Acta Orientalia Academiae Scientiarum Hung. Tomus XLVI (2–3), 339–346 (1992/93)

PRINCIPLES OF THE EASTERN CLASSIFICATION OF THE VOWELS IN CHAĠATAY DICTIONARIES

ERGASH A. UMAROV (Tashkent)

During my research into medieval Chaġatay dictionaries it turned out that they offer an impeccable classification of the Eastern Turkish vowel system. This classification seems to be more complete than those found in later works. More specifically, its peculiarity lies in the more accurate definition of the vocalic phonemes in terms of two pairs of distinctive features: long or short, and wide or narrow. The medieval authors of dictionaries of the Chaġatay gave an individual definition to each grapheme of each entry. Thus, an *alif* rendering *a* is classified in terms of length. This classification is based on the principles of the Arabic writing system.

Badā'i al luġat[1] of the 15th century is the first known Chaġatay dictionary. In order to characterize the quality of the vowels marked by the Arabic letter *alif* Ṭāli' Īmānī, the author of this dictionary, combined the terms *fatḥa* "open" and *mamdūda* "long" or *maqsur* "short". From this dictionary it is evident that *alif* with *fatḥa* marks the vowel *a* of normal length:

ارباغ (5a)[2] *ba fatḥ-i alif wa sukūn-i rā'i muhmala wa ġayn-i mu'jama afsūn-i mārrā gūyand*

arbaġ — the alif is open, there is *sukūn* over the *r*, the letter *ġ* has the dot, meaning "snake's charm".

Ṭāli' Īmānī defines long *ā* with the term *alif-i mamdūda*:

آچون (8b) *ba alif-i mamdūda wa jim-i fārsi jahānrā nāmad*

āchun — *alif* is long, *jim* is Persian (i.e. *ch*), meaning "world".

In his *Badā'i al luġat* the expression *alif-i maqsūra* "short alif" marks a reduced *a*:

ارا (5a) *ba alif-i maqṣūra ba ma'nī dar miyān*

ara — *alif* is short, meaning "between; in the middle"

[1] See A. K. Borovkov, *Badā'i al luġat*. Moskva 1962.

[2] The numbers in brackets indicate the leaves of the manuscript.

Akadémiai Kiadó, Budapest

From among the aforementioned three vowels, the normal *a* written by *alif* with *fatḥa* represents a phoneme opposed to the two allophones, the long and the reduced variants of the "marked" phoneme, as shown in the following scheme:

	+ *fatḥa* "open" = *a*
ا-(*alif*)	+*mamdūda* "long" =*ā*
	+*maqsūr* "short" =*a*

An analogous system can be seen in Mahdixan's 18th century Sanglax,[3] where the vowel *a* with normal length (i.e., short, prosodically "unmarked") is defined by the term *fatḥa*:

تیرلاماک (192r) *ba sukūn-i rā wa fatḥ-i lām ʿaraq kardan*

terlamak — there is *sukūn* over the *r*, the *lām* is with *fatḥa*, meaning "to sweat". Mahdixan marks long *ā* in two ways. He puts a *madda* over the *alif* for the initial, while in the middle of a word he uses the term *kashīda* "prolonged, held on":

آبوسون (109r) *ba madd-i alif wa ẓamm-i bā'i abjad wa sīn-i muhmala du hamshīrarā gūyand*

ābusun — *alif* with *madda*, *b* with *ẓamma*, *s* without diacritics, meaning "two sisters"

تاراماق (152v) *ba fatḥ-i rā ba alif-i kashīda shāna kardan*

tarāmaq — *r* with *fatḥa*, alif held on, meaning "to comb".

The reduced short *a* is represented in Sanglax by the letter *hā' i hauwaz*. In most cases it occurrs in the final position after the consonants *ġ* (غ), q (ق), *ch* (چ), *j* (ج), *n* (ن), *r* (ر), *m* (م), *l* (ل), *s* (س):

جیغه (218v) *ba fatḥ-i ġayn ba hā zada ba maʿnī par bāshad*

jiġa — *ġ* with *a*, meaning "feather".

The following table is to formulate the rule:

	+*madda* or *kashīda* ="long" *a*, held on
ا-(*alif*)	+*fatḥa* ="open" =*a*
	ha'i hauwaz (ه) marks reduced *a*

[3] See Sir Gerard Clauson, *Sanglax. A Persian Guide to the Turkish Language by Muhammad Mahdi Xan*. London 1960.

