A Comparative Analysis of Hungarian Folk Songs and Sanskrit Literature Using Motif Similarity Matrices

PETER Z. REVESZ
Department of Computer Science and Engineering
University of Nebraska-Lincoln
Lincoln, NE 68588-0115
USA
revesz@cse.unl.edu, http://cse.unl.edu/~revesz/

Abstract: - This paper shows connections between the texts of Hungarian folk songs and ancient Sanskrit literature that both have been transmitted orally for a long time before being recorded in writing. The comparative analysis is based on a set of ten important motifs that are shared among the studied texts. We build a similarity measure between any pair of texts based on the number of their common motifs. The surprisingly close similarities found suggest extensive contacts between Proto-Hungarian and Proto-Indo-Iranian people. The paper also investigates the chronology of these contacts using word similarities and places the last phase of them after the dissolution of the West-Ugric branch of the Uralic language family.

Key-Words: - computational linguistics, data mining, folk songs, similarity matrix, textual analysis, Sanskrit

1 Introduction

People have a remarkable tendency to repeat songs without much attention to the exact meaning of their words. That remarkable tendency is probably due to two reasons. First, the primary goal of singing is to evoke an emotional effect rather than to convey some knowledge. This emotional effect can be achieved by a somewhat cryptic text with the same ease as a clearer text. Second, changing the text of a song is difficult while preserving the rhythm and the melody of the song. Hence the text of folk songs can preserve some phrases even more precisely than folk tales and folk ballads can. Nevertheless, it is usually difficult to tell whether a particular phrase in an individual folk song is ancient. Distinguishing between ancient and later phrases becomes easier when an original song survives in several versions because the versions can be compared with one another. In general, the commonalities can be presumed to be ancient, while the disparities can be presumed to be later additions.

We can use textual comparison in combination with word etymology developed by linguists and music history developed by musicologists to deduce the hypothetical original forms of folk songs. While such reconstructions cannot be expected to be precise, the reconstructed forms can be assumed to be close to the true originals. Moreover, the reconstructed forms allow an interpretation of the original songs in terms of pagan mythology and religion.

The rest of this paper is organized as follows. Section 2 describes the data sources, including references to the studied motifs and folk songs. Section 3 presents the results of our analysis of the folk songs in terms of the motifs. Section 4 considers linguistic borrowings between Indo-European and Uralic languages. These borrowings corroborate the close relationship that existed between Proto-Hungarian and Proto-Indo-Iranian people. Finally, Section 5 gives some conclusions and directions for future work.

2 Data Sources

Section 2.1 describes the peacock and related motifs that we found by consulting the literature on peacocks in the myths and art of India. Section 2.2 presents four Hungarian folk songs with English translations that contain some peacock motif.

2.1 Peacock and related motifs

Our literature search of the role of peacocks and related motifs in the art and myths of India and neighboring cultures reveals that these motifs can be categorized into ten groups as follows:

1. The peacock is a *god* or symbol of a god.
   The name of the god is:
   a. Murugan's mount *Paravani* (Tamil)
   b. *Tammuz* (Sumerian)
   c. Malik-i-*Ta'us* (Yezidi)
   d. *Tari* (Khond)

2. The peacock is the *ancestor* of a people and/or the *name* of a people derives from the peacock.
   a. *Mayura, Mayurbhanja* (India)
   b. *Magyar* (Hungarian)

3. The peacock is a *lover*.

4. The peacock is a *carrier of souls* of the dead.

5. The name of heaven is:
   a. Abode of the 33 devas (Hindu)
   b. Trayastriṃśa (Buddhist)

6. The name of some important gods:
   a. Yama, god of death (Hindu/Buddhist)
   b. Yima, god of death (Zoroastrianism)
   c. Juma, god of sky (Finno-Permic)
   d. Juno, wife of sky god Jove (Latin)

7. Fire is associated with marital love.

8. Fire is extinguished at death.

9. A horse is sacrificed at full moon.

10. The horse is an object of love.

Motif 1 occurs in southern India among the Tamils who consider their supreme god, Murugan, a deified peacock according to Nair [15]. In Sumer, a peacock was associated with Tammuz, a life-death-rebirth deity. Finally, this motif also occurs among the Yezidis whose high god is Malik-i-Ta'us, which can be translated as "king peacock" (see Lal [12]). It can also be found among the Khonds, a tribal people in India, who represent the Earth-Goddess, called Tari, as a peacock (see Robertson [25]).

Motif 2 occurred in Mayurbhanja, India, where the chiefs claimed their lineage from a peacock [12]. The Sanskrit name of peacock is *mayura*. The name *mayura* may have some connection with *magyar*, which is the self-name of the Hungarians. This self-name originally referred to only the tribe of Hungarians from which came the House of Árpád, an early medieval ruling dynasty. In Hungarian myths, the ancestor of the Hungarians is a magical bird called *Turul*, which seems to derive from a Turkish name for hawk [34]. It is possible that the name of *mayura* was replaced by *turul*

under Turkish influence in the ancestor myths, while the self-designation modified only slightly to *magyar*. That is a reasonable assumption given that whatever name people had, they usually were assumed to have had an eponymous ancestor.

Motif 3 occurs in a folk song in Gujarat, where the wife of a prince falls in love with a peacock (Lal [12], p. 32).

Motif 4 was apparently believed in the Indus Valley Civilization because a peacock is shown carrying a dead body in its stomach on a Harappan pottery [12].

In addition to the four peacock motifs, we found some other interesting frequently occurring motifs.

Motif 5 refers to the second heaven in Buddhist cosmology, called *Trāyastriṃśa*, which derives from an adjective formed from the numeral 33 and can be translated in English as "belonging to the thirty-three [devas]" described in ancient Vedic literature [32].

