

ANNA V. DYBO
Institute of Linguistics, RAS, Moscow
adybo@mail.ru

ON THE REFLEXES OF PROTO-TURKIC VOWEL LENGTH IN THE TURKIC LANGUAGES

*... Parmi ces lambeaux de phrases presque effacés,
les flots ont respecté quelques chiffres: malheureusement,
la longitude...*

— On s'en passera! s'écria le jeune garçon.

Jules Verne

The 129th volume of *Studia Linguistica Universitatis Iagellonicae Cracoviensis* contained an annotated edition of a previously unpublished paper by one of the “fathers” of modern Altaic studies, W. Kotwicz, written, it appears, around 1938 (K. Stachowski 2012). The edition is a critical one, and an attempt at conveying to the reader as accurately as possible what Kotwicz had intended to publish, i.e. what he might have considered to be the completed text. In my opinion, the editor has fully succeeded in this.

The publication is doubtless of great historical-linguistic importance. So is the history of its creation that the publisher gave in the introduction, for it is a novel contribution to our understanding of the evolution of one of the problems of Turkic and Altaic reconstruction that was debated throughout the second half of the 19th and the 20th century, the problem of “primary” long vowels in the Turkic languages. The views on the Manchu-Tungus and the Mongolic long vowels that Kotwicz had expressed in his paper, have generally remained unchanged till this day; with the Turkic, this is not the case.

In this paper, I attempt to give a more detailed “contextual commentary” to Kotwicz’s article, i.e. to present the wider context of his research, and what the Turkic historical-comparative linguistics has finally settled on; that is to say, the views that Turkologists-comparativists currently adopt on the shortness : longness opposition in Proto-Turkic, and how they were influenced by Ligeti’s paper mentioned in the editorial commentary. I hope that the reader, not necessarily an expert on

historical-comparative grammar of the Turkic languages, will be able to use this presentation to evaluate Kotwicz's assumptions and conclusions, and eventually, to understand why he has never published his paper. In particular, we will see the role played by methodological flaws and lack of certain factual data which were not clarified until much later (especially those on vowel length in the languages of the Sayan group, referred to as "Uryankhai" in Kotwicz's paper).

As is known, Boetlingk (1851: 159–169) was the first to suggest phonological vowel length in Proto-Turkic. He compared Yakut long vowels and diphthongs with the long vowels of "Nizhny Novgorod Tatar", i.e. Mishar Tatar. At any rate, the data he collected were sufficiently relevant to assume the length opposition for **a* and **o*.

Boetlingk's material makes it also possible to isolate secondary (contracted) long vowels and diphthongs in Yakut, i.e. the cases where other Turkic languages have a vowel-consonant sequence, such as in Yak. *ȳs* 'smoke': Kazan *ijis*, Nizhny Novgorod *īs*. In this way, Boetlingk anticipates Grønbech's objections; see below on his hypothesis.

The idea of ancient long vowels was not immediately recognised. Budenz (1865) extended Boetlingk's analysis of Yakut lengths with lengths in Uzbek and Chuvash. We know today that in Uzbek, traces of long vowels can be found in dialects, in particular in the Oghuz Khwarezmian dialect, and in the dialects of Ikan and Karabulak (see below), although regular recordings have practically not been conducted. In Chuvash, Budenz saw the reflex of the original length in a different phenomenon, namely in the preservation of an old rising diphthong. Although it does not accurately correlate with vowel length,¹ this misconception is to be seen even in modern Turkological literature, in the works of Ligeti, Räsänen, and Doerfer, and later ones (see Schönig in Janhunen 2005: 409).

Foy (1900) established Proto-Turkic length based on the relations between the vowels in the Orkhon-Yenisey, Yakut, and Turkish. Later studies of Orkhon-Yenisey

¹ The rising diphthong *ia*, both long and short, was introduced into the Proto-Turkic reconstruction after the work of Vladimircv and Poppe (1924). For PT *a* and *ā*: Chuv. *jo-* (the more archaic form of the Upper dialect, Viryal) when the word begins with a vowel, and *o* with a palatalisation of the preceding guttural or dental postconsonantal consonant otherwise (PT **Kiān* 'blood', Trkm. *Gān*, Khal. *qān*, Sal. *Ga(:)n*, Yak. *xān*, Chuv. *jon*; PT **Kiār* 'snow', Trkm. *Gār*, Yak. *xār*, Chuv. *gor*; PT **Kiār-* 'to mix', Trkm. *Gār-*, Chuv. *gor-*, *jur-*; PT **Kiārȳ* 'abdominal fat', Yak. *xaha*, Chuv. *gor-var* 'modest food'; PT **diāl* 'stone', Trkm. *dās*, Khal. *tās*, Yak. *tās*, Chuv. *čol*; PT **ciāk* 'time', Trkm. *čāG*, Yak. *sax*, Chuv. *čox*; PT **siāk-* 'to think', Trkm. *sāGyn-*, Yak. *āx-*, Chuv. *šoyš*; PT **siāryg* 'yellow, white', Trkm. *sāry*, Khal. *sāruy* 'orange', Yak. *arayas*, *arȳ* 'butter', Chuv. *šorš*; PT **iak-* 'to flow', Trkm. *aq-*, Chuv. *jox-*). Presently, it is widely accepted to treat this rising diphthong as a strictly Chuvash innovation and continuation of the Proto-Turkic long **ā*. One can see why this attitude is incorrect; cf. the cases where the quite reliable Proto-Turkic **ā* corresponds to Chuv. *u* (*o*) with no palatalisation at all: PT **Kār* 'goose', Trkm. *Gāz*, Yak. *xās*, Chuv. *xor*; PT **Kāpuk* 'bark', Trkm. *GābyG*, Chuv. *xobš*; PT **dāla-* 'to bite', Trkm. *dāla-*, Chuv. *tola-*, recipr. *tola-š-*; PT **sāl* 'raft', Trkm. *sāl*, Yak. *āl*, Chuv. *solš*; PT **ār-* 'to stray', Trkm. *āz-*, Khal. *hāz-*, Yak. *ās-*, Chuv. *or-*; PT **āt-* 'to tread', Trkm. dial. *āt-*, *ādym*, Yak. *atyllā-*, Chuv. *odšm*, *ot-*. On the other hand, there are examples where a diphthong or palatalisation in Chuvash corresponds to short vowels in Turkmen and Yakut (as above in 'abdominal fat' and 'to flow'), and in effect, a short diphthong needs to be reconstructed. See also Dybo 2007: 60.

do not confirm the possibility of distinguishing long from short vowels in these attestations (see Kormušin 2008); length in literary Turkish is entirely secondary and due to the dropping of the final **y*; finally, Foy's data (1900: 214) on length in Turkish dialects is incomplete and inconsistent.

