information

Article

Minoan Cryptanalysis: Computational Approaches to Deciphering Linear A and Assessing Its Connections with Language Families from the Mediterranean and the Black Sea Areas

Aaradh Nepal [1,*] and Francesco Perono Cacciafoco [2,*]

[1] School of Computer Science and Engineering (SCSE), Nanyang Technological University (NTU), Singapore 639798, Singapore
[2] Department of Applied Linguistics (LNG), School of Humanities and Social Sciences (HSS), Xi'an Jiaotong-Liverpool University (XJTLU), Suzhou 215123, China
* Correspondence: aaradh001@e.ntu.edu.sg (A.N.); francesco.perono@xjtlu.edu.cn (F.P.C.)

Abstract: During the Bronze Age, the inhabitants of regions of Crete, mainland Greece, and Cyprus inscribed their languages using, among other scripts, a writing system called Linear A. These symbols, mainly characterized by combinations of lines, have, since their discovery, remained a mystery. Not only is the corpus very small, but it is challenging to link Minoan, the language behind Linear A, to any known language. Most decipherment attempts involve using the phonetic values of Linear B, a grammatological offspring of Linear A, to 'read' Linear A. However, this yields meaningless words. Recently, novel approaches to deciphering the script have emerged which involve a computational component. In this paper, two such approaches are combined to account for the biases involved in provisionally assigning Linear B phonetic values to Linear A and to shed more light on the possible connections of Linear A with other scripts and languages from the region. Additionally, the limitations inherent in such approaches are discussed. Firstly, a feature-based similarity measure is used to compare Linear A with the Carian Alphabet and the Cypriot Syllabary. A few Linear A symbols are matched with symbols from the Carian Alphabet and the Cypriot Syllabary. Finally, using the derived phonetic values, Linear A is compared with Ancient Egyptian, Luwian, Hittite, Proto-Celtic, and Uralic using a consonantal approach. Some possible word matches are identified from each language.

Keywords: Linear A; Minoan; cryptanalysis; computational linguistics; language decipherment

Citation: 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* **2024**, *15*, 73. https://doi.org/10.3390/info15020073

Academic Editor: Peter Revesz

Received: 27 December 2023
Revised: 21 January 2024
Accepted: 23 January 2024
Published: 25 January 2024

1. Introduction

In 1900, Sir Arthur Evans, while excavating the Knossos Palace in Crete, unearthed clay tablets with unknown scripts on them. The writings belong to a family of scripts used in Crete, Mainland Greece, and Cyprus [1]. Among the two, which showed many similarities, the older one, Linear A, was used between 1700–1450 BCE and is yet to be deciphered [2]. The other script, Linear B, which seems to be the grammatological offspring of Linear A, was deciphered in 1952 by Michael Ventris [3]. Linear A also served as a model for another script near the end of the Bronze Age: Cypro-Minoan, which was used by the pre-Greek inhabitants of Cyprus. Cypro-Minoan, in turn, served as a model for the Cypriot syllabary, a script used by the locals to write their own dialect of Greek [1].

Sir Evans' choice of the name 'Linear' stems from the fact that both Linear A and B consist of only lines inscribed in clay [4]. Since their discovery, however, Linear A inscriptions have also been found on artefacts such as vases, jewelry, and other objects in different locations including Cyprus, mainland Greece, Turkey [5], and other Aegean islands (Kea, Kythera, Melos, and Thera) [6]. The corpus, altogether, currently consists of about 7150 signs inscribed on 1427 artefacts [1].

Most decipherment attempts begin by attributing Minoan, the language behind Linear A, to a known language family. Scholars have hypothesized links between Minoan and the Indo-European languages, the Semitic languages, the Tyrsenian languages, and the Uralic languages, among others. However, most such arguments are met with skepticism, as these attributions only yield a limited number of meaningful results [7]. Furthermore, a major ('fatal') challenge in the decipherment process is our inability to 'read' Linear A. Although there exist reasonable justifications to assign Linear B phonetic values onto Linear A for this, such an approach still produces meaningless words and has not been proven to be reliable.

Recently, novel attempts to decipher the script have emerged, which usually involve an algorithmic component. In this paper, we propose the combination of two such approaches, by Loh and Perono Cacciafoco [7] and Revesz [8], and the reasons for this are two-fold. Firstly, we aim to overcome the limitations involved in provisionally assigning Linear B phonetic values to Linear A and, secondly, we aim to shed more light on the possible connections between Linear A and other writing systems and languages from the Mediterranean and the Black Sea areas.

This paper is organized as follows. Section 2 outlines the main challenges with deciphering Linear A, along with some past attempts. Section 3 introduces and gives an overview of our proposed approach. Section 4 consists of the methodological details of this approach. Section 5 presents the results obtained when comparatively assessing the writing systems (Carian Alphabet and Cypriot Syllabary) and the languages (Ancient Egyptian, Luwian, Hittite, Proto-Celtic, and Uralic), and Section 6 discusses their implications. Section 7 concludes the entire paper and invites further work.

2. Some of the Past Decipherment Attempts

The main challenge with deciphering Linear A begins with our inability to 'read' it. Since its phonetic values are unknown, analytical attempts are likely to be unproductive. To address this, one approach, as aforementioned, has been to assign the phonetic values of Linear B to Linear A. Not only is Linear B largely derived from Linear A, but there exist visual similarities among signs in these two systems. Hence, it is reasonable to approach the decipherment of Linear A by assigning to it the phonetic values of Linear B. However, as discussed, although this allows for the 'reading' of Linear A, it has not proven to be very reliable, as it has, so far, only produced meaningless words [9].

