A New Approach to the Decipherment of Linear A, Stage 2

Cryptanalysis and Language Deciphering: A "Brute Force Attack" on an Undeciphered Writing System

Loh Jia Sheng Colin · Francesco Perono Cacciafoco

Abstract. This paper discusses the attempt of an algorithmic approach to contribute to the decipherment of Linear A. With the assistance of software developed in Python, Linear A clusters can be compared to various dictionaries of languages respecting a certain degree of chronological and geographical compatibility, as a "brute-force attack" for the reconstruction of new clusters of Linear A symbols.

1. Introduction

Linear A is a writing system of the Ancient Aegean Minoan Civilisation of Crete that was in use between approximately 1850 and 1450 BCE (Olivier, 1986). Linear B is a syllabic writing system partly derived from Linear A that was used to transcribe Mycenaean Greek. Both Linear A and Linear B have been discovered by British archaeologist Sir Arthur John Evans during excavations between 1886 and 1901 (Chadwick, 1967). The terms "Linear A" and "Linear B" were formulated by Evans based on the linear structure of inscriptions found on the tablets, in contrast to the pictographic writings that were used extensively during the same time period. Linear A samples were discovered to occur in

---

This study has been funded by the MOE AcRF Tier 1 Research Grant 2017-T1-002-193 *Giving Voice to the Minoan People: The Decipherment of Linear A* (Perono Cacciafoco).

Loh Jia Sheng Colin 0000-0001-7009-5007
School of Physical and Mathematical Sciences,
Nanyang Technological University, Singapore
jloh035@e.ntu.edu.sg

Francesco Perono Cacciafoco 0000-0002-0977-063X
Linguistics and Multilingual Studies Programme,
School of Humanities,
Nanyang Technological University, Singapore
fcacciafoco@ntu.edu.sg

Y. Haralambous (Ed.), *Grapholinguistics in the 21st Century 2020. Proceedings*
Grapholinguistics and Its Applications (ISSN: 2681-8566, e-ISSN: 2534-5192), Vol. 5.
Fluxus Editions, Brest, 2021, pp. 927–943. https://doi.org/10.36824/2020-graf-cacc
ISBN: 978-2-9570549-7-8, e-ISBN: 978-2-9570549-9-2

various locations such as Crete, Aegean islands of Kea, Kythera, Melos and Thera, and even mainland Greece (Olivier, 1986).

Currently, the Linear A corpus consists of approximately 1,400 artefacts with Linear A inscriptions, with signs appearing over 7,400 times. A large majority of these Linear A signs was used for administrative documentations found on tablets, roundels and seals (Schoep, 2002). Hypotheses regarding the origin of Linear A script and of the Minoan language include the Luwian Hypothesis, the Semitic Hypothesis and more. These will be discussed in later sections of this paper.

Despite numerous attempts by scholars and glyph-breakers, Linear A continues to remain undeciphered, as researchers' attempts to attribute a language family relation with Linear A have provided only a limited amount of meaningful results. This paper explains the use of a programme written in Python programming language (Eu, Xu, and Perono Cacciafoco, 2019) to isolate potential Linear A clusters for further analysis, with the intention of identifying potential family relations with Linear A and other dictionaries of languages.

2. Selected Literature Review

Among the attempts made to decipher Linear A, many involve comparison with Linear B (Petrolito, Petrolito, Perono Cacciafoco, and Winterstein, 2015). Linear B, a syllabic writing system used in Crete, has been deciphered by the architect and philologist Michael Ventris in 1952, with the assistance of the linguist John Chadwick (Chadwick, 1967). They discovered that Linear B was used to transcribe Mycenaean Greek. Given that a large portion of Linear B was a derivation of Linear A, coupled with similarities between signs in both writing systems, there were reasonable justifications to support the conclusion of provisionally assigning Linear B phonetic values to Linear A signs that appear graphically similar. However, attempts to use Linear B directly for deciphering Linear A were unsuccessful. The results appear to be inconclusive as large amounts of 'meaningless' Linear A words were generated (Godart, 1984). As such, the direct use of Linear B alone might prove to be insufficient to decipher Linear A, thus motivating the studies of a variety of languages that are chronologically and geographically mutually compatible as possible relations to Linear A.

