

ЧАСТОТА ОСНОВНОГО ТОНА И ДЛИТЕЛЬНОСТЬ КАК ОСНОВНЫЕ ФАКТОРЫ ВАХАНСКОГО УДАРЕНИЯ

Иванов В.Б., Силантьева Л.Г.

*Институт стран Азии и Африки Московского государственного университета
имени М.В. Ломоносова, Москва, e-mail: iranorus@mail.ru, lyubovsilantieva@yandex.ru*

Для акустико-глоттографического анализа словесного ударения в ваханском языке было организовано два эксперимента. Первый эксперимент предполагал запись речи четырёх носителей таджикского варианта ваханского языка (двух мужчин и двух женщин) только по акустическому каналу. Во втором эксперименте речь трёх носителей ваханского языка из Таджикистана (двух мужчин и одной женщины) записывалась в двухканальном режиме с помощью программно-аппаратного комплекса Real-Time EGG: в первый канал фиксировался сигнал микрофона, а во второй – сигнал глоттографа. Использование глоттограммы как дополнительного канала информации значительно повышает точность сегментации речевого потока в сравнении с его одноканальной записью. Записи обоих каналов были обработаны с помощью программы речевого анализа Praat, затем выделенные релевантные параметры были обчислены статистическим пакетом SPSS по многомерной модели. В результате исследования было обнаружено, что наиболее значимыми параметрами при маркировании ударного слога в ваханском языке являются частота основного тона и длительность. Интенсивность тоже повышается в ударном слоге, но менее регулярно. Статистически значимая связь ларингальных параметров с просодией не была зафиксирована. На основании этого можно сделать вывод, что ударение в ваханском языке является квантитативно-тоническим.

Ключевые слова: ваханский, акустический анализ, глоттографический анализ, просодия, ударение, ударный слог, гласный, частота основного тона, длительность, интенсивность, ларинкс

PITCH AND DURATION AS MAIN FACTORS OF WAKHI STRESS

Ivanov V.B., Silanteva L.G.

*Institute of Asian and African Studies, Moscow State University n.a. M.V. Lomonosov,
Moscow, e-mail: iranorus@mail.ru, lyubovsilantieva@yandex.ru*

Wakhi is one of the Pamir languages in the Eastern Iranian branch of Indo-European languages. Native speakers of Wakhi live in Tajikistan, Afghanistan, Pakistan and China, as a koiné language Wakhi occupies a dominant position in Wakhan. Phonetically, positional quantitative changes in vowels are typical for Wakhi, and vowels are characterized by inherent duration. Before this research the Wakhi word stress had not been studied experimentally. The purpose of the work is to provide a complex instrumental analysis of the Wakhi word stress in order to determine its stress type. For the acoustic-glottographic study of the word stress two experiments were held. In the first one, a recording of the speech of four native speakers of Tajik variant of Wakhi was conducted (only the acoustic channel). In the second experiment, the speech of three native speakers was recorded in two-channel mode with the use of Real-Time EGG – hardware and software complex. In this case the microphone signal was recorded in the first channel, the glottographic signal – in another one. After that the channels were treated by Praat – the program for speech analysis. All the relevant parameters were processed by the SPSS statistical package with a multidimensional model. Pitch and duration appeared the most significant parameters for marking a stressed syllable in Wakhi. Intensity rises in a stressed syllable, but less regularly. We did not discover any statistically significant connection between laryngeal movements and prosody in Wakhi. Thus, we can define the Wakhi stress as quantitative-tonic one.

Keywords: Wakhi, acoustic analysis, glottographic analysis, stress, pitch frequency, duration, intensity, larynx

Wakhi is one of the Pamir languages in the Eastern Iranian branch of Indo-European languages. Native speakers of Wakhi live in Tajikistan, Afghanistan, Pakistan and China. The total number of Wakhi-speakers is about 70 thousand people [1]. The language does not have unified writing system, literary norm and educational status, but as a koiné language it occupies a dominant position in Wakhan. Although the Wakhan language is officially non-written, several types of alphabets have been created: Wakhis from Tajikistan prefer Cyrillic alphabet, in Afghanistan Arabic system is harnessed, in Pakistan and China Latin script is applied. Russian scientists traditionally use Latin transcription with additional Cyrillic letters. Latin script is also applied in this research.

