CHAPTER 8

INVESTIGATIONS INTO THE LANGUAGE(S) BEHIND CRETAN HIEROGLYPHIC AND LINEAR A

Brent Davis

In a recent article[1] I employed a new method of analysis to evaluate the likelihood that Linear A and the Phaistos Disk encode the same language. This new method of analysis relies on *syllabotactics*[2] – the linguistic constraints that dictate how syllables can be combined to form words. All languages have such syllabotactic constraints, and these constraints are language-specific – that is, they differ from language to language. Thus in the case of Aegean syllabic scripts such as Linear A, Cretan Hieroglyphic, Linear B and Cypriot Syllabic, the syllabotactic constraints on the language behind each script have the effect of limiting the range of *syllabograms* that can sit side by side within a word in each script – that is, these constraints limit the range of *word-internal pairs*[3] of syllabograms in each script; and they do so in similar ways in scripts that encode the same language, and in dissimilar ways in scripts that encode different languages.[4] Furthermore, there is a strong and growing body of evidence indicating that in the Aegean family of syllabic scripts, *homomorphs* in any two of these scripts (i.e. syllabograms that closely resemble each other in any two of these scripts) are most often also *homophones*, or nearly so (i.e. they represent the same or very similar phonetic values).[5] These characteristics of Aegean syllabic scripts thus provide us with a valuable method of *syllabotactic analysis* for evaluating the probability that any two of these scripts encode the same language – i.e., one can (1) identify a set of homomorphs in the two scripts; (2) tabulate the ways in which those homomorphs form word-internal pairs in each script; and (3) evaluate the degree of similarity between the two tables. If the two tables are

[1] Davis 2018.

[2] The term was originally coined by Alamolhoda (2003). Though the study of syllabotactics has not yet made significant inroads into linguistics, syllabotactic approaches have proven very useful in the field of automatic language and speech recognition (e.g., Antoine *et al*. 2004; Zhu *et al*. 2005; Zhu and Adda-Decker 2006a; 2006b; Hieronymus *et al*. 2009; Kordek 2012; González 2015).

[3] As an illustration of what I mean by *word-internal pair*: a four-sign word ABCD contains three word-internal pairs of signs: AB, BC and CD.

[4] For a detailed illustration involving Linear A and Linear B, see Davis 2018: 374–6.

[5] For example: (1) Linear B *vs* Linear A: Davis 2014: 189; (2) Linear B *vs* Cypriot Syllabic: Woodard 1997; (3) Linear A *vs* the Phaistos Disk: Davis 2018.

similar to a statistically significant degree, then the likelihood is that both scripts encode the same language; but if the two tables are *not* similar to a statistically significant degree, then the likelihood is that the two scripts encode different languages.

In this chapter, I use this method of analysis to address the question, 'What is the likelihood that Cretan Hieroglyphic and Linear A encode the same language?' The structure of this syllabotactic analysis is identical to the structure of the analysis in Davis 2018, in that it consists of two experiments: (1) a control experiment; and (2) a main experiment. Both experiments involve evaluating a *target text* (from the Cypriot Syllabic corpus in the control experiment, and from the Cretan Hieroglyphic corpus in the main experiment) for its syllabotactic similarity to two large *benchmark texts* (one from the Linear B corpus, the other from the Linear A corpus).[6]

(1) In the **control experiment**, the Cypriot Syllabic target text is evaluated for its syllabotactic similarity to the Linear B and Linear A benchmark texts.[7] This experiment is designed to illustrate the validity of this method of analysis, in that if the method is valid and productive, we should expect the Cypriot Syllabic target text to show a significant degree of syllabotactic similarity to the Linear B benchmark text, but not to the Linear A one.
(2) In the **main experiment**, the Cretan Hieroglyphic target text is evaluated for its syllabotactic similarity to the Linear B and Linear A benchmark texts. In this experiment, we should expect the Cretan Hieroglyphic target text to show an *insignificant* degree of syllabotactic similarity to the Linear B benchmark text, while the degree of syllabotactic similarity between the Cretan Hieroglyphic target text and the Linear A benchmark text will serve as an indicator of the likelihood that both texts encode the same language.