The other challenge lies in the fact that the language behind the Linear A signs is unknown. Attempts to link so-called 'Minoan' to other known languages have remained unsatisfactory. Among the numerous hypotheses, there appear to be four main language families that scholars argue have a connection with Minoan: Indo-European, Semitic, Tyrsenian, and the Uralic language families. Vladimir Ivanov Georgiev, one of the scholars who suggests an Indo-European connection, posits that Linear A tablets, specifically the ones from Hagia Triada, encode Ancient Greek. He also believes that the other Linear A documents were transcribing Hittite–Luwian [10]. Other scholars similarly suggest an Indo-European connection. Gregory Nagy, for instance, conducted a comparative analysis of Linear A and Linear B by looking into the visual compatibilities between them, demonstrating Minoan's Indo-European-like features [11]. Similarly, Gareth A. Owens, by using phonetic values from Linear B and the Cypriot Syllabary, postulated that Minoan could be related to Sanskrit or Latin [12]. Leonard R. Palmer, another prominent scholar, suggested the possibility of Minoan being an Anatolian language linked to Luwian [13]. Palmer's theory stemmed from the two historically reconstructed invasions of Crete and Greece by the Luwians during the time when Linear A was adopted. Furthermore, statistical techniques applied to the frequency analysis of symbols and grammatological comparisons have also been considered for studying Indo-European links. Most notably, Hubert La Marle used such techniques to derive conclusions that suggest an Indo-Iranian connection for Minoan [14–17]. These theories, however, have remained controversial and unproven. Palmer's work, specifically, was criticized for relying on his subjective interpretation of the tablets, which led to varied interpretations [7]. La Marle's work, similarly, has been

contested by John Grimes Younger due to questionable comparisons with various writing systems from different origins [18].

Other scholars argue for possible connections between Minoan and the Semitic language family. Cyrus H. Gordon, one of the first to propose this link, also assigned Linear B phonetic values to Linear A signs and discovered words in Linear A that appeared to be similar to words from the Semitic language family [19]. However, Gordon's approach was also met with skepticism. Critics have argued that because the matches identified were mainly vocabulary items, the reliability of the language family connection is compromised, as they could be Semitic lexical borrowings rather than examples of Linear A. Additionally, Gordon's methodology involved associating elements to several Semitic languages, such as Akkadian and Canaanite. The fact that the comparison was not carried out with one specific language led scholars to consider the Semitic hypothesis unsuccessful [20]. Jan Best's attempts at postulating Phoenician as the ancestor of Linear A were similarly countered by scholars who highlighted the lack of linguistic evidence supporting the Semitic link [21]. Eu Min et al. also investigated the plausibility of a Semitic link with their study of Linear A libation tables [20]. Although their research pointed towards a possible connection, the result was not significant enough—indeed, they produced negative results, which, in their conclusions, led to their exclusion of a Semitic option.

The third language family that has received consideration for its possible connections with Minoan is the Tyrsenian one. Helmut Rix was the scholar who theorized this language family's existence, which would include, according to him, Etruscan (spoken in central Italy between around 700 BC and 50 AD), Lemnian (spoken on the island of Lemnos around the VI Century BC), and Rhaetic (spoken in the Eastern Alps between the I millennium BC and the III century AD) [22]. Giulio Mauro Facchetti, one of the first to propose the connection, hypothesized relationships between Etruscan, Lemnian, and Minoan [23]. Facchetti suggested that Minoan could be the ancestor of the proto-Tyrsenian branch of languages from which Etruscan, Lemnian, and Rhaetic were derived. This also meant, then, according to Facchetti, that Minoan would be the ancestor of the Eteocretan branch, which he assumes is different from the other Cretan branch [24]. James Mellaart extended this work by positing a connection between Etruscan, Lemnian, and Rhaetic and pre-Indo-European Anatolian languages by studying Anatolian place names [25]. However, due to a lack of proper verification of the plausibility of connections among Etruscan, Lemnian, and Rhaetic [21], the Tyrsenian argument remains disputed.

Another approach to using Linear B phonetic values to attempt to decipher Linear A has involved an analytical interpretation of the symbols shared between the two writing systems by John Grimes Younger. Younger attempted to discover words and names by assigning the Linear B phonetic values to Linear A. He was able to recognize a few possible Linear A toponyms with a comparison with Mycenaean Greek along with a positional and frequency study of place names in Linear B (and Linear A) tablets [26]. However, this comparative examination between Linear A and Linear B, although logical and well-grounded, did not yield decisive results, unfortunately.

A recent decipherment approach [7] proposes comparing Linear A with other language families according to their grammatological elements through a 'brute force attack'. The method, originating from cryptanalysis, involves assigning Linear B phonetic values to Linear A and then comparing the consonant clusters of Linear A with the consonant clusters of other languages from the region. A set of dictionaries of various languages stored in spreadsheet files is used as the input for a Python program which performs this comparison. The 'consonantal' approach is declared to be effective, because consonant clusters are, presumably, more stable and consequently allow for the easier analysis of the morphological parts of a language.