Motif 6, *Yama*, the god of death of Hindu and Buddhist religion, may have some connection to the common Finno-Ugric name of the chief god, Jumala, according to Steblin-Kamenskij [26].

The next two are fire and horse related motifs that we also found in various Hungarian folk songs and ancient Sanskrit texts, as we will see in the examples in Section 2.2.

2.2 Hungarian folk songs and Sanskrit texts
For our comparative analysis of motifs, we selected nine texts as shown in Table 1. Six of the texts were those of Hungarian folk songs and three were ancient Sanskrit sacred texts. The first four folk songs are from the collection of Kodály [10], who gave their numberings, which we also use, while folk songs D and M are two other popular songs. Since the subject of the analysis in this paper is only the text of these folk songs, we omit their scores, which can be looked up in Kodály's book or other collections. On the right side of Table 1, we also give our English translations of each text, except for the Aupasana Agni for which we give a short English summary instead of a translation.

3 A Comparative Textual Analysis

3.1 Identification of common motifs
We give an interpretation of each text and identify the motifs in them. It needs to be remembered that these songs and hymns spread by oral traditions that kept modifying the text over centuries.

Table 1. The texts selected for a comparative analysis of motifs.

Folk Song 1 (Kodály [10])	English Translation
Leszállott a páva vármegye házára. De nem ám a rabok szabadulására.	The peacock flew down to the courthouse. But not to free the prisoners.

Folk Song 95 (Kodály [10])	English Translation
Szomorú fűzfának harminchárom ága Arra reá szállott harminchárom páva Ki zöldbe, ki kékbe, ki földig fehérbe, Csak az én édesem tiszta feketébe. Szólítottam volna, szántam búsítani Egy ilyen efiát megszomorítani.	A weeping willow tree has thirty-three branches Thirty-three peacocks flew onto it Some are dressed in green, some in blue, some in white to the ground Only my beloved is completely black I would have called him, but I felt sorry to tell him And sadden such a man.

Folk Song 98 (Kodály [10]) and alternative 3rd line (*3)	English Translation
Megkötöm lovamat szomorú fűzfához Lehajtom fejemet két első lábához. Lehajtom fejemet a babám ölébe Hullajtom könnyeim rózsás kötényébe. (*3) Lovamat eloldom mikor a hold felkel	I tie my horse to a weeping willow tree. I bend my head to its two front feet. I bend my head to my lover's lap. I shed my tears into her roses-embroidered apron I untie my horse when the moon rises

Folk Song 101 (Kodály [10])	English Translation
Leszállott a páva vármegye házára, Szájába visz vizet a rabok számára.	The peacock flew down to the courthouse. In its beak takes water for the prisoners.

Folk Song D	English Translation
Zöld erdőben, zöld mezőben sétálgat egy páva Kék a lába, zöld a szárnya, aranyos a tolla. Hívott engem útitársnak, el is megyek véle, Vásárhelyen nincsen olyan legény, sej, aki nekem kéne. Édesanyám, kedves anyám, menjen az erdőbe, Keresse meg azt a pávát, kérdezze meg tőle, Nem látta-e a kedvesem abban az erdőben, Kiért fáj a, fáj a gyönge szívem, sej meg is halok érte.	In a green forest, on a green meadow walks a peacock He has blue legs, green wings, and golden feathers He called me as a travel companion. I'll go with him I cannot find any lad whom I like at Vásárhely Mother, my mother, go to the forest Find that peacock and ask him Whether he saw my beloved in that forest For whom my weak heart aches, hej, I will die for her

Folk Song M (collected by A. Péczely)	English Translation
Megrakják a tüzet, mégis elaluszik. Nincs az a szerelem, ami el nem múlik. Rakd meg, babám, rakd meg, lobogó tüzedet, Hadd melegítsem meg, gyönge kezeimet. Szól a kakukkmadár, talán megvirrad már. Isten veled, rózsám, magad maradsz immár.	They set well the fire, but it still burns out. There is no love, which does not pass. Kindle your fire, my love, kindle your fire Let me warm my tender hands. The cuckoo sings, it already dawns, Good bye, my rose, you'll remain alone.

Ashvamedha	English Translation
Harivamsa Purana, Bhavishya Parva 3.5.11-12 http://mahabharata-resources.org/harivamsa/bhavishyaparva/hv_3_005.html see also Rig Veda book 1, hymn 162 [6]	"After the passage of some time, king Janamejaya, who offers plenty of tributes (in sacrifices) observed the horse sacrifice as ordained. Devi Vapushtama, the daughter of the king of Kashi, went and slept with the slain horse, according to the ritual as prescribed."

Aupasana Agni	Description [33]
See [33]	"The Aupasana Agni lit at the time of the grooms wedding … These two fires are to be preserved throughout the individual's life. … At the time of the individual's demise, cremation is done with the fires that have been preserved during his lifetime, and then the deceased individual's fires are extinguished."

Vena Bird	English Translation (Griffith [6])
Rig Veda book 10, hymn 123 [6]	"Performing sacrifice they reached the river: for the Gandharva found the immortal waters. The Apsaras, the Lady, sweetly smiling, supports her Lover in sublimest heaven. In his Friend's dwelling as a Friend he wanders: he, Vena, rests him on his golden pinion. They gaze on thee with longing in their spirit, as on a strong-winged bird that mounteth skyward; On thee with wings of gold, Varuna's envoy, the Bird that hasteneth to the home of Yama."

In addition, references to pagan religious concepts were veiled in a Christian culture. Many times the veiled references became unintelligible to the singers too. Therefore, many of the original details were lost.