For each experiment, this process of evaluation consists of four steps:

(1) **Defining the sets of homomorphs**: (1a) For the control experiment: identifying a set of homomorphs that exist in all three scripts used in that experiment (Cypriot Syllabic, Linear B and Linear A); and (1b) for the main experiment: identifying a set of homomorphs that exist in all three scripts used in that experiment (Cretan Hieroglyphic, Linear B and Linear A);

[6] The two target texts and the two benchmark texts are defined in detail later in this chapter.
[7] Importantly: in the control experiment, Cypriot Syllabic is treated as an *undeciphered script* – that is, the experiment is conducted using Cypriot Syllabic texts as originally inscribed, with no reference to the phonetic values of the signs.

(2) **Defining the four texts, and tabulating word-internal pairs they contain**: (2a) For the main experiment: defining the Linear B and Linear A benchmark texts and the Cretan Hieroglyphic target text, and tabulating the ways in which the Linear B, Linear A and Cretan Hieroglyphic homomorphs identified in Step 1b form unique word-internal pairs with each other in these three texts (and by 'unique', I mean that duplicates are not counted); and (2b) for the control experiment: defining a Cypriot Syllabic target text analogous to the Cretan Hieroglyphic target text in terms of the number of unique word-internal pairs that it contains, and tabulating the ways in which the Linear B, Linear A and Cypriot Syllabic homomorphs identified in Step 1a form unique word-internal pairs with each other in the Linear B and Linear A benchmark texts and the Cypriot Syllabic target text;

(3) **Scoring the target texts for their syllabotactic similarity to the benchmark texts**: (3a) In the control experiment: for the Cypriot Syllabic target text, determining the number of its unique word-internal sign pairs (as tabulated in Step 2b) whose *Linear B and Linear A homomorphs* are also attested in the Linear B and Linear A benchmark texts, with those two numbers then serving as the syllabotactic similarity scores for the Cypriot Syllabic target text vs the Linear B and Linear A benchmark texts; and (3b) in the main experiment: for the Cretan Hieroglyphic target text, determining the number of its unique word-internal sign pairs (as tabulated in Step 2a) whose *Linear B and Linear A homomorphs* are also attested in the Linear B and Linear A benchmark texts, with those two numbers then serving as the syllabotactic similarity scores for the Cretan Hieroglyphic target text vs the Linear B and Linear A benchmark texts; and finally;

(4) **Evaluating the scores**: (4a) In the control experiment: evaluating where the two scores for the Cypriot Syllabic target text vs the Linear B and Linear A benchmark texts each sit relative to the average score that we would expect to be produced by chance alone; and (4b) in the main experiment: evaluating where the two scores for the Cretan Hieroglyphic target text vs the Linear B and Linear A benchmark texts each sit relative to the average score that we would expect to be produced by chance alone. In both experiments: any score that sits two or more standard deviations above the average score produced by chance alone is deemed to indicate a *statistically significant degree*[8] *of syllabotactic similarity* between the relevant

[8] In this chapter, 'statistical significance' is defined as including scores that sit two or more standard deviations above the average. All definitions of 'statistical significance' are ultimately subjective, but this definition is by far the most widely used one in the literature.

 https://doi.org/10.1017/9781009490122.010

target text and benchmark text, strongly implying that both texts encode the same language; while any score that sits less than two standard deviations above the average score produced by chance alone is deemed to indicate a *statistically insignificant degree of syllabotactic similarity* between the relevant target text and benchmark text, strongly implying that the two texts encode different languages.

8.1 Step 1: Identifying Homomorphs

Step 1a *(For the Control Experiment): Identifying a Set of Cypriot Syllabic/Linear B/Linear A Homomorphs*

This step was effectively completed in Davis 2018,[9] resulting in a set of ten trios of Cypriot Syllabic/Linear B/Linear A homomorphs. Table 8.1 shows these ten trios, together with the AB numbers of the Linear A and Linear B signs in the bottom row:

Table 8.1 *Ten trios of Cypriot Syllabic/Linear B/Linear A homomorphs*

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
CS	𐠀	𐠒	𐠙	𐠞	𐠡	𐠨	𐠩	𐠭	𐠯	𐠰
LB	𐀀	𐀨	𐀙	𐀞	𐀡	𐀭	𐀮	𐀅	𐀴	𐀵
LA	𐘇	𐘴	𐘅	𐘂	𐘊	𐘞	𐘈	𐘀	𐘠	𐘄
AB	*08	*60	*06	*03	*11	*31	*09	*01	*37	*05

Step 1b *(For the Main Experiment): Identifying a Set of Cretan Hieroglyphic/Linear B/Linear A Homomorphs*

In identifying a set of Cretan Hieroglyphic/Linear B/Linear A homomorphs, I have conservatively relied entirely on the established suggestions of past scholars,[10] resulting in a set of twenty-one trios of

[9] Davis 2018: 380–1, tables 9 and 13.

