Bāla'y-balan

Benjamin Trigona-Harany

July 19, 2012

1 Introduction

Bāla'y-balan is the name of the language created by the Ottoman mystic Mehmed ibn Fethullah ibn Ebü Tâlib in the mid-16th century and described in his *Kitâb-ı Bâleybelen* [Book of Bāla'y-balan].[1] The author, who is more commonly known as "Muhyî-i Gülşenî", was born in Ottoman Edirne to a family that traced its origins to Shiraz in Persia. Most of Muhyî's life, however, was spent in Egypt, where he was an active member of the Gülşenî sufi order.[2] Muhyî's mysticism was the motivation for the creation of Bāla'y-balan, which he envisioned as a language of devotion rather than, say, a replacement for the languages spoken in the Ottoman Empire.

Presently, the only monograph on Bāla'y-balan is *Bâleybelen: İlk Yapma Dil* by Mustafa Koç (Istanbul: 2005). Besides providing a transcription of all the Bāla'y-balan texts, Koç includes orignal research on Muhyî's life, the Gülşenî order, the creation of Bāla'y-balan and references to it in contemporary texts. Koç also provides his own summary of the language's grammar and has compiled the glosses from the orignal manuscripts into a 10,000-word Bāla'y-balan–Turkish lexicon and its Turkish–Bāla'y-balan reverse.

The name Bāla'y-balan is formed from the noun **bāl** 'language' and the infinitive **balam** 'to give life; to revive'. The linking suffix **-a** is attached to the governing noun **bāl**, giving **bāla**. The definite article **y-** precedes the goverened noun **balan**, which is formed from the verb stem **bal-** plus the active participle suffix **-an**, creating the word 'the one who gives life; the reviver'. As a whole Bāla'y-balan translates as 'the language of the one who gives life' or simply 'the language of the reviver'. The Arabic word *muhyî* (deriving from the Arabic root حيّ) also translates as **balan**, so an alternate reading of Bāla'y-balan could be 'the

[1] Bāla'y-balan transcribed into the Latin alphabet will appear in bold face. Arabic, Persian and Turkish will all appear in italics using the standard transcription system for Ottoman. Square brackets indicate IPA notation. English equivalents for these words will appear in single quotes when it is necessary to distinguish them from the rest of the text: **ẓafam** [zafam] *yazmak* 'to write'.

[2] For more on the Gülşenî, see John Curry's *Transforming Muslim mystical thought in the Ottoman Empire*.

language of Muhyî'.

2 Phonology and Orthography

Bāla'y-balan is written using the modified Arabic script also employed for Ottoman Turkish and Persian. Because Muhyî did not describe how his language was to be pronounced in *Kitâb-ı Bâleybelen*, we must rely uniquely on orthography, and it is, therefore, not possible to reconstruct the ponetics of Bāla'y-balan with complete certainty. However, I have done my best to determine how he *probably* intended Bāla'y-balan to be pronouced by evaluating Muhyî's decisions with respect to his language's orthography. The following section describes how I came up with my proposed phonology of Bāla'y-balan; readers not interested in the rationale for my decision can skip directly to the section on the Bāla'y-balan alphabet.

2.1 Ottoman Turkish Orthography

Orthographically, Bāla'y-balan is heavily influenced by Ottoman Turkish, and the modified Arabic alphabet used to write it.[3] While Arabic orthography distinguishes between three short and three long vowels, Turkish has four back vowels (*a*, *ı*, *o*, *u*) and four front vowels (*e*, *i*, *ö*, *ü*). To accommodate these extra sounds, Ottoman Turkish makes use of the fact that Turkish phonology does not distinguish between certain consonant pairs found in Arabic; the letters س and ص, for example, are both pronounced [s] in Turkish. When appearing in Turkish words, however, the letters show whether surrounding vowels should be pronounced in their front or back form. Thus, the conditional suffixes *-sa* and *-se* were written using the Arabic consonants ص and س respectively, where only these letters indicate how the following vowel should be pronounced.[4] The consonant pairs which serve this function in Ottoman Turkish were ط-د, ط-ت and ص-س. Two more,

[3] The same changes to the standard Arabic script are also present in Persian.