Peter Z. Revesz [8] proposed another approach, which involves comparing Linear A to other writing systems visually, by using a feature-based similarity measure. This novel algorithm was employed to develop a new phonetic grid for Linear A, which was then used to generate a Minoan–Uralic dictionary. According to Revesz, he was able to translate

twenty-eight Linear A inscriptions. More recently, Revesz also pointed to archaeogenetic evidence that suggests Minoans may have originated in the Danube Basin and the Western Black Sea coastal area [27], which could further suggest a Minoan–Uralic connection, given earlier and newer publications that argue that the Proto-Uralic people once lived in the Black Sea area [28–31].

Another important challenge in deciphering Linear A, however, is simply that the corpus is very small. There are currently about 7150 Linear A signs inscribed on 1427 artefacts [1]. This is in contrast with the larger corpus of Linear B, comprising about 30,000 signs at the time it was deciphered [26].

3. Proposed Approach

First, Linear A is compared with other writing systems in the region using the feature-based similarity measure proposed by Revesz [8]. Second, the phonetic values of those writing systems that are visually similar are assigned to Linear A. These two steps are to ensure that the potential limitations involved with provisionally assigning Linear B phonetic values to Linear A are accounted for. Finally, Linear A is compared with other languages using the consonantal approach/'brute force attack' proposed by Loh and Perono Cacciafoco [7].

3.1. Writing Systems Compared with Linear A

The writing systems that will be compared visually to Linear A include the Cypriot Syllabary and the Carian Alphabet. Some scholars have previously assigned phonetic values from the Cypriot Syllabary onto Linear A for its analysis. Most notably, Owens [12] used phonetic values from the Cypriot Syllabary and Linear B to hypothesize possible links between Minoan and the Indo-European language family. Since Linear A was used as a model for Cypro-Minoan which, in turn, was used to model the Cypriot Syllabary, this paper aims to further explore the relationships between the two.

The Carian Alphabet, similarly, is argued to be linked to the Cretan Scripts' family which, among other writing systems, includes Linear A and Linear B [32]. Revesz discusses the possible connections between Old Hungarian and the Carian Alphabet using a feature-based similarity measure and postulates that the Carian Alphabet is an ancestor of Old Hungarian. Therefore, a possible link between the Carian Alphabet and Linear A is considered.

3.2. Languages Compared with Linear A

Adopting the 'brute-force attack' proposed by [7], the languages/language clusters compared using the consonantal approach include Ancient Egyptian, Luwian, Hittite, Proto-Celtic, and Uralic, which belong, largely, to three language families: Indo-European, Afro-Asiatic, and Uralic. Since a considerable number of decipherment attempts suggest the possibility of Minoan belonging to the Indo-European language family, this paper aims to explore this further, with Luwian and Hittite. With Ancient Egyptian, it aims to investigate possible connections between the Minoans and the Egyptians. Sir Arthur Evans posited that the interaction between Crete and Egypt began during the third millennium BC [33]. Archeological evidence also strongly suggests a link between the two. Thus, we propose a further analysis of the possible connections between their languages. We also include a comparison with Proto-Celtic, which, although it does not have an apparent relation to Linear A, allows us to leverage the unbiased and universally applicable 'brute-force' nature of the consonantal approach. Finally, we also aim to further explore the Minoan–Uralic connection mentioned above.

4. Methods

4.1. Deriving the Phonetic Values

We use the feature-based similarity measure proposed by Revesz [8] to derive a new phonetic grid for Linear A. It has the following components:

1. Similarity Function

To compute the similarity between any two symbols, we let $S = \{s_1, s_2, s_3, \ldots, s_n\}$ be a set of n symbols, $F = \{f_1, f_2, f_3, \ldots, f_m\}$ be a set of m elementary features, and $T:(S, F) \rightarrow \{0,1\}$ be a function that maps a symbol–feature pair with either 0 or 1, depending on whether that symbol has that feature. Then,

$$\mathrm{sim}(s_i, s_j) = \sum_{k=1, T(s_i, f_k)=T(s_j, f_k)}^{m} w_k$$

where w_i is a weight function that maps any feature i to a real number (the weight assigned to that feature).

2. Elementary Feature Set

The elementary feature set describes the feature of each symbol using a set of descriptors. Each feature corresponds to elements found in the symbols. Table 1 shows the elementary feature set used for this paper, which is based on the one developed by Revesz [8].

Table 1. Elementary features with their corresponding weights.

Elementary Feature	Weight
The symbol contains some curved lines	0.01
The symbol encloses some regions	0.01
The symbol has a slanted straight line	0.01
The symbol contains parallel lines	0.02
The symbol contains crossing lines	0.02
The symbol's top is a wedge	0.12
The symbol's bottom is a wedge	1.00
The symbol's right side is a wedge	0.33
The symbol contains a stem, that is, a straight vertical line that runs down the middle	0.03
The symbol's bottom has two legs	0.06
The symbol's bottom has three legs	0.09
The symbol contains a hair, a small line extending from an enclosed space	0.04
The symbol contains two triangles	0.33

3. Weight of Each Feature

In Revesz [8], the weight of all features is 1. However, in this study, a different set of weights for each feature is used. The weight of each feature is the inverse of its frequency of occurrence across all symbols in Linear A. In other words, a feature that exists in most symbols will have a lower weight compared to a feature that only exists in some. This means that sharing a rarely occurring feature is given more importance than sharing a commonly occurring one.

Table 1 illustrates the weight of each feature in the elementary set based on a frequency analysis performed for all features across all standard simple signs in Linear A (A001–A371) from GORILA (the Linear A corpus by Louis Godart and Jean-Pierre Olivier).