There have been studies carried out to decipher Linear A through potential association of Linear A with other language families. Vladmir Ivanov Georgiev believed that Linear A has connections with Greek. Based on his work published in 1952, Georgiev speculated that the Linear A inscriptions found on Haghia Triada tablets were transcribing Greek. Several studies have also highlighted that Linear A was associated with the Indo-European language family. Studies by Gareth A.

Owens (1999) have postulated the possibility that Linear A might belong to the Indo-European language family with relations to Greek, Sanskrit and Latin. Leonard R. Palmer (1961) proposed that Linear A could be an Anatolian language, possibly Luwian. Palmer posited his hypothesis due to a possible historical event which indicates the invasion of Crete and Greece by Indo-European peoples around the second millennium BC (ibid.), necessitating the migration over to Crete. Similarly, Gregory Nagy (1963) proposed a hypothesis that Linear A was largely similar to Luwian, a language that belongs to the Anatolian branch of the Indo-European language family.

Theories that propose a relationship between Linear A and Luwian have been constantly challenged and remain controversial in the academic community. Critics have indicated that Palmer's work is highly dependent on the interpretations of the Linear A tablets that can entail varied interpretations attributed to a limited understanding of Linear A orthography.

Linear A has also been associated with the Semitic language family. Semitic languages originate from the East of the Mediterranean and are speculated to be in use from around 3800 to 3500 BCE (Olivier, 1986). Based on chronological and geographical facts, Linear A is regarded to be compatible with the Semitic language family. Cyrus H. Gordon (1982), a scholar with an extensive knowledge of Semitic languages, was one of the first to attempt deciphering Linear A through comparison with the Semitic language family. By assigning Linear B phonetic values onto Linear A signs, Gordon was able to identify certain words present in Linear A that appeared to be largely similar to that of those belonging to the Semitic language family, such as Hebrew and Akkadian. Thus, he drew the conclusion that Linear A might be connected to Semitic languages, and West Semitic ones, in particular.

In support of the Semitic Hypothesis, archaeological evidence reveals trade practices by the Minoans over Semitic-speaking languages like Eastern Mediterranean and findings of Minoan-style artefacts discovered in areas that include Cyprus and Canaan (Bradley, 2014). This evidence supports the hypothesis that there was language contact between the language of Linear A and Semitic language family.

Nevertheless, many scholars continue to remain sceptical about Linear A having a language contact with the Semitic language family. Sceptics have argued that Gordon's work on the comparison of Linear A and the Semitic language family has major flaws. Firstly, the matches which Gordon has identified appear to be mainly vocabulary terms. Next, Gordon has adopted a methodology of associating elements from various Semitic languages, like Akkadian and Canaanite, for comparison with Linear A. However, the act of doing so has suggested that the Semitic hypothesis might not be conclusive due to the lack of linguistic evidence.

While some scholars continue to argue against Semitic being a possible family for the language represented by Linear A script, there are still recent studies on the decipherment of Linear A suggesting a possible Semitic connection. Eu, Perono Cacciafoco, and Cavallaro (2019) studied Linear A libation tables, in an effort to identify Semitic roots present in recurrences found in these tables. Eu Min et al. obtained very limited and sporadic matches that do present evidence for the relation of Linear A to the Semitic language family (Eu, Perono Cacciafoco, and Cavallaro, 2019).

The small number of such matches could be attributed to the small sample size of Linear A artefacts to study on, as well as the possibility that Linear A was used to encode a ritual language, involving a more complicated writing style.

American scholar John G. Younger set up a Web application[1] in 2000 dedicated to his work on Linear A, in an effort to provide better access to Linear A resources to other scholars. With the use of Linear B Syllabary, Younger has managed to transcribe most Linear A signs. Furthermore, Younger has attempted a reconstruction of Linear A based on the analytical interpretation of possible shared symbols between Linear A and Linear B. Through such forms of development, Younger was able to locate possible toponyms in Linear A that are pre-Greek but compatible with Linear B transcriptions.

Apart from the above-mentioned hypotheses, there are also other possible connections proposed by scholars. Margalit Finkelberg (2001) compared the phonological and morphological profiles of Minoan with those of other languages, in order to narrow down the range of possible languages associated with Linear A. The morphological profile of Minoan was compared to Greek, Lydian, Hittite, Luwian and Lycian. This comparison was based on the preliminary readings of certain Minoan texts. Consequently, Finkelberg proposed that Linear A might be an ancestor of Lycian or possibly a distant relative of it.

