A Machine-Readable Dictionary for the Corpus of Linear-A Script

MAVRIDAKI A.[1], ZACHARIS N.[1], GIACHOS I.[2*], PAPAKITSOS E.[3]
[1] Department of Informatics and Computer Engineering
University of West Attica
Egaleo Park Campus, Ag. Spyridonos Str., 12243 Egaleo, Athens
GREECE
[2] Department of Industrial Design & Production Engineering
University of West Attica
Ancient Olive Grove Campus, Thivon 250, 12241 Egaleo, Athens
GREECE
[3] School of Physics
University of the Witwatersrand
Private Bag 3, Wits 2050, Johannesburg
SOUTH AFRICA

Abstract: - The present paper describes a machine-readable dictionary for the corpus of Linear-A script, which is part of a relevant software platform, developed to facilitate scholars in deciphering Linear-A script. Although the script of Linear-A is conventionally called "Minoan language" quite often, the actual underlying language still remains unknown, without a certain linguistic affinity. The entire project aims at providing scholars/researchers standardized digital resources for the study of Linear-A that has been discovered more than a century ago, inscribed in a total of 1644 artefacts. Yet, despite of this relatively large quantity of artefacts, only a small number is useful for linguistic purposes (i.e., for deciphering attempts), something that makes the deciphering process even more difficult. Several databases of Linear-A exist, mainly in the form of images, none of them being machine-readable, i.e., able of being directly accessible by a programming language for computational processing. The presented herein machine-readable dictionary has been developed following the methodology described in a previous series of works, which includes the method of Syllabic Grouping for allowing the search-engine of the software platform to discover possible cognates of Linear-A words to the other contemporary languages of the platform. Its design and how this facilitates the study of Linear-A corpus are presented and discussed.

Key-Words: - Linear-A, Minoan Language, Aegean Scripts, Bronze Age Crete, Cretan Protolinear Script, Syllabic Grouping, Machine-Readable Dictionary

Received: April 27, 2025. Revised: June 15, 2025. Accepted: July 2, 2025. Published: July 31, 2025.

1 Introduction

Linear-A script is a syllabary (or logosyllabary) [1], [2] that was used in Bronze Age Crete from 1800 to 1450 BCE [3], firstly discovered in 1900 CE at the excavations of the Minoan palaces by Sir Arthur Evans [4]. Since then, more inscriptions were discovered in Greece, both around the mainland and the Aegean Sea islands, in Italy, Israel, the Black Sea coasts of Turkey, and as far as Margiana of Central Asia [5], [6], [7], [8].

Linear-A is composed of phonetic signs (simple or more composite ones), logograms (symbols of domesticated animals, plants, or everyday tools and commodities), numerals, metrical signs and ligatures. Each phonetic sign renders a syllable of one preceding consonant followed by a vowel (e.g., "DA", "NO", "PI", "TE", etc.) or a simple vowel (e.g., "A", "E", "I", "O", etc.), thus named syllabograms. Originally, more than 75 syllabograms were discovered, while recently 97 signs are registered [9], in more than 1427 inscriptions mainly on clay tablets, but also on ceramics, roundels, vessels, silver and gold hairpins, offering tables and stone [3]. Nowadays, the database in [10] contains a total of 1644 artefacts, mainly of commercial use for palatial or religious writings [11] that contain approximately 7,360 repeated signs [12].

Conventionally, Linear-A is regarded as having common origin with Cretan Hieroglyphics [13] or being a descendant of them. It has been repeatedly argued though that both scripts were derivatives of the original Cretan Protolinear Script (CPS), where

the former is a surviving subset [14], while the latter is its ritual and ornamental version [15].

Apart from symbols of commodities, numerical values, toponyms, anthroponyms and theonyms (all of them being readable), the script remains formally or unanimously undeciphered. Even the proper names above have unknown etymological origins. Consequently, Linear-A attracts a lot of interest from scholars and researchers around the world, with many contemporary languages suggested as being the underlying ones [16], [17], [18], [19], [20], [21], including those of the Middle East, Eastern Mediterranean, Anatolia, Balkans, Eastern and Central Europe [22], [23], [24], [25], [26], Caucasus [27], even as far as Japan [28], [29]. The lack of consensus on the conveyed language has also led to suggestions about a disappeared local language, named Minoan [30], as well as Proto-Greek. Of course, nowadays, the decipherment attempts include computational tools and methods, as well, that will be briefly described next.

Accordingly, the research herein describes the development of a machine-readable dictionary (MRD) of Linear-A corpus, implemented in Microsoft Excel spreadsheets, in order to facilitate the decipherment of the script and the study of the information contained in the artefacts of the corpus.