Most of the Wakhis are multilingual: the second language in Tajikistan is Tajik and the third language is Russian. A majority of Wakhis in Afghanistan speaks Dari and Pashto. Wakhis of Northern Pakistan speak Sheena and Burushaski as well as the official languages (English and Urdu). For native speakers of Wakhis in China Uighur is the second language and Chinese is the third one.

The Tajik variant of Wakhi has two main dialects: upper and lower; sometimes the central dialect is distinguished as an intermediate one. Over the past decades, differences in dialects have become more clearly visible, and they emerge in lexical borrowings [2].

In China the Wakhi language has features of the upper dialect with subtle differences presented in the lexical aspect. According to the

available data, due to the remote location and presumably a number of borrowings from non-Iranian languages, Xinjiang variant of Wakhi appears to be a distinct dialect of the Wakhi language, but due to the lack of reliable relevant materials, it is impossible to form a clear idea of the extent of their divergence.

Phonetically, positional quantitative changes in vowels are typical for Wakhi, and vowels are characterized by inherent duration. There are 6 vowel phonemes distinguished: *i, ə, a, o, u, i*, with *i* marked out by increased variability and the ability to realize in forms *ü* and *ö*. Besides, vowel *e* appears in borrowings from Tajik. Vowels are characterized by the inherent duration, which was studied by Sokolova V.S., Pakhalina T.N., Lashkarbekov B.B., Grünberg A.L., Steblin-Kamensky I.M. Duration intervals for Wakhi vowels were computed by Ivanov V.B.

Purpose of the study

Before this research, the Wakhi word stress has not been studied experimentally. It was considered to be expiratory (dynamic) [3, p. 175], i.e. a stressed syllable marker was supposed to be intensity. The purpose of the work is to provide a complex instrumental analysis of the Wakhi word stress in order to determine a stress type in this language and the importance of such parameters as intensity, pitch and duration.

Methodology and course of study

All the vowels acted as syllable nuclei during our experiments, in which the parameters of syllable nuclei in stressed and unstressed syllables were compared. Phonetic environment of syllable nuclei can be both voiced (sonorous) and unvoiced, and consequently can vary in terms of vocal cords work. Since when searching for an acoustic correlate it is important to have the same set of parameters for each item studied, phonetic environment was not regarded, and we only dealt with syllable nuclei themselves.

In the course of the first experiment, the speech of four Wakhi speakers from Tajikistan (students, two men and two women) recorded solely via the acoustic channel was analyzed. Two- and three-compound numerals were taken as a lexical material studied, such as:

cibir – four;
dasiw – eleven;
dasbuy (*dasətbuy*) – twelve;
dastruy (*dasətruy*) – thirteen;
descəbir – fourteen;
daspanz (*dasətpanz*) – fifteen;
dasšad – sixteen;
dasib – seventeen;
dasat (*dashať*) – eighteen;
dasnaw – nineteen;
truđas – thirty;
 etc.

The segmentation of the voice signal and the acoustic analysis of the recorded realizations were carried out using Praat at the Laboratory of Experimental Phonetics of Institute of Asian and African Studies, Moscow State University. In each syllable nucleus duration (T) was measured as an independent parameter.

The parameters of intensity (I) and pitch frequency (F_0) in normal speech are correspondent via subglottal pressure, i.e. with an increase in subglottal pressure both these parameters enhance at the same time (but to a different degree). In vocal speech these parameters are separate, since vocalists are able to control them being independent of each other. We proceed from the fact that the interaction of these three main parameters depends on the prosodic system of the language and in each language occurs in a different way.

In addition to these main parameters, derivative integral parameters were examined. F_0 -area parameter represents the area of the figure, bounded by the curve and the time axis. This parameter is connected to both: pitch frequency (F_0) and duration (T), and consequently increasingly correlates with them. The structure of intensity area (I-area) parameter, which depicts the area of the figure under the intensity curve, is similar. Volume (V) parameter is a three-dimensional figure, limited by I-curve, F_0 -curve and center line (see Fig. 1).