[4] An additional Arabic consonant, ث, is also pronounced [s] in Turkish but appears only in Arabic loan words and does not influence the pronunciation of surrounding vowels.

ك-ق and گ-غ, indicate a pronunciation difference in the consonant in addition to the vowel: for example, ك is [c] while ق is [k]. The words اولمك and اولمق differ only in their final consonant but are pronounced *ölmek* and *olmak* respectively.

There are additionally numerous Arabic letters which are never used to write Turkish words and whose pronunciation is identical to a native Turkish consonant. For example, the letters ح, خ and ه are all pronounced [h], while only the latter appears in words of Turkish origin. Similarly, the letters ز, ذ, ض and ظ are all realised [z] in Turkish. A number of these consonants also influence the pronunciation of neighbouring vowels, but since these only appear in loan words these sound changes do not have any systematic semantic meaning.

Muhyî appears to have explicitly avoided using consonants which are not native to the Turkish language in function words and affixes (which are used extensively in Bāla'y-balan for syntactic purposes). Consonants which are used in Ottoman Turkish orthography only to mark vowel pronunciation are used, but only in lexical words (nouns, verbs, adjectives and adverbs). Certain consonants, namely ذ, ض and ظ, are rarely encountered at all in the language.[5] The purpose of eliminating these letters from grammatical suffixes an function words appears to be to either mimic Turkish usage or to eliminate confusion between phonemes due to their identical pronunciation. If these letters were given their full Arabic values, there would be no reason to restrict their usage in Bāla'y-balan.

Further evidence comes in the letter names Muhyî created for Bāla'y-balan since the standard Arabic names of the letters would have been ambiguous: حاء, خاء and هاء are all *hâ* in Turkish. Assuming that all three letters were pronounced identically in Bāla'y-balan, each would need a distinct name, which is exactly what Muhyî did: [ħa:], [xav] and [hir] respectively. The same is done for all other consonant groups ([d], [s], [t] and [z]) where confusion would have occurred. Note that since Turkish distinguishes between [c] and [k], the letters ك and ق do not require this distinction and are known as **qin** and **kin** respectively.[6]

It is also reasonable to assume that Muhyî would select consonants which are

[5] These letters are also some of the rarest in Arabic.

[6] This distinction is not reflected in the spelling system of modern Turkish, in which both are *k*. Azeri Turkish, and other Turkic languages of the former Soviet Union that have adopted the Latin alphabet, do make use of *q* to distinguish the two sounds in writing.

more readibly producable by speakers of Persian, Turkish and other languages which do not share the full range of Arabic consonantal phonology. Many of the sounds which I argue have been excluded from Bāla'y-balan would not even have been pronounced identically in the various dialects of Arabic spoken in the Ottoman Empire.

Since Muhyî has included consonants which, in Ottoman Turkish, influence the pronunciation of vowels in the same word, the second question is whether Bāla'y-balan has all eight vowels of Ottoman or the same three present in Arabic. In my transcription, I have opted to only use the short [a], [u] and [i] and long [aː], [uː] and [iː] that can be definitively elicited from Muhyî's writing. I have done this with the caveat that he may have intended the short vowels in Bāla'y-balan to be influenced by neighbouring consonants. Under normal circumstances, the Arabic short vowels [a], [u] and [i] should be read [e], [i] and [y] in Ottoman; these usually change to [a], [ɯ] and [u] under the influence of the following consonants: ح, خ, ص, ض, ط, ظ, ع, غ and ق. The sounds [o] and [ø] are usually only present in the initial syllable of Turkish words, and I have not used them in my Bāla'y-balan phonology.

Words in Bāla'y-balan are not required to have vowel harmony, a distinguishing feature of the Turkish language.[7] Words such as **qaydak** *kapı* 'door, gate', which has both a back **q** and a front **k**, would not be possible in a pure Turkish word. This point has further significance when we consider Bāla'y-balan's suffixes, which will not harmonise with the vowels in the root to which they are attached.