2 Problem Formulation

The problem of deciphering Linear-A scripts involves two main aspects: the linguistic and the computational one, along with the peculiarities of the script itself, as expressed below to their corresponding subsections.

2.1 The Linguistic Aspect

As mentioned previously in the Introduction section, the candidate languages for deciphering Linear-A are many, the most prominent being those contemporary ones of geographic proximity to start with. The initial study herein was focused in twelve (12) languages from Minor Asia, Italy and the Balkans. The largest group of them belong to the Anatolian linguistic family of the Indo-European phylogenetic tree that comprise seven (7) languages. All of these twelve languages will be briefly presented next, in alphabetic order, before commenting on the peculiarities of Linear-A script.

2.1.1 Carian

The Carian language belongs to the Anatolian linguistic family [31] and was spoken in south-western Minor Asia. The used vocabulary herein is only 374 words, written in an alphabetic script of 34 letters with the direction being from left to right [32]. Its written peculiarity was the large consonantal clusters, because the short vowels were omitted. For example, the word roughly pronounced as “masanorada” was written as “msnordś”, retaining just the long vowel “o”.

2.1.2 Etruscan

The Etruscan language was spoken in nowadays Tuscany, Italy, the ancient Etruria [33]. Conventionally, it is classified in a small group of lost languages, although several studies have demonstrated its affinity to the Anatolian family [34], [35], [36]. The used vocabulary herein consists of 563 words, written in an alphabetic script of Greek origin that consists of 26 characters [37].

2.1.3 Hattic

The Hattic language was spoken in Minor Asia, attested until the 2nd millennium BCE in 360 survived fragments. It is not an Indo-European language. The used herein small vocabulary consists of 303 words, with not a well-established meaning.

2.1.4 Hittite

The Hittite language is the most significant and well-known language of the Anatolian family, attested in 30,000 fragments or tablets of cuneiform script, dated from the 17th to the 12th centuries BCE [38]. The vocabulary used herein consists of 301 words.

2.1.5 Hurrian

The Hurrian language belongs to the Vannic linguistic family, which affinity (e.g., Kartvelian, Indo-European, etc.) is disputed. It is attested from 2300 BCE to 1000 BCE, in the most eastern Minor Asia, in bilingual inscriptions with Hittite, written in cuneiform [39]. Extended linguistic work argues in favour of Hurrian as the rendered language by Linear-A [40]. The used herein vocabulary consists of 352 Hurrian words.

2.1.6 Luwian

Luwian is a language that belongs to the Anatolian linguistic family, spoken in southern Minor Asia in the 2nd millennium BCE and attested in two scripts: the Hieroglyphic and the Cuneiform Luwian [41], [42]. The used herein vocabulary consists of 180 words.

2.1.7 Lycian

Lycian is a language of the Anatolian family, spoken in southwestern Minor Asia by the Lycians,

who are considered among the Peoples of the Sea (12th-11th centuries BCE). It is attested on inscriptions of tombs, coins and some longer texts [43], [44], [45], written in alphabetic script related to the Greek one. The used herein vocabulary consists of 394 words.

2.1.8 Lydian

The Lydian language of the Anatolian family was spoken in western-central Minor Asia (ancient Lydia), attested in 100 inscriptions [46], [47] and written in an alphabetic script of 26 letters, similar to the Greek one. The used vocabulary herein consists of 225 words.

2.1.9 Palaic

The Palaic language of the Anatolian family was spoken in northern Minor Asia during the 2nd millennium BCE [48]. It is attested on cuneiform tablets from the capital city of Hittites (i.e., Hattusa), being a liturgical language [49]. The used herein vocabulary consists of 133 words.

2.1.10 Phrygian

Phrygian is an Indo-European language, being most related to Ancient Greek [50] or, according to an alternative theory, to the Italo-Celtic languages [51]. It is attested in inscriptions of Old Phrygian, written in an alphabet of nineteen (19) letters similar to the Greek one, and of New Phrygian, written in the Greek alphabet [52]. It was spoken in western Minor Asia [53] and the used herein vocabulary consists of 356 words.

2.1.11 Thracian

Thracian is an Indo-European language spoken in Southeast Europe until the 6th century CE. It is poorly attested on a small number of inscriptions that contain glosses in Greek writings and proper names [54]. The used herein vocabulary consists of 246 words.

2.1.12 Urartian

Urartian is a language of the Vannic linguistic family, along with Hurrian (see sub-subsection 2.1.5), spoken around Lake Van in the most eastern Minor Asia. It is attested in approximately two hundred inscriptions, written in cuneiform-type script [55]. The herein used vocabulary consists of 253 words.