Folk song 1 in its present form is extremely cryptic. It helps to remember the historical context that prisoners were often executed. Therefore, when the peacock appears, then it appears to await the execution of the prisoners and then to carry their souls to heaven. Such an appearance of the peacock is not for the physical release of the prisoners, which is the reason the song says that the peacock did not come to free the prisoners. The text of this song as recorded by Kodály [9] seems more original than the apparently later versions where the first line urges the peacock to fly, or second line claims that the peacock came to free the prisoners. The latter small change in text with opposite meaning could have occurred at some revolutionary period of history, when the people's sympathies were with the prisoners. In summary, Folk song 1 contains peacock motifs 1 and 4.

Folk song 95 is about a wife dying. As she dies, her soul is taken to the heaven by peacocks. The thirty-three references the thirty-three divines of Hindu mythology, which are mentioned in the Rig Veda, and from which *Trayastriṃśa*, the name of the second heaven of Buddhist mythology, derives [32]. As her soul flies over the village where she lives, she sees the people dressed in various colors except her husband, who is dressed in black as a sign of mourning. She would like to tell him something, but she is afraid to scare him. In summary, Folk song 95 contains the following motifs: 1 because peacocks symbolize the thirty-three divines, 4 because peacocks carry the wife to heaven, 5 because the number thirty-three also references heaven.

Folk song 98 is about a horse that is tied to a weeping willow tree. As we already saw in Folk song 95, the weeping willow tree is a signal of approaching death. Hence tying the horse to a weeping willow tree implies that the horse is sacrificed. As the alternative third line, which survives in some regions, suggests, the horse is untied upon death, which takes place when the moon rises. The song also suggests a physical attraction to the horse in the line "I bend my head to its two front feet." In the following two lines, the song seems to abruptly jump to the singer's human lover. However, these lines may have changed. The alternative third line seems to be the original third line because it does not give an abrupt transition to a human lover. Therefore, Folk song 98 contains motifs 9 and 10.

Folk song 101 is clearly just a corrupted version of Folk song 1. The original understanding of what the peacock symbolizes was completely lost. Hence a new function of the peacock was found, that is, to carry water for the prisoners. Rational thinking would disallow this as a good reason. First, peacocks cannot carry water in their beaks, especially enough for a group of prisoners. Second, there surely were some mean wardens in the past. However, any warden who prematurely killed the prisoners by not giving them water would have incurred the wrath of the authorities. The authorities intended the prisoners' executions to be public events as a deterrent to the people. The premature death of the prisoners would have prevented that. Therefore, Folk song 101 can be assumed to be a corrupted version of Folk song 1 and originally also containing motifs 1 and 4.

Folk song D is about a woman falling in love with a peacock. The peacock is very special, far better than anyone she could find in her village. The peacock asks the woman to come with him, and the woman agrees. The earthy lover of the woman misses her and asks his mother to inquire from the peacock about the woman. He is so heartbroken that he almost dies from sorrow. In summary, Folk song D contains the following motifs: 1 is hinted at because of the magic qualities of the peacock, 3 because the woman falls in love with the peacock, and 4 because the journey's destination is heaven, that is, the woman has to die to accompany the peacock on their journey.

Folk song M is about a woman kindling a fire, which is extinguished at the death of her lover. The death is implied by the arrival of a cuckoo, but presumably the cuckoo may have been originally a peacock. The peacock implies the arrival of death as in Folk song 95. Folk song M contains the following motifs: 4 because a bird, originally likely a peacock, implies the arrival of death, 7 because fire is associated with love, and 8 because it is extinguished at the end of love and death.

The *Ashvamedha* ritual is a horse sacrifice that is described in the Rig Veda in book 1, hymn 162. It is elaborated in the Mahabharata, from which two lines are translated in Table 1. The first translated line makes clear that the horse is sacrificed, while the second implies a physical attraction between a human and the horse. In addition, the Ashvamedha Parva mentions that the horse sacrifice by King Yudhisthira takes place at full moon. Therefore, the Ashvamedha ritual contains motifs 9 and 10.

The *Aupasana Agni* is a fire that is lit at the wedding and is extinguished at the time of the groom's death. Therefore, the Aupasana Agni ritual contains motifs 7 and 8.

The *Vena bird* is mythical bird described in the Rig Veda (book 10, hymn 123). While the Vena bird is not explicitly described as peacock, its name may have some connection with *Paravani*, the peacock mount of the god Murunga (motif 1a). The Vena bird brings the heavenly singer Gandharva

and his wife Apsaras to the heaven, where they become immortal. The hymn implies that previously they died (motif 4). According to the hymn, they gaze longingly at Vena (motif 3).

In Hindu mythology Yama, the god of death, may be related to the Finno-Permic Juma, the god of the sky motif 6). Therefore, the Vena hymn's phrase "hasteneth to the home of Yama" recalls the singer in the reconstructed folk song saying that she will die for him.

Table 2. The motifs distribution in the texts of selected Hungarian folk songs and Vedic scripts.

Text	1	2	3	4	5	6	7	8	9	10
Folk song 1	✔			✔						
Folk song 95	✔			✔	✔					
Folk song 98									✔	✔
Folk song 101	✔			✔						
Folk song D	✔		✔	✔						
Folk song M				✔			✔	✔		
Ashvamedha									✔	✔
Aupasana Agni							✔	✔		
Vena Bird	✔		✔	✔	✔	✔				

Table 3. The motifs similarity matrix of the Hungarian folk songs and Vedic scripts with green highlighting of significant values between Hungarian folk songs and Vedic scripts.