In the dictionary Abushqa, compiled by Mustafā bin Sadiq (its oldest copy, the Vienna manuscript was completed in 1552 A.D.), two kinds of *a* are distinguished: long *ā*, and *a* with normal length. Initial long *ā* is marked by *madda* over *alif*; it occurs in the following three entries: آت-*āt* "name", or "horse", or "throw", آتاغلیغ-*ātagliq* "respected; noted"

The vowel *a* with normal length is defined by Mustafā bin Sadiq as *maftūḥa*. This vowel is found in words اقسوم-*aqsum* "pity", ابوشقه-*abushqa* "old man", listed in chapter *ḥarf-i alif-i maftūḥa* "Words beginning with *alif*". Thus, according to the rule in Abushqa:

	+*madda* ="long" *ā*
ا (*alif*)	+*maftūḥa* =*a*

In the *Luġat-i turkī*[4] compiled in 1779 in India by Fazlullah Khan there are also two kinds of *a*: the long *ā*, and normal *a*. The long *ā* is marked by the term *madda*:

آولاماق (6b) *ba madd-i alif wa sukūn-i wāw wa fatḥ-i lām sikār kardan*

āwlamaq — at first long *ā*, and over *wāw sukūn* and *lām* with *fatḥa*, meaning "to do hunting". Normal *a* is marked by *fatḥa*:

بارماک (8b) *ba fatḥ-i bā wa sukūn-i rā'i muhmala raftan*

barmak — *b* with *fatḥa*, and *r* with *sukūn*, meaning "to walk".

Thus

	+*madda* =*ā*
ا (*alif*)	+*fatḥa* =*a*

'Ali bin Shamlu's *Luġat-i Nāvā'i*[5] completed in 1599, has the same two kinds of vowel in question. The long *ā* is rendered by *alif* with *madda* as in آتلیق -*ātliq* "horseman" and آلماق-*ālmaq* "to take", while the *a* with normal length has the *alif* with no mark, as in اتابیک-atabek "atabek" and ارلات-*arlat* "name of a Chaġatay tribe".

So the rule here is

	+*madda* =*ā*
ا (*alif*)	+with no mark =*a*

In the *Kitāb-i luġat-i atrākīya*[6] compiled in 1860 long *ā* appears marked in two ways. First, the words with an initial long *ā* are collected in the chapter

[4] MS of the British Library. I am grateful to Professor Devin DeWeese for the opportunity of using his copy of this manuscript.

[5] MS of the University of Teheran.

[6] MS of the University of St. Petersburg.

of this initial *bāb-i alif-i mamdūda*. Second, *madda* is put over the *alif*. This vowel occurs:

آندا ق (14a) *ba sukūn-i nūn ba ma'ni misl*

āndaq — *sukūn* over the *nūn*, meaning "similar"

آسراسون (14b) *ba sukūn-i sīn-i be nuqta wa ẓamm-i sīn-i digar ʿamp az nigāh dāshtan āsrasun* — *sukūn* without a dot over the *s*, the other *s* with *ẓamma* (is) the imperative form *nigāh dāshtan* meaning "to protect, guard".

The mark for the normal length is the *fatḥa*:

الاو (36b) *ba fatḥ-i awwal shul'ai ātashrā gūyand*

alav — at the beginning *a* with *fatḥa* meaning "flame"

انجه (37a) *ba fatḥ-i awwal wa sukūn-i nūn ya'nī anqadar*

anja — initial *a*, *sukūn* over the *nūn*, that is "that much".

The rule is as follows:

	$+mamdūda = ā$
ا (*alif*)	$+fatḥa = a$

Thus, according to the dictionaries of the 15th and 16th centuries Chaġatay had three kinds of vowel *a*, viz. long *ā* normal *a* (i. e., with normal length), and short reduced *a*. Only the normal *a* is a phoneme (performing the function of semantical distinction), the two other sounds being allophones.

