The Vowels of Standard Albanian

Sylvia Moosmüller[†] and Theodor Granser[†]
† Acoustics Research Institute, Vienna, Austria
E-mail: sylvia.moosmueller@oeaw.ac.at, drgranser@nextra.at

ABSTRACT

The vowel system of Standard Albanian comprises 7 vowels. Read speech of 9 Standard Albanian speakers from 3 traditionally defined (sub)dialect regions within the Republic of Albania has been used for the current investigation. All vowels in prosodically strong position have been subjected to acoustic phonetic analysis. All vowels show statistically significant regional differences. The two central vowels /a/ and the schwa additionally reveal a high, phonetically unmotivated range of variation from front to back articulation for /ə/ and from central to back articulation for /a/.

1. INTRODUCTION

The notion of a Standard Albanian language variety is sometimes the subject of an emotional and even political debate. The roots of this debate lie in the existence of two main dialect varieties – Gheg and Tosk – and previous efforts to establish a unified Standard Albanian language. Historically, two literary variants coexisted for Gheg and Tosk. In 1923, the Elbasan version of South Gheg was proclaimed the official Albanian language [1]. The reasons for choosing the Elbasan version were understandability for both North Gheg and Tosk speakers as well as the location of the newly founded capital Tiranë [2]. However, this variety never succeeded in becoming the basis for a standard language. As a consequence, the two main varieties (North Gheg and Tosk) remained as two separate entities [1]. At the National Conference of Orthography in 1953 the way was paved for establishing a standard language based on the Tosk variety. In 1972, this uniform standard language became obligatory. After the dissolution of the communist regime, the language debate arose anew. Some Gheg speakers felt betrayed by the toskicization [3] of Standard Albanian language. Nevertheless, in 1995, after a summer seminar on Albanian language and literature, it was acknowledged that Standard Albanian is an irreplaceable means of national unity [1]. Moreover, Standard Albanian, the official language since 1972, has now established itself in North (=Gheg) Albania (most notably, Gheg members of the cultural elite have adjusted their speech behaviour to the Standard variety [3]).

The vowel system of this standard variety based on the Tosk variety comprises seven vowels:

- 3 high vowels (front unrounded, front rounded and back rounded)
- 3 mid vowels (front unrounded, central unrounded and back rounded)
- 1 low, central to back vowel.

2. METHOD

9 male subjects (3 each from North Albania = North Gheg, Middle Albania = South Gheg and South Albania = Tosk) were asked to read a list of words, a list of sentences, a literary text and to speak spontaneously. For the current investigation, the stressed vowels of the list of read sentences have been analyzed.

The recorded material was digitzed at 22,05 kHz, 16 Bit using the acoustic work station STX. All prosodically strong vowels (1170 in total) have been labeled and frame by frame formant frequency contours have been produced. The first three formants were calculated using LPC with 26 coefficients, a pre-emphasis of 0.9, a frame width of 46ms and a 2ms frame shift. Linear time standardization was performed. For speaker normalization, a readjustment of the Lobanov normalization procedure, as proposed by Disner [4: 260, FN1], was chosen. One-way ANOVAs were calculated for each vowel.

3. RESULTS

Differences between the three regionally defined varieties proved statistically significant for all vowels except for /e/ (see Figure 1).

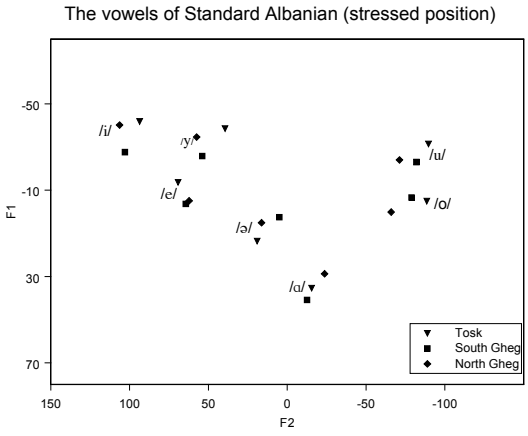

Figure 1: Normalized F1/F2 plot of the vowels according to the three investigated dialectal regions.