In recent years, studies on Linear A as well as the decipherment attempts made by scholars have included an algorithmic approach, with the introduction of high-computational power. Perono Cacciafoco (2017) has proposed the possibility of having the Linear A inscriptions analysed beyond the grammatical level, but through comparisons with other languages and language families according to the Linear A grammatological elements. This novel approach of deciphering Linear A could result in the reconstruction and recombination of new clusters of Linear A, based on the languages compared with the Minoan language. Peter Z. Revesez (2017) has proposed the language of Linear A to be connected with the Uralic language family. In his study, Revesez has introduced an algorithm to obtain the 'syllabic values of Linear A

1. http://people.ku.edu/~jyounger/LinearA/.

signs'. With these 'syllabic values', Revesez constructed a dictionary of Uralic-Minoan language that translates Linear A documents from the corpus: published by Louis Godart and Jean-Pierre Olivier in the 70s and 80s of the 20th century, GORILA is a Linear A corpus containing Linear A inscriptions. GORILA contains five volumes (Godart and Olivier, 1976a,b; 1979; 1982; 1985) and was made available digitally on the Web in the 21st century[2]. Like many previous attempts of deciphering Linear A, Revesez's work was heavily scrutinised as cross-family comparisons were not conducted, and the pertinent issue of varied interpretations of the 'syllabic values of Linear A' was not properly addressed in Revesez's work.

While it may appear that deciphering Linear A is an insurmountable challenge, it is often beneficial to consider other perspectives in approaching such problems (Tan, 2018), as an alternative means to attain a solution.

3. Methodology and Preparation of Materials

In contrast to past attempts to decipher Linear A, a comprehensive analysis and comparison of Linear A clusters should be conducted beyond the grammatological level, through methods originating in cryptanalysis. By analysing the combinatory data and comparison frequencies with language families of different natures, new possible clusters of Linear A words can be reconstructed and recombined. This project adopts and expands on the method of a "brute-force attack" to attempt deciphering Linear A, through the use of a programme developed in Python programming language (Eu, Xu, and Perono Cacciafoco, 2019).

For the comparison of Linear A with other languages, a set of documents containing dictionaries of various languages and language families has been generated digitally as input into the programme functions. These documents have data stored in spreadsheet files in order to be easily editable by a human. First, a compiled master list of transcribed Linear A words has been created out of the collections of text and Linear A samples contained in Olivier and Godart's Corpus of Inscriptions in Linear A, namely GORILA vols. 1 to 5 (Godart and Olivier, 1976a,b; 1979; 1982; 1985). The transcription of Linear A words has involved assigning phonetic values of Linear B symbols that appear compatible with the Linear A characters. Symbols in Linear A that do not resemble graphically to any of the symbols present in the Linear B syllabary have been replaced by three-digit numbers. For example, in Haghia Triada tablet HT1, there is a Linear A word transcribed by KU-[AB056]-NU, in

2. http://mnamon.sns.it/index.php?page=Risorse&id=19.

which the three-digit number '056' is used to denote that unique Linear A symbol (see Fig. 1, where KU, AB056 and NU correspond to 𐙁, 𐘯 and 𐘮 resp.).

FIGURE 1. Pictorial representation for Linear A transcriptions of the HT1 tablet. Source: GORILA vol. 1 (Godart and Olivier, 1976a, p. 3)

A newly reconstructed digital corpus of Linear A signs is used to complement the database of Linear A signs and symbols present in GORILA vols. 1 to 5 (Godart and Olivier, 1976a,b; 1979; 1982; 1985). This allows more rigorous statistical analysis of Linear A signs and sign sequences, as comparative analysis will become much more efficient as compared to the printed version of Godart and Olivier's corpus.

In the future, dictionaries of languages from different language families will be required for comparison with the Linear A clusters. Examples of such dictionaries would include Luwian, Anatolian, Hamito-Semitic and Hittite.

4. Overview of the Programme Functions

Our programme aims at the implementation of a "brute-force attack" method as an attempt to decipher Linear A. In the context of cryptography and cryptanalysis, a "brute-force attack" is defined to be "an exhaustive search method in order to recover the secret key in a cryptosystem by testing all possible combinations" (Verdult, 2015). Through clever optimisations and after a sufficient amount of computation, we aim to associate a possible language family to the Linear A writing system. The "brute-force attack" on the Linear A symbols can help in constructing phoneme *n*-gram databases with respect to the languages of comparison. Consequently, the Python programme aims to develop an automatic procedure for evaluating language sources that would be of "best fit" to Linear A. For the development of the programme, the 'pandas' and 'PyQt' Python modules have been used extensively for the purpose of data analytics and the creation of a graphical user interface (GUI) for the programme. In particular, the programme can be segmented into two portions—the specific decipherment approach (§ 4.3) and the general decipherment approach (§ 4.1). Aside from the aforementioned two portions, we have also incorporated the use of the Linear A fonts into the study of potential clusters of Linear A words found in various Linear A tablets.