Radloff suggested that those long vowels and diphthongs in Yakut that do not evidently originate from contractions of vowel-consonant sequences (still partially preserved in mediaeval texts), are products of secondary changes enforced by the necessity to create a semantic distinction between phonetically similar words, such as Yak. *tās* 'stone' and *tas* 'outer part'. This explanation, positively abominable for comparative linguistics as it is, was put forward in Radloff 1882–1883, 1901 (against Foy), and 1908.

Grønbech, apparently, had initially identified the phenomenon of "Oghuz voicing" as the appearance of voiced intervocalic obstruents in Turkish after those vowels which correspond to long vowels in the analogous Yakut words (Grønbech 1902). Nonetheless, in the same book, he opposed the reconstruction of Proto-Turkic length, and proposed to explain all cases of Yakut long vowels and diphthongs through the syncope of intervocalic consonants – some of which have been preserved in other Turkic languages, and others have not – and offered quite speculatively such developments as Yak. *tüört* 'four' < **töbärttä*, Yak. *xān* 'blood' < **kajan*, etc. Pedersen (1903) challenged this solution on methodological grounds.

The debate continued with mixed success; there were attempts to include data from various Turkic languages in the comparison, but those proposals were easy to rebut as none of them was supported by a relevant number of comparisons.

As it seems, the first to establish the phonological character of vowel length in Turkmen, and to demonstrate its regular correspondence to the reflexes of the original length in Yakut, was Polivanov in his two papers of 1923 and 1927. In the earlier one, he contests Radloff's assumptions (1882–1883, §106):

В. В. Радлов, приводя объяснения якутских долгих гласных из стяжения двух слогов, склонен, видимо, отрицать наличие долгот в обще-турецкую эпоху—на том основании, что в прочих турецких языках (кроме якутского) он находит соответствующие краткие гласные. Но это происходит, очевидно, потому, что Радлову были незнакомы туркменские факты. На самом деле, соответствия якутской и туркменской долгот могут быть непосредственно истолкованы в пользу восстановления праязыковых обще-турецких долгих гласных (впоследствии сокращенных другими тур. языками). Приведу хотя бы следующие примеры: якутск. *ta:s* «камень»—туркм. *da:š* (отличается долготой от *daš* «далекий»); якут. *sa:s* (< *ja:z*)—туркм. *ja:z* (отличается долготой от *jaz* «пиши»); якут. *ba:r* «есть»—туркм. *ba:r* (отличается долготой от *bar* «иди»). (Polivanov 1923: 157)²

² Radloff, by suggesting to explain Yakut long vowels through contraction of two syllables, is apparently inclined to negate the existence of lengths in the Proto-Turkic epoch on the grounds that in the other Turkic languages (except Yakut), he finds the corresponding vowels short. But this is clearly because Radloff was not familiar with Turkmen facts. In point of fact, the correspondences between Yakut and Turkmen lengths can be immediately interpreted in favour of reconstruction of primary long vowels in Proto-Turkic (which were subsequently shortened in the other Turkic languages). Let us cite at least the following examples: Yak. *ta:s*

Polivanov adduced also one case of appearing inconsistency: long in Yakut, short in Turkmen: Yak. *at* ‘name’ : Trkm. *ad*. His explanation for this case, based on phonetic typological research, was that in Yakut, a secondary lengthening is possible before an old voiced consonant. I should mention that this last example is cited erroneously in Polivanov’s work: in Turkmen ‘name’, *āt* with a long vowel, the voicedness of the final consonant only appears in the non-final position (*ādi* ‘his name’). One might conclude from this observation that Polivanov had not yet been at that time acquainted with the work of Grønbech, the discoverer of the phenomenon of “Oghuz voicing” (i.e. of the appearance of voiced intervocalic stops in Turkish after vowels corresponding to long ones in Yakut; Grønbech 1902). In his later works on Oghuz dialects of Uzbek and on the classification of Uzbek dialects, i.e. after 1929, Polivanov is already well acquainted with the phenomenon of “Oghuz voicing” and treats it as an established fact.

Polivanov’s later paper (1927) is more extensive. He endorses Poppe’s (1925) view that the Yakut-Mishar correspondences with Chuvash, together with the evidence of ancient written monuments, necessitate the reconstruction of quantitative differences in the Proto-Turkic vowel system. He extends this conviction with remarks on the notation of length in the runic Orkhon Turkic monuments (this hypothesis was first proposed by Foy (1900) but Polivanov does not mention this) and in Mahmud al-Kashgari’s Middle Uighur, and he particularly insists that Turkmen (assuming that length is reflected in all of its dialects), and northern Uzbek dialects be introduced into the comparison. Polivanov produces etymologies of fifteen one- and two-syllabic words with primary length, plus of seven words with primary short vowels which constitute minimal pairs for some of the fifteen long ones. He makes a series of suggestions: a) that the comparison might be extended with lengths observed by him in one of the Karachai dialects (this was not confirmed by later studies); b) that the ancient long *ō* and *ō̄* changed to *u* and *ü* in the northern dialect of Kumyk, while short *o* and *ö* have been preserved (this was not confirmed by later studies); and c) that old lengths are also reflected in Radloff’s (actually, Castrén’s) recordings of the Karagas language (nowadays it is believed that length is positional in Karagas, and low vowels are lengthened before a high vowel in the following syllable, although contradictory evidence from Khakas dialects has recently appeared [see Dybo 1986, 2010b] which may confirm the existence of an indirect correspondence to the ancient lengths). Polivanov’s paper ends so:

Вышеприведенные примеры, конечно, — только иллюстрация соответствий, повторяющихся *passim* по всему словарному материалу. Добавлю, что в настоящее время Центроиздатом Народов Востока издается мой фонетический словарь туркменского языка, с помощью которого будет возможно сопоставляя туркменские формы с якутскими глос. мещерскими и т.д. определять количество праязыковых гласных для любого примера. (Polivanov 1927: 153)³

‘stone’ : Trkm. *da:š* (distinguished by length from *daš* ‘distant’), Yak. *sa:s* (< *ja:z*) : Trkm. *ja:z* (distinguished by length from *jaz* ‘to write’), Yak. *ba:r* ‘there is’ : Trkm. *ba:r* (distinguished by length from *bar* ‘to go’).