[10] Trios 5 and 6 were suggested by *Docs*² (33, fig. 6), while trio 10 was suggested by Younger (2013: Sign 010); all the other trios were suggested by *CHIC* (19). Thus, with the exception of trio 10 (which was suggested years ago), all these homomorphs were suggested decades ago. Note that a few other homomorphs suggested by these scholars have been excluded from Table 8.2: (1) *Docs*² (ibid.) suggest that CH 046 [sign] and A *301 [sign] are homomorphs, but these signs have no homomorph in Linear B; (2) *Docs*² (ibid.) also suggest that the Cretan Hieroglyphic 'catface' sign [sign] is a homomorph of A *80 [sign] / B *80 [sign], but *CHIC* (14, n. 37) treat the Cretan Hieroglyphic sign as a 'decoration' rather than a syllabogram, an assessment that I

Cretan Hieroglyphic/Linear B/Linear A homomorphs. Table 8.2 shows these twenty-one trios, together with the *CHIC* numbers of the Cretan Hieroglyphic signs in the top row, and the AB numbers of the Linear A and Linear B signs in the bottom row:

Table 8.2 *Twenty-one trios of Cretan Hieroglyphic/Linear B/Linear A homomorphs*

	1	2	3	4	5	6	7	8	9	10	11
CHIC	042	094	008	038	007	012	052	095	031	010	070
CH											
LA											
LB											
AB	*08	*38	*28	*57	*73	*23	*24	*60	*27	*53	*02

	12	13	14	15	16	17	18	19	20	21
CHIC	092	019	025	049	041	005	017	006	069	040
CH										
LA										
LB										
AB	*26	*31	*04	*37	*54	*79	*85	*48	*76	*86

8.2 Step 2: Defining the Four Texts and Tabulating the Word-Internal Pairs They Contain

Step 2a *(For the Main Experiment): Defining the Linear B and Linear A Benchmark Texts and the Cretan Hieroglyphic Target Text, and Tabulating the Word-internal Pairs They Contain That are Formed from the Twenty-one Trios of Homomorphs in Table 8.2*

The Linear B benchmark text was defined as the Linear B corpus as transcribed by Aurora.[11] From that corpus, I tabulated the total number of unique word-internal pairs[12] consisting solely of the twenty-one Linear B

can find no justifiable reason to reject; and (3) *CHIC* (19) suggest that CH 024 and 035 are homomorphs of A *30 / B *30 and A *58 / B *58, respectively, but as the Cretan Hieroglyphic corpus contains no instances of either of these Cretan Hieroglyphic signs forming secure word-internal pairs with any of the other Cretan Hieroglyphic signs in Table 8.2, there was no point in including these two trios in this analysis.

11 *DĀMOS*.

12 This and all other statements regarding word-internal pairs in Linear B have been checked against *DĀMOS*, which contains a searchable corpus of all Linear B inscriptions published to date.

signs in Table 8.2, resulting in a list of 219 unique word-internal pairs, as shown in Table 8.3:

Table 8.3 *The twenty-one Linear B signs in Table 8.2 form 219 different word-internal pairs in the Linear B benchmark text*

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]									

The Linear A benchmark text was defined as the Linear A corpus as transcribed by Younger.[13] From that corpus, I tabulated the total number of unique word-internal pairs[14] consisting solely of the twenty-one Linear A signs in Table 8.2, resulting in a list of 140 unique word-internal pairs, as shown in Table 8.4:

[13] Younger 2020.

[14] This and all other statements regarding word-internal pairs in Linear A have been checked against Younger 2020, which contains a searchable corpus of all Linear A inscriptions published to date.