With the elementary feature set, a feature map is first computed for each symbol in Linear A, the Carian Alphabet, and the Cypriot Syllabary. The feature map demonstrates the existence of specific elementary features in the symbol. Each symbol in Linear A is then compared with each symbol in the Carian Alphabet and the Cypriot Syllabary to derive their similarity scores. This expectedly results in a large output. Hence, after the comparison, some criteria are necessary to keep only those symbol matches that are strongly correlated. In this paper, the following criteria are employed:

- The threshold for the similarity value given by the similarity function is set to 2.05. This means only two symbols whose similarity values are above or at 2.05 are considered potential matches;
- If there are multiple matches with the same similarity value, the tie is broken by manually analyzing the symbols;
- Matches that meet the threshold, but are visually dissimilar upon manual analysis are also not considered. Such a case could arise due to the limited number of features considered or because of the interdependence among features in the elementary feature set. For instance, for a symbol to contain a hair it must also enclose some 'region'.

Additionally, since the phonetic grid is derived via visual comparison, allographs in these writing systems are important for consideration. The Carian Alphabet, specifically, is composed of a few of them. For instance, 𐊂, 𐊏, 𐊍, and 𐊖 all have the same phonetic value. In this paper, all variants of symbols in the Carian Alphabet specified in the Unicode Standard are examined independently for comparison. In the case of both the Cypriot Syllabary and Linear A, only the standard signs are used.

4.2. Consonant Cluster Comparison

After assigning the new phonetic values, the comparison between the languages is performed using a Python program developed by a research team led by Francesco Perono Cacciafoco and Colin Loh at Nanyang Technological University (NTU), Singapore [34]. The program works by using two CSV files as the input. The first CSV file is a Linear A master list with three columns: 'Source' (the artefact that contains the Linear A word), 'New Format' (the Linear A word with phonetic values derived from the feature-based similarity measure), and 'Linear A' (the Linear A word with the vowels removed). The second CSV file contains a single column with all the dictionary words of the language being compared. The program then removes vowels from the words of all the dictionary words of the language being compared and compares each of them with words from the Linear A master list. It finally produces a list of exact matches found between Linear A consonant clusters and the consonant clusters of the language which is compared. These matches are finally compared manually, in turn, to dictionary entries in the selected language, to see whether a meaning can be assigned to them. If a meaningful entry is found, this is cross-referenced with the original Linear A tablet and a judgment is made as to whether it allows us to 'read' the tablet itself, or part of it.

5. Results

5.1. Phonetic Values for Linear A

Using the feature-based similarity measure, each symbol in Linear A is compared with every symbol in the Cypriot Syllabary and the Carian Alphabet, to derive the possible phonetic values of Linear A. Table 2 shows a sub-set of these comparisons, filtered using the criteria outlined in the Methodology. The last column indicates the writing system that the matched sign is assumed to belong to ('CS' denotes the Cypriot Syllabary, 'CA' denotes the Carian Alphabet). The phonetic values are transcribed using Latin/Roman letters. For the Carian Alphabet, they are based on the transcription system posited by Ignacio J. Adiego [35].

It is important to note that Linear A, being—plausibly—a syllabary, is likely not composed of pure consonants, unlike the Carian Alphabet. This poses a challenge with using the Carian Alphabet to derive the phonetic values of Linear A. Revesz [8] proposes that if a Linear A symbol corresponds to a Carian Alphabet symbol with the phonetic value /C/ (a consonant), then the Linear A symbol for some vowels will have a phonetic value of /CV/ (consonant/vowel). For instance, the Linear A sign 𐘀, which could match with the Carian Alphabet sign 𐊆, would have a phonetic value of /L/+/V/. Revesz then derives the value for this /V/ by searching for the "appropriate word to describe the meaning of the Linear A symbol" [8] in Uralic, Finno-Ugric, and Ugric vocabulary lists.

Table 2. Feature-based similarity scores for a sub-set of symbol pairs.

Linear A Sign		Matched CS/CA Sign		Phonetic Value	Similarity Score	Assumed Origin
AB01		TA		TA	2.07	CS
AB02		LO		LO	2.07	CS
AB03		PA		PA	2.07	CS
AB07		UUU2		Y	2.07	CA
AB08		UUU3		Y	2.07	CA
AB09		SE		SE	2.07	CS
AB11		B		B	2.07	CA
AB13		NE		NE	2.06	CS
AB17		RA		RA	2.07	CS
AB20		TI		TI	2.05	CS
AB22		U		U	2.07	CS
AB24		LD		L	2.05	CA
AB31		U		U	2.07	CA
AB34		D		D	2.07	CA
AB37		A		A	2.06	CA
AB39		E		E	2.06	CS
AB44		A		A	2.07	CS
AB46		X		C	2.06	CA
AB48		NG		NG	2.05	CA
AB50		S		S	2.06	CA
AB51		SU		SU	2.07	CS
AB54		UU		W	2.06	CA
AB55		E2		E	2.06	CA
AB59		R		R	2.06	CA
AB65		D2		D	2.07	CA
AB70		JA		JA	2.07	CS
AB77		Q		QU	2.07	CA
A302		RI		RI	2.07	CS
A304		TI		TI	2.07	CS
A306		XE		XE	2.06	CS
A309A		O		O	2.07	CA
A311		TT2		CH	2.07	CA
A312		L		L	2.05	CA
A314		MB		MB	2.07	CA
A318		G2		G	2.07	CA
A319		LD		L	2.07	CA
A325		T		T	2.07	CA
A326		NN		N	2.07	CA
A330		KU		KU	2.06	CS
A339		LE		LE	2.07	CS
A349		ST2		Z	2.07	CA
A351		PE		PE	2.07	CS
A355		KI		KI	2.07	CS

In this paper, no particular vowel was concatenated with the pure consonants, as it was assumed that they could have any value. Since this comparison of Linear A with other languages involves a consonantal approach, the lack of specific vowels does not entirely render the phonetic values obsolete.