The dictionary was never published.

In his later works, Polivanov freely uses Turkmen and South Khwarezmian lengths as a criterion for the classification of Uzbek dialects.

Prior to Polivanov, the only scientific records of Turkmen were those by Ilminskij (1863) and Samojlovič (1912), although in the latter lengths are not marked. They are marked in Pocolujevskij's 1929 book *Руководство для изучения туркменского языка (с приложением краткого туркменского словаря)*, and we also know from his letter to Samojlovič (from June 1928, see Blagova 2008: 253) that Pocolujevskij was planning to publish a scientific dictionary of Turkmen in collaboration with Boriev. The Turkmen linguist was the head of both the Turkmen Cultural Institute, and of the Turkestan publishing house, and also the president of the Turkestan State Council of Scientific Research in the years 1925–1930. He had decided in 1928, that publishing a practical dictionary is an important issue, and in 1929 a Russian-Turkmen dictionary appeared, authored by Aliev and Boriev. According to his contemporaries, Aliev was the principal author. Lengths are carefully marked in this work.

Later the idea was picked up by Räsänen (1937) and Ligeti (1937, 1938) who immediately began to contest their priority. It needs to be said that they both cite Aliev and Boriev's (1929) Russian-Turkmen dictionary as the source of their Turkmen material, the first dictionary to consistently mark vowel length, and they both certainly knew about Polivanov's 1927 paper. In 1938 Ligeti wrote:

Avant de partir pour l'Afghanistan, comme je comptais y étudier le turcoman, du moins dans ses grandes lignes, je m'étais procuré l'ouvrage de A. M. Pocoluevskij, intitulé *Rukovodstvo dlja izutchenija turkmenskogo jazyka* (Achkhbat, 1929). Dès que je l'ai eu entre les mains, au mois de mai 1930, j'ai communiqué à l'Institut de turcologie de l'Université de Budapest mon observation selon laquelle dans le turcoman on trouve régulièrement des voyelles longues et que celles-ci pourraient bien être en rapport avec les longues du yakoute. A l'automne de cette même année j'ai adressé d'Afghanistan à M. J. Németh une lettre où je me suis prononcé d'une façon encore plus explicite sur la liaison que je croyais voir entre les longues du turcoman et celles du yakoute. On trouve du reste les voyelles longues dans les dialectes turcomans d'Afghanistan aussi, et – autant que j'ai pu le constater – elles coïncident exactement avec les données des sources russes. Indépendamment de moi, M. Räsänen a également établi la connexion des longues turcomanes et yakoutes, ainsi que le prouve son étude intitulée *Über die langen Vokale der Turkischen Lehnwörter im Ungarischen*, dans *FUF*, t. XXIV, pp. 246–55, qui fut publiée au printemps de 1937, mais dont je n'ai pu prendre connaissance qu'à mon retour d'Afghanistan, c'est-à-dire au cours de l'automne de cette même année. Quant au travail de M. E. D. Polivanov, *K voprosu o dolgikh glasnykh v obchtchetureckom prajazyke*, il m'est resté malheureusement inaccessible. Très vraisemblablement d'autres turcologues aussi, indépendamment les uns des autres, sont arrivés à la même conclusion à propos des langues du turcoman.

³ Of course, the above examples are just an illustration of the correspondences which recur throughout the lexical material. I should add that the Centroizdat Narodov Vostoka publishing house is currently producing my phonetic dictionary of Turkmen which will make it possible to compare Turkmen forms with the Yakut, resp. Mishar etc. ones, and to determine for each example the quantity of the vowels in the protolanguage.

Ainsi, je viens de remarquer que, dans les derniers fascicules de son grand dictionnaire yakoute, M. Pekarskij donne déjà quelques rapprochements yakouto-turcomans à voyelles longues. (Ligeti 1938: 178–179)⁴

In 1937 Räsänen wrote:

Vor einigen Jahren bekam ich aus Russland ein neues russisch-turkmenisches Wörterbuch Aliev-Boriev, Ashh. 1929, das einen ziemlich sorgfältigen und zuverlässigen Eindruck macht. Unter anderem sind die langen und kurzen Vokale genau unterschieden... Je mehr ich das Buch studierte und die Längen mit den jakutischen Längen verglich, desto mehr bemerkte ich zu meiner grossen Überraschung, dass diese besonders in den einsilbigen Wörtern einander vollständig decken... Beim Korrekturlesen dieses Artieckels sehe ich in den «Studien zu einer osttürkischen Lautlehre» von Gunnar Jarring (Lund 1933) p.44 note 3, dass auch E.D. Polivanov in seinem, leider mir nicht zugegangenen Aufsatz «К вопросу о долгих гласных в общетуецком праязыке» eine Anzahl Wörter anführt, vor allem aus Turkmenischen Dialekten, die er als urtürkische Längen enthaltend ansieht, also ist er in der Hauptsache zu demselben Resultat gekommen wie ich. (Räsänen 1937: 249)

In (1949: 59), he wrote:⁵

Вопрос был решен, кажется, в положительном смысле только после появления русско-туркменского словаря Алиева-Бориева, в котором были ясно обозначены различия и благодаря которому выяснилось, что эти различия исключительно точно соответствуют различиям по долготе в якутском языке.⁶

and then he referred to the works of Polivanov (1927), and also of Räsänen (1937), and Ligeti (1937, 1938). In the footnote on p. 59, he answers Ligeti:

О приоритете этого открытия спорить бесполезно. Относительно примечания Л. Лигети [...] я хочу заметить, что работа автора этих строк по вине типографии вышла значительно позднее, чем я ее написал. Дело в том, что многие исследователи независимо друг от друга пришли к одинаковым результатам, так что этот спорный вопрос решен теперь окончательно.⁷

Since the works of Ligeti and Räsänen, Turkic length has acquired the status of an established truth in the linguistic community.