The sounds *i*, *e*, *u* and *o* denoted by the graphemes *yā* and *wāw*, respectively, are classified in the monuments as wide vs. narrow. In Chaġatay texts *wāw* can mark both the narrow round vowel *u* and the half-narrow *o*. In order to distinguish between these two sounds the lexicographers used the combinations of *wāw* with *ẓamma* and *fatḥa*. In Ṭāli Imānī' s *Badā'i al luġat* the combination of *wāw* with *ẓamma* stands for the narrow round *u*:

کومور (75a) *ba ẓamm-i kāf-i'arabī zuġal*

kumur — *arabic k* with *ẓamma*, meaning "coal"
while *wāw* with *fatḥa* marks the half-narrow *o*:

شوشه (63a) *bil fatḥ . . . yax-i . . . shāx-i daraxt*

shosha — with *fatḥa*, meaning "ice on the twigs of trees".
This can be formulated as

	$+fatḥa = o$
و (*wāw*)	$+ẓamma = u$

The same contrast of wide vs. narrow may be observed in Mustafā bin Sadiq's description where the graphemes اوز with no mark represent the word meaning *kand* "self" (oz), but the same sequence of graphems marked *imāla* "narrow" corresponds to the word meaning "hundred" (*uz*).

This suggests that the graphemes او with no mark denote the wide round vowel, while the same symbols with the mark stay for the narrow round vowel. Similarly اوی with the mark *imāla* is to be read *uy* "cow", but it is *oy* "house; room" if inmarked.

These data show that in Abushka the graphems و ا in *anlaut* with no mark =*o*; او in *anlaut*, with the mark *imāla* =*u*.

In Mahdīxān's *Sanglax* the term *ishbā'* "narrow" is used for the *wāw* rendering the narrow round vowel *u*:

کول (308v) *ba ishbā'-i ẓamma xākistarrā gūyand* . . .

kul — marked as narrow, meaning "ash" . . .

The term *bilā ishbā'* "not narrow" shows that the vowel is the half-narrow *o*:

کول (308v) . . . *bilā ishbā' makān-i bāshad, ki āb* . . .

kol- . . . with *o*, meaning "lake" . . .

قولاووز (289v) . . . *bilā ishbā' dū ma'nī dārad, awwal dalīlu balad wa rahnamārā gūyand*

qolavuz — with no narrow, has two meanings: the first (is) "a guide", the second (is) "a road-guide".

That is:

	+*bal ishbā'* =*u*
و (*wāw*)	+*bilā ishbā'* =*o*

In Fath 'Alixan's *Kitāb-i luġat-i atrākīya* these sounds are disingushed by the aid of the symbols (and terms) *ẓamma* and *fatḥa*. The first with *wāw* marks the narrow sound:

بورون (109a) *ba ẓamm-i bā-i abjad wa rā-i qarashat ya'nī awwal wa muqadam*

burun — *b* with *ẓamma* and *r*, meanings "at the beginning" and "earlier".

A skilled lexicographer Fath 'Alixan accurately marks the half-narrow round *o* by the combination of *wāw* and *fatḥa*:

بوغماق (109b) *ba fatḥ-i bā-i abjad wa sukūn-i ġayn-i nuqtadar ya'nī xaufi kardan*

boġmaq — *b* with *fatḥa* (i.e., with *o*) *sukūn* and a dot over the *ġ*, meaning 'to strangulate".

In his dictionary he also distinguishes these round vowels according to their length. The long, half-narrow *o*: has the mark *wāw-i ma' lūm*:

تو قسان (158a) *ba fatḥ-i tā-i qarashat wa wāw-i ma'lum wa sukūn-i qāf ʿadād-i navadrā gūyand*

to : *qsan* — *t* with long *o*: and *sukūn* over *q*, meaning "ninety"

For the short vowel *u* he uses the term *wāw-i majhūl*:

تون (156b) *ba ẓamm-i tā-i qarashat wa wāw-i majhūl ba ma'nī shab*

tun — *t* with *ẓamma* and *wāw-i majhūl*, meaning "night".

These facts can be summarized in the following scheme:

	+*fatḥa* =*o*
	+*fatḥ-i wāw-i ma'lūm* =long *o*:
و (*wāw*)	+*ẓamma* =*u*
	+*wāw-i majhūl* =short *u*

The grapheme (ی) *yā* rendered both the unrounded narrow *i* and the half-narrow *e* of Chaġatay texts. To show the distinction between these sounds in the words they occurred, the lexicographers combined the name of this grapheme with those of the bound graphemes *kasra* and *fatḥa* in the descriptive part of the given entry.

In the Badā'i al luġat the narrow unrounded vowel is described by the term *bil kasr*:

بیلاک (42b) *ba kasr-i al ba wa ba fatḥ-i lām band i sā'idrā gūyand*

bilak — *b* with *kasra*, *lām* with *fatḥa*, meaning "wrist".