The Alphabet

Because Bāla'y-balan's phonetics are uniquely tied to its orthography, the latter will be presented first. There are 34 letters proper used in Bāla'y-balan, with **amam** not usually considered a letter in its own right. One letter, **anā** is used as a vowel marker only, and two more, **var** and **yam**, can be used as either consonants or vowels.

[7] The definite article **y-** does in fact harmonise with the following word, but this is unlike the harmony found in Turkish. See section 4.2.

In the transcription, a macron over a vowel indicates lengthening. Other letters with diacritics may or may not represent a pronunciation difference between them and their base form; refer to the IPA column for guidance on pronunciation.

Each letter in the Arabic alphabet may have up to four different forms depending on its position in the word (initial, medial, final or isolated). Only the isolated form is presented in the table below. Letters in Bāla'y-balan's alphabet also have numerical values which are used to write numbers, as explained in section 5.

Transcription	Name	Arabic	IPA
ā	anā	ا	[aː]
ʾ	amam	ء	[ʔ]
b	bar	ب	[b]
p	pī	پ	[p]
t	tir	ت	[t]
s	sam	ث	[s]
c	cā	ج	[dʒ]
ç	çī	چ	[tʃ]
ḥ	ḥā	ح	[h]
x	xav	خ	[h]
d	dav	د	[d]
ẕ	ẕav	ذ	[z]
r	ran	ر	[ɾ]
z	zan	ز	[z]
j	jī	ژ	[ʒ]
s	sir	س	[s]
ş	şir	ش	[ʃ]
ṣ	ṣad	ص	[s]
ż	żad	ض	[d]
ṭ	ṭī	ط	[t]
ẓ	ẓī	ظ	[z]
ʿ	ʿī	ع	[ʔ]

ġ	ġī	غ	[g]
f	fin	ف	[f]
q	qin	ق	[k]
k	kin	ك	[c]
g	gī	گ	[ɟ]
ñ	gim	ڭ	[ɲ]
l	lā	ل	[l] [ɬ]
m	mā	م	[m]
n	nā	ن	[n]
ū, v	var	و	[uː] [ʊ]
ä, h	hir	ه	[a] [h]
ī, y	yam	ى	[iː] [j]

Table 1: Bāla'y-balan alphabet

There are several other additional orthographic devices used in Bāla'y-balan. In the table, diacritics are shown in conjunction with the letter ف, except **qam**, which may only appear above the letter ا.

Transcription	Name	Arabic	IPA
ā	qam	آ	[aː]
-	qaqnam	فّ	-
-	ʿazam	فْ	-
a	kavam	فَ	[a]
u	cumam	فُ	[u]
i	ṣiyam	فِ	[i]

Table 2: Bāla'y-balan diacritics

Notes

1. In word-initial position, [aː] is represented by ا with the horizontal **qam** diacritic: آ. In medial and word-final positions, **ā** is written with a plain ا and a **kavam** diacritic above the preceding consonant.

2. In word-initial position, [uː] is represented by the digraph اُو. In medial and word-final positions, **ū** is written with a plain و and a **cumam** diacritic above the preceding consonant.

3. In word-initial position, [iː] is represented by the digraph اِی. In medial and word-final positions, **ī** is written with a plain ی and a **ṣiyam** diacritic below the preceding consonant.

4. Bāla'y-balan's short vowels, [a], [u] and [i], are indicated only through the use of the three diacritics, **kavam**, **cumam** and **ṣiyam**. In word-initial position, these diacritics are placed above or below the letter ا. In medial position, the diacritics are placed above or below the consonant that precedes the vowel.[8] In figure 1, **kavam** is used to indicate long and short vowels.[9] ذَاتَ يوَكَشَا و حَاتَ يَيْكَشَا **ẕāta'ya-vkaşā** *ve* **ḥāta'ya-bkaşā**.[10] Word-final [a] is indicated by a ه and will be transcribed here by **ä**. Short [u] and [i] do not occur word-finally in Bāla'y-balan. A suffix which consist of a single short vowel are marked by the addition of a diacritic as in the case of the **-a** suffix on **ẕāt** and **ḥāt** in the example above.