2.2 The Computational Aspect

A recent study [30] presents a review of computational tools for Linear-A script prior to 2024, along with the rest of the Aegean scripts (Linear-B, Cretan hieroglyphics, Cypriot-Greek, and the Cypro-Minoan syllabaries). The conclusion is that "… there is an obvious lack of a standardized digital dataset of Bronze Age Aegean and Cypriot inscriptions, as most of the existing approaches use internal datasets that are not available online" [30]. These tools will be classified herein in two (2) major categories:

- Eleven (11) databases;
- six (6) software applications known (at least).

They will be briefly presented next.

2.2.1 Databases

The following databases are well organized and complete, yet, they cannot be directly accessed by a programming language for processing purposes, except the last one. They mainly include images of the inscribed artefacts and explanatory information; therefore, they cannot be considered as MRDs (except the last one).

a) The original and complete corpus of GORILA (**G**odart and **O**livier **R**ecueil des **I**nscriptions en **L**inéaire **A**) [56] has been digitalized, containing scanned images of the inscribed artefacts; it is a main source for other implementations.
b) The original and complete corpus of CHIC (Corpus Hieroglyphicarum Inscriptionum Cretae) [57] has been digitalized, containing scanned images of the inscribed artefacts; it is a main source for other implementations.
c) Professor Emeritus J.G. Younger's website [58] has been one of the most comprehensive resources, yet not an MRD; it is accessible online.
d) AIDA (Ancient Inscription Database and Analytics) is a database with inscriptions from ancient scripts and languages, including Linear-A, no longer accessible online at the given URL [59].
e) SMID (Studies in Mycenaean Inscriptions and Dialect) is a reference tool of the University of Texas that contains plenty of cultural, linguistic, religious, archaeological and historical information in the form of scanned pages [60].
f) ERC INSCRIBE (INvention of SCRIpts and their BEginnings) is a database available online [61] that contains 3D representations of artefacts and inscriptions.
g) SigLA (**Sig**ns of **L**inear **A**) is a recent database available online [1] that includes an abundance of information (signs,

phonetic transcriptions, time period, places of origin, types of artefacts, etc.).

h) The Linear-A/Minoan digital corpus of Petrolito et al. [63], now not available on the URL given by [30].
i) The PA-I-TO ongoing project of the Sapienza Università di Roma, Università IULM and Heraklion Archaeological Museum [63]; the Linear-A part is not available online yet.
j) The Linear-A Explorer, a database with images and information of inscribed artefacts [10].
k) The multilingual MRD of the vocabulary of Linear-A and twelve (12) contemporary target languages by Mavridaki et al. [64] that is implemented in MS Excel spreadsheets, thus, directly accessible by a programming language; it is part of a software platform that also includes an interactive interface and a search-engine, implemented in C# and designed to facilitate both decipherment purposes and self-learning; it can be accessed online, although this facility is not available yet; the Linear-A's section contains 1150 single words in the middle column in their conventional alphabetic transcription to Latin alphabet (Table 1), with the designation of their corresponding inscribed artefact(s) on their left column and their corresponding designation of Syllabic Grouping [65] for the search-engine [66] on their right column.

Table 1. A part of Linear-A's MRD of vocabulary.

HT 116a.1	U-TA-LO	UDL
HT 117a.4	U-DI-MI	UDM
KH 16.2	U-TA-NO-CƏ	UDNC
KH 7b.2	U-TA-NO-SE	UDNS

2.2.2 Software Applications

The below six (6) software applications aim at assisting the decipherment of Linear-A or/and its self-learning (deliberately by the last one).