ISBN 1-876346-48-5

As concerns the three geographical regions, the speakers of North Gheg turned out to be the most homogeneous group, followed by the speakers of South Gheg. It has to be added, however, that the geographical distances between the speakers of North Gheg (Kryezi and Shkodër) were small as compared to the speakers from South Gheg (Durrës, Elbasan and Lin) and Tosk (Pogradec, Korçë and Butkë).

Irrespective of the investigated regions, all vowels expose a more or less high degree of variability. Although the variability as concerns height/openness (F1) is higher than the variability on the front/back dimension (F2) and shows considerable regional differences, the latter is of greater significance since points of contact with the Gheg dialect varieties might cause overlapping. Affected by such contact are the schwa and the vowel /a/. Additionally, the amount of variability on the front/back dimension is evenly distributed across the three regions and no regional differences can be found (see Figure 2).

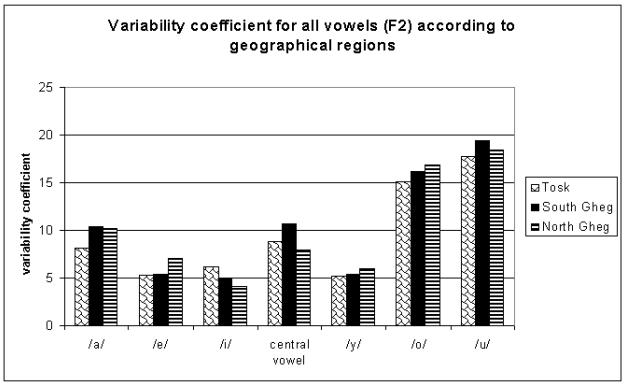

Figure 2: Degree of variability for all vowels (F2) according to the three defined regions.

Variability is highest in the back, rounded vowels /u/ and /o/. This high variability is solely due to the phonetic context. A sequence like /tona/ *tona* "our f." or /kʊʃ/ *kush* "who" starts with a high F2, changes slowly towards a very low F2 (target) and then rises towards the next segment. Therefore, the variability expressed in Figure 2 refers solely to the variability within one segment and is phonetically motivated.

However, the variability of the schwa cannot be explained so easily. The variability of the schwa affects both the height/openness and the front/back dimension. Although the variability in height/openness is greater than the variability on the front/back dimension, the latter is of greater importance because the change of phonemic schwas is predominantly described along F2 [5]. Additionally, schwa shows a tendency towards the peripheral articulation [6, 7]. In the current investigation, the articulation of the schwa reaches from a back articulation [ʌ] to a front articulation [ɛ]. It has to be emphasized that the back articulation is [-round] and [-nasal] (except for nasal environment), therefore the standard articulation for all speakers clearly differs from the Gheg dialect variety, which shows a back, rounded, nasal vowel instead [8].

Figure 3 shows the distribution on the front/back dimension (F2) for the schwa according to three defined articulation zones:

- back articulation: F2 ≤ 1400 Hz
- central articulation: F2 = 1400 Hz – 1600 Hz
- front articulation: F2 ≥ 1600 Hz

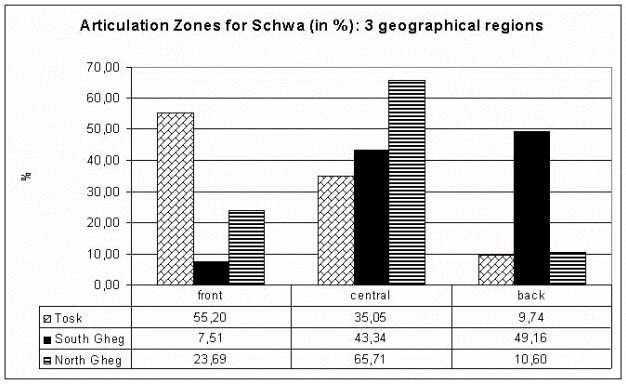

Figure 3: Distribution of phonemic schwa according to three articulation zones.

All three regional varieties are represented within each articulation zone, differing only in the amount of realizations within a certain zone. Therefore, the difference between the three regional varieties is more of a quantitative than of a qualitative nature. From the investigated speakers, the Tosk speakers predominantly realize a front vowel, the South Gheg speakers a back vowel, whereas the North Gheg speakers clearly favor a central articulation for the schwa. According to these results, all speakers show a high variability range (see Figure 4):

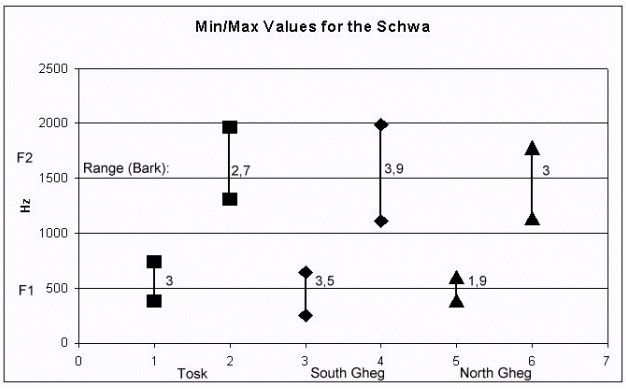

Figure 4: Variability range of the schwa for F1 and F2.

Contrary to the back rounded vowels, the variability of the schwa is not connected with the phonetic context and consequently, is not phonetically motivated. The same speaker can realize the schwa in the same CəC-sequence either as a front, a central or a back vowel (examples see Table 1):

	Example	token 1	token 2	Translation
S4 SG	thënë	[θɑn]	[θɛ̃ː]	said (m.)
S3 NG	lëng	[lɪ̃ŋ]	[lɜ̃ŋ]	liquid
S7 Tosk	u bë	[ubə]	[ubɒ]	became

Table 1: Examples of interspeaker variability of schwa-realizations. S = speaker, SG = South Gheg, NG = North Gheg.

Figure 5 shows the variability of schwa articulation according to front, central and back articulation within the same item.

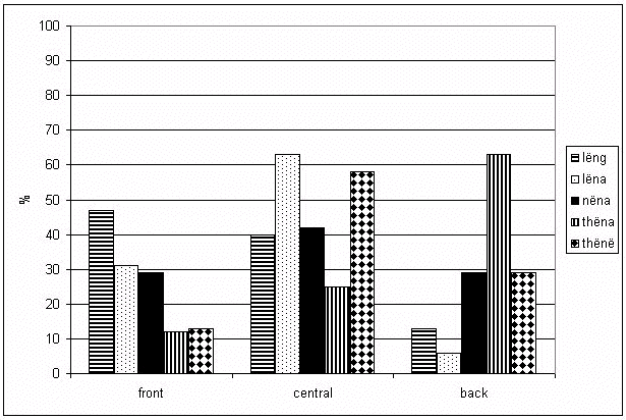

Figure 5: Within-item variability of schwa-realizations according to front, central or back articulation (in %).

All items exhibit a high within-item variability, and, moreover, although pairs of items have the same consonantal environment with respect to place of articulation (except for *lëng* – "liquid"), no preference can be observed across items. For example, *thëna* "said (f.)" shows a tendency towards back articulation, whereas *thënë* "said (m.)" shows a preference for central articulation. It follows from these results, that the variability of the schwa is phonetically unmotivated and can therefore not be explained by the phonetic context.

From these results the question arises as to whether this phoneme is a schwa at all or whether it is rather a back or a front vowel. However, since the within-item variability is relatively large and arbitrary (see Figure 5), since graduality from front to back articulation is given among all speakers and since no really pronounced back or front articulation can be observed, it can be assumed that the phoneme is still a central vowel, which has to be clearly distinguished from the Gheg dialect [9].

A similar high amount of variability can be found within the vowel /a/, again both for the high/low and for the front/back dimension. Like the schwa, the front/back dimensions are of greater importance, because in the Gheg dialect, a front and a back vowel /a/ are discerned phonologically [8]. In the Gheg dialect, frontness correlates with short vowel duration, whereas back vowels are long. Consequently, according to the measurements given in Beci [8], the vowel /a/ in *plak* [plak] "old man" has a front vowel (F2 1705 Hz), whereas in *plakë* [plɑːk] "old woman" a rather back vowel [ɑ] (F2 1294 Hz) is articulated. It should be mentioned that in most cases the long vowel is the consequence of compensatory lengthening caused by the deletion of the final schwa.

In our data, however, back articulation is not limited to long vowels. For determining back or front articulation, again three articulation zones have been defined. The average F2 value of all /a/ – vowels across all speakers (= 1342 Hz) was taken as the starting point for the definition of the articulation zones:

- back articulation: F2 ≤ 1300 Hz
- central articulation: F2 = 1300 Hz – 1400 Hz
- front articulation: F2 ≥ 1400 Hz

As can be seen in Figure 6, high agreement as regards back or front articulation can only be observed for three items: the vowel /a/ in *bardhë* "white" is articulated as a back vowel in 89% of all realizations, whereas in *larg* "far" and *ata* "they (m.)" it is articulated as a front vowel (92% and 88% respectively).

All the other vowels do not show a clear preference with respect to front or back articulation. Even a nasal context does not trigger a back articulation (as it is the case in Gheg dialect [8]), this can be easily seen from the examples *Lana* "Lana (a river name)", *malit* "mountain-ABL" and *natë* "night".

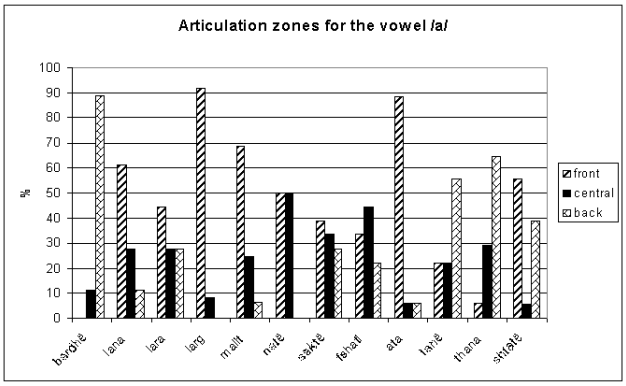

Figure 6: Within-item variability of the vowel /a/ according to front, central or back articulation (in%).

As regards the three investigated geographical regions, a tendency towards front articulation can be observed for the Tosk speakers, whereas the Gheg speakers show a tendency towards back articulation (see Figure 7). Again, these differences are statistically significant.

ISBN 1-876346-48-5

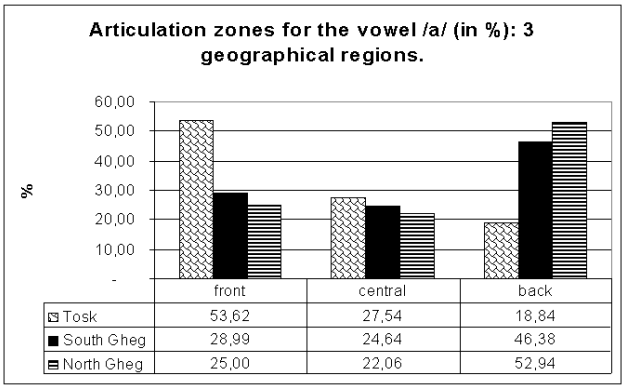

Figure 7: Distribution of the vowel /a/ according to geographical regions and according to front, central or back articulation.

However, just as for the schwa, the statistical significance is less important than the fact that there is no clear division between back and front articulation, i.e. that there is a gradual progression from front to back articulation (for the Tosk speakers) or from back to front articulation (for the Gheg speakers). This graduality together with the arbitrary assignment pattern in the actual realizations points out that there is only a single phoneme for the vowel /a/.