⁴ The remark about Pekarskij is inaccurate. In the 1928 edition of his dictionary, we find forms described as “тюркские”: *dūz* ‘salt’ (possibly Turkmen), *tōn* ‘clay’ (Oirat, the length is secondary), but at the same time *dōrt* ‘four’ (Turkish? It is *dōrt* in Turkmen) – see Pekarskij (1959).

⁵ Quotes after the Russian 1953 edition.

⁶ It seems that the problem was only positively resolved after the appearance of Aliev-Boriev’s dictionary in which the differences were clearly marked, and thanks to which it became clear that these differences very accurately correspond to length differences in Yakut.

⁷ It is useless to argue the priority of this discovery. As regards Ligeti’s remark [...] I want to note that the work of this author was published significantly later than it had been written, through the fault of the printing house. The point of the fact is that many researchers have independently reached the same conclusions, and so the controversial issue is now finally resolved.

It was, however, Ligeti who was credited with the discovery of Yakut-Turkmen correspondences in vowel length after T. Tekin had quite liberally attributed it to him in his 1967 paper *Determination of Middle-Turkic long vowels through 'arūd*:

However, it was not until the appearance of the very instructive et inspiring article de L. Ligety, that the systematic existence of primary long vowels in Proto-Turkic was established definitely. In fact, Ligety was the first to notice the systematic preservation of the primary long vowels in another Turkic language, namely in Turkmen spoken in West Turkestan, an area far away from that of Yakut. (Tekin 1967: 152)

One should also note Tekin's remark that "[...] Ligety was also the first who noticed the preservation of Proto-Turkic long vowels in Middle Turkic, by Kashgari." (Tekin 1967 : 153).

Indeed, Ligeti wrote that:

C'est ainsi que nous sommes arrivés à Mahmūd al-Kāšyarī, source de la plus haute importance du turc du xi^e siècle... J'ai été surtout frappé de l'esprit de suite de Kāšyarī dans la notation de la quantité dans les couples de mots du type *at* « cheval » ~ *āt* « nom ». Il groupe les mots turcs selon son modèle arabe et, par là, automatiquement, les mots à voyelle longue se trouvent rangés dans un même chapitre. Seulement, comme on sait, les mots ayant le même nombre de radicaux doivent être cherchés, selon qu'ils commencent par une voyelle ou par une consonne ou qu'ils sont des verbes ou des substantifs, etc., dans des endroits différents de son livre. (Ligeti 1938: 182–183)

But here, too, Tekin is not quite correct about the priority. Not to mention earlier works, already Polivanov had written in his 1927 paper that "... из древних памятников, кроме уйгурских и орхонских, надо назвать словарь Махмуда Кашгарского, последовательно дифференцирующий долгие и краткие" (Polivanov 1927: 151).⁸

The history of the publication of Aliev and Boriev's dictionary has already been discussed, but there is still the following remark by Samojlovič to be added: "С древних пор существуют долгие гласные, которые, за исключением якутского и туркменского языков, не привлекают к себе особого внимания" (Samojlowitsch 1931: 958).⁹ All this leads to the conclusion that it was in fact Polivanov who first highlighted the importance of Turkmen lengths for Turkology, and that consistent marking of vowel length in Turkmen materials did not start until after his 1923 paper.

Since the works of Ligeti and Räsänen, Turkic length has acquired the status of an established fact in the linguistic community, but the data supporting the reconstruction of Proto-Turkic long vowels remained fragile. A number of researchers, less experienced in the usage of the historical-comparative method,

⁸ ... of the ancient monuments, apart from the Orkhon and Uighur ones, one must mention the dictionary by Maḥmūd al-Kāšyarī where long and short vowels are consistently marked.

⁹ From the ancient times, long vowels have been preserved which, apart from Yakut and Turkmen, do not attract particular attention.

allow for example conclusions about the proximity of Yakut to the Oghuz group, based on the general feature of preservation of primary vowel length, and thereby implicitly suggest that these lengths are a common Yakut-Oghuz innovation – which, to be sure, is wrong.

Presently, our research group (see SIGTJa 1997/2000, 2002, 2006) has adopted the following rules for Proto-Turkic length. In the first syllable, long vowels are reconstructed based on Turkmen-Yakut correspondence, but Oghuz voicing is also taken into account. These data correlate with the presence (in the case of short vowels), and absence (in the case of long vowels) of pharyngealization in Tuvinian (sporadically marked in the orthography, and usually noted in Bičeldej 2001) and Tofalar (after Rassadin 1971, 1995). This correlation was first observed by Illič-Svityč in 1963. On the occasional preservation of primary long **ā* and **ō* in Gagauz, see below and SIGTJa (2002: 23–24). Moreover, primary length is reflected in Azeri in the different continuants of low and high **e*, and in Turkmen in the reflexes of vowel-consonant sequences (*av* > *ov* while *āv* has been preserved). We consider the tripartite opposition of length in Khalaj (Doerfer et al. 1971) to be non-phonological (cf. the numerous variants in the case of half-length). Nonetheless, Khalaj overlong vowels correlate quite consistently with the remaining Turkic data.

Räsänen's hypothesis that Proto-Turkic length opposition was preserved in Proto-Bolghar (whence borrowings in Hungarian), is fully confirmed. The opposition is indirectly reflected in the Slavonic borrowings from Danube Bolghar (see Dybo 2010a,c), and also partly in the orthography of Volga Bolghar epitaphs (Erdal 1993: 149–151). Further research is required into how the Proto-Turkic length opposition is reflected in Old Turkic monuments in various writing systems. As of now, exact rules have generally not yet been established. Examples of Proto-Turkic minimal (and quasi-minimal) pairs for length:

- **āt* 'name' (Yak. *āt*, Trkm. *āt*) : **āt* 'horse' (Yak. *at*, Trkm. *at*),
- **ōt* 'fire' (Yak. *uot*, Trkm. *ōt*) : **ot* 'grass' (Yak. *ot*, Trkm. *ot*),
- **ēt-* 'to do' (Trkm. *et-*, *īdeg* 'supervision') : **et* 'meat' (Yak. *et*, Trkm. *et*),
- **āl* 'front' (Trkm. *āl-yn*) : **āl* 'bottom' (Yak. *al-yn*, Trkm. *al-t*),
- **āč* 'hunger' (Yak. *ās*, Trkm. *āč*) : **ač-* 'to open' (Yak. *as-*, Trkm. *ač-*),
- **ūn* 'sound' (Yak. *ūn*, Trkm. *ūn*) : **ün* 'flour' (Yak. *ün*, Trkm. *ün*).