4.1. General Decipherment (GD) Approach

In the GD function, users can input any spreadsheet file available on their computers, provided data are stored only in the first column of the spreadsheet file and the file is in CSV (comma-separated values) format. Using the pandas module, the programme carries out similarity comparisons between the words present in the uploaded file with those of the Linear A master list. Results generated through the GD function are displayed in a clear table format consisting of four columns, titled "Identical Matches," "Linear A word," "Original Word" and "Source". An example can be seen in Figure 2, in which there are 19 matches obtained from the programme's GD function. In particular, there are 6 matches for character "r," each of which originates from a different source, including ZA011b, HT27b and HT85b.

The results of word comparison can be downloaded locally as CSV files. Allowing an indiscriminate comparison between words from various language dictionaries and lexical lists and the Linear A master list, we can provide a large-scale brute-force attack on the current Linear A corpus while the decipherment of Linear A would involve a more statistical approach.

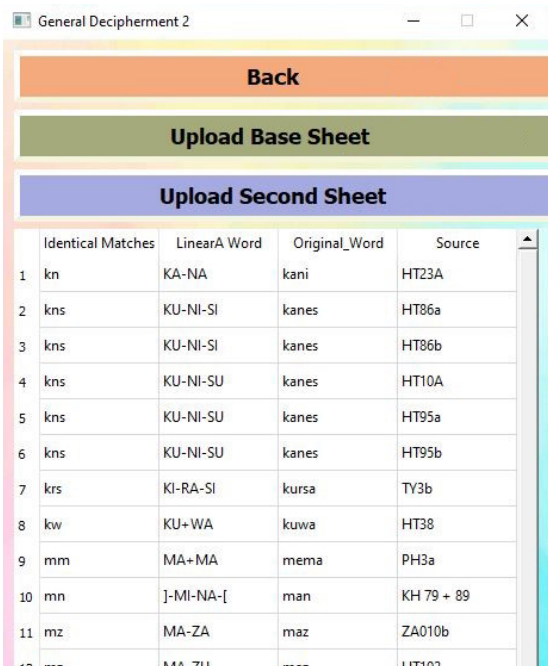
4.2. Modifications to GD

A modified version of GD, named "General Decipherment 2" (GD 2), will be implemented in the programme. By using the GD 2 function, users will be able to make dynamic changes in their comparisons and alter, not only the dictionary list, but also the Linear A master list—these

	Identical Matches	LinearA Word	Original_Word	Source
1	r	RA2	ara	ZA011b
2	r	RE	ara	HT27b
3	r	RO	ara	HT85b
4	r	RO	ara	Wa 1032-1121
5	r	RO	ara	Wb 2002
6	r	]RU	ara	MA 10
7	mn	]MINA[	man	KH 79 + 89
8	nw	NUWI	nawa	HT115b
9	ns	NASI	nis	AP ZA 2
10	ns	NASI	nis	HT28b
11	ns	NUSE[	nis	KN Wa 33b
12	pr	PARA	pari	HT128a
13	pr	PARA	pari	PH3a
14	pr	PARE	pari	HT 4
15	pr	PURA	pari	HT116a
16	pr	PURE	pari	PK ZA 11
17	srr	SARARA	sarra	HT30
18	srr	SIRERE	sarra	PH31a
19	ws	WISA	wasu	HT113

FIGURE 2. GD results of an unknown dictionary

two will be referenced to as “Comparison Sheet” and “Base Sheet” respectively. This form of comparison will be more beneficial in carrying out a more comprehensive analysis on the clusters of Linear A words, by incorporating the findings in the frequency analysis of the 3-digit numbers available in the SD portion of the programme. The frequency analysis of these numbers will be discussed in later sections of this paper. The only restrictions on the “Base Sheet” is that it is stored in CSV format. In the CSV file, the programme only uses the data stored in the first three columns: “Source,” “New Format” and “Linear A word”. Results obtained from this function will be available for download in CSV format. An example of the results obtained from GD 2 can be seen in Figure 3, where there are 5 matches for the string “kns,” obtained from the programme's GD 2 function. These matches originate from different sources, including HT86a, HT86b and HT10A.



