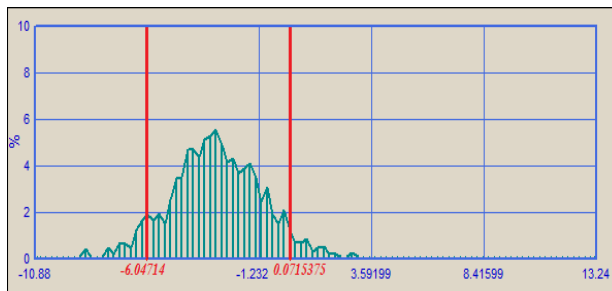


Fig. 3 Distribution of pitch change speed and its typical interval for Tuvinian male speakers

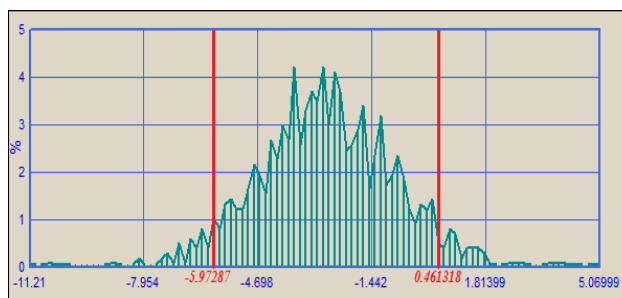


Fig. 4 Distribution of pitch change speed and its typical interval for Tuvinian female speakers

Both male and female distributions of pitch speed values demonstrate closeness to normal distribution, as well as distributions of maximum F0 and F0 interval values. Other parameters values are also distributed approximately normally.

Value ranges of pitch parameters obtained for Tuvinian were compared with value ranges for Uzbek and Azerbaijani (male and female speaker values were compared separately). Value ranges for 49 female and 35 male Uzbek speakers and 13 female and 45 male Azerbaijani speakers were measured using the same method and the same software as for Tuvinian speech analysis and processing. The comparison of the three languages is demonstrated in Table I (male speakers) and Table II (female speakers). A cross-gender analysis can easily be conducted on the basis of the data from these tables: it is clear that all the pitch parameter values are significantly higher for female voices in comparison with male voices, while the pitch change speed values for male and female voices are close.

TABLE I
RANGES OF TYPICAL PITCH PARAMETER VALUES FOR UZBEK, AZERBAIJANI AND TUVINIAN MALE SPEAKERS

Parameters	Value ranges (Uzbek male speakers)	Value ranges (Azerbaijani male speakers)	Value ranges (Tuvinian male speakers)
Maximum value, Hz	162..192	125..178	134..174
Minimum value, Hz	70..88	64..84	87..105
Interval, Hz	66..112	56..96	38..80
Interval, st	10.5..14.5	10..14	6..10
Average value, Hz	106..137	93..117	108..129
Pitch change speed, st/sec	-8..-3	-9..-5	-6..0.07

TABLE II
RANGES OF TYPICAL PITCH PARAMETER VALUES FOR UZBEK, AZERBAIJANI
AND TUVINIAN FEMALE SPEAKERS

Parameters	Value ranges (Uzbek female speakers)	Value ranges (Azerbaijani female speakers)	Value ranges (Tuvinian female speakers)
Maximum value, Hz	250..330	274..348	216..336
Minimum value, Hz	135..155	161..184	147..177
Interval, Hz	120..175	128..190	67..137
Interval, st	9.5..13.5	10..13	6..11.5
Average value, Hz	185..225	219..237	183..221
Pitch change speed, st/sec	-6.5..-2.5	-6.5..-2.5	-5.9..0.4

The comparison of the three languages shows that pitch values for Uzbek and Azerbaijani are closer to each other than to those for Tuvinian. A further cross-linguistic analysis of pitch value ranges for Tuvinian leads to the conclusion that the parameters with the most different typical value ranges are F0 interval (measured in both Hz and semitones) and pitch change speed (the parameter that shows how steep the pitch changes of the speech signal are). Other parameters show more similarity between the three languages. Such a conclusion is valid for both male and female Tuvinian speech (see Tables I and II).

Language-specific data comparison shows that the F0 interval of Tuvinian speech is noticeably narrower than that of Uzbek and Azerbaijani. The values of pitch change speed ranges are lower for Tuvinian than for the two other languages. Also, the minimum F0 value range for Tuvinian male speech tends towards higher frequencies than that for Uzbek and Azerbaijani.

A brief analysis of the ranges of pitch parameter values for Tuvinian in comparison with average language-independent pitch value ranges typical for male and female voices also shows a difference in F0 interval: 67..137 Hz (Tuvinian female speakers) vs. 89..208 Hz (average female voice); 38..80 Hz (Tuvinian male speakers) vs. 57..137 Hz (average male voice), and pitch change speed: -5.9..0.4 St/sec (Tuvinian female speakers) vs. -8.4..0.5 St/sec (average female voice); -6..0.07 St/sec (Tuvinian male speakers) vs. -10..-0.2 (average male voice).

To sum up, taking into account the results above, a general conclusion about the sound of Tuvinian speech can be made: its pitch frequencies lie in a lower pitch range than in Uzbek and Azerbaijani and the pitch change in Tuvinian tends to be more sloping and smooth.

V. CONCLUSION

The results of this cross-gender study made on the basis of Tuvinian speech in comparison with Uzbek and Azerbaijani speech lead us to the following conclusions.

First, the data shows pitch value differences between male and female voices on the basis of Tuvinian speech: women's speech tends to have higher pitch and a higher percentage of rising intonation patterns than men's speech. However, pitch change speed in Tuvinian does not vary much depending on the speaker's gender, neither does the common way of

intonating declarative utterances, which is the falling tone (and sometimes the level tone) for both men and women.

Statistical analysis of pitch characteristics of Tuvinian speech not only provided the pitch value ranges required for the purpose of automatic speaker identification, but also made it possible to compare the three Turkic languages. The following conclusions were made: first, Uzbek and Azerbaijani pitch value ranges appeared to be closer to each other than to the pitch value ranges obtained for Tuvinian. Such parameters as F0 interval (measured in Hz and semitones) and pitch change speed show the most significant difference between the languages, which is the case for both genders. The importance of these three parameters is also confirmed by the results of a comparison of Tuvinian and average language-independent data on pitch characteristics values.

Such studies on the basis of different languages have a great practical significance, firstly, for the purpose of the speaker's native language identification, and, secondly, as a way of describing intonation features of a certain language in its cross-gender and cross-linguistic aspects.

REFERENCES

- [1] S. L. Koval, P. V. Labutin, T. S. Pekhovskiy, E. A. Proshchina, N. S. Smirnova, A. O. Talanov, "Composite Forensic Speaker Identification Method," in *Dialog-2007*, <http://www.dialog-21.ru/digests/dialog2007/materials/html/39.htm>.
- [2] N. S. Smirnova, "Speaker identification based on the comparison of utterance pitch contour parameters," in *Dialog-2007*, <http://www.dialog21.ru/digests/dialog2007/materials/pdf/77.pdf> H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, ch. 4.
- [3] Yu. Matveev, "The Problem of Voice Template Aging in Speaker Recognition Systems", *LNCS/LNAI*, vol. 8113, pp. 345-353, 2013.
- [4] A. P. Simpson, "Phonetic differences between male and female speech", *Language and Linguistic Compass*, vol. 3, March 2009.
- [5] E. Pepiot, "Voice, speech and gender: male-female acoustic differences and cross-language variation in English and French speakers," http://halshs.archives-ouvertes.fr/docs/00/76/48/11/PDF/Article_Pepiot_RJC_2012_-_Pre-Print.pdf.
- [6] M. Khitrov, L. Beldiman, A. Vasiliev, "Statistical Language Aspects of Intonation and Gender Features Based on the Lithuanian Language," *Speech and Computer: Lecture Notes in Computer Science*, vol.8113, pp. 327-332, 2013.
- [7] J. D. O'Connor, G. F. Arnold, "Intonation of Colloquial English," Longman, London, 1973.
- [8] E. A. Bryzgunova, "Intonation", in "Russian Grammar", vol. 1, pp. 96 - 123, 1980 "Russian Grammar", vol. 1, pp. 96 - 123.