5. The **qaqnam** diacritic indicates that a consonant should be doubled; except in the case of affixation, Bāla'y-balan does not allow two identical consonants to be written side-by-side.

[8] In Arabic and Turkish, vowel diacritics are usually omitted, but Muhyî was diligent about providing diacritics when writing Bāla'y-balan.

[9] Note that in *Kitâb-ı Bâleybelen*, Muhyî wrote Bāla'y-balan in red ink and Ottoman Turkish in black. Diacritic marks for Bāla'y-balan are, nevertheless, written in black.

[10] Koç transcribes these as *ẕâte'y-vekeşâ ve ḥâte'y-bekeşâ*, which shows a different reading of the definite article and the first syllable in each of the following nouns. The text, especially in the second of the two compounds, would seem to support my reading. See the discussion of ˚**azam** in point 6 and section 4.2.

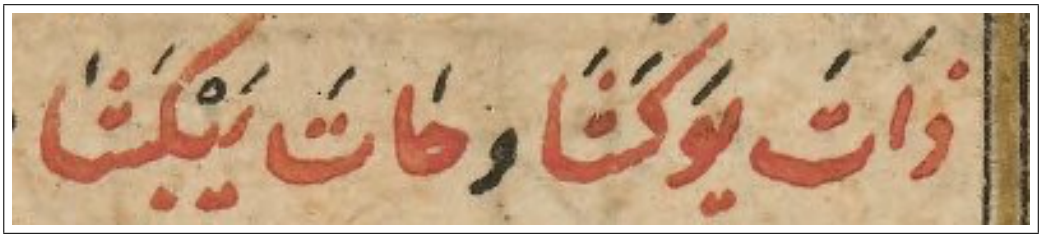

Figure 1: **ẕāta'ya-vkaşā** ve **ḥāta'ya-bkaşā**

6. The ˋ**azam** diacritic is used to indicate that a consonant has no following vowel.

3 Morphology

4 Syntax

4.1 Coordinating Suffix

4.2 Definite Article

The definite article **y** is placed before the noun that it marks. In Bāla'y-balan the article is attached directly to the noun, while in transcription, a dash will be used to separate the two.

There is disagreement on how the article should be formed in certain instances. According to Koç, the article is **y** when following a word ending in a vowel or preceding a noun beginning in a vowel: **y-anā**. Elsewhere, the article requires a buffer vowel between it and its noun: **ya-bar**.

The original manuscript, however, suggests that the rules are somewhat more complex. The main differences is that the article absorbs the first vowel from nouns with an open first syllable. The steps for selecting the appropriate form for an article in Bāla'y-balan can be summarised as:

1. if the noun starts with a vowel, no buffer is necessary
2. if there is *not* a vowel immediately preceding the article, use a buffer vowel
3. if the noun consists of a single syllable, no buffer is necessary
4. if the noun starts with a closed syllable, no buffer is necessary
5. use a buffer vowel

If the buffer vowel is necessary, the first syllable of the noun loses a short vowel, which will have been represented by a diacritic, and that same diacritic is placed over the **y**-, giving it the form **ya**-, **yu**- or **yi**-.[11] *In the case of a long vowel, the diacritic is reduplicated in both the article, which will have a short vowel, and the first syllable of the noun, which will retain its long vowel.*

The following examples from a single passage in *Kitâb-ı Bâleybelen* demonstrate the application of these rules.[12] Koç does not follow this pattern, preferring **y**- in all cases except **ya'ḫşenâ**, but no explanation is given for this deviation.[13]

1. رَگِوْزَاوَ يُنْشَا
   ra-givzāva'yu-nşā
2. فَمِيمَ يْمَفْنَا
   fā-mīmā'y-mafnā
3. اَ يَنْفَمْ
   a ya-nfam
4. جَامَ يْنَنْشَنَا اَ يَخْشَنَا
   cāma'y-nanşanā a ya'xşanā

[11]Note that the first syllable in **balan** is closed, so ***bāla'ya-blan** is not the correct form of Bāla'y-balan.

[12]All diacritics are from the original: *Kitâb-ı Bâleybelen.*

[13]Page 188.