a) A system based on bioinformatics analysis was presented by Revesz [67], comparing the signs of seven (7) scripts that included Linear-A (the others being Linear B, Cretan Hieroglyphics, Greek, South Arabic, Phoenician, and Old Hungarian); a hypothetical evolutionary tree was constructed through the ClustalW2 phylogenetic algorithms, concluding that the above scripts had a common ancestor.
b) Another system based on bioinformatics analysis was presented by Revesz [68], comparing this time the signs of nine (9) scripts that included Linear-A (those of 2.2.2.a plus Cypriot-Greek and Tifinagh); by proposing a measure to calculate the similarity between pairs of signs, the conclusion was that Phoenician, Greek and South Arabic alphabets belong to one branch, while the rest scripts to a different one.
c) Yet another system by Revesz [23] that proposed a measure of feature-based similarity, in order to compare signs from various scripts, based on 13 distinct features (e.g., having curved or parallel lines); a new phonetic grid for Linear-A has been constructed so, and then a Minoan-Uralic-English dictionary.
d) A method proposed by Karajgikar et al. [69] uses various natural language processing techniques (e.g., exploratory knowledge mining, exploratory n-gram analysis and weighted frequency) to determine the distribution of signs in Linear-A's corpus. The results of their analysis were inconclusive on whether Linear-A actually represents a language or not.
e) The system of Minoan Cryptanalysis [70] combines two computational methods, i.e., a feature-based similarity measure and the consonantal approach, to initially compare Linear-A with the Cypriot-Greek syllabary and the Carian alphabet, as well as with Proto-Celtic, Hittite, Luwian, Ancient Egyptian and Uralic at a later stage; the aim of this approach was to avoid the supposed by the authors biases of assigning Linear-B phonetic values to Linear-A, and resulted in a few word-matches between these languages, without any conclusive outcome.
f) The software platform of Mavridaki et al. [64] consists of an interface and an underlying search-engine that access the connected database of 2.2.1.k, in order to discover possible cognates between the Linear-A words and the target languages of the database (i.e., the multilingual MRD; the interface is based on the theory of the Cretan Protolinear Script (CPS) [14], which is the most comprehensive theory on the origins of the Aegean scripts, where Linear-A is considered a derivative of CPS; the search is based on the method of Syllabic Grouping [65], which is a semi-consonantal

approach similar to the previous software application (see 2.2.2.e), while the results are also recorded in an external text-file and evaluated manually.

2.3 The Linear-A Script

The Linear-A script, combined with the searching method of Syllabic Grouping, pose four particular problems to be discussed that prevent the complete automation of the computational process and eventually require manual assessment: the potential under/over-spelling of Linear-A's words, the similar specific inaccuracies of the scripts and other limitations of the target languages, the misleading size of the Linear-A's corpus, and the combinatorial explosion of the retrieved by the searching results.

Before discussing the above problems, a few words are needed about how the method of Syllabic Grouping works. Having the character string (i.e., the letters) of a word, every vowel of the word is erased, unless it defines alone a syllable. Then, the remaining letters (both consonants and vowels) are grouped in corresponding phonetic families, and substituted by a characteristic letter for every group, in order to form the searching designation of this method. For example, looking at the first word of Table 1 (U-TA-LO), originally it consists of three (3) Linear-A's syllabograms that have been transliterated to their Latin form (i.e., U, TA and LO). The first syllable "U" is a stand-alone vowel that remains as it is, having itself as designation (the other letter of this group is "O" that, when discovered as a single syllable, is substituted by "U"). The second syllable "TA" loses its vowel "A" and the remaining consonant "T" is substituted by "D", which is the characteristic letter for the dental group of consonants, where "T" belongs to. Similarly, the third syllable "LO" loses "O", while "L" remains, since it is the characteristic letter of the liquid group of consonants (the other being "R", substituted by "L", whenever discovered). Thus, the designation "UDL" is formed. The same process is applied to all the words of the target languages, and the search-engine retrieves all of them having the same designation (i.e., UDL), as potential cognates of the Linear-A's word, to be evaluated manually. Although this method has been tested to verify the well-known linguistic affinity of the Anatolian languages (see section 2.1) with interesting results [71], it is neither a trivial one nor it is optimized yet (planned to be for the near future).

2.3.1 Under/Over-spelling

In the case of Linear-B, which is considered by the CPS theory as its derivative for Greek, an under-spelling of words is observed, i.e., adjacent consonants can be omitted, according to specific rules. For example, the Greek word "sperma" (seed) is written as PE-MA (i.e., with two syllabograms), by omitting the initial "s" in front of "p" and "r" in front of "m". In addition, final consonants were systematically omitted, as in the case of the word "Knossos" (name of the well-known capital city) that was written KO-NO-SO (i.e., with three syllabograms). Moreover, in the written name of this city over-spelling is also observed, since a sound "O" is inserted after the initial "K". Because of the limited usage of Linear-B as merely a commodity and proper-name record-keeping script of the palatial archives, under/over-spelling was not a problem for determining words of commodities, persons (anthroponyms), divinities (theonyms) or places (toponyms). Yet, since this may well be applied to Linear-A as well (also being a record-keeping script), finding cognates in the target languages might be inaccurate enough to require manual assessment.

2.3.2 Inaccuracies & Limitations

Similar problems to under-spelling (see 2.3.1 above) also exist in some target languages, like Carian (see 2.1.1 previously), but also over-spelling, because the script of some target languages depict more phonemes (especially long vowels) than necessary. In addition, the vocabulary of the target languages is small, therefore, not finding a cognate to a Linear-A word doesn't necessarily means that the target language lacks one.