According to our measurements, a table was compiled which was processed by the SPSS statistical package, using a linear multidimensional model. The statistical connection of stress with the above-mentioned parameters of syllable nucleus was verified. Along with the parameters' absolute values, its' relative derivatives were considered and calculated in percent (%). The maximum value of the parameter within the phonetic word was taken as 100%. In total 120 syllables were examined (58 stressed and 62 unstressed).

The results of statistical processing showed a highly significant connection between all the parameters and the stress ($p < 0,001$), which is exceptional for the Iranian languages: in other Iranian languages stress is either quantitative or tonic [4]. So far, the stress marked by all the parameters mentioned was found exclusively in one isolated language – Burushaski. It is worth drawing attention to the fact that in the north of Pakistan all the Wakhi speakers are multilingual, and one of the variants is Burushaski-Wakhi bilingualism. Although we worked with the Tajik Wakhi speakers, it is possible that owing to the migration and interaction, the prosodic system of Burushaski has influenced the prosodic system of the Tajik variant of Wakhi. This problem requires special research.

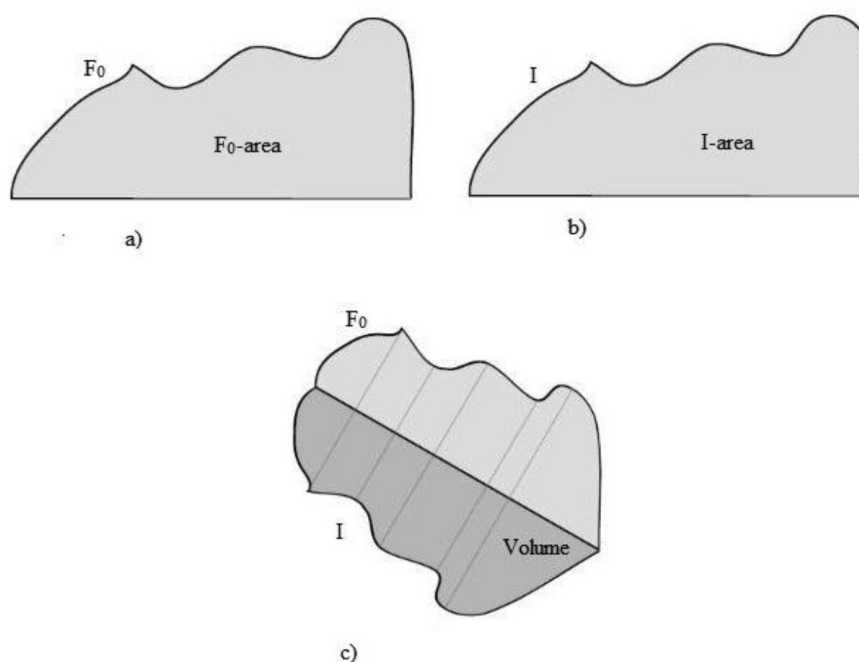


Fig. 1. Integral parameters of syllable nucleuses:
a) representation of F_0 -area; b) representation of I-area; c) representation of Volume

At this stage, we could draw a conclusion that the stress in the Wakhi language was multicomponent – dynamic, tonic and quantitative. In order to clarify the role of various parameters in Wakhi prosody, the second experiment was planned.

During the second experiment, the speech of three other native speakers of Wakhi (two men and one woman) was recorded at the Laboratory of Experimental Phonetics of IAAS MSU, using hardware and software complex Real-Time EGG in two-channel mode: the microphone signal was captured in the first channel, and the glottographic signal – in the second one. Informants were asked to read out loud the text of the Wakhi folk tale, recorded by T.N. Pakhalina during her field research in 1965 according to Makhmudov Khanjarbek (20 y.o., village Vrang, Tajikistan). After that the phonetic words, in which the syllable nucleuses were in suitable positions and no phonetic process affected the accuracy of the experiment, were selected, with all of them being two-syllable:

a-bu: a – a, bu – two;
arzuq – orzuk, naan bread;
a-ska: a – a, ska – this;
bowar – faith;
disivd – show;
kishun – hang;
pacan – prepare;
potsho – tsar, king;

pid-i – legs;
rachan – go;
ruz-i – once;
sayish – you;
sapo – our;
sawor – on horseback;
shafshish – hair;
tuwatk – was;
tisha – provisions;
vadək – road;
wizit – come;
wizimd – bring;
xoli – only, empty;
yaw-an – he/she has;
imit – (he/she) has.