	Identical Matches	LinearA Word	Original_Word	Source
1	kn	KA-NA	kani	HT23A
2	kns	KU-NI-SI	kanes	HT86a
3	kns	KU-NI-SI	kanes	HT86b
4	kns	KU-NI-SU	kanes	HT10A
5	kns	KU-NI-SU	kanes	HT95a
6	kns	KU-NI-SU	kanes	HT95b
7	krs	KI-RA-SI	kursa	TY3b
8	kw	KU+WA	kuwa	HT38
9	mm	MA+MA	mema	PH3a
10	mn	]-MI-NA-[	man	KH 79 + 89
11	mz	MA-ZA	maz	ZA010b

FIGURE 3. GD 2 results of an unknown dictionary and modified Linear A master list

Allowing the users with dual choices of file input into the programme will provide a higher degree of freedom for more efficient research and analysis to be conducted on any possible cluster of Linear A words.

4.3. Specific Decipherment (SD) Approach

Under the SD function, numerous dictionaries from different language families are incorporated into the programme for comparison with the Linear A words present in the master list. Using a 'consonantal' approach for analysis, in which vowels occurring in words in the dictionaries will be provisionally removed, the programme will carry out a search in the master list to have a one-to-one character match between the Linear A clusters and the modified strings in the dictionaries.

This comparison adopts a consonantal approach that has been introduced for use with the Semitic family languages and Afro-Asiatic lan-

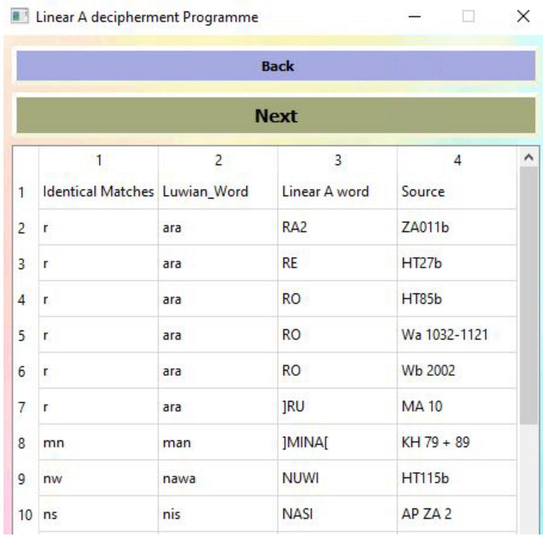
FIGURE 4. Digital recreation of the Linear A tablet HT 95b. The Linear A word "KU-NI-SU" is circled in red (Perono Cacciafoco and Cavallaro, 2020)

guages such as Ancient Egyptian. The approach has also been adopted for comparison for Indo-European languages. This is because consonantal clusters are more stable and consequently it is easier to highlight morphological parts of lexical items of a language and of their roots. In this aspect of the programme, identical character matches will be collated and displayed in a clear table format with the same four columns as the GD functions. An example of the result can be seen in Figure 5. Through such a comparison, dictionaries with high incidence of similarities can be quickly identified and isolated for further analysis.

As shown in Figure 5, the programme reads the results in four columns. Upon a one-to-one character match between the modified Luwian words as well as the Linear A clusters, the first column would output the results under 'Identical Matches'. According to the results, the SD function would select the original Luwian word, Linear A words as well as the source of the Linear A word. These would be available in the next three columns, respectively.

Aside from the direct comparison with words between Linear A and dictionaries, the programme function also conducts a series of frequency analyses. By allowing the 'triplets' present in Linear A clusters to use any possible character ranging from A to Z, the programme will analyse the frequency in which a character replacing the 3-digit number would achieve a one-to-one character match. An example of the results of the analysis of 3-digit numbers can be seen in Figure 6.

In Figure 7, we have the analysis of the triplets' one-to-one correspondence with the characters ranging from A to Z. It indicates that between the Thracian dictionary and Linear A master list, the triplet '029' has a one-to-one match with <a> one time, <b> three times, <c> five times, etc.