The Yakut-Dolgan vowel system preserved the ancient length opposition. At some point in its development, long mid vowels in the first syllable changed into diphthongs (**ē* > *ie*, **ō* > *uo*, **ō̄* > *üō̄*), while the remaining ones were preserved unchanged. As a result of the contraction of vowel-consonant sequences, new long mid vowels arose, and the system of diphthongs became fully symmetric (the diphthong *ya* appeared as a harmonic counterpart for *ie*).

It was observed long ago (Räsänen 1937) that Yakut is only consistent in the preservation of reflexes of primary length in monosyllabic words. The issue was investigated by M. Stachowski (1993: 42–45). He has shown that vowels were consistently shortened in disyllabic stems and where primary long vowels appear to have been preserved in the first syllable, they are in fact due to secondary, i.e. internal

Yakut, word formation. Thus, it can be assumed that the presence in the initial syllable of a mono- or polysyllabic Yakut word, of a phonological unit which is identical to the reflex of a primary length, is grounded for reconstruction of a long vowel in the Proto-Turkic etymon. On the other hand, the absence of such a unit from a mono- or polysyllabic word can be considered evidence of a primary short vowel, if it cannot be interpreted as a result of Old Yakut shortening of Proto-Yakut polysyllabic stems.

Reflexes of primary long vowels are more stable in Turkmen than in Yakut. In Turkmen, the primary Oghuz vowel length was preserved. On the phonetic (not phonological!) level, high vowels evolved into diphthongal sequences and so merged with secondary diphthongs ([uv], [üj], [yj], [ij]). Sample minimal pairs for primary length: *at* ‘horse’ – *āt* ‘name’, *ot* ‘grass’ – *ōt* ‘fire’, *öl* ‘die!’ – *ōl* ‘wet’, *gurt* ‘dried curd’ – *gūrt* ‘wolf’, *düš* ‘fall!’ – *dūš* ‘sleep!’, *gyz* ‘heat!’ – *gŷz* ‘girl’, *bil* ‘know!’ – *bīl* ‘waist’.

The phonemes short *e* and long *ā* require special consideration. Strictly speaking, they both have their respective pairs for length in modern literary Turkmen, but their use is limited. Short *ā* appears in borrowings and two Turkic stems, albeit with an irregular development: *ākel-* ‘to bring’ (< **alyp kel-*) and *ākit-* ‘to raise’ (**alyp kit-*). Long *ē* can be found in two words which also arose through contraction: *bēr* ‘maybe s/he will give’ and *gēr* ‘maybe s/he will come’ (from *berer* and *gerer*, respectively). Peripheral as they are, the two phonemes are usually not included in the phonological system (see Dmitriev 1955: 185). The ancient (Proto-Oghuz) **ē* yielded *ī* in the literary language, as in *bīl* ‘waist’ mentioned above, and merged with the reflex of long **ī*: *dīs* ‘tooth’.

Secondary shortening occurs in Turkmen in the following situations:

1. In the case of illabial high, and front labial vowels: a) dissimilative elision of the second element of the rising diphthong with initial *j*–: *jilān* ‘snake’, *jir-* ‘to split’, *jüz* ‘face’, ‘hundred’ (Yak. *syl-* ‘to creep’, *sīr-*, *sūs*), or b) merging of this element with the following *j* of the root: *dij-* ‘to prevent’, *dŷ-yi* ‘stop’, *ij-* ‘to send’, *tūj* ‘wool’, *üvi-*, *üvüt-* ‘to grind’ (**tŷd-*, Yak. *tŷt-*, **ŷd-*, Yak. *ŷt-*, **tūk*, Yak. *tū*, **ūk-*). Apparently, this process occurred in general Turkmen; cf. the reflexes of **ē* with dissimilation in the opposite direction: *ij-* ‘to eat’, *ijmiš* ‘fruit’ < *jēmilč* (the change to *ī* is not general Turkmen, see below).
2. In all the other cases: these are almost exclusively closed monosyllabic verbal stems (often there also exist semantically divergent derivatives coined from them, or other non-verbal monosyllables, in which the root syllable is open but the length is preserved): *sūr* ‘to drive’ (Yak. *ūr-*, Khal. *sīr-*), *Guč-* ‘to hug’ (*Gužaq* ‘hug’ – an associated and regular derivative; the length aligned to the verb; “Oghuz voicing” remains), *et-* ‘to do’ (*īdeg* ‘supervision’ – a dissociated derivative), *ber-* (Yak. *bier-*, Khal. *vīer-*) ‘to give’, *bol-* ‘to be’ (Khal. *uol-*, Yak. *buol-*), *mün-* ‘to ride on horseback’ (**būn-*, dial. *mīn-*), *tez-* ‘to run’ (*tīz* ‘quick’). Cf. the absence of shortening in monosyllabic closed roots within disyllabic stems: *dūjr-le-* ‘to twist’, *būš-la-* ‘to be the first to report good news’, *āt-le-* ‘to walk’, *jān-ža-* ‘to upbraid’, *iber-* (**jēp-*) ‘to send’.

Apparently, at the common Oghuz stage, or nearing the moment of disintegration of the Oghuz community, several examples of secondary length have arisen, which are typically reflected in Turkmen in the same way as primary length, and result in “Oghuz voicing” in the remaining Oghuz languages: *-iji-, *-yji- > *ī* (Tksh. *iğ*, Acc. *iği* ‘spindle’, Gag. *ī*, Acc. *iji*, Az. *ij*, Acc. *iji*, Trkm. *ik*, -*yi* < **ijik*; Tksh. *čig*, Acc. *čiği* ‘raw’, ‘dew’, Gag. *čī*, Acc. *čiji*, Az. *čij*, Acc. *čiji*, Trkm. *čyγ*, Acc. *čyγγ* < **čyj-γk*, cf. Az. *byγ*, Acc. *byyy* ‘moustache’ < **byjyk*); *-ubu- > *-ū- (Tksh. *ut*, -*du* ‘shame’ < **uvut*, cf. **uvt-an*- > Tksh., Gag., Az., Trkm. *utan*-; Tksh. *jut*- ‘to swallow’, *judum* ‘sip’, Gag. *jut*-, *judum*, Az. *ud*-, but Trkm. *juvut*-).