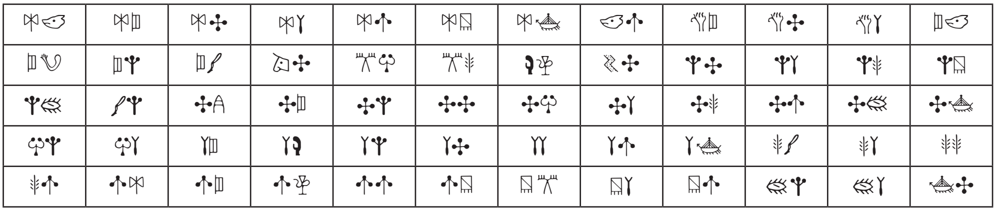
Table 8.4 *The twenty-one Linear A signs in Table 8.2 form 140 different word-internal pairs in the Linear A benchmark text*

The Cretan Hieroglyphic target text was defined as the Cretan Hieroglyphic corpus as transcribed by Olivier and Godart and augmented by Younger.[15] From that corpus, I tabulated the total number of unique word-internal pairs[16] consisting solely of the twenty-one Cretan Hieroglyphic signs in Table 8.2, resulting in a list of sixty unique word-internal pairs, as shown in Table 8.5:

[15] The starting-point for the Cretan Hieroglyphic corpus used in this analysis was the 'Index des Signes' in *CHIC* (319–79), from which I first excluded the following items: (1) italicised (i.e. insecure) readings of signs; (2) asterisked (i.e. secondary) readings of inscriptions; (3) inscriptions marked 'o' (indicating that it is impossible to determine which sign begins the inscription); (4) inscriptions marked '>?' (indicating that a left-to-right direction of reading is only probable), with one exception (#101.a); and (5) inscriptions marked '><' (indicating that the direction of reading is uncertain), with nine exceptions (#074.a, #154, #163, #168, #222.b, #224.α, #276.α, #297.β1, #310.δ). The exceptions to exclusions (4) and (5) include inscriptions with CH 008 or 042 at one end or the other; these signs are homomorphs of Linear A/B vowel signs A *28 / B *28 and A *08 / B *08, respectively, and indeed, the Cretan Hieroglyphic signs behave like vowel signs as well, in that in inscriptions containing secure instances of these signs and with a known direction of reading, CH 008 is always word-initial, while CH 042 is word-initial most of the time. Thus, in each of the ten exceptions to exclusions (4) and (5), the occurrence of CH 008 or 042 was taken as the beginning of the word. (It is worth noting that the effect of these exceptions was actually quite minimal, in that together, they contributed just three of the word-internal pairs in Table 8.5). As a final step: to this amended corpus I then added the Cretan Hieroglyphic documents that have been found since *CHIC* was published: (1) the documents from Petras (Younger 2010); four miscellaneous documents from Malia, Pyrgos and Kato Syme (Younger 2016b: MA/V Yb 03, MA/V Yb 04, PYR Yb 01 and SY Hf 01); and (3) twelve seals and seal-impressions (Younger 2016a: 'Additions since (or not in) *CHIC*').

[16] This and all other statements regarding word-internal pairs in Cretan Hieroglyphic have been checked against *CHIC* and Younger 2010, 2016a and 2016b.

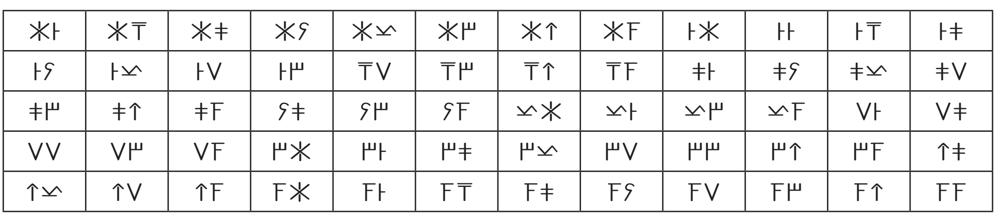
Table 8.5 *The twenty-one Cretan Hieroglyphic signs in Table 8.2 form sixty different word-internal pairs in the Cretan Hieroglyphic target text*



***Step 2b** (For the Control Experiment): Defining the Cypriot Syllabic Target Text, and Tabulating the Word-internal Pairs It and the Linear B and Linear A Benchmark Texts Contain That are Formed from the Ten Trios of Homomorphs in Table 8.1*

The Cypriot Syllabic target text was defined so as to contain sixty unique word-internal pairs formed from the ten Cypriot Syllabic signs in Table 8.1, thus making it analogous to the Cretan Hieroglyphic target text in terms of the number of different word-internal pairs that it contains.[17] Table 8.6 shows those sixty Cypriot Syllabic pairs:

