

2018

# Computational Linguistics Techniques for the Study of Ancient Languages

Peter Z. Revesz

*University of Nebraska-Lincoln*, prevez1@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/csearticles>

---

Revesz, Peter Z., "Computational Linguistics Techniques for the Study of Ancient Languages" (2018). *CSE Journal Articles*. 193.  
<https://digitalcommons.unl.edu/csearticles/193>

This Article is brought to you for free and open access by the Computer Science and Engineering, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in CSE Journal Articles by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Computational Linguistics Techniques for the Study of Ancient Languages

Peter Z. Revesz<sup>1,a</sup>

<sup>1</sup> Department of Computer Science, University of Nebraska-Lincoln, Lincoln, NE, 68588, USA

**Abstract.** This paper presents a grammatical comparison of the Minoan language with the proto-Ugric and proto-Hungarian languages. Recent research showed that these languages are closely related, but this paper presents a novel grammatical comparison. The grammatical comparison shows the Minoan language to have an agglutinative type of grammar, with a CVCV type root word structure. The Minoan language also features assimilation between the word roots and the suffixes and a possessive phrase structure that is similar to that in Hungarian.

## 1 Introduction

The Minoan Cretan Hieroglyph writing [19, 20, 21] and Linear A [7, 8] have been recently deciphered and shown to be closely related to the Ugric branch of the Finno-Ugric languages [12, 13, 14, 15]. The corpus of Cretan Hieroglyph texts includes the Phaistos Disk [1], the Arkalochori Axe and the Malia Altar Stone inscriptions [14]. The Phaistos Disk and the Arkalochori Axe are sometimes described as separate scripts, but the Malia Altar Stone is listed as item number 328 in the *Corpus Hieroglyphicarum Inscriptionum Cretae* (CHIC) [8].

The key to the decipherments came from the application of phylogenetic algorithms [10, 16] applied to ancient scripts. Using these clustering algorithms, Revesz [11] analyzed the evolutionary relationships within the *Cretan script family*, which includes the following scripts: Cretan Hieroglyph, Linear A, Linear B [3], Cypriot, Greek, Phoenician, South Arabic, Tifinagh, and Old Hungarian [4, 5], which is also called *rovásírás* in Hungarian and also written sometimes as *Rovas* in English language publications. To these scripts, Revesz [15] added the Carian alphabets [18], which shows a particularly strong relationship with the Old Hungarian alphabet. In fact, the Carian alphabet is either an intermediary or very close to some intermediary between the Linear A and the Old Hungarian scripts.

The Carian and Old Hungarian phonetic values were used to find the syllabic phonetic values of the Linear A symbols [15] under the assumption that CV type Linear A phonetic values later simplified to C phonetic values during the gradual evolution of writing within the Cretan Script Family. (Here, as usual, C denotes a consonant and V denotes a vowel.) Moreover, there was enough circumstantial evidence of the relationship between the Minoan and the Ugric

languages, which includes Hungarian, Khanty, and Mansi. Therefore, it seemed logical to assume that each Linear A sign represented some syllabic value CV and originally depicted a picture of an object or action that in the common proto-Ugric language had a name that started with the same CV syllable. In other words, searching for the phonetic values is a constraint problem solving [6, 9] application. In particular, the syllabic value CV of any Linear A symbol  $\sigma$  has to satisfy the following constraints:

1.  $\sigma$  depicts the picture of an object or action, which is described by some word  $\omega$ .
2.  $\omega$  starts with the same syllable CV.
3.  $\omega$  is Proto-Ugric.
4. C is the phonetic value of the Carian and Old Hungarian letters that are visually closest to  $\sigma$  according to an objective mathematical measure of visual similarity.

Applying the constrained acrophonic principle, it was possible to establish the syllabic values of close to fifty Linear A symbols. The substitution of the syllabic values yielded texts that were similar to some early medieval Hungarian texts. The above description underestimates the role that grammatical considerations played in an early stage of decipherment. This paper illustrates some of the grammatical considerations.

This rest of this paper is organized as follows. Sections 2 and 3 describe a set of Minoan and Hungarian language and script similarities, respectively. Finally Section 4 gives some conclusions and directions for future work.

<sup>a</sup> Corresponding author: [revesz@cse.unl.edu](mailto:revesz@cse.unl.edu)

**Table 1.** Some blocks of the Phaistos Disk and the Arkalochori Axe inscriptions arranged to reveal repeatedly used suffixes and word roots. The Arkalochori Axe symbols are transliterated into the Phaistos Disk symbols as shown in [14].

Block	Possible Root (red)	Possible Suffix (blue)	Block	Possible Root (red)	Possible Suffix (blue)
12					
45					
51					
59					
Ark. 3					
3					
6			2		
9					
10			7		
20					
22					
24					
27					
61					
34					
Ark. 1					
28					
29			38		
33			40		
43					
47					
49					
37					
39					
52			30		
25					
26					
50					
60					
36			44		
53					
58					
Ark. 2					
8					
46					

## 2 Minoan and Hungarian grammar similarities

In this section, we list some grammatical similarities between the Minoan and the Hungarian languages. Our grammatical comparisons could be applied to any writing, not just the Minoan writing.

### 2.1 Minoan is also an agglutinative language

Agglutinative languages append suffixes to word roots without changing those roots. Table 1 displays some blocks of the Phaistos Disk and the Arkalochori Axe. Such a table can be built either manually or by a computer program, which would be preferable when the number of words in the corpus of a script is large. We built such a table manually because the Phaistos Disk and the Arkalochori Axe were previously considered as a separate corpus of texts from those of the Linear A and the Cretan Hieroglyph corpuses, which later were shown to be actually closely related variants for writing the same Minoan language [12, 14]. We notice in Table 1 the following:

1. Table 1 shows eight different endings that each occurs at least two different times.
2. Some endings are apparently optional. For example,   is optional because it occurs in block 6 but does not occur in block 2. Similarly,  is optional because it occurs in block 36 but does not occur in block 44.
3. Some endings are replaceable with another ending. For example, blocks 29 and 38 have the same apparent root but end with  and , respectively. Similarly, blocks 33 and 40 have the same apparent root but end with  and , respectively.
4. Whenever the endings are attached to a root, the root does not change. Table 1 indicates by red some of the apparent roots.

The above items imply that the Minoan language is an agglutinative language.

### 2.2 CVCV root structure

Linear A has been suspected to be a mostly syllabic writing with CV type syllables, where C is a consonant and V is a vowel, similar to Linear B, which is its immediate descendant.

The CV type syllables fit well with proto-Uralic word roots that generally have two syllables with a CVCV structure. For example,

- \*jäñke > \*jäñe > jég (ice),
- \*kala > hal (fish),
- \*käte > kéz (hand),
- \*lovu > ló (horse),
- \*mete > méz (honey),
- \*pesä > fész-ek (nest),
- \*pukta > \*puta > fut (run),
- \*uje > úsz (swim),
- \*vaske > \*vasa > vas (iron),

where in the above words /ä/ is a more open /e/.

Words with a CVCV structure can be represented by two symbols that stand for CV syllables. This fact may have influenced the Linear A script to develop as a syllabic script. Table 1 already shows several roots that contain two Linear A symbols. These include for example   and   and  .

### 2.3 Suffixes ending in -k

A problem with a pure syllabic script is that many suffixes do not fit into a CV structure. For example, consider the following Hungarian suffixes that end with a /k/ phoneme. We also give some examples, as they appear in the earliest Hungarian language documents. One of the frequently consulted documents is the *Halotti Beszéd (Funeral Sermon)* [2], which will be referenced as HB below.

1. /-ak, -ek, -ok/ is the plural of words that end in a consonant. The vowel that is chosen according to vowel harmony rules. For example, *hal-ak* (fishes) and *kez-ek* (hands) and *ablak-ok* (windows).
2. /-k/ is the plural of words that end in a vowel. For example, *faluk* (villages) or *kapuk* (gates).
3. /-k/ is the 1st person singular present tense verbal suffix in the indeterminate case.
4. /-juk, -jök/ is the 1st person plural present tense verbal suffix in the determinate case. For example, *tümet-jük* (we bury) appears in HB. As another example, present Hungarian uses *számol-juk a pénzt* (we count the money).
5. /-juk/ is also the 3rd person plural possessive suffix. For example: *kutyájuk* (their dog).
6. /-muk/ is the 1st person plural present tense verbal suffix in the indeterminate case. This appears as *vogy-muk* in HB. This suffix appears to be simply the composition of the /-om, em/ first person singular verbal suffix in the

determinate case and the plural /-k/, ex: *olvas-unk* (we read a book).

7. /-nak, -nek/ is the third person plural present tense verbal suffix in the indeterminate case, for example, *esz-nek* (they eat).
8. /-nak, -nek/ is also a marker of the possessor of an object. For example, a *ló-nak* a *lába* (the horse's leg).
9. /-nak, -nek/ is also a lativusz suffix. For example, *fal-nak megy* (goes to a wall), *hegy-nek fordul* (turns towards a mountain).
10. /-omk/ is the 1st person plural possessive suffix, which appears as *uromc* [ur-omk] (our lord) in HB. Etymologically, this suffix appears to be simply the composition of the first person possessive /-om, -am, -em/, as in *ház-am* (my house) and the plural /-k/. Today, this suffix appears as /-unk, -ünk/, as in *ház-unk* (our house).

The Minoan  symbol represents not a syllable but some single phoneme because it used only at the end of the words with one exception. According to Table 1, in the Minoan language about half of the suffixes end with a  symbol. Remarkably, about half of the suffixes end in /k/ in Hungarian. Therefore, it is tempting to associate Minoan  with Hungarian /k/.

Moreover, the above Hungarian suffixes could be grouped into three groups: (1-3), which have the form /-Vk/, where the vowel V is optional, (4-5), which have the form /-jVk/, and (6-10), which probably had the form /-mVk/ assuming that m > n or n > m changes in some cases. These groups seem to match a natural grouping of the Minoan words into those that end with , with  and with , respectively.

Old Hungarian contains two letters that denote the /k/ phoneme:  and . According to some researchers one letter was used only within the words and the other was used only at the end of words. When carving the symbols into wood, a diamond is a convenient simplification of a circle, which may have denoted a human head [18]. Therefore, the shapes of the Minoan  symbol and the Old Hungarian  symbol have a connection. Moreover, the Minoan symbol depicts the head of a man with prominent hair. The Mansi word for man is /kom/, while the ancient Greek word for hair was /komi/, which may have been borrowed from Minoan. This shows a /k/ or a /ko/ phonetic connection between the two symbols.

## 2.4 Conjunction

Table 2 shows another pair of blocks that allows us to

suspect that the symbol  is a conjunction symbol, meaning “and,” a disjunction symbol, meaning “or,” or it is some prefix. The shape of this symbol read from left-to-right suggests that it may denote two paths that merge together, that is, a conjunction.

When rotated ninety degrees, the symbol also reminds one of the Old Hungarian  symbol, which denotes the /ʃ/ phoneme and occurs in the Hungarian words *s* and *és* that both mean “and.”

**Table 2.** Possible Minoan conjunction or affix.

Block	Conjunction or Prefix	Root	Suffix
6			
31			

## 2.5 Assimilation by consonant doubling

Table 3 shows the doubling of some symbols before the hypothetical suffixes.

**Table 3.** Two blocks contain a doubling of some symbols right before possible suffixes.

Block	Possible Root	Doubling	Possible Suffix
3			
49			

The doubling of consonants before suffixes is common in Hungarian. These doublings result from assimilation between the last consonant of the root and the beginning consonant of the suffix. Table 4 shows some examples.

**Table 4.** Hungarian assimilation with consonant doubling.

Root + juk Suffix	Assimilation
<i>mosjuk</i> (we wash)	/ʃ ʃ/
<i>úszjuk</i> (we swim)	/s s/
<i>főzjük</i> (we cook)	/z z/
<i>hagyjuk</i> (we let)	/j j/
<i>hunyjuk</i> (we close [eyes])	/p p/
<i>bátyjuk</i> (their older brother)	/c c/

Therefore, the Minoan  and  symbols behave similarly to the Hungarian doubled consonants and likely denote one of the consonants that is doubled in Hungarian except /ʃ/, which we already associated with .

## 2.6 Assimilation by palatalization

Assimilation can occur without a doubling in case of some consonants. Table 5 gives some examples from Hungarian.

The palatalized sounds in Table 5 may not have been originally used in the Hungarian language within word roots, but they tend to occur naturally with the addition of suffixes that start with /j/. It is likely that in the Minoan language the palatalized sounds also first occurred as a result of assimilation.

**Table 5.** Hungarian assimilation without doubling.

Root + juk Suffix	Assimilation
mondjuk (we say)	/j/
fonjuk (we weave)	/j/
futjuk (we run)	/c/

Table 6 shows that in block 22 a palatalization during assimilation can be suspected because the apparent assimilation yields a symbol that is rarely used. Moreover, it is never used at the beginning or the end of words, where palatalization is absent. It is also noticeable that it occurs only before  , which we already associated with the /-juk, -jük/ suffix. Compare Phaistos Disk block 22 with the Arkalochori block 3, where there is no assimilation sound in a similar context before  which we associated with the /-Vk/ suffix.

**Table 6.** Possible Minoan assimilation without doubling.

Block	Root	Assimilation	Suffix
22	 		 
Ark. 3	  		

### 3 Minoan and Hungarian Script Similarities

The grammatical comparisons in Section 3 already enable the identification of the phonetic values of some of the Phaistos Disk Minoan symbols as shown in the first two columns of Table 7. From Table 7, it is apparent that the Old Hungarian alphabet has a strong connection to the Minoan symbols. After such a realization, the logical step is the thorough comparison of all Minoan and Old Hungarian symbols to identify possible phonetic values of the Minoan symbols. The details of this step are described in Revesz [11].

**Table 7.** Minoan and Old Hungarian script similarities.

Symbol	Grammatically Identified Phoneme	Old Hungarian Letter	Old Hungarian Phoneme
	/k/		/k/
	/j/		/j/ or /ɲ/
	/ʃ/		/ʃ/
	/s/, /z/, /j/, /ɲ/ /c/		/s/, /z/
	/s/, /z/, /j/, /ɲ/ /c/		/z/

### 4 Conclusions and future work

In the future we plan extend our analysis and to give an explicit parser for the Minoan grammar.

### References

1. W. Achterberg, J. Best, K. Enzler, L. Rietveld, and F. Woudhuizen, *The Phaistos Disc: A Luwian Letter to Nestor*, Publications of the Henri Frankfort Foundation **13** (Amsterdam, Dutch Archaeological and Historical Society, 2004)
2. L. Benkő, *Az Árpád-kor magyar nyelvű szövegeimlékei* (Budapest, ELTE, Régi Magyar Irodalomtudományi Intézet 1980)
3. J. Chadwick, *The Decipherment of Linear B*, (Cambridge University Press, 1958)
4. S. Forrai, *The Old Hungarian Writing from Ancient Times to the Present*, (in Hungarian), (Antológia Kiadó, Budapest, Hungary 1994)
5. G. Hosszú, *Heritage of Scribes: The Relation of Rovas Scripts to Eurasian Writing Systems*, (Rovas Foundation Hungary, Budapest, Hungary 2013)
6. P.C. Kanellakis, G.M. Kuper, P.Z. Revesz, J. of Comp. and Sys. Sciences, **51**, 1, 26-52 (1995)
7. J.-P. Olivier, Cretan writing in the second millennium B.C., *World Archaeology*. **17**, 3, 377–389 (1986).
8. J.-P. Olivier, L. Godart and J.-C. Poursat, *Corpus Hieroglyphicarum Inscriptionum Cretae Études Crétoises* **31**, (De Boccard, Paris 1996)
9. P.Z. Revesz, *Introduction to Databases: From Biological to Spatio-Temporal*, (Springer, New York, 2010)
10. P.Z. Revesz, Proc. 4th ACM Int. Conf. on Bioinfo. and Comp. Bio., 731-734, (ACM Press, New York, USA, 2013)
11. P.Z. Revesz, Bioinformatics evolutionary tree algorithms reveal the history of the Cretan script family, *Int. Journal of Applied Mathematics and Informatics*, **10**, 67-76, (2016)
12. P.Z. Revesz, A computer-aided translation of the Phaistos Disk, *Int. Journal of Computers*, **10**, 94-100, (2016)
13. P.Z. Revesz, A computer-aided translation of the Cretan Hieroglyph script, *Int. J. Signal Proc.*, **1**, 127-133 (2016)
14. P.Z. Revesz, A translation of the Arkalochori Axe and the Malia Altar Stone, *WSEAS Transactions on Information Science and Applications*, **14**(1), 124–133 (2017)
15. 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*, **14**(1), 306–335 (2017)
16. M. Shortridge, T. Triplet, P.Z. Revesz, M. Griep, R. Powers, *Comp. Bio. and Chem.*, **35**, 1, 24-33, (2011)
17. G. Varga, *Bronzkori Magyar Írásbeliség*, Irástörténeti Kutató Intézet, 1993.
18. Wikipedia, “Carian alphabets,” downloaded April 5, 2017. Available: [https://en.wikipedia.org/wiki/Carian\\_alphabets](https://en.wikipedia.org/wiki/Carian_alphabets)
19. F. Woudhuizen, *The Earliest Cretan Scripts*, (Innsbrucker Beiträge zur Kulturwissenschaft 2006)
20. J. G. Young, *Minos* 31-32 (1996-1997[1999]) 379-400.
21. P. Yule, *Early Cretan Seals: A Study of Chronology*, *Marburger Studien zur Vor und Frühgeschichte* **4** (Mainz 1981)