In the same work the grapheme *yā* with *fatḥa* describes the half-narrow *e*:

کییک (73b) *ba fatḥ-i kāf wa kasr-i yā az āhu ki dar kūh bāshad* . . .

keyik — *k* with *fatḥa* (=half-narrow *e*), *yā* with *kasra*, meaning "gazelle living in the mountains".

These facts can be formulated as follows:

	+*bil kasr* =*i*
ی (*yā*)	+*fatḥa* =*e*

To distinguish between these sounds Mahdīxān combines *yā* with the term *ishbā'* "denoting thus the vowel *i*:

بیل (149v) . . . *ba ishbā'i kasra 'amrist az dānistan*

bil — with *ishbā'* (it is) the imperative form of the (verb) "to know" Combining *yā* with the term *bilā ishbā'* he describes the half-narrow *e*:

این (117v) . . . *bilā ishbā' ba ma'nī 'arẓ wa pahna buvad* . . .

en — without narrowing (the word) meaning "width" and "wideness", "breadth".

Thus in his system

	$+bal\ ishbā' = i$
ی (*yā*)	$+bilā\ ishbā' = e$

A particular system appears in Fath 'Alixan's *Kitāb-i luġat-i atrākīya*. Here *yā* marked with the term *kasra* denotes *i*:

بیلیک (117b) *ba kasr-i bā-i abjad wa lām dānānda wa fahmrā gūyand*

bilik — *b* with *kasra*, and *lām*, meaning "knowledge" and "intellect".

The combination of *yā* with *fatḥa* stays for the half-narrow vowel *e*:

بیرکیت (119a) *ba fatḥ-i bā-i abjad wa kasr-i kāfi tāz-i 'amir az mahkam kardan*

berkit — *b* with *fatḥa*, Arabick *k* with *kasra* (is) the imperative form (of the verb meaning) "to hold firmly" (mahkam kardan). Thus

	$+bil\ kasr = i$
ی (*yā*)	$+fatḥa = e$

Old Uzbek vowels are described by 'Mir Alisher Nāvā'i in an analogous way. In his *Muḥākamat al-Luġatain* he destinguishes two kinds of the vowel *a*, wherein one with the normal length is rendered by *alif* (ا), while a short one, *a* is marked by *hā'i hauwaz* (ه). However there is also a long *ā* that appears in the script as *alif* with *madda* (آ). Thus it may be said there were three kinds of vowel *a* in Old Uzbek, the one denoted by *alif*, and only this played a real phonological role, the other two, the long *ā* and the short *a*, were allophones thereof.

As for the round vowels, the poet classifies them as long and short: تور-*tor* (with normal *o*) "snare, trap", تور-*tor* (with short vowel) "a stand on which birds roost", تور-*tor* (with shorter vowel) "the place of honour in the house", and تور-*tor* (with the shortest vowel) "roll up (imperative of *tormak*) the curtain or cloth of a tent entrance".[7]

[7] See *Mir 'Alisher Nāvāi's Works in Ten Volumes*, vol. 10 (Tashkent 1970), p. 116 (translation by A. N. Malekhova) and *Muḥākamat al Lugatain* by Mir 'Ali Shir, Introduction, Translation and Notes by Robert Devereux, Leiden 1966, p. 16. I accepted Devereux's interpretation of the passage in question.

These data give evidence of the four kinds of Chaġatay *o* distinguished by length.

Likewise the poet classifies the kinds of *i* according to their length.

"However, even for the *yā*, there are no more than three kinds of vocalic reading: بیز — is what is called by the Sarts قدود *qadud* "gland"; بیز — Persian — ما *mā*, Arabic ن — *nahn* "we"; and finally: بیز — meaning the object called by the Sarts درفش *darafsh* "awl"[8].

These data suggest that there existed three kinds of *i* differentiated according to length. The first is *i*^{e} (*bi*^{e}*z* "gland", now in the Uzbek language — *bez*), the second is *i* with normal length (*biz* "we") and the third long *i*: (*bi* : *z* "awl", now in Uzbek dialects — *biyiz*, in the standard language *bigiz*).

This how the authors of Chaġatay dictionaries classified the vowels: according to length and according to width. An advantage of this classification is the accurate distinction between the actual allophones of the phonemes that existed in Chaġatay.[9]

[8] Tashkent edition, p. 118.

[9] I would like to thank G. Kara for translating this article from Russian and David Tyson for editing the English text.