2.3.3 Misleading Size

The database in [10] (see 2.2.1.j previously), which has been used to construct the MRD herein, contains a total of 1486 (maximum) inscribed artefacts of Linear-A. Nevertheless, only 89 of them are directly useful for decipherment purposes, and another 368 (maximum) could be useful under conditions. The misleading size of the original existing corpus is a serious limitation to decipherment attempts that requires a comparison to the equivalent in content and structure Linear-B artefacts, for discovering meaningful correlations, as suggested by other authors [72], [73].

2.3.4 Combinatorial Explosion

An artefact of Linear-A, designated PHZb4, contains two consecutive words: CƏ-MA NO-JA-TE. If the searching returns (e.g.) six (6) potential cognates for the first word and four (4) for the second in a given target language, then there will be 24 possible combinations of meanings for

assessment. Artefacts with many words must be studied in this way, creating a vast number of combinations to be manually assessed. Therefore, the construction of a MRD for the corpus of Linear-A was considered necessary, in order to facilitate the manual assessment.

3 Problem Solution

Since Syllabic Grouping is a semi-consonantal method (see subsection 2.3), over-spelling of vowels both in Linear-A and the target languages (see 2.3.1 and 2.3.2 previously) can be dealt with effectively. The combinatorial explosion problem (see 2.3.4 previously) cannot be dealt with automatically and thus auxiliary structures are required to facilitate the manual work of the researchers. Such a structure is the very first MRD of the entire useful corpus of Linear-A, described below.

The MRD of the corpus of Linear-A consists of three (3) spreadsheet files, implemented in MS Excel that can be directly accessed by modern computer programming languages. These files are the main one and two (2) indexing files that allow the indirect access from the main file to the MRD of the vocabulary of Linear-A (see 2.2.1.k previously). The following description is restricted to the main file, in order to avoid unnecessary and indifferent technicalities.

It was mentioned previously (see 2.3.3) that only a small part of the Linear-A's 1579 inscribed artefacts is useful for linguistic purposes. Indeed, the afore-mentioned artefacts have been classified in eight (8) categories [64]:

i. 41 fragmented artefacts that are not readable and, therefore, completely useless;
ii. 988 clay tokens of a single syllabograms that are also completely useless;
iii. 91 fragmented artefacts with a part of a single word that could be useful at a third phase of decipherment, only in combination with full words and then with larger texts, although this would have a very limited value, in general;
iv. 103 complete artefacts that contain a single full word in 122 instances (i.e., 19 words can be spelled in two ways), which may mean anything, but could be useful at a second phase of decipherment in combination with larger texts;
v. 17 complete artefacts with a few words in line, i.e., small full phrases not of record-keeping structure, that can be useful;
vi. 30 fragmented artefacts of the previous category (v) with partial phrases, useful only at a second phase of decipherment, when complete single words will be isolated from them and therefore falling to category {iv};
vii. 72 complete clay tablets of genuine record-keeping, in good condition;
viii. 144 fragmented/damaged clay tablets of the previous category (vii), partially useful if completed records will be determined and isolated, in comparison to the structure of the full tablets (vii).

The categories {v} and especially {vii} contain enough information to start a decipherment attempt, at a first phase.

Accordingly, in the main file of the MRD of the corpus of Linear-A, the last five categories (vi-viii) have been implemented, each occupying a corresponding sheet. The internal structure of every sheet is the same for every category, namely, every word(s) of a single artifact is/are placed in a single row. An example is given in Table 2, below, with the first part of the clay tablet designated HT8a (category {vii}).

Table 2. The first part of a clay tablet

8	HT8a	JeDi	(OLE+Ci)	[10]	BoCaLaCo

The first column contains a serial number for the internal reference of the MRD. The second column contains the formal designation of the particular artefact (HT8a). From the third column and afterwards the words and the other symbols of the tablet are placed in a sequence same as on the original tablet. Specifically:

- a word (JeDi) occupies the third column, being the first word of the actual tablet, which is either an anthroponym or a toponym;
- then, there are two symbols of commodities in parentheses (OLE+Ci) in the fourth column, already of known meaning (not to be deciphered);
- the numbers that denote the quantities of commodities are enclosed in square brackets, in this case [10], at the fifth column (already known, not to be deciphered);
- another word (BoCaLaCo) follows in the sixth column, and so on.

It should be noted here that the words in this MRD are written not with the conventional/usual styles (i.e., JE-DI or *je-di*), but with the style designed by Evangelos Papakitsos, which avoids having dashes inside the words to denote the separate syllabograms (the word JE-DI is composed of two syllabograms).

Here, the dashes inside a word denote a syllabogram that is not readable on the artefact, because of a damaged surface. Therefore, the preceding consonant of a syllabogram is written in the capital letter, while the following consonant in small-case one. Vowels that denote a syllable alone are written in capital letters. For example, the word conventionally written as U-TA-LO (see Table 1) with three syllabograms, herein it is written as "UTaLo", where the capital letters denote the beginning phoneme of a syllable and the small-case letters denote the next/closing phoneme of the syllable.

The various possible cognates of a single word from a particular target language are placed aligned vertically in the cells right below the word. This way, the tablet can be read as a single sentence, comparing the various combinations of possible cognates together, and therefore trying to extract reasonable meanings, if any.

4 Conclusion

In order to deal with the problem of lacking standardized digital resources for the decipherment of Linear-A, the study herein presented and briefly described a MRD of the linguistically useful corpus of Linear-A. Implemented in spreadsheets, in order to have the content directly accessible both by modern computer programming languages and by interesting individuals easily, the purpose of this MRD is to facilitate the scholars in their attempts to decipher the original script and potentially discover the conveyed language (or maybe languages).

The future research directions, apart from actually using this tool for decipherment, include the optimization of the searching method of Syllabic Grouping, in order to achieve more accurate results, and the addition of more target languages to be enquired.

References:

[1] Salgarella E. and Castellan S., SigLA: The signs of Linear A: A palæographical database, *Grapholinguistics and Its Applications*, Vol. 5, 2021, pp. 945–962.

[2] Salgarella E., *Linear A*, Oxford University Press (Oxford Classical Dictionary), 2022.

[3] Christidis A.-F., *History of the Ancient Greek language* (2nd reprint, in Greek), Institute of Modern Greek Studies, 2010.

[4] Chadwick J., *Linear B and Related Scripts*, University of California Press,1987.

[5] Finkelberg M., Uchitel A. and Ussishkin D., A Linear A Inscription from Tel Lachish (LACH Za 1), *Tel Aviv*, Vol. 23, 1996, pp. 195–207.

[6] Oren E., Olivier J.-P., Goren Y., Betancourt P.P., Myer G.H. and Yellin J., A Minoan Graffito from Tel Haror (Negev, Israel), *Cretan Studies*, Vol. 5, 1996, pp. 91-117.

[7] Sarianidi V.I., *Margiana and Protozoroastrism*, self-published, 1998.

[8] Woudhuizen F.C., *Minoan and Mycenaean Oversea's Contacts: The Epigraphic Evidence*, ÉTUDES tome LIII, 2009.

[9] Tan K.M.W.Y., *Understanding Linear A through the lens of maritime history during the Bronze Age* (Master's thesis), Nanyang Technological University, 2022.

[10] Hogan R. and Douros G., *Linear A Explorer*, available online at: lineara.xyz

[11] Davis B., *Introduction to the Aegean Pre-Alphabetic Scripts*, Center for Classics and Archaeology - University of Melbourne, 2010.

[12] Schoep I., The administration of neopalatial Crete: A critical assessment of the Linear A tablets and their role in the administrative process, *Minos: Revista de Filología Egea*, Vol. 17, 2002, pp. 1–230.

[13] Owens G.A., The common origin of Cretan hieroglyphs and Linear A, *Kadmos Bd.*, Vol. 35, No. 2, 1996, pp. 105–110.

[14] Papakitsos E.C., The Linear-A Syllabary in the Context of Cretan Protolinear Theory, *Bulletin of the Georgian National Academy of Sciences*, Vol. 15, No. 2, 2021, pp. 154-162.

[15] Papakitsos E.C. and Kenanidis I.K., Cretan Hieroglyphics: The Ornamental and Ritual Version of the Cretan Protolinear Script, *Anistoriton Journal*, Vol. 15 Essays, 2016, pp. 1-12.

[16] Brown E.L., The Linear A Signary: Tokens of Luvian Dialect in Bronze Age Crete, *Minos*, Vol. 27-28, 1992-1993, pp. 25-54.

[17] Gordon C., *The Semitic Language of Minoan Crete*, Bono homini donum, 1981.

[18] Owens G., Pre-Hellenic Language(s) of Crete: Debate and Discussion, *The Journal of Indo-European Studies*, Vol. 28, No. 1 & 2, 2000.

[19] Owens G., *Labyrinth: Scripts and Languages of Minoan and Mycenaean Crete*, Centre for Cretan Literature, 2007.

[20] Woudhuizen F.C., *Middle Bronze Age Luwian Hieroglyphic and Its Ramifications to Crete*, Proceedings of the Vth International Congresks of Hititology, 2002.