The vowels /e, i and y/ show only a small amount of variability. Their range of variation however, – though smaller than for the central vowels – is still considerable (see Table 2):

Range of variation	e	i	y	a	schwa
Bark	1,9	1,9	1,6	3,2	3,2

Table 2: Range of variability (min/max values, difference in Bark) for the front vowels as compared to the central vowels

However, these vowels do not convey any regional or social information and consequently, just as for the back, rounded vowels, the variation within these vowels is not salient.

4. CONCLUSIONS

Although all vowels show statistically significant regional differences, the weak points in the vowel system of Standard Albanian are definitely the two central vowels /a/ and /ə/. The variation within these vowels not only carry regional information, but may also carry social information. Gheg dialect and Standard Albanian can be more easily discerned in the vowel /ə/, since this vowel has to be described as a back, rounded and nasalized vowel in the Gheg dialect [8], whereas the Standard Albanian articulation lacks roundness and nasalization. Therefore, the difference between the regions is limited to a quantitative difference: a rather back articulation in the regions traditionally described as Gheg, a rather front articulation in the regions traditionally described as Tosk. On the qualitative level, however, the regions do not differ; all expose front and back articulation for the schwa.

The case is not as clear for the vowel /a/. The difference between the Gheg dialect and Standard Albanian is not a qualitative one, but rather a distributional one: According to the measurements given in Beci [8] for the Gheg dialect, all long vowels as well as all vowels in nasal context are realized as back vowels, whereas the short vowels are front. In Standard Albanian, however, the distribution of front and back realization is an arbitrary one, where the speakers from the North show a preference for back articulation and the speakers from the South a preference towards front articulation.

Faced with such results, evaluation tests should be performed in order to find out whether the whole range of variation within the /a/ and the /ə/ vowels are assessed as being Standard Albanian articulation. Acceptance of the whole range of variability would speak for a high tolerance as regards Standard Albanian vowel articulation.

REFERENCES

[1] X. Lloshi, "Albanian" in *Handbuch der Südosteuropa-Linguistik*, U. Hindrichs, Ed., pp. 277–299. Wiesbaden: Harrassowitz Verlag, 1999.

[2] G.-D. Nehring, "Albanisch" in *Lexikon der Sprachen des europäischen Ostens*, M. Okuka and G. Krenn, Ed., pp. 47–65. Klagenfurt/Celovec: Wieser Verlag, 2002 (= Wieser Enzyklopädie des Europäischen Ostens, Band 10).

[3] A. Vehbiu, "Standard Albanian and the Gheg Renaissance: A Sociolinguistic Perspective," *International Journal of Albanian Studies*, vol. I, pp. 1–14, 1997.

[4] S. F. Disner, "Evaluation of vowel normalization procedures," *Journal of the Acoustical Society of America*, vol. 67 (1), pp. 253–261, 1980.

[5] S. Björsten and O. Engstrand, "Swedish "Damped" /i/ and /y/: Experimental and Typological Observations," *Proceedings of the XIVth International Congress of Phonetic Sciences*, vol. 3, pp. 1957–1960, 1999.

[6] P. Ladefoged and I. Maddieson, *The Sounds of the World's Languages*, Oxford, Cambridge: Blackwell, 1996.

[7] E. Zee and W.-S. Lee, "An acoustical analysis of the vowels in Beijing Mandarin." *Proceedings of the 7th European Conference on Speech Communication and Technology, Eurospeech 2001*, pp. 643–646, 2001.

[8] B. Beci, *Të folmet veriperëndimore të Shqipërisë dhe sistemimi fonetik i së folmes së Shkodrës*, Tiranë, Akademia e Shkencave e RSh, 1995.

[9] T. Granser and S. Moosmüller, "Phonemic Schwa – a challenge for Natural Phonology? The case of Albanian" in *Future Challenges for Natural Linguistics*, K. Dziubalska-Kołaczyk and J. Weckwerth, Ed., pp. 165–178. München: Lincom Europa, 2002.