“Oghuz voicing”:

- Az. *ad* – *adym* ‘name’ vs. *at* – *atym* ‘horse’;
- Tksh. *at* – *adym* ‘name’ vs. *at* – *atym* ‘horse’.

This phonetic process was entirely consistent but its effects in the Oghuz languages are often obscured by morphological process of paradigmatic unification. This is particularly clear in nominal stems.

As a matter of fact, the recording of length opposition in Turkic written monuments has not yet been investigated. While describing these monuments, researchers list the alleged cases of marked length if they correspond to Proto-Turkic long vowels, but they neither examine those cases which correspond to Proto-Turkic short vowels, nor those Proto-Turkic long vowels which are not marked as long. Clearly, the principal task is to systematically describe the orthographic systems of these monuments, to establish the statistics of spelling variants, and based on that, of the opposition. Such analysis has not been carried out yet. Indeed, traditional Turkology often attempts to solve the problem of Proto-Turkic length based on random examples from the monuments and completely ignores the clear descriptions of phonological oppositions in modern languages. It will be worthwhile to once again remind Dmitriev’s remark from his work on the importance of Old Ottoman monuments for the historical dialectology of Turkish:

Направление всех наших изысканий должно идти «снизу вверх», т. е. от данных современности к прошлому, данные которого мы должны получить путем реконструкции. Идти от большого и известного материала к скудному и мало известному, а не наоборот. Это — настолько обычный методологически путь, что говорить о нем было бы, пожалуй, излишне, если бы среди представителей современной туркологии (науки молодой) не имелось лиц, которые (по крайней мере, в устной беседе) не держались бы обратного мнения. (Dmitriev 1928: 419)¹⁰

¹⁰ The direction of all our research should be from the bottom up, i.e. from modern data to the past, to the data we ought to be obtaining through reconstruction. Setting out from the well-known material towards the scanty and unfamiliar, and not the other way round. This is so much the natural way, that it would be perhaps superfluous to talk about it, had it not been for those names in contemporary Turkology (a young science) who, at least in speech, hold the opposite view.

The most systematic descriptions of Old Turkic writing systems, regarding the marking of vowel length, are:

- Tuna (1960), where a failed attempt is made to correlate plene notations in the Orkhon runic script with Proto-Turkic length,
- Tekin (1967), a study of the correlation between plene notation in the Qutadyu Bilig poem (11th c.) in the Arabic script, and the rhythmic of aruz (a traditional Perso-Arabic rhythm); also with precarious results (see the methodological criticism in Boeschoten 1990: 187–188 and further).

The newest description of Old Turkic (Erdal 2004: 46–50) clearly states that Turkic dialects did not have phonological length already in the 10th century (when the most relevant part of Old Turkic monuments was created).

An interesting exception, although possibly only due to their low number, are words ascribed by Maḥmūd al-Kāšyarī to Oghuz, in which length is marked with rare regularity (it should be noted that – apparently, contrary to popular belief – other Old Turkic materials do not refer to the Oghuz group):

- *ā in an open syllable, short: *dāvā* < **dābā* ‘camel’ (Az. *dāvā*, Trkm. *düje*, Yak. *taba* ‘deer’),
- *e in a closed syllable, short: *yäng* ‘to conquer’ (**jeŋ-*, Az. *jen-*, Trkm. *jeŋ-*, Khal. *jeŋg-*); *yärdä:s* ‘fellow countryman’ (**jer* ‘land’, Az. *jer*, Trkm. *jer*, Khal. *jer*, Yak. *sir*),
- *ā in an open syllable, long: *ke:p* ‘how’ (**gāp* ‘shape, example, image’, Trkm. *gāp*, Chuv. *kap*, Yak. *kiep*),
- *ē in a closed syllable, long: *ye:r* ‘to loathe’ (**jēr-*, Trkm. *jer-*, Az. *jer-*, Trkm. *ir-*, Yak. *sir-*). Cf. a trace of Oghuz voicing: *oba* < PT **ōpa* ‘village, clan’, Trkm. *ōba*.

Fully systematic traces of Proto-Turkic quantitative opposition can be found in the Bolghar group. This in itself already gives evidence of its Proto-Turkic origin. It has been mentioned above that linking Chuvash diphthongoid reflexes of *a*-type vowels to Proto-Turkic length, a procedure that has been present in the literature since Budenz (1865), is erroneous. Nonetheless, Chuvash conserved the oppositions of initial *ō – *o (*o > *vi-*, *ō > *vu-*), and *ō – *ö in the first syllable (*ō > *va*, *-ōva*, *ö > *vi-*, *-u/-ü-*), see SIGTJa (2006: 166–168, 173–174) where Mudrak proposes updates to the traditional reconstruction.

In Dybo (2010a,c) we attempted to specify the origin of Turkic borrowings in Hungarian and the Slavonic languages, and have reached the following conclusions.

Proto-Turkic length is consistently reflected:

1. In Hungarian loanwords from Old Bolghar (5th–7th c.): long vowels behave in the same way as Old Hungarian long vowels, and short ones as Old Hungarian short ones: *pödör*- ‘to twist between fingers’: PT **püt-ir*- ‘to weave’ vs. *gyül-öl* ‘to be angry’: PT **jül-*; *ökör* ‘bull’: PT **ökür* vs. *szün-ik-* ‘to wilt’: PT **sön-* ‘to vanish, to wane’; *kos* ‘ram’: PT **Koč* ‘ram’ vs. *turó* ‘cottage cheese’: PT **tōrak*; *karó* ‘nail, spike’: PT **Kař-guk* vs. *szál* ‘raft’: PT **sāl*.
2. In Hungarian loanwords from Danube Bolghar (9th–11th c.) where the Proto-Turkic oppositions *o vs. *ō and *ö vs. *ō had yielded the apparently already Bolghar

- diphthongisation: *apró* ‘small’ < PT **opra-k* vs. *váj-* ‘to dig’ < PT **ōj-*; *kár* ‘harm’ < PT **Kōr* ‘harm, damage’ (initial clusters are prohibited in Hungarian); *seprő* ‘yeast, sediment’ < PT **čōprā-g* vs. *ver-* ‘to plait’ < PT **ōr-*; *lék* (non-alternating length), dial. *vék* ‘ice-hole’ < PT **ōkü* ‘ice-hole’; *kék* ‘blue’ < PT **gōk* ‘blue, green’.
3. In Southern Slavonic loanwords from Danube Bolghar (8th–10th c.) where the PT **a* : **ā* and the PT **ä* : **ā* oppositions are reflected as *o* : *a*: *тоязъ* ‘stick, staff’ < PT **daja-* ‘to prop, to lean’ vs. *самъчии* ‘manager’ < PT **sā-m*; *ковригъ* ‘type of bread’ < PT **gābrāk*, *озаръ* ‘hound’ < PT **ākār* vs. *канъ* ‘idol’: PT **gāp*. This gives evidence that the **ā* > *a* change that is reflected in Chuvash, had already taken place in the Bolghar language, i.e. in the source of these borrowings. The length opposition was preserved, and mirrored in the Slavonic opposition of quality, in accordance with the Slavonic system. The PT **o* : **ō* opposition reflects both the completed process of diphthongization, and also the diversification of vowels with respect to length: *сокачии* ‘butcher’ < PT **sok-* ‘to slaughter’ vs. *кваръ* ‘harm’ < PT **Kōr* ‘harm, damage’, *тварогъ* ‘cottage cheese’ < PT **tōrak*. Cf. also **ō* : **ō̄*: *колимогъ* < PT **kōlūngū* ‘wagon’ vs. *веруза* < PT **ōr-ūk*.
 4. In Eastern Slavonic loanwords from Volga Bolghar (10th–12th c.) which reflect the PT **o* : **ō* opposition: *воръ* < PT **ogry* vs. *ватага* ‘marquee’ < PT **ōtag*. Cf. also *ковъръ* < PT **Kābiř* ‘carpet’ and *Хонужьское море* ‘Caspian sea’ < PT **kapug* ‘gate’.

Reflexes of original lengths can also be found in the Turkic languages of southern Siberia. S.V. Kodzasov discovered during the 1985 MSU expedition that Tuvinian pharyngealization has the phonetic feature of broken phonation, from creaky to lax, and as a result, pharyngealized vowels are pronounced distinctly longer. (This result was never published, but cf. Bičeldej 2001: 30–36, about broken tone and longitude of the pharyngealized vowels). The Latin orthography of Tuvinian mirrors this phenomenon through doubled spelling. The correspondence between this “length” and Proto-Turkic shortness was observed by Poppe (1929). The sound law and its explanation (analogous to the “Oghuz voicing”) was first formulated by Illič-Svityč ([1963: 53–55]: **āt* ‘name’ > **ad* > *at*, *adym*; **at* ‘horse’ > **at^h* > *a”t*, *a”dym*). Thus, “pharyngealization” only appears after the ancient short vowels when the morpheme ends with an obstruent. Described in the works of Rassadin, Tofalar “pharyngealization” (phonetically – lax voice) fully corresponds to its Tuvinian counterpart and follows Illič-Svityč’s law (**āt* ‘name’ > **ad* > *at*, *adym*; **at* ‘horse’ > **at^h* > *a”t*, *a”tym*). The languages of the Sayan group have in addition a secondary phonological length originating from contraction. Bičeldej’s 2001 work shows that in absolute numbers, the duration of pharyngealized vowels lies between the duration of Tuvinian short vowels and that of the contracted long ones.

Long vowels observed in Mishar by Boetlingk to correspond to Yakut lengths, have not been confirmed by later descriptions. In 1985 I conducted a pilot hearing of Mishar continuants of Proto-Turkic minimal pairs for length. We found no diversification in monosyllabic shapes of the corresponding lexemes; in disyllabic words, however, the differentiation was clear. For the Russian ear, it seemed as if the

stress shifted between the first and the second syllable: *at* ‘horse’ – *atým*, *at* ‘name’ – *átým*; *ot* ‘grass’ – *otým*, *ot* ‘fire’ – *ótým*. Apparently, in the system of this particular informant, ancient long vowels reveal their length (perceived as stress) before a high vowel in the second syllable, while ancient short vowels display nothing, and the standard Turkic stress on the final syllable becomes audible. Certainly, these findings need to be subjected to a more rigorous test.

Abbreviations and references

Az. = Azeri | **Chuv.** = Chuvash | **Gag.** = Gagauz | **Khal.** = Khalaj | **PT** = Proto-Turkic | **Sal.** = Salar | **Tksh.** = Turkish | **Trkm.** = Turkmen | **Yak.** = Yakut

- Aliev – Boriev 1929. *Russko-turkmenskij slovar’*. Ashkhabad.
- Bičeldej K.A. 2001. *Faringalizacija v tuvinskom jazyke*. Kyzyl, Moskva.
- Blagova G.F. (ed.). 2008. *Aleksandr Nikolaevič Samojlovič. Naučnaja perepiska. Biografija*. Moskva.
- Boeschoten H. 1990. Vokalquantität in Rabghuzis Poesie. – *CAJ* 34: 187–213.
- Boetlingk O. 1865. *Über die Sprache der Jakuten*. St. Petersburg.
- Budenz J. 1865. Khivai táatarság. – *Nyelvtudományi Közlemények* 4: 269–331.
- Dmitriev N.K. 1928. Materialy po osmanskoj dialektologii. *Fonetika „karamalickogo” jazyka*, I–II. – *Zapiski kollegii vostokovedov* 3.2: 417–458.
- Dmitriev N.K. 1955. Dolgie glasnye v turkmenskom jazyke. – *Issledovaniya po sravnitel’noj grammatike t’urkskix jazykov*. 1. *Fonetika*. Moskva.
- Doerfer G., Heschke W., Scheinhardt H., Tezcan S. 1971. *Khalaj Materials*. Bloomington, The Hague.
- Dybo A.V. 1986. K voprosu o xakasskoj prosodii. – *Fonetika jazykov Sibiri i sopredel’nyx regionov*. Novosibirsk: 117–120.
- Dybo A.V. 2007. *Lingvističeskie kontakty rannix t’urkov. Prat’urkskij period*. Moskva.
- Dybo A.V. 2010a. Vokalizm rannet’urkskix zaimstvovanij v vengerskom. – *Gedenk von E.A.Helmsky. Finnisch-Ugrische Mitteilungen* 32/33: 71–120.
- Dybo A.V. 2010b. Sud’ba «pervičnyx» dolgot v t’urkskix jazykax. – *V prostranstve jazyka i kul’tury: Zvuk, znak, smysl. Sbornik v čest’ 70-letija V.A. Vinogradova*. Moskva: 598–622.
- Dybo A.V. 2010c. Bulgars and Slavs: phonetic features in early loanwords. – Mańczak-Wohlfeld E., Podolak B. (eds.). *Studies on the Turkic World. A Festschrift for Professor Stanisław Stachowski on the Occasion of His 80th Birthday*. Kraków: 21–40.
- Erdal M. 1993. *Die Sprache der wolgabolgarischen Inschriften*. Wiesbaden.
- Erdal M. 2004. *A Grammar of Old Turkic*. Leiden.
- Foy K. 1900. Türkische Vocalstudien, besonders das Köktürkische und das Osmanische betreffend. – *Mitteilungen des Seminars für Orientalische Sprachen. Westasiatische Studien* III: 180–217.
- Grønbech V. 1902. *Forstudier til tyrkisk Lydhistorie*. København.
- Illič-Svityč V.M. 1963. Altajskie dental’nye: t, d, δ. – *Voprosy jazykoznanija* 6: 23–45.
- Ilminskij N.I. 1863. Über die Sprache der Turkmenen. – *Mélanges Asiatiques* IV (1860–1863). St.-Petersbourg: 63–74. [Russian translation: *Voprosy jazykoznanija* 6 (2005): 97–114].
- Janhunen J. (ed.). 2005. *The Mongolic Languages*. London, New York.
- Kormušin I.V. 2008. *T’urkskie enisejskie epitafii. Grammatika, tekstologija*. Moskva.
- Ligeti L. 1937. A török hosszú magánhangzók. – *Magyar Nyelv* 34: 65–76.

- Ligeti L. 1938. Les voyelles longues en Turc. – *Journal Asiatique* 230 (avril-juin): 177–204.
- Pedersen H. 1903. Türkische Lautgesetze. – *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 57.3: 535–561.
- Pekarskij E.K. 1959. *Slovar' jakutskogo jazyka*. [vols. 1–3]. Moskva. [phototypic edition].
- Pocelujevskij A.P. 1929. *Rukovodstvo dl'a izučenija turkmenskogo jazyka (s priloženiem kratkogo turkmenskogo slovar'a)*. Ašxabad.
- Polivanov E.D. 1923. K voprosu ob obšče-tureckoj dolgote glasnyx. – *B'ulleten' 1-go Sred-neaziatskogo gos. univ.* [Taškent 1924 (april)] 6: 157.
- Polivanov E.D. 1927. K voprosu o dolgixe glasnyx v obšče-tureckom prajazyke. – *Doklady Rossijskoj Akademii Nauk* B: 151–153.
- Poppe N.N. 1925. Čuvaškij jazyk i ego otnošenje k mongol'skomu i tureckomu jazykam IV. – *Izvestija Rossijskoj Akademii Nauk* 6: 405–426.
- Poppe N.N. 1929. Zametki o fonetike tannu-tuvinskogo jazyka v sv'azi s voprosom ob alfavite. – *Kul'tura i piš'mennost' Vostoka* IV. Baku: 52–53.
- Radloff W.W. 1882–1883. *Phonetik der nördlichen Türkischen Sprachen*. Leipzig.
- Radloff W.W. 1901. *Zur Geschichte des türkischen Vokalsystems*. – *Bulletin de l'Académie Impériale des Sciences de St.-Pétersbourg* 14.4: 425–462.
- Radloff W.W. 1908. *Die jakutische Sprache in ihrem Verhältnisse zu den Türkischen Sprachen (= Mémoires de l'Académie Impériale des Sciences de St.-Pétersbourg. VIII^e série VIII.7)*. St.-Pétersbourg.
- Räsänen M. 1937. Über die langen Vokale der türkischen Lehnwörter im Ungarischen. – *Finnisch-Ugrische Forschungen* 24: 246–255.
- Räsänen M. 1953. *Materialien zur Lautgeschichte der Türkischen Sprachen*. Helsinki 1949. [Russian translation: 1953].
- Rassadin V.I. 1971. *Fonetika i leksika tofalarskogo jazyka*. Ulan-Ude.
- Rassadin V.I. 1995. *Tofalarsko-russkij, russko-tofalarskij slovar'*. Irkutsk.
- Samojlovič A.N. 1912. Turkmenskije zagovory. – *Živaja starina* 21.1: 117–124.
- Samojlowitsch A.N. 1931. Türkische Sprachen. – *Enzyklopädie des Islam* IV. Leipzig: 956–963.
- SIGTJa 1997/2000 = *Sravnitel'no-istoričeskaja grammatika t'urkskix jazykov: Leksika*. Moskva.
- SIGTJa 2002 = *Sravnitel'no-istoričeskaja grammatika t'urkskix jazykov: Regional'nye rekonstrukcii*. Moskva.
- SIGTJa 2006 = *Sravnitel'no-istoričeskaja grammatika t'urkskix jazykov: Prat'urkskij jazyk-osnova. Kartina mira prat'urka*. Moskva.
- Stachowski K. 2012. W. Kotwicz's unpublished study *Les voyelles longues dans les langues altaïques* (1938). – *Studia Linguistica Universitatis Jagellonicae Cracoviensis* 129: 245–315.
- Stachowski M. 1993. *Geschichte der jakutische Vocalismus*. Kraków.
- Tekin T. 1967. Determination of Middle-Turkic long vowels through 'arūd. – *Acta Orientalia Academiae Scientiarum Hungaricae* 20: 151–178.
- Tuna O. 1960. Köktürk yazılı belgelerinde ve Uygurcada uzun vokaller. – *TDAYB*: 213–282.
- Vladimircev B.Ja., Poppe N.N. 1924. Iz oblasti vokalizma mongolo-tureckogo prajazyka. – *Doklady Rossijskoj Akademii Nauk* B (aprel'-i'un'): 33–35.