Linear A decipherment Programme

Back

Next

	1	2	3	4
1	Identical Matches	Luwian_Word	Linear A word	Source
2	r	ara	RA2	ZA011b
3	r	ara	RE	HT27b
4	r	ara	RO	HT85b
5	r	ara	RO	Wa 1032-1121
6	r	ara	RO	Wb 2002
7	r	ara	]RU	MA 10
8	mn	man	]MINA[	KH 79 + 89
9	nw	nawa	NUWI	HT115b
10	ns	nis	NASI	AP ZA 2

FIGURE 5. SD comparison results between Luwian dictionary and Linear A master list

FIGURE 6. Digital recreation of the Linear A tablet HT 27b. The Linear A cluster "RE" is circled in red (Perono Cacciafoco and Cavallaro, 2020)

The study of such frequency analyses can serve to justify the replacement of the 'triplet' with the character of the highest frequency of identical match. This process enables the reconstruction of possible new clusters of Linear A words, beyond the available clusters from inscriptions.

	1	2	3	4	5	6	
1	Triplets	a	b	c	d	e	f
2	056	2	1	0	0	0	1
3	021	1	0	2	0	0	1
4	329	0	0	0	0	1	0
5	301	3	4	4	0	0	3
6	029	1	3	5	0	0	1
7	188	1	3	1	1	0	0
8	076	1	3	3	2	2	2
9	123	1	0	4	0	0	1
10	305	1	1	0	1	0	1
11	312	0	1	1	0	0	1
12	100	0	0	1	1	0	0

FIGURE 7. SD 3-digit number analysis between a Thracian dictionary and the Linear A master list

4.4. Introduction of Linear A Fonts

The Linear A fonts introduced into the programme were created by Sabrina Soo Su-Ann, a fellow member in the Nanyang Technological University Linear A decipherment team. They expand the Unicode chart of Linear A, incorporating variants of Linear A signs.

The introduction of Linear A fonts will be represented by a function known as `LinAfontConv` in the programme. Through it, users can input their own files in Microsoft Excel Open XML Spreadsheet (XLSX) file format, provided they contain four columns labelled "Identical matches," "Linear A Word," "Original Word" and "Source" as shown in Figure 6. The function "LinAfontConv" will consider information under the "Linear A word" column, and display the output as the Linear A characters under another column titled "Linear A fonts". Similarly to the other functions in the programme, users can download the output file, which now contains the additional column "Linear A fonts," locally into their computers, in XLSX format.

Through such forms of linguistic analysis, we are able to identify and study potential clusters using the Linear A characters present in the tablet inscriptions. The incorporation of Linear A fonts will provide a better understanding of how Linear A was used by the Aegean Minoan people as a writing system when compared with various languages and language families.

1	Identical M	Linear A Word	Original Word	Source	Linear A fonts
2	pr	PARA	pari	HT128a	𐘂𐘴
3	pr	PARA	pari	PH3a	𐘂𐘴
4	pr	PARE	pari	HT 4	𐘂𐘙
5	pr	PURA	pari	HT116a	𐘫𐘴
6	pr	PURE	pari	PK ZA 11	𐘫𐘙
7	srr	SARARA	sarra	HT30	𐘞𐘴𐘴
8	srr	SIRERE	sarra	PH31a	𐘤𐘙𐘙

FIGURE 8. Results after input into "LinAfontConv" function

As seen in Figure 8, data from the input file should be ordered based on "Identical match," "Linear A word," "Original Word" and "Source". The programme will then output the data in the "Linear A word" column into Linear A fonts in the last column.

4.5. Preliminary Results and Discussion

With the assistance of the programme, we were able to attest identical matches between Linear A transcriptions and other dictionary lexical items. Figure 9 shows the examples of two consonantal clusters "PR" and "SR". The consonantal cluster of "PA" is derived from the Linear A cluster "PA-RE," located in HT4 of GORILA vol. 1 (Godart and Olivier, 1976a). Among the words from Hittite, Thracian, Anatolian, Luwian and Hamito-Semitic language families, with the same syllabic structure, a possible phonetic counterpart could be "PARI" from the Luwian dictionary (Melchert, 1993). The Luwian word "PARI" represents "forth or away" according to the dictionary. Similarly, "PA-RA" in Linear A has a possible counterpart in the Hamito-Semitic word "PARA" (Orel, 1994); "SI-RU" has a possible counterpart in the Thracian word "SIRIU" (Paliga, 2006); "SU-RE" with the Pre-Basque form "SUR" (Trask, 2008). Going through more examples of the consonantal clusters obtained by the programme, we hope to obtain a better understanding of what Linear A words can represent when compared with other language families.