5.2. Comparing Linear A with Other Writing Systems and Related Languages

With the phonetic values derived in Table 2, Linear A was compared with Ancient Egyptian, Luwian, Hittite, Proto-Celtic, and Uralic by using the consonantal approach proposed by Loh and Perono Cacciafoco [7]. The results derived from the operation of the Python program developed by the two scholars for this task, highlighted in the Methods section, are presented in Tables 3–7. Since the Python program yields a lot of matches for each of the languages, the results presented have been filtered manually, to ensure that only matches with a high likelihood of plausibility are kept for consideration.

Table 3. Python program results for Ancient Egyptian.

Matched Consonants	Linear A Cluster	Egyptian Word	Linear A Source	Meaning '|' Separates Different Meanings '?' Indicates that the Meaning Is Uncertain
nr	NE-RA	iner	HT10A	shell of an egg | gravel, stone
p	PA-[],]-PA	ipA	KN32b, KH 91	to make to fly, to fly | house, dwelling, harem
pr	PA-RA-[ ],]-PA-R	aper	ZA006b, KH 79 + 89	to be equipped, to be provided with, furnished (of a house) | a boat equipped with everything necessary and a crew
r	RA	ArA	ZA009	to go up, to embark in a boat, to bring, to be high
rp	R-PA	irp	HT104	wine | wine plant, vine | to rot, to decay, to ferment
rr	RI-R	irr	HT30	deaf (?) | grapes, grape seeds | a wine jar
ry	RI-Y	ary	HT28a, HT28b	he who goes up | light, fiery one | the name of a Dekan | a kind of fish | breeze, wind
yS	Y-SE	AyS	HT132, HT81, HT93a, HT85a	truce

Table 4. Python program results for Luwian.

Matched Consonants	Linear A Cluster	Luwian Word	Linear A Source	Meaning '|' Separates Different Meanings '?' Indicates that the Meaning Is Uncertain
ll	LO-LO	lalai	KE Wc 2b	take
p	PA-[],]-PA	pa	KN32b, KH91	protect (?)
r	RA	ura	ZA009	great
ry	RI-Y	ariya	HT28a, HT28b	raise | check, restrain (?)
t	TA	ta	HT86a, Wa 1031	step | arrive
tn	TA-NE	taini	HT95a, HT95b	of oil, oily
w	W	wi	HT98a, Wc 3019, HT97a	see | appear (?)

Table 5. Python program results for Hittite.

Matched Consonants	Linear A Cluster	Hittite Word	Linear A Source	Meaning '|' Separates Different Meanings '?' Indicates that the Meaning Is Uncertain
ll	LO-LO	lulu	KE Wc 2b	evenness, steadiness, stability, security
lr	L-R	luri	HT10B	loss, shortfall, decimations | loss of standing, comedown, disgrace, degradation
p	PA-[],]-PA	apa	KN32b, KH 91	that (one) | he, she, it | the one in question | thy, thine, your(s)
pr	PA-RA-[ ],]-PA-R	puri	ZA006b, KH 79 +89	lip | rim, edge, border
prl	PA-R-L	parala	HT122a, HT94b	something of wood used on sacrificed cattle, nom
prt	PA-R-TA	parta	PH31a	side, siding, partition
ps	PA-SE	pus	HT18, HT27b	diminish, fade, be eclipsed | be small, act petty, be pusillanimous
r	RA	ara	ZA009	belonging (or: proper) to one's own social group, communally accepted or acceptable, congruent with social order
rp	R-PA	arp	HT104	bad luck, setback, misfortune

Table 6. Python program results for Proto-Celtic.

Matched Consonants	Linear A Cluster	Proto-Celtic Form	Linear A Source	Meaning '|' Separates Different Meanings '?' Indicates that the Meaning Is Uncertain
lr	L-R	*liro	HT10B	sea (?)
n	[]-NE	*ne	ZA020	not
nr	NE-RA	*nero	HT10A	hero (?)
rr	RI-R, []-RI-R	*eriro	HT30	eagle
ry	RI-Y	*aryo	HT28a, HT28b	free man
sny	SU-NE-Y	*sniyo	HT19	spin, weave
t	TI, TI-[, TI-[]	*eti	HT28a, KH90, Wc 3015b	yet, still, but | beyond | also
tn	TA-NE	*tini	HT95a, HT95b	melt
wy	W-Y	*way	HT94b	woe, oh, alas
y	Y	*yo	HT85b, We 1023/ Wd 1024	which
yr	Y-R-[]	*yaro	ZA009	chicken, hen

Table 7. Python program results for Uralic.