	Folk song 1	Folk song 95	Folk song 98	Folk song 101	Folk song D	Folk song M	Ashvamedha	Aupasana Agni	Vena Bird
Folk song 1	2	2	0	2	2	1	0	0	2
Folk song 95	2	3	0	2	2	1	0	0	3
Folk song 98	0	0	2	0	0	0	2	0	0
Folk song 101	2	2	0	2	2	1	0	0	2
Folk song D	2	2	0	2	3	1	0	0	2
Folk song M	1	1	0	1	1	3	0	2	1
Ashvamedha	0	0	2	0	0	0	2	0	0
Aupasana Agni	0	0	0	0	0	2	0	2	0
Vena Bird	2	3	0	2	2	1	0	0	5

Table 3. Textual parallelisms of folk song 95 and folk song D highlighted by different colors.

#	Folk song 95	Folk song D	#
1	Szomorú fűzfának harminchárom ága	Zöld erdőben, zöld mezőben	1A
2	Arra reá szállott harminchárom páva	sétálgat egy páva	1B
3	Ki zöldbe, ki kékbe, ki földig fehérbe,	Kék a lába, zöld a szárnya, aranyos a tolla.	2
5A	Szólítottam volna,	Hívott engem útitársnak, el is megyek véle,	3
5B	szántam búsítani	Kiért fáj a, fáj a gyönge szívem,	8A
6	Egy ilyen efiát megszomorítani.	sej meg is halok érte.	8B

Table 4. An English summary of the parallel elements in folk song 95 and folk song D.

#95	#I	#95 element	#D element	Explanation
1	1A	willow tree	forest	forest substitutes for tree with many branches
2	1B	peacock	peacock	
3	2	blue, green, white	blue, green, gold	white substitutes for gold
5A	3	call	call	two synonym words for "call"
5B	8A	feel sorry [for someone]	heart aches	feelings of sorrow is common element
6	8B	husband mourns wife	lover mourns beloved	mourning is common element

The Vena bird flies all the way to the heaven, which is described in Vedic script as the abode of thirty-three devas (motif 5). Therefore, the Vena bird hymn contains motifs 1, 3, 4, 5 and 6.

3.2 Text similarity measures

Table 2 summarizes the motifs distribution in the selected Hungarian folk songs and Vedic scripts.

We define the *motif similarity* of two texts A and B, denoted $sim(A, B)$, as the number of motifs on which they agree. For example,

$$sim(Folk\ song\ 95, Vena\ Bird) = 3$$

Let $motifs(A)$ be the set of motifs that text A contains. The *Jaccard similarity* of two texts A and B, denoted $J(A, B)$, is the defined as follows:

$$J(A, B) = \frac{|\ motifs(A)\ \cap\ motifs(B)\ |}{|\ motifs(A)\ \cup\ motifs(B)|}$$

For example,

$$J(Folk\ song\ 95, Vena\ Bird) = 0.6$$

While the motif similarity measure is a raw measure of the similarity, the Jaccard similarity also shows the significance of the similarity. We consider *significant* a Jaccard similarity of 0.6 or higher. Table 3 shows with green highlighting the significant Jaccard values between Hungarian folk songs and Sanskrit scripts.

3.3 Text parallels in Folk songs 95 and D

Here we show that Folk songs 95 and D are also closely related by showing that they have significant line-by-line text parallels.

Tables 3 and 4 show the parallels in Hungarian and English, respectively. Some lines were broken into parts A and B. The parallelism is most obvious when the same words are used, like "peacock" in both line 2 of folk song 95 and line 1B in folk song D, or the words "blue" and "green" in line 3 of folk song 95 and line 2 of folk song D. The parallelism is also clear when only synonym words are used, such as the two synonym words for "call" in line 5A of folk song 95 and line 3 of folk song D. The willow tree with thirty-three branches in line 1 of folk song 95 can be also recognized to parallel the forest in line 1A of folk song D. The other parallels in Tables 3 and 4 are similarly plausible.

3.4 Reconstruction of the common original form of Folk song 95 and Folk song D

There are enough variations between folk songs 95 and D to make it unlikely that one is simply a later version of the other. There would be no real reason to change the text from one word to its synonym. Instead, the situation seems to be more likely to have resulted from two different translations of the same original song into Hungarian. The hypothetical original song probably contained some version of all six parallel elements of the lines and half lines shown in Table 3. For each parallel element, the reconstruction needs to decide whether to prefer the version in folk song 95, the version in folk song D or some combination.

For the first parallel element in Table 1, line 1 of Folk song 95 seems more original because "willow tree" is more concrete than "green forest" and the occurrence of the number thirty-three seems an archaic feature as a reference to *Trayastriṃśa* [32].

For the second parallel element in Table 1, the two folk songs agree that the peacock appears. However, the folk songs disagree about the way the appearance occurs.

Folk song 95 claims that the peacock flew onto the previously mentioned willow tree, which we already chose as more likely to be the original element. In contrast, folk song D claims that the peacock was walking in the "green meadow" that seems to be only an added phrase after "green forest" perhaps as a replacement of the number thirty-three. Hence "flew onto it" seems the original form.

For the third parallel element in Table 3, several colors are mentioned. The colors refer to people's clothes in folk song 95, and to the leg, the wing and the feathers of the peacock in folk song D. It seems that the reference to the peacock's feathers is the original reference, while the other references are later additions. Although there is a superficial agreement on the blue and the green colors, their order is different and they refer to different things. In addition, blue and green are natural associations with peacocks that could be added at any later stage and contain no surprise element for the listeners. The differences in the details and the lack of a surprise element suggest that the blue and green references were not part of the original song. It seems that white in folk song 95 substitutes for gold in folk song D, and reference to clothes substitutes for feathers. Hence in our reconstruction, we kept the reference to the golden feathers of the peacock.

For the fourth parallel element in Table 3, it makes more sense for the peacock to call rather than the singer to call her husband. However, the peacock calling the woman simply for a journey (implicitly to heaven) is more likely than calling for

a travel companion. The word "útitárs" (travel companion) likely developed out of "út" (journey) after the awareness that this journey is a journey to heaven was lost.