[21] Woudhuizen F.C., The Language(s) of Linear A: An Updated Review Article, *DO-SO-MO:*

Fascicula Mycenologica Polona, Vol. 6, 2005, pp. 95-121.
[22] Schrijver P., *Talking Neolithic: The case for Hatto-Minoan and its relation to Sumerian*, Proceedings of the Workshop on Indo-European Origins Monograph No. 65 (Journal of Indo-European Studies), 2019.
[23] Revesz P.Z., 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, pp. 306–335.
[24] Revesz P.Z., *Computational linguistics techniques for the study of ancient languages*, MATEC Web of Conferences 210, 03014, 22nd International Conference on Circuits, Systems, Communications and Computers (CSCC 2018), 2018.
[25] Revesz P.Z., *Computational Linguistics and Ancient Inscriptions*, University of Helsinki, 2019.
[26] Revesz P.Z., *Minoan and Finno-Ugric regular sound changes discovered by data mining*, 24th International Conference on Circuits, Systems, Communications and Computers (CSCC), 2020.
[27] Kvashilava G., *On Decipherment of the Inscriptions of Linear A in the Common Kartvelian Language: ku-ro and ki-ro*, ATINER's Conference Paper Proceedings Series LNG2018-0153, 2019.
[28] Gretchen E.L., *Fifteen, Linear -Script Signs with Correspondences in Japanese Kanji and Hiragana*, Academia.edu, 2018.
[29] Kuroda M., *Ancient Minoans and the Japanese (Linear A is Nihongo)*, Academia.edu, 2021.
[30] Braović M., Krstinić D., Štula M. and Ivanda A., A Systematic Review of Computational Approaches to Deciphering Bronze Age Aegean and Cypriot Scripts, *Computational Linguistics*, Vol. 50, No. 2, 2024, pp. 725-779.
[31] Adiego I.J., *The Carian Language*, Brill 2007.
[32] Shafer Robert, A Break in the Carian Dam, *L'antiquité Classique*, Vol. 34, No. 2, 1965, pp. 398-424.
[33] Haynes S., *Etruscan Civilization*, J. Paul Getty Museum, 2005.
[34] Szałek B.Z., *Tartessian, Etruscan, Linear A, Sumerian and related problems in the light of heuristics and cryptology*, Szczecin, 2015.
[35] Thomopoulos I., *Pelasgics* (commented edition, in Greek), Rigopoulos, 1994.
[36] Woudhuizen F.C., *Etruscan As a Colonial Luwian Language: The Comprehensive Version*, Publications of The Henri Frankfort Foundation (Dutch Archaeological and Historical Society), 2019.
[37] Benelli E., *Alphabets and language*, Etruscology - Walter de Gruyter, 2017.
[38] *Hittite language*, Encyclopedia Britannica, 2016.
[39] Gernot W., *Hurrian*, Cambridge University Press (The Ancient Languages of Asia), 2008.
[40] Soesbergen Peter G. van, *The Decipherment of Minoan Linear A* (Second revised edition, Volume II), INGRAMSPARK, 2022.
[41] Melchert C., *The Luwians*, Brill, 2003.
[42] Vertegaal A.J.J., *Voices in stone: Studies in Luwian historical phonology*, LOT (LOT dissertation series), 2020.
[43] Drews R., *Greater Anatolia and the Indo-Hittite Language Family*, Papers Presented at a Colloquium Hosted by the University of Richmond (Journal of Indo-European Studies, Anatolian languages), 2001.
[44] Dupont-Sommer A., *La stèle trilingue récemment découverte au Létôon de Xanthos: le texte araméen*, Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres (Année), 1974.
[45] Melchert H.C., *Lycian language*, Encyclopedia Britannica, 2011.
[46] *Anatolian languages summary*, Encyclopedia Britannica, 2021.
[47] Laflı Ergün and Kan Şahin Gülseren, *Archaeology and history of Lydia from the early Lydian period to late antiquity (8th century B.C.-6th century A.D.)*, EKVAM Colloquia Anatolica et Aegaea Congressus internationales Smyrnenses IX (Dokuz Eylül University – DEU The Research Center for the Archaeology of Western Anatolia), 2017.
[48] Van den Hout T.P.J., Melchert H.C. and Houwink P.H.J. ten Cate, *Anatolian languages*, Encyclopedia Britannica, 2015.
[49] Adiego I.-X. et al. (eds.), *Luwic dialects and Anatolian*, Institut del Pròxim Orient Antic (IPOA) (BARCINO MONOGRAPHICA ORIENTALIA, Vol. 12, Series Anatolica et Indogermanica), 2019.
[50] Obrador-Cursach B., On the place of Phrygian among the Indo-European languages, *Journal of Language Relationship*, Vol. 17, No. 3–4, 2020, pp. 233–245.
[51] Hamp E.P., On Some Gaulish Names in -Ant- and Celtic Verbal Nouns, *Ériu*, Vol. 27, 1976, pp. 1-20.
[52] Obrador-Cursach Bartomeu, *Lexicon of the Phrygian Inscriptions*, University of Barcelona