Matched Consonants	Linear A Cluster	Uralic Form	Linear A Source	Meaning '|' Separates Different Meanings '?' Indicates that the Meaning Is Uncertain
n	[ ]-NE	une	ZA020	sleep, dream
nr	NE-RA	nure	HT10A	to press

Table 7. *Cont.*

Matched Consonants	Linear A Cluster	Uralic Form	Linear A Source	Meaning '|' Separates Different Meanings '?' Indicates that the Meaning Is Uncertain
pr	PA-RA-[ ]	para	ZA006b, KH 79 + 89	good
ps	PA-SE	pese	HT18, HT27b	to wash (head?)
r	RA	ora	ZA009	awl | squirrel
rp	R-PA	orpa	HT104	melt
sr	SU-[ ]-RA	sira	ZA018a	a k. of relative
t	TA	ta	HT86a, Wa 1031	this, that
t	TI, TI-[, TI-[]	tE	HT28a, KH 90, Wc 3015b	you
tn	TA-NE	tana	HT95a, HT95b	birch bark
w	W, W-[]	owe	HT98a, Wc 3019	door
wl	W-L	wElE	HT38	to understand

6. Discussion

Although the results could suggest possible links between Linear A and Ancient Egyptian, Luwian, Hittite, Proto-Celtic, and Uralic, the matches found are insufficient to yield conclusive evidence of any connection. For each compared language, the matches appear sparse and spread across multiple tablets. Additionally, the number of matches across the languages are similar, with certain Linear A words matching with words in all the languages, suggesting that the result is coincidental rather than indicative of concrete links.

The limited number of matches could be due to the phonetic values used for the comparison. The feature-based similarity measure, with the parameters utilized in this paper, was only successful in producing 43 matches for comparing Linear A with other languages. In contrast, since Linear A and B potentially share 92 similar signs, naturally the phonetic grid based on Linear B includes more signs. There are several reasons for the derived phonetic grid being small. Firstly, it could simply indicate a lack of concrete links between the scripts. Secondly, while the feature-based similarity measure allows for an analysis of different writing systems, it is not without its limitations. The method depends highly on the elementary feature set, and since we only had a few features, it is plausible to assume that certain important features may have been missed during the analysis. Additionally, a small feature set also increases the probability of finding multiple matches for any symbol with the same similarity scores, and breaking the tie becomes a challenging decision. In Revesz [8], for instance, the tie is broken by choosing the symbol that is earlier in the standard ordering of symbols.

It is important to note, additionally, that the limited number of matches could simply indicate a lack of connections between the languages. Most connection hypotheses, as discussed previously, have shown to be unsuccessful due to reasonable justifications. Considerations such as the temporal and spatial relations of the writing systems and languages are undeniably important factors. For instance, a Minoan and Luwian or Hittite (Indo-European) link could be considered unlikely due to temporal gaps, if the emergence of the Minoan civilization is believed to predate the arrival of Indo-European speakers to Anatolia.

The Combined Approach

Our approach aimed to leverage different characteristics of two computational methods of decipherment, in the effort to interpret Linear A. The feature-based similarity measure, for instance, has been considered effective for visual comparisons of writing systems. In [36], Barla et al. performed a feature analysis of Indus Valley- and Dravidian-connected scripts. They propose a novel elementary feature set consisting of six additional features on top of the one employed in this paper and generate heat maps for the different writing systems. Comparing their approach to our approach in this paper, we chose to use the same feature set as Revesz [8]. However, this choice is arbitrary and evidently influences the results obtained post analysis. Selecting a good elementary feature set is not straightforward and requires experimentation and further analysis. This suggests that although the approach seemingly aims to provide an objective way to compare writing systems, it is still subjective to an extent. For this paper, however, the approach has allowed us to account for biases that arise while assigning Linear B phonetic values to Linear A, which is also inherent in the so-called 'consonantal approach'. It is important to note, however, that visually similar symbols may not necessarily share the same phonetic values [37].

After the derivation of the phonetic values, the consonantal approach has enabled us to attempt a statistical analysis of a new rendition of Linear A with other languages from the region, resulting in a few matches. In [7], Colin Loh and Francesco Perono Cacciafoco outlined preliminary results using this consonantal approach with Linear B phonetic values, and they found matches across Hittite and Luwian. They posited that the Linear A cluster "PA-RE", from the document HT4 of *GORILA*'s volume 1, could be a possible match with "PARI" from the Luwian dictionary, which represents "forth", or "away". Due to the limited number of phonetic values derived in this paper, it is difficult to assess the effectiveness of the consonantal approach for the comparative analysis of languages. A limitation inherent in such an approach is the loss of information resulting from the removal of vowels. The matches may just be loanwords or purely coincidence. Furthermore, filtering the large number of matches generated by the Python program is not arbitrary and requires further consideration and study. The approach's effectiveness in performing a 'brute-force' analysis, however, is evident.

Overall, the combined approach is effective in a cross-language and cross-script analysis, albeit with some limitations inherent in the two approaches that have been combined. It is also worth noting that there are possible limitations with the combination as well. Due to the dependence of the consonantal approach on the feature-based similarity measure, it may be difficult to determine the plausibility of links between languages using this approach. Obtaining a low number of matches, for instance, could indicate a lack of connection between the languages, compared to using the consonantal approach or when the writing systems are compared visually. Hence, the combined approach necessitates a stepwise assessment of the results. If there are low matches when comparing the writing systems visually, the decision of whether these writing systems are appropriate for use with the consonantal approach must be made first.