For the fifth parallel element in Table 3, in folk song 95 the phrase "fáj a szívem" (my heart aches) seems original but curiously the word "fáj" (ache) is repeated. In folk song D, the words "szán" (feel sorrow for someone, pity) and "búsul" (feel sorrowful) are synonyms for heartache. Hence it seems likely that the original song, like folk song D, contained two different synonym words for sorrow. As a compromise approach, in our reconstruction, we picked both "búsul" from folk song D and "fáj a szívem" from folk song 95 once.

For the sixth parallel element in Table 3, the death of the singer is common. Folk song 95 implies that the husband would also die of sorrow if he would know about the death of his wife (that's why the wife is afraid to call him). That is more indirect and contrived than folk song D's statement of "meg is halok" (I will die). Therefore folk song D's phrasing is chosen for the reconstruction.

Based on the above arguments, we reconstruct the original song as follows:

> *Szomorú fűzfának harminchárom ága,*
> *Arra reá szállott arany tollú páva*
> *Hívott engem útra, el is megyek véle,*
> *Búsul, fáj a szívem, meg is halok érte.*

The English translation of the reconstructed song, which we call Folk song R, is the following:

> *A weeping willow tree has thirty-three branches.*
> *A golden-feathered peacock flew onto it.*
> *He called me to a journey. I'll go with him.*
> *My heart aches from sadness. I'll die for him.*

Folk song R contains all the motifs that folk song 95 and folk song D have. In other words, it appears that the original song had motifs 1, 3, 4 and 5. Later motif 3 dropped out of folk song 95 and motif 5 dropped out of folk song D. Note that the Jaccard similarity of folk song R and the Vena Bird hymn is even higher that either of folk song 95 or folk song D with the Vena Bird hymn. In fact, we have the following

$$J(Folksong\ R, Vena\ Bird) = 0.8$$

4 Linguistic Borrowings

Linguistics identified a set of words that reflect borrowings between various branches of the Uralic [9] and the Indo-European language families [3].

These linguistic borrowings corroborate the comparative textual analysis of Hungarian folk songs and Sanskrit literature. The people who brought these words with them also may have brought with them the peacock motifs and other motifs that we found to be common in the previous sections.

Table 1 shows some of the borrowed words. In the first column, Table 1 describes the English meaning, and in the other columns the cognate words in the Hungarian, other Uralic languages, Greek, Proto-Indo-European or Avestan or Sanskrit, and Sumerian.

The Uralic and Proto-Indo-European cognates are based on the *Mansi Dictionary of Munkácsi and Kálmán* [13], and the etymological dictionaries of Rédei [17] and Zaicz [34]. We also found some ancient Greek cognates using the etymological dictionary of Beekes [1, 2]. For the Greek words the superscript PG indicates that Beekes identified it as 'Pre-Greek.' Many of the words that were not identified as PG were also weakly argued to be PIE in origin. The Sumerian cognates are generally based on Parpola [16], whose relevant Sumerian entry number is shown in the last column. The dash --- indicates that no corresponding entry was found in Parpola [16] and the Sumerian cognate word in that case is our addition. The most interesting cognate word that we found is 'labyrinthos' whose Uralic origin is discussed in the next section.

In Table 1 and in the rest of this paper, when x and y are words, then the notation $x \sim y$ indicates that words x and y are cognates, $x > y$ indicates a derivation from x to y and $*x$ indicates a hypothetical form that is not attested in writing. The notation $x^L\ (m)$ denotes that x occurs in language L and means m in English. The similar consonant sounds are highlighted in red, inserted glide consonants are highlighted in blue, and omitted sounds are indicated by underscores.

The study of linguistic borrowings is interesting because it can suggest a chronology of the contacts between various people. To arrive at a valid chronology, we have to identify the likely time of borrowing for each borrowed word. Table 1 uses the following color convention:

Layer	Color
Proto-Uralic	blue
Proto-Finno-Ugric	green
Proto-Ugric	yellow
Proto-West-Ugric	orange
Proto-Hungarian	pink

The Proto-Uralic, Proto-Finno-Ugric, and Proto-Ugric layers are traditional classifications in linguistics [9, 17]. Koivulehto [11] suggests that Proto-Indo-European and Proto-Uralic were neighbors and several words were exchanged at that time. The words he identified are indicated by a PIE superscript. Others assume that these words are better explained as coming from a common ancestor of the Uralic and the Indo-European languages.

Wiik [30] proposed and Kallio [8] supports the idea that there is a Uralic substrate in Germanic languages. However, Wiik based his theory purely on phonological considerations and failed to demonstrate actual borrowed words.

Revesz [22, 23] recently identified the Proto-West-Ugric layer as a group that contains Euphratic, an ancestor of Sumerian, Hattic, Hungarian and Minoan. The Minoan language's inclusion in this group is strengthened by translations of the Phaistos Disk [19], and some Cretan Hieroglyphic [20, 21], and Linear A scripts inscriptions [22], and the noted similarities between the Minoan script symbols and the Old Hungarian alphabet [4, 5, 18], which is also called *rovásírás* in Hungarian and is sometimes translated as *Rovas* or *Rovas*h in English language publications [7].

Word borrowings occurred between Proto-West-Ugric and Proto-Indo-Iranian as evidenced by Avestan and Sanskrit borrowed words in all the branches of West-Ugric considered in Table 1, that is, Hungarian, Greek with PG markings as a proxy for Minoan, and Sumerian.

In the last layer we see borrowed words at a stage after the Proto-West-Ugric branch has separated because Sumerian no longer contains any of the Indo-Iranian words but Greek and Hungarian do. The common Greek and Indo-Iranian words may be explained as innovations on a branch of the Indo-European family tree before the separation of the Greek and the Indo-Iranian branches.