In the work of Hussain Q., Mielke J. [5] the Pakistan variant of Wakhi was examined using a glottograph. Since we worked with the Tajik variant of Wakhi, this part of the study as well as the first experiment has undoubted novelty. Both idioms can be considered different dialects. Although native speakers of different variants speak the same language, Tajik informants note that they no longer understand the speech of Pakistan Wakhi speakers since they begin to communicate with each other. This fact points to the conclusion that in the Wakhi language local dialects co-exist with koiné forms. Native speakers are aware of these differences and choose the appropriate register depending on the «friend-or-foe» situation.

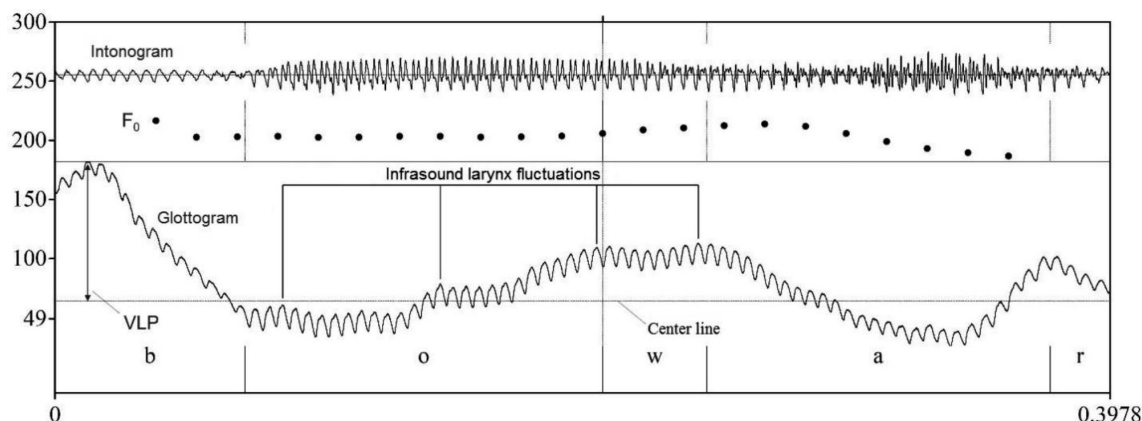


Fig. 2. Segmentation of the word *bowar* 'faith' performed by informant M.: intonogram; F_0 -graph; glottogram; infrasound larynx fluctuations; VLP

Using a glottograph, frequency of vocal cords vibrations (F_0) and coefficient of glottis openness (Q – quotient), i.e. ratio of the time of glottis openness to the entire period of vocal cords fluctuation (calculated in %), are registered. Another definition of quotient is possible but the numerator must be less than the denominator. With that type of phonation when the duration of open position is longer than the duration of closed one, the voice turns hoarse.

The common approach is when vertical larynx position (VLP) is recorded and measured, which allows to assess the vocal tract's length. The rise of larynx was clearly visible in the glottogram and was measured in relative units in relation to zero line. The laryngeal movements in a vertical plane can be characterized by an average infrasound frequency, which in our case was determined in the range of 7–37 Hz. With such a low frequency, only 1-2 fluctuations within a syllable nucleus could be detected.

Two more parameters of syllable nucleus were added to those in the first experiment: subsonic frequency (Sub) and vertical larynx position (VLP). In total 88 syllables (44 stressed and 44 unstressed) were analyzed. An additional channel of information – a glottogram – significantly increased the accuracy of speech segmentation in comparison to the single-channel recording because when using two-channel recording, the harmonic component and the noise were presented separately.