7. Conclusions

Among the numerous attempts to decipher Linear A, some recent approaches involve a computational component. This paper aimed to combine two such novel methods to firstly account for the biases inherent in provisionally assigning Linear B phonetic values to Linear A and, secondly, to shed more light on the possible connections between Linear A and other writing systems and languages of the Mediterranean and the Black Sea areas. This paper also aimed to highlight some limitations inherent to such approaches. The first step in the combined approach involved a feature-based similarity measure to visually compare writing systems and the second involved using a consonantal approach to compare different languages. Although the writing system still remains largely undeciphered, by employing the combined approach some Linear A signs were found to be similar to signs from both the Cypriot Syllabary and the Carian Alphabet. Applying the phonetic values of those

similar signs to Linear A and comparing it with Ancient Egyptian, Luwian, Hittite, Proto-Celtic, and Uralic resulted in a few word matches between the languages. Although these could suggest possible connections, they are not significantly conclusive, due to the limited number of matches. Along with some limitations inherent to the combined approach, the small corpus still poses a major challenge in deciphering Linear A. However, our approach can be applied and used to compare any known writing system and language possibly connected to Linear A, removing our dependence on assigning Linear B phonetic values to Linear A and allowing for an unbiased analysis. Further research could investigate the use of the combined approach with other scripts and languages.

Author Contributions: Conceptualization, A.N. and F.P.C.; methodology, F.P.C. and A.N.; software, F.P.C. and A.N.; validation, F.P.C.; formal analysis, A.N.; investigation, F.P.C. and A.N.; resources, F.P.C.; data curation, F.P.C.; writing—original draft preparation, A.N. and F.P.C.; writing—review and editing, F.P.C. and A.N.; supervision, F.P.C.; project administration, F.P.C.; funding acquisition, F.P.C. All authors have read and agreed to the published version of the manuscript.

Funding: The tools and research developed for this paper were funded by a Singapore Ministry of Education (MOE) AcRF Tier 1 Research Grant (Grant Number 2017-T1-002-193—Principal Investigator: Dr Francesco Perono Cacciafoco), "Giving Voice to the Minoan People: The Decipherment of Linear A", 2018–2021.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data collected and used in this study are safely stored in physical data discs and in an online (private) data repository. They can be made freely and promptly available to any scholar interested in it upon request by email to the authors. The Python program mentioned in our study can be found at the following GitHub page: https://github.com/L-Colin/Linear-A-decipherment-programme (accessed on 21 November 2023).

Acknowledgments: We would like to acknowledge Colin Loh (National University of Singapore—NUS, Singapore) for all his work and help in the development of the software (https://github.com/L-Colin/Linear-A-decipherment-programme, accessed on 21 November 2023) used and implemented in the project which is the origin of this paper. This project was supported by Nanyang Technological University (NTU), Singapore, under the URECA Research Programme.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Davis, B. Introduction to the Aegean Pre-Alphabetic Scripts. *Kubaba* **2010**, *43*, 38–61.
2. Cadogan, G. *Palaces of Minoan Crete*; Barrie and Jenkins: London, UK, 1976; ISBN 978-0-214-20079-3.
3. Chadwick, J. *The Decipherment of Linear B*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2014; ISBN 978-1-107-69176-6.
4. Robinson, A.; Robinson, A. *Writing and Script: A Very Short Introduction*; Very Short Introductions; Oxford University Press: Oxford, NY, USA, 2009; ISBN 978-0-19-956778-2.
5. Mycenean Artifacts Found in Bodrum. Available online: https://www.hurriyetdailynews.com/mycenean-artifacts-found-in-bodrum--74114 (accessed on 21 November 2023).
6. Perono Cacciafoco, F. Linear A and Minoan: The Riddle of Unknown Origins. In Proceedings of the LMS Fieldwork and Language Analysis Group (FLAG) Meeting, School of Humanities (SoH), Nanyang Technological University (NTU), Singapore, 10 June 2014.
7. Loh, J.S.C.; Perono Cacciafoco, F. A New Approach to the Decipherment of Linear A: Coding to Decipher Linear A, Stage 2 (Cryptanalysis and Language Deciphering: A 'Brute Force Attack' on an Undeciphered Writing System). In *Proceedings of the Grapholinguistics in the 21st Century—2020, Part II*; Fluxus: Brest, France, 2020; pp. 927–943. ISBN 978-2-9570549-7-8.
8. Revesz, P. Establishing the West-Ugric Language Family with Minoan, Hattic and Hungarian by a Decipherment of Linear A. *WSEAS Trans. Inf. Sci. Appl.* **2017**, *14*, 306–335.
9. Godart, L. Du Lineaire A Au Lineaire B. In *Aux Origines de L'hellénisme: La Crète et la Grèce. Hommage à Henri van Effenterre*; Histoire Ancienne et Médiévale; Publications de la Sorbonne: Paris, France, 1984; pp. 121–128.
10. Georgiev, V.I. Les Deux Langues des Inscriptions Crétoises En Linéaire A. *Linguist. Balk.* **1963**, *7*, 1–104.
11. Nagy, G. Greek-like Elements in Linear A. *Greek Rom. Byzantine Stud.* **1963**, *4*, 181–211.
12. Owens, G. The Structure of the Minoan Language. *J. Indo Eur. Stud.* **1999**, *27*, 15–56.
13. Palmer, L.R. *Mycenaeans and Minoans: Aegean Prehistory in the Light of the Linear B Tablets*, 1st ed.; Alfred A. Knopf: New York, NY, USA, 1962.