The chronology that emerges from the study of word borrowings (see Fig. 1) corroborates the comparative textual analysis of the previous section. The strong connection between Hungarian folk songs and ancient Sanskrit literature suggests a relatively late contact between Proto-Hungarian and Proto-Indo-Iranian groups. The weaker connection between Hungarian folktales and Sumerian literature [28] supports an earlier separation of the Euphratic (Proto-Sumerian) and Hungarian languages.

4.1 The Uralic origin of the word 'labyrinth'
Scholars have extensively debated the mysterious origin of "labyrinth" [1]. Below we propose a Uralic origin of this word.

First consider two Mansi words that have a Proto-Uralic origin [17, 34]. These words are:

khåli^{Mansi} (die)
pēri^{Mansi} (twist, turn)

The latter word is also associated with wrapping something [13]:

jel-*pēri^{Mansi} (wrap)

The composition of these two words yields the following:

khålå-*pårau^{Mansi} (a weed)

While the exact type of weed is not specified in the dictionary [13], the component words suggest that it is some kudzu-like weed that goes in all directions with its tendrils and leaves wrapping around other plants that are eventually killed. This Mansi word can be compared with the following word that Beekes [1] identifies as Pre-Greek:

γλαβρήνη^{PG} (a Cretan plant)

Arguably, the Mansi and the Pre-Greek words are cognates. The first correspondence is the following:

khålå^{Mansi} > γλα^{PG}

Here the first vowel seems omitted and the initial consonant becomes voiced due to the following liquid /l/. The second correspondence is the following:

*pårau^{Mansi} > *βυρήνη^{PG} > βρήνη^{PG}

Here also the first vowel seems omitted and the initial consonant becomes voiced due to either being between vowels or before a liquid /r/. Apparently this is a general tendency in Pre-Greek. The /n/ may be a hiatus filling insertion between two vowels.

Therefore, the Pre-Greek word likely also denotes some weed with twisty tendrils going in all directions. The tendril structure of such a weed was like the hallway pathways of a labyrinth. We suggest that this similarity is not accidental because the word:

λαβὺρινθος^{PG} – labyrinth

Table 1. Uralic (blue), Finno-Ugric (green), Ugric (yellow), West-Ugric (orange) layer cognates with Greek, PIE or Indo-Iranian and Sumerian

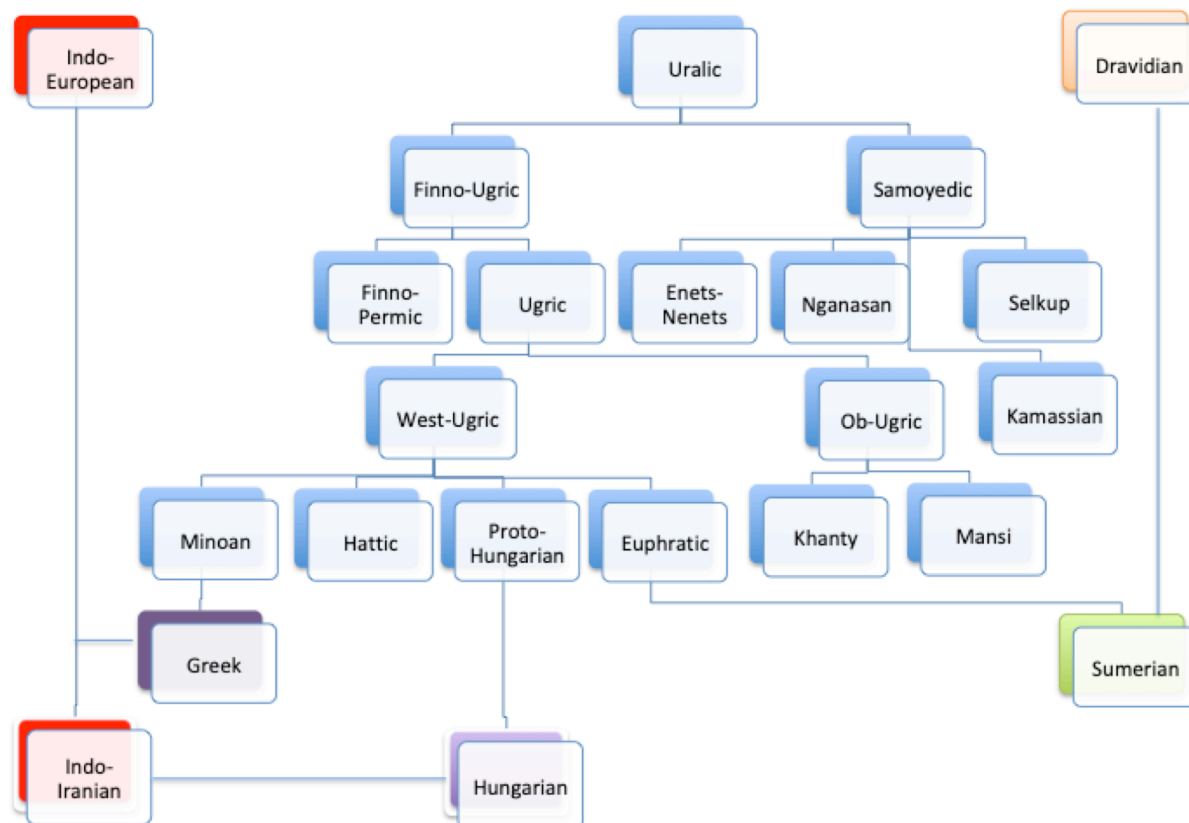| English	Hungarian	Other Uralic or Finno-Ugric	Greek	Proto-Indo-European Avestan, Sanskrit	Sumerian	#
axe	fejsze	päćt[Mansi], pīči[Selkup]		wāćī[Proto-Aryan] (knife)	pa-a-šu	---
fear (v.)	fél	pĕl[Khanty], pelkää[Finnish]		*pelh₁[PIE]	bu-luh	356
drill (v.)	fúr	pura[Finnish]	πριων[?] (saw)	*bhr(H)-[PIE]	bùru	379
braid, weave[1]	fon	păn[Khanty] (yarn), panne[Saami] (spin)	υφαινειν[1]	*pnH-e/o-[PIE]	pan	1952
twist, turn wrap die, death weed a Cretan plant	forog hal	pēri[Mansi] jel-*pēri[Mansi] khåli[Mansi] khålå-*påråu[Mansi]	κήρ[PG] γλαβρήνη[PG]		u₂-kul	---
raven, eagle[1]	holló	kolāk[Mansi] kulé[Selkup]	_ŏρνις	kurara[Sanskrit]	hurin[1]	1192
to die god (of death) god (of heaven) heaven	menny	jamme[Saami] jumala[Finnish], jumo[Mari] jen[Zyrian] ḿeńel[Mordvin]		yama[Sanskrit] juno[Latin]	nun me	1914 1681
wash, purify[+]	mos	moška[Mari]	νιζειν	*mozg-eh₂-ye/o[PIE]	maš[+]	1654
road, street	út	āχt[Mansi], ηut[Yurak]	οδος	adhvan[Sanskrit]	tilla₂	---
gold ornament	_arany	zarńi[Zyrian]		zaranya[Avestan]	še-er-hu-lum	---
price, value	ár	arvo[Finnish]		arəĵa[Avestan]	ár	136
thousand,	ezer	śurs[Zyrian]		sa-hásram[Sanskrit] hazaŋra[Avestan]	šár (3600)	2307
to milk	fej	ped'a[Mordvin] (also: drain)	βδάλλω[PG]	payō[Avestan] (milk, n.)	gima-sim (sieve)	---
sword	kard	kēr[Mansi] (iron), kärki[Finnish] (blade)		kard[Ossetian]	ĝiri	1079
hundred	száz	sat[Khanty], sata[Finnish]		śatám[Sanskrit]	me-a-ad	1684
castle, cover, weir	vár	varu[Finnish] βērü[Mansi]	φρούριον	vara[Avestan] vāra[Sanskrit]	uru	2827
maternal relative	ara	år[Mansi], orti[Khanty]		ærwadæ[Ossetian] (brother)	uruₓ (maid)	2806
lady, woman	asszony	khåusä nē[Mansi] (whimsical w.)	χsīn[Alan]	ka-ša-an	---
come!, foot+	gyere	jērte[Mansi] (soon)			ĝiri₃[+]	---
meat	hús (meat)	kös[Mansi]		ɣuš[Avestan]	uzu	2888
to warm, heat, soften by heat	olvad (melt)	tal[Mansi], sil[Zyrian] (melt)	θάλπω[PG]		zal	2912
bury	temet	tåw[Mansi]	θάπτω		túm	2597
palsy, stroke, kill lame	guta			ghāta[Sanskrit]	gud₆	879
river, stream	ügy		ὔδωρ (water)	udakám[Sanskrit] (water)	id	1211
wax			κηρος[PG] κήρωμα[PG]		eškurum	---
soon+, at same time	hamar[+]		_ἁμαρτῆ			
watch	őriz (guard)		όράω			
cow, heifer+	tehén		δάμαλις[+]	dhénā[Sanskrit]		
milk	tej		/thínjon/	dháyati[Sanskrit] (suck)		
garment, canvas+	vászon[+]			vásman[Sanskrit]		

Fig. 1. A partial diagram showing the Dravidian (yellow), Indo-European (red) and the Uralic (blue) language families. Greek (dark purple), Hungarian (light purple) and the Sumerian (green) descend from two different language families. The colors and their shades are chosen to suggest the combinations and the degrees of affiliations with each family.

can be obtained by omitting the initial γ and adding a common −νθος suffix. The process of the omission of γ is not clear. It may have been prompted by a folk etymology due to confusion with:

$$\lambda\tilde{\alpha}\alpha\varsigma^{PG} - stone$$

In particular, people may have felt that a stone structure has to start with the word for 'stone.'

5 Conclusions and Future Work

The peacock motif already caught the attention of Kodály [10] as something special in Hungarian folk music. The present work traced the origin of peacock and other motifs back to some Rig Vedic hymns and other ancient Sanskrit literature. It would be interesting to also study these motifs in Hungarian folk ballads, folktales and folk art in relation to Indian parallels. Finally, it would be interesting to compare the melodies of Hungarian folk songs and Indian music as well.

References:

[1] R. S. P. Beekes, *Etymological Dictionary of Greek*, Brill NV, Leiden, Netherlands, 2009.

[2] R. S. P. Beekes, *Pre-Greek Phonology, Morphology, Lexicon*, Brill NV, Leiden, Netherlands, 2014.

[3] R. Bouckaert, Ph. Lemey, M. Dunn, S. J. Greenhill, A. V. Alekseyenko, A. J. and Drummond, R. D. Gray, M. A. Suchard, and Q.

D. Atkinson, Mapping the Origins and Expansion of the Indo-European Language Family, *Science*, Vol. 337, No. 6097, 2012, pp. 957-960.

[4] S. Daggumati, P. Z. Revesz, Data mining ancient script image data using convolutional neural networks, *Proc. 22nd International Database Engineering and Applications Symposium,* New York, ACM Press, 2018, pp. 267-272.

[5] S. Daggumati and P. Z. Revesz, Data mining ancient scripts to investigate their relationships and origins, *Proceedings 23rd International Database Engineering and Applications Symposium*, New York, ACM Press, 2019, pp. 209-218.