– Faculty of Philology – Department of Classical, Romance and Semitic Philology, 2018.
[53] *Phrygia*, Encyclopedia Britannica, 2023.
[54] Soesbergen Peter G., Thracian Personal, Ethnic and Topographic Names in Linear A and B, *Kadmos*, Vol. 18, No. 1, 1979, pp. 26-39.
[55] Gernot W., *Urartian*, Cambridge University Press (The Ancient Languages of Asia Minor), 2008.
[56] Godart L. and Olivier J.-P., *Recueil des inscriptions en Linéaire A: Addenda, corrigenda, concordances, index et planches des signes, 5*, Etudes crétoises, 21, 1985.
[57] Olivier J.P., Godart L. and Poursat J.-C., Corpus Hieroglyphicarum Inscriptionum Cretae, *Études Crétoises*, Vol. 31, 1996, pp. 1–447.
[58] Younger J.G., *Linear A*, 2023, available online at: https://web.archive.org/web/20240125202540/http://people.ku.edu/~jyounger/LinearA/
[59] Revesz P.Z., Parvez Rashid M. and Tuyishime Y., *The design and implementation of AIDA: Ancient Inscription Database and Analytics system*, Proceedings of the 23rd International Database Applications & Engineering Symposium (IDEAS '19), pp. 1–6, 2019 (https://doi.org/10.1145/3331076.3331117).
[60] SMID, *Program in Aegean Scripts and Prehistory*, Department of Classics, The University of Texas at Austin, 1979-1994 (https://repositories.lib.utexas.edu/items/084b49a3-8c8e-48eb-b582-68b066548817).
[61] Ferrara S., Kelley K., Montecchi B., Ottaviano L., Valério M., Santamaria A., Cartolano M. and Corazza M., *INSCRIBE 3D Interactive Web Viewer*, 2023b, available online at: https://www.inscribercproject.com/3d_viewer_home.php
[62] Petrolito T., Petrolito R., Perono Cacciafoco F. and Winterstein G., *Minoan linguistic resources: The Linear A digital corpus*, Proceedings of the 9th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH), pp. 95–104, 2015 (https://doi.org/10.18653/v1/W15-3715).
[63] Greco A., Flouda G. and Notti E., *The pa-i-to epigraphic project*, 2023, available online at: https://www.paitoproject.it/en/home-2/
[64] Mavridaki A., Zacharis N. and Papakitsos E.C., Enhancing a Software Platform for the Decipherment of Linear-A Inscriptions, *WSEAS Transactions on Information Science and Applications*, Vol. 22, 2025, pp. to be announced.
[65] Papakitsos E.C., *Lexical Data in Multilingual Context: Seeking Cognates through Syllabic Grouping*, The 14th International Scientific Conference "eRA 2021 - The SynEnergy Forum: in the field of Industry 4.0" (University of West Attica), 2021.
[66] Mavridaki A., Galiotou E. and Papakitsos E.C., Developing a software application for the study and learning of Linear a script, *Review of Computer Engineering Research*, Vol. 8, No. 1, 2021, pp. 8-13.
[67] Revesz P.Z., *A computational study of the evolution of Cretan and related scripts*, Mathematical Models and Computational Methods (Joint Proceedings of AMCSE-MMMAS-EAS), INASE Press, pp. 101–105, 2015.
[68] Revesz P.Z., 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.
[69] Karajgikar J., Al-Khulaidy A. and Berea A., Computational pattern recognition in Linear A, *hal-03207615*, 2021, pp. 1–18.
[70] Nepal A., Perono Cacciafoco F., Minoan Cryptanalysis: Computational Approaches to Deciphering Linear A and Assessing Its Connections with Language Families from the Mediterranean and the Black Sea Areas, *Information*, Vol. 15, 2024, p. 73 (https://doi.org/10.3390/info15020073).
[71] Mavridaki A., Papakitsos E.C. and Zacharis N., A Software Application for Enquiring the Affinity of Anatolian Languages, *Journal of Software Engineering and Simulation*, Vol. 9, No. 12, 2023, pp. 59-67.
[72] Vallance Janke R. and Solcà A., *High Correlation Linear A—Linear B vocabulary, grammar and orthography in Linear A*, KONOSO Press, n.d.
[73] Vallance Janke R., Mycenaean Linear B Rosetta Stone for Linear A Tablet HT 31, *Archaeology and Science*, Vol. 12, 2016, pp. 75-98.