5. Conclusion

Overall, to improve the efficacy of the Python programme, there will be only three functions available for users. These three functions are the aforementioned SD function, GD 2 function as well as the "LinAfontConv" function.

Consonantal Cluster	Linear A word	Dictionary Word
PR	PA-RE (𐘂𐘙)	Luwian: PARI
	Found in HT 4	(Defintion: forth, away)
PR	PA-RA (𐘂𐘴)	Hamito-Semitic: PARA
	Found in HT128a, PH3a	(Definition: Equid, Onager)
SR	SI-RU (𐘤𐘘)	Thracian: SIRIU
	Found in HT55a, TR ZA 1, IO ZA 15, PK ZA 10	(Definition: Vrancea region (Romania))
SR	SU-RE (𐘲𐘙)	Pre-Basque: SUR
	Found in HT 32	(Definition: Pour out)

FIGURE 9. Analysis of results obtained from the programme's SD function

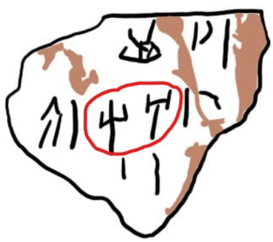


FIGURE 10. Digital recreation of the Linear A tablet HT55a. The Linear A cluster "SI-RU" is circled in red (Perono Cacciafoco and Cavallaro, 2020)

The programme will be provided as an executable. Using the PyQt Python module, the programme will have a simple GUI to enable easy access to the functions. A shortcut will be created to enable quick distribution of the programme. This will also allow users to have direct access to the programme even without having the necessary dependencies and Python modules installed on their devices.

The Python programme has the potential to be further developed to enable possible image and text recognition as well as natural language processing through the use of the TensorFlow module. The incorporation of Linear A fonts in the programme can also allow the prediction of potential clusters of Linear A signs for further frequency analysis.

While the programme is able to conduct statistical and comparative analysis of dictionaries of various languages compared to Linear A, there are still certain areas of the programme functions that can be further improved.

Furthermore, with just the use of PyQt and pandas Python modules, the programme is unable to "read" and "render" the Linear A fonts. As a result, the programme will have to employ the use of an external software, such as Microsoft Excel or TEX, to ensure clear data rendering of the results using the Linear A fonts. As an improvement to this constraint, the Python module matplotlib will be introduced to render the Linear A fonts in the programme directly.

No doubt, the task of deciphering Linear A continues to be daunting to many researchers as the ancient Aegean Minoan writing system remains an unsolved enigma (Tan, 2018). The systematic and multidisciplinary approach consisting in using such a programme to decipher Linear A would be a shift away from the past philological attempts, as it involves methods from comparative linguistics and cryptolinguistics. This is a novel approach to tackle Linear A script and the elusive language behind it, and more research in this field will be able to be carried out.

A brute-force attack to the Linear A texts with our programme allows simultaneous comparison between Linear A and other languages from language families such as the Semitic and the Indo-European families, and provides us with a better understanding of the Minoan language. This could serve as a strong foundation for the crypto-linguistics approach of deciphering Linear A, where further work on Linear A can be carried out to facilitate the identification of a language family to which the Minoan language may belong.

6. Acknowledgements

We wish to acknowledge the funding support for this project from Nanyang Technological University under the URECA Programme.

We also wish to acknowledge the support and help from Dr Xu Duoduo at Nanyang Technological University, School of Humanities, History Programme, as well as Mr Hu Man Keat and Ms Yu Ying Yao under the Nanyang Research programme (NRP) at Nanyang Technological University.

References

Bradley, P. (2014). *The Ancient World Transformed*. Cambridge: Cambridge University Press.

Chadwick, J. (1967). *The decipherment of linear B*. Cambridge: Cambridge University Press.

Eu, N., F. Perono Cacciafoco, and F. Cavallaro (2019). “Linear A Libation Tables: A Semitic Connection Explored.” In: *Annals of the University of Craiova: Series Philology, Linguistics / Analele Universității Din Craiova: Seria Științe Filologice, Linguistică*. Vol. 41. 1/2, pp. 51–63.

Eu, N., D. Xu, and F. Perono Cacciafoco (2019). “Coding to Decipher Linear A.” In: *Proceedings of the 2019 Pacific Neighbourhood Consortium Annual Conference and Joint Meetings (PNC)*. Singapore: Nanyang Technological University, pp. 44–48.