5. رَ يْكَرْفَنَا رَايَا يَعْشَنَا
rāʾy-karfanā rāyā ya-ʿşanā

5 Numerals

5.1 Cardinals

Numbers in Bāla'y-balan are based on the decimal system as in Arabic, Persian and Turkish.

1	**ad**	10	**yā**
2	**baz**	20	**kar**
3	**cil**	30	**liş**
4	**dum**	40	**mud**[14]
5	**hin**	50	**nih**
6	**vus**	60	**sav**[15]
7	**zā**	70	**ʿaz**
8	**ḥī**	80	**fī**
9	**hid**	90	**ṭay**[16]

0	**baḥlā**	100	**qā**	1000	**ġan**

Table 3: Bāla'y-balan numbers

The 10 digits can be combined to create larger numbers although they are not always arrange in the order of most to least significance. For numbers between 11 and 101, the numbers 1 through 9 precede the numbers 10 through 100: **hidyā** 'nine-ten'; **bazkar** 'two-twenty'. Compare these to: **qā adyā** 'hundred-one-ten'; **qā dumnih** 'hundred-four-fifty'. For the number 101, both orderings are possible: **adqā** = **qaʾad** 'one-hundred'. The reversed form may be used when using multiples of 10 plus 100: **qabaz** = **bazqā** 'twenty-hundred'.[17]

Most digits have shortened forms used in compounds, fractions and ordinals that are created by dropping the final consonant. Shortened forms are possible

[17]Note that **qā** can lose its long vowel in compounds.

for the numbers 11 through 17: **adyā** = **ayā** 11; **bazyā** = **bayā** 12; **cilyā** = **ciyā** 13; **dumyā** = **duyā** 14; **hinyā** = **hiyā** 15; **vusyā** = **vuyā** 17; **zāyā** = **zayā** 1; **ḥīyā** 18; **hidyā** 19. This pattern repeats itself for 21 through 27, 31 through 37 and so on. In the rare case, of inversion and multiples of 10, reduced forms are also possible: **yaqā** 110; **kaqā** 120.

Hundreds are formed by adding the multiplier before **qā** 100: **cilqā** 'three-hundred'; **vusqā zāliş** 'six-hundred-seven-thirty'. Again, shortened forms are possible for the hundreds, although in this case from 200 to 800: **bazqā** = **baqā** 200; **cilqā** = **ciqā** 300; **dumqā** = **duqā** 400; **hinqā** = **hiqā** 500; **vusqā** = **vuqā** 600; **zāqā** = **zaqā** 700; **ḥīqā** = **ḥaqā** 800; **hidqā** 900.[18]

Thousands are formed in the same manner as the hundreds, and shortened forms : **cilġan** = **ciġan** 'three-thousand'; **vusġan ḥīqā hifī** 'six-thousand-eight-hundred-five-eighty'. Thousands and hundreds may be combined for larger numbers: **dukar ġan qā ciṭay** 'four-twenty-thousand-hundred-three-ninety' 24,193. A shortened form, **ġa**, may be used in the place of **ġan**: **ġanad** = **ġaʾad** 1,001; **ġanbaz** = **ġabaz** 1,002.

Other numeral orders are possible when they do not produce ambiguity: **ciyā qaġan** 'three-ten-hundred-one thousand' 1,113. Contrast this with **qā ciyā ġan** 'hundred-three-ten-thousand' 113,000.

5.2 Fractions

Fractions are formed in several different manners. The standard means is to separate the numerator from the denominator with the prefixes **da** or **fa**: **ad facil** '⅓'. It is also possible to omit the prefix in thes constructs and use simple juxtaposition: **adcil**. The shortened digits may also be used in the numerator: **acil**. Finally, it is also possible to reverse the order of the numerator and denominator: **cilad**.

[18]The number **hid** has no reduced form in Bāla'y-balan.