14. La Marle, H. *Linéaire A: La Première Écriture Syllabique de Crète: Essai de Lecture*; Librairie Orientaliste Paul Geuthner: Paris, France, 1996; ISBN 978-2-7053-3641-7.
15. La Marle, H. *Linéaire A, la Première Écriture Syllabique de Crète: Éléments de Grammaire*; Librairie Orientaliste Paul Geuthner: Paris, France, 1997; ISBN 978-2-7053-3642-4.
16. La Marle, H. *Linéaire A, la Première Écriture Syllabique de Crète: L'histoire et la Vie de Crète Minoenne: Textes Commentés*; Librairie Orientaliste Paul Geuthner: Paris, France, 1998; ISBN 978-2-7053-3643-1.
17. La Marle, H. *Linéaire A: La Première Écriture Syllabique de Crète. Signes Rares, Textes Brefs, Substitutions 4*; Librairie Orientaliste Paul Geuthner: Paris, France, 1999; ISBN 978-2-7053-3644-8.
18. Younger, J.G. Linear A: Critique of Decipherments by Hubert La Marle and Kjell Aartun. Available online: https://people.ku.edu/~jyounger/LinearA/LaMarleAartun.html (accessed on 7 December 2023).
19. Gordon, C.H. *Forgotten Scripts: Their Ongoing Discovery and Decipherment*; Revised and Enlarged edition; Dorset Press: Dorchester, UK, 1987; ISBN 978-0-88029-170-5.
20. Eu Min, N.C.; Perono Cacciafoco, F.; Cavallaro, F.P. Linear A Libation Tables: A Semitic Connection Explored. *Ann. Univ. Craiova Ser. Philol. Linguist. Analele Univ. Din Craiova Ser. Ştiinţe Filol. Linguist.* **2019**, *41*, 51–63.
21. Perono Cacciafoco, F. Linear A and Minoan: Some New Old Questions. *Ann. Univ. Craiova Ser. Philol. Linguist. Analele Univ. Din Craiova Ser. Ştiinţe Filol. Linguist.* **2017**, *39*, 154–170.
22. Rix, H. *Rätisch und Etruskisch*; Institut für Sprachwissenschaft der Universität Innsbruck: Innsbruck, Austria, 1998; ISBN 978-3-85124-670-4.
23. Facchetti, G.M.; Negri, M. *Creta Minoica: Sulle Tracce Delle più Antiche Scritture d'Europa*; Archivum romanicum Biblioteca dell'Archivum romanicum; Ser. 2, Linguistica; Olschki: Firenze, Italy, 2003; ISBN 978-88-222-5291-3.
24. Facchetti, G.M. *Appunti di Morfologia Etrusca: Con Un'appendice Sulla Questione Delle Affinità Genetiche Dell'etrusco*; Biblioteca dell'"Archivum Romanicum"; L.S. Olschki: Firenze, Italy, 2002; ISBN 978-88-222-5138-1.
25. Mellaart, J. *The Neolithic of the Near East*, 1st ed.; Scribner: New York, NY, USA, 1975; ISBN 978-0-684-14483-2.
26. Younger, J.G. Linear A Texts & Inscriptions in Phonetic Transcription & Commentary. Available online: https://people.ku.edu/~jyounger/LinearA/ (accessed on 7 December 2023).
27. Revesz, P.Z. Data Mining Autosomal Archaeogenetic Data to Determine Minoan Origins. In Proceedings of the 25th International Database Engineering & Applications Symposium, New York, NY, USA, 7 September 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 46–55.
28. Krantz, G.S. *Geographical Development of European Languages*; P. Lang: Berlin, Germany, 1988; ISBN 978-0-8204-0800-2.
29. Makkay, J. An Archeologist Speculates on the Origin of the Finno-Ugrians. *Mank. Q.* **2003**, *43*, 235–272. [CrossRef]
30. Wiik, K. The Uralic and Finno-Ugric Phonetic Substratum in Proto-Germanic. *Linguist. Ural.* **1997**, *33*, 258–280. [CrossRef]
31. Revesz, P. Inscription on a Naxian-Style Sphinx Statue From Potaissa Deciphered as a Poem in Dactylic Meter. *Mediterr. Archaeol. Archaeom. MAA* **2023**, *23*, 1–15.
32. Revesz, P.Z. The Cretan Script Family Includes the Carian Alphabet. *MATEC Web Conf.* **2017**, *125*, 5019. [CrossRef]
33. Marinatos, N. The Indebtedness of Minoan Religion to Egyptian Solar Religion: Was Sir Arthur Evans Right? *J. Anc. Egypt. Interconnect.* **2010**, *1*, 22–28. [CrossRef]
34. Loh, J.S.C. Linear-A-Decipherment-Programme. Available online: https://github.com/L-Colin/Linear-A-decipherment-programme (accessed on 21 November 2023).
35. Adiego Lajara, I.-J. *The Carian Language*; Handbook of Oriental studies; Section one, The Near and Middle East; Brill: Leiden, The Netherlands; Boston, MA, USA, 2007; ISBN 978-90-04-15281-6.
36. Barla, S.S.; Alamuru, S.S.S.; Revesz, P.Z. Feature Analysis of Indus Valley and Dravidian Language Scripts with Similarity Matrices. In Proceedings of the International Database Engineered Applications Symposium, Budapest, Hungary, 22 August 2022; ACM: Budapest, Hungary, 2022; pp. 63–69.
37. Yao, Y.; Perono Cacciafoco, F.; Cavallaro, F.P. The Minoan Challenge: An External Analysis Approach to the Linear A Decipherment. *Ann. Univ. Craiova Ser. Philol. Linguist.* **2022**, *44*, 456–475. [CrossRef]