Fig. 2 represents the implementation of the word *bowar* 'faith', performed by informant M. (woman). The total duration is 0.3978 seconds and is presented on the x-axis. The y-axis measures pitch frequency (F_0) in Hz. The F_0 -graph is presented in the bold dots form. The lower vertical lines that extend from the x-axis

indicate the boundaries between segments. The horizontal central line of the glottogram passes above.

The sawtooth curve that crosses it is a glottogram. It is captured by the electrodes, placed around the informant's throat. When the vocal cords close, the skin's resistance decreases, which reflects in an upward shift in the curve. Consequently, each «tooth» of the sawtooth curve corresponds to one closing of the vocal cords.

Furthermore, the skin's resistance diminishes when the larynx moves up. This reflects in large waves of the sawtooth curve. The value of the larynx upward deflection is indicated in the figure by the VLP parameter. The Praat program measures it in relative units. The peaks of the larynx's upward displacement waves are indicated by vertical lines above the glottogram. These irregular waves correspond to infrasound larynx fluctuations, measured in Hz.

It might also be pointed out that the wavy infrasound curve in the graph is modulated by the frequency of the vocal cords' vibrations. The microphone signal is represented as an intonogram in the upper part of the Fig. 2.

Similarly to the first experiment, all the absolute parameter values were converted to relative values in percent (%). The table obtained was processed by the SPSS statistical package using a multidimensional linear model. The presence of syllable stress (0 or 1) was taken as an independent parameter while the other eight parameters were considered dependent.

Statistical analysis has shown that the parameters of duration (T) and pitch frequency (F_0) are significantly related to the syllable nucleus's stress ($p < 0,001$). The integral parameters F_0 -area and Volume (V), into which these parameters are jointly included as components,

appeared to the same extent substantial. Intensity (I) ($p = 0,001$) and its' derivative parameter I-area ($p = 0.002$) are found to be slightly less relevant. Laryngeal parameters of infrasound frequency (Sub) and vertical larynx position (VLP) appeared insignificant for the prosody ($p = 0.838$ and $p = 0.897$, respectively).

Results and conclusion

According to the results of both experiments, the most significant parameters for indicating a stressed syllable appear pitch frequency (F_0) and duration (T), which mark stressed syllables in almost all the cases. Intensity (I) and its integral modification (I-area) also rise in the stressed syllable but this occurs less regularly. Statistically significant connection between stress and laryngeal parameters (subsonic frequency (Sub) and vertical larynx position) was not discovered at this stage of the study. Thus, the stress in the Wakhi language can be defined as quantitative-tonic.

The question of how much the intensity factor is contrasted to other acoustic character-

istics of stress requires further analysis. This can be done by analysis through speech synthesis followed by listening. In addition, the problem of the connection of laryngeal movements with prosody and intonation also requires further study. In order to do this, it is necessary to consider the participation of such a parameter as Quotient in speech formation.

References

1. Gulomaliev Sh.L. Dialects of Wakhi language of Wakhan corridor of Tajikistan. Bulletin of the Tajik national university. Series of philological sciences. 2017. No.4/4. P. 35–39.
2. Lashkarbekov B.B. Izucheniye vahanskikh obshchin v usloviyakh yazykovogo sdviga: sociolingvisticheskaya situatsiya v Sintzyane (KNR) i severnykh provintsiyakh Pakistana. Vestnik rossijskogo gumanitarnogo nauchnogo fonda. Moscow, 2013. № 1 (70). P. 102–111.
3. Pakhalina T.N., Lashkarbekov B.B. Vahanskij yazyk. Yazyki mira. Iranskie yazyki. T. 3. Vostochnoiranskie yazyki. Moscow, 2000. P. 174–186.
4. Ivanov V.B. Udarenie v iranskih yazykah. Moscow, 2012. P. 49–51.
5. Hussain Q., Mielke J. Stop laryngeal contrasts of endangered languages of Northern Pakistan. Journal of the Acoustical Society of America (JASA). 2018. Vol. 143. № 3. P. 1754.