5.3 Ordinals

Ordinals can be regularly formed with the **-(y)am** suffix attached to the last element in the number: **adam** ‘first’; **yāyam** ‘tenth’; **qāyam** ‘hundredth’. With certain single digit ordinals, the sufix **-am** can optionally be fixed to the first syllable of the number; others have irregular forms: **ad + am > am**[19]; **baz + am > bam**; **cil + am > cam**; **dum + am > dam**; **hin + am > ham**; **vus + am > vam**; **zā + am > zam**; **hī + am > ham**; **hid + am > hidam**.

5.4 Distributives

Distributive nouns can be formed in Bāla’y-balan by replacing the vowel of the last syllable with **ā**: **ād** ‘one each’; **ġān** ‘a thousand each’.

Distributive adjecties are formed in the same manner with an additional suffix **i**.[20] A **y** is inserted between a word-final vowel and the suffix **i**: يَایَ اَوَرْ **yāyi avar** ‘a hundred horses each’.

Using reduplication, adverbial constructs such as **bazābaz** ‘two-by-two’ can be formed. Note that when the numeral ends in a long vowel, the first loses its length and a buffer **v** is used: **qā + v + qā > qavāqā** ‘a hundred at a time’.

5.5 *Abjad*

Each base letter[21] of the Bāla’y-balan alphabet has a numerical value which is used to write numbers. The system, in Arabic *abjad*, completely replaces the standard Indic digits which are regularly used in languages which use the Arabic alphabet.[22]

In *Kitâb-ı Bâleybelen*, Muhyî also presented variant characters to replace the standard letters for writing Bāla’y-balan’s numbers. There is a one-to-one re-

[19]Note that vowel-initial words are considered to start with a glottal stop represented by the letter **anā**, so combining اَد and ـَم produces اَم.

[20]The suffix is written with a **ṣiyam** diacritic.

[21]The base letters are only those found in the Arabic alphabet and exclude the additional forms used for Persian and Ottoman Turkish. Persian letters take the value of the original letter.

[22]The Indic numbers are: ٥, ١, ٢, ٣, ٤, ٥, ٦, ٧, ٨, ٩.

lationship between the new characters and the standard Arabic letters, so it is possible that these symbols could have additionally been used to write Bāla'y-balan.[23] Bāla'y-balan's numerical symbols are in fact partially copies of Arabic's Indic digits (seven standard forms and two variants used in Persian: ۴ and ۵) and one letter (ظ). The remainder are Muhyî's own invention. There is some precedence for alphabet creation at the time; Koç mentions that a contemporary sufi is said to have written a "strange tongue" in an "unusual alphabet".[24] The work in question, however, is lost, and the new characters in *Kitāb-ı Bâleybelen* are only used numerically.

ا	ب	ج	د	ه	و	ز	ح	ط	
١	٢	٣		٥	٦			٩	
1	2	3	4	5	6	7	8	9	
ى	ك	ل	م	ن	س	ع	ف	ص	
						٧	٨		
10	20	30	40	50	60	70	80	90	
ق	ر	ش	ت	ث	خ	ذ	ض	ظ	غ
			۴	۵				ط	
100	200	300	400	500	600	700	800	900	1000

Table 4: *Abjad* values in Bāla'y-balan

Numbers 1999 and below are formed from the *Abjad* symbols by arranging them in a descending order: هشغ (1000-300-5) = 1305.[25] It should be noted that despite the Arabic alphabet being read right-to-left, the highest order digits in Bāla'y-balan numbers still appear to the left of lower order numbers.

For numbers in the thousands, Bāla'y-balan suffers from the fact that غ always equals 1000 and is not used as a thousand marker: جنظغسر (200-60-1000-900-

[23]Table 4 is currently incomplete as many glyphs are not available. Consequently, I will use the Arabic alphabet in all examples.

[24]See Koç, *Bâleybelen*, 53-54.

[25]Under normal circumstances the Arabic letters should be joined as normal; however the symbols Muhyî devised for his numbering system do *not* connect, so they will appear as such in this document.

50-3) = 261,953. Muhyî was aware of this difficulty and announced that "200,000 is written 200 and 60,000 is written 60", with context differentiating between the number which are and are not in the thousands. He does not, however, provide the means of writing numbers such as 2,000; the most reasonable assumption would be غب (2-1000), but this is not stated.

6 Glossary