[6] R. T. H. Griffith, translator, *The Rig-Veda*, Evinity Publishing Inc., Santa Cruz, CA, USA, 2009. available: https://en.wikisource.org/wiki/The_Rig_Veda

[7] G. Hosszú, *Heritage of Scribes: The Relation of Rovas Scripts to Eurasian Writing Systems*, Rovas Foundation Hungary, 2013.

[8] P. Kallio, Uralic substrate features in Germanic, *Journal de la Société Finno-Ougrienne*, Vol. 87, 1997, pp. 123-130.

[9] J. Kiss and F. Pusztai, eds., *A Magyar Nyelvtörténet Kézikönyve (Handbook of Hungarian Language History)*, Tinta Publ., Budapest, Hungary, 2018.

[10] Z. Kodály, *A Magyar Népzene,* Editio Musica Budapest, Hungary, 1951.

[11] J. Koivulehto, The earliest contacts between Indo-European and Uralic speakers in the light of lexical loans, In: C. Carpelan, A. Parpola, P. Koskikallio, eds., *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations,* Helsinki: Société Finno- Ougrienne, 2001, pp. 235–263.

[12] K. Lal, *Peacock in Indian Art, Thought and Literature,* Abhinav Publications, New Delhi, India, 2006.

[13] Mansi Dictionary of Munkácsi and Kálmán, downloaded June 22, 2018. Available: http://www.babel.gwi.uni-muenchen.de/munka/index.php

[14] K. L. Muttarayan, Sumerian, Tamil of the First Cankam, *Journal of Tamil Studies, Vol.* 8, 1975, pp. 40-61.

[15] P. T. Nair, The peacock cult in Asia*, Asian Folklore Studies,* **33** (2), pp. 93-170, 1974.

[16] S. Parpola, *Etymological Dictionary of the Sumerian Language, Vols. 1 and 2,* Helsinki, Finland*,* Foundations for Finnish Assyriological Research, 2016.

[17] K. Rédei, ed., *Uralisches Etymologisches Wörterbuch,* Budapest, Akadémiai Kiadó, 1988.

[18] P. Z. Revesz, Bioinformatics Evolutionary Tree Algorithms Reveal the History of the Cretan Script Family, *International Journal of Applied Mathematics and Informatics*, Vol. 10, No. 1, 2016, pp. 67-76. Available: http://www.naun.org/main/UPress/ami/2016/a182013-133.pdf

[19] P. Z. Revesz, A Computer-Aided Translation of the Phaistos Disk, *International Journal of Computers*, Vol. 10, No. 1, 2016, pp. 94-100. Available:http://www.naun.org/main/NAUN/computers/2016/a282001-455.pdf

[20] P. Z. Revesz, A Computer-Aided Translation of the Cretan Hieroglyph Script, *International Journal of Signal Processing*, Vol. 1, No. 1, 2016, pp. 127-133. Available: iaras.org/iaras/filedownloads/ijsp/2016/003-0017.pdf

[21] P. Z. Revesz, A Translation of the Arkalochori Axe and the Malia Altar Stone, *WSEAS Trans. on Information Science and Applications*, Vol. 14, No. 1, 2017, pp. 124-133. Available: http://www.wseas.org/multimedia/journals/information/2017/a285909-084.pdf

[22] P. Z. Revesz, Establishing the West-Ugric language family with Minoan, Hattic and Hungarian by a decipherment of Linear A, *WSEAS Transactions on Information Science and Applications*, Vol. 14, No. 1, 2017, 306-335.

[23] P.Z. Revesz, Sumerian contains Dravidian and Uralic substrates associated with the Emegir and Emesal dialects, *WSEAS Transactions on Information Science and Applications*, Vol. 16, 2019, pp. 8-30.

[24] P. Z. Revesz, M. P. Rashid, and Y. Tuyishime, The design and implementation of AIDA: Ancient Inscription Database and Analytics system, *Proceedings 23rd International Database Engineering and Applications Symposium*, New York, ACM Press, 2019, pp. 292-297.

[25] J. M. Robertson, *Pagan Christs – Studies in Comparative Hierology,* (Watts & Co., London, UK, 1911)

[26] M. I. Steblin-Kamenskij, Ariysko-ural'skie svyazi mifa o Yima, In: *Rossiya i Vostok: Problemy Vzaimodeystviya* (1995).

[27] *The Pennsylvania Sumerian Dictionary*, Available: http://psd.museum.upenn.edu

[28] Varga, Z. *Az Ősmagyar Mitológia Sumir és Ural-Altáji Öröksége,* Hermit Könyvkiadó Budapest, 2017.

[29] G. Whittaker, The case for Euphratic, *Bulletin of the Georgian National Academy of Sciences*, Vol. 2, No. 3, 2008, pp. 156-168.

[30] K. Wiik, The Uralic and Finno-Ugric phonetic substratum in Proto-Germanic, *Linguistica Uralica*, Vol. 33, No. 4, 1997, pp. 258-280.

[31] Wikipedia, Phoenix (mythology), available at: https://en.wikipedia.org/wiki/Phoenix_(mythology), (retrieved: May 2, 2019)

[32] Wikipedia, Trāyastriṃśa, available at: https://en.wikipedia.org/wiki/Trāyastriṃśa, (retrieved: May 3, 2019)

[33] Wikipedia, Yajna, available at: https://en.wikipedia.org/wiki/Yajna (retrieved: May 31, 2019)

[34] G. Zaicz, chief editor, *Etimológiai Szótár: Magyar Szavak és Toldalékok Eredete, (Etymological Dictionary: Origin of Hungarian Words and Affixes),* Tinta Könyvkiadó, 2006.