Finkelberg, M. (2001). “The Language of Linear A: Greek, Semitic, or Anatolian?” In: *Greater Anatolia and the Indo-Hittite Language Family*. Ed. by R. Drews. Vol. 38. Journal of Indo-European Monograph: Series, pp. 81–105.

Godart, L. (1984). “Du linéaire A au linéaire B.” In: *Aux origines de l’hellénisme: La Crète et la Grèce*. Paris: Publications de la Sorbonne, pp. 121–128.

Godart, L. and J.-P. Olivier (1976a). *Recueil des inscriptions en linéaire A. 1. Tablettes éditées avant 1970*. Vol. 21. Études crétoises. Paris: Librairie orientaliste Paul Geuthner.

——— (1976b). *Recueil des inscriptions en linéaire A. 3. Tablettes, nodules et rondelles édités en 1975 et 1976*. Vol. 21.3. Études crétoises. Paris: Librairie orientaliste Paul Geuthner.

——— (1979). *Recueil des inscriptions en linéaire A. 2. Nodules, scellés et rondelles édités avant 1970*. Vol. 21.2. Études crétoises. Paris: Librairie orientaliste Paul Geuthner.

——— (1982). *Recueil des inscriptions en linéaire A. 4. Autres documents*. Vol. 21.4. Études crétoises. Paris: Librairie orientaliste Paul Geuthner.

——— (1985). *Recueil des inscriptions en linéaire A. 5. Addenda, corrigenda, concordances, index et planches des signes*. Vol. 21.5. Études crétoises. Paris: Librairie orientaliste Paul Geuthner.

Gordon, C. H. (1982). *Forgotten scripts: Their ongoing discovery and decipherment*. New York: Basic Books.

Melchert, Craig H. (1993). “Cuneiform Luvian Lexicon.” In: vol. 2. Lexica Anatolica. Chapell Hill, NC: self-published.

Nagy, G. (1963). *Greek-like Elements in Linear A*. Durham: William H. Willis.

Olivier, J.-P. (1986). “Cretan writing in the second millennium BC.” In: *World Archaeology* 17.3, pp. 377–389.

Orel, V. (1994). “On the ancient contacts between Hamito-Semitic and north Caucasian.” In: *Folia Linguistica Historica* 15.1–2.

Owens, G. A. (1999). “The Structure of the Minoan Language.” In: *J. Indo-Eur. Stud.* 27.1/2, pp. 15–56.

Paliga, Sorin (2006). *Etymological Lexicon of the Indigenous (Thracian) Elements in Romanian*. București: Editura Evenimentul.

Palmer, L. R. (1961). *Myceneans and Minoans: Aegean pre-history in the Light of the Linear B tablets*. Faber & Faber.

Perono Cacciafoco, F. (2017). “Linear A and Minoan: Some New Old Questions.” In: *Annals of the University of Craiova: Series Philology, Linguistics / Analele Universității Din Craiova: Seria Științe Filologice, Linguistică*. Vol. 39. 1–2, pp. 154–170.

Perono Cacciafoco, F. and F. Cavallaro (2020). “Linear A Corpus.” https://blogs.ntu.edu.sg/linear-a/.

Petrolito, T. et al. (2015). “Minoan linguistic resources: The Linear A Digital Corpus.” In: *Proceedings of the 9th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences and Humanities (LaTeCH), Beijing, China*, pp. 95–104.

Revesez, P. (2017). “Establishing the West-Ugric Language Family with Minoan, Hattic and Hungarian by a Decipherment of Linear A.” In: *WSEAS Trans. Inf. Sci. Appl.* 14, pp. 306–355.

Schoep, I. (2002). “Social and Political Organisation on Crete in the Proto-palatial Period: The case of Middle Minoan II Malia.” In: *J. Mediterr. Archaeol.* 15.1, pp. 101–132.

Tan, K. M. W. Y. (2018). *The Minoan Engima: Deciphering Linear A*. Tech. rep. http://hdl.handle.net/10356/76547. Singapore.

Trask, R.L. (2008). *Etymological Dictionary of Basque*. Sussex: University of Sussex.

Verdult, R. (2015). “The (in)security of proprietary cryptography.” PhD thesis. Radboud University, The Netherlands and KU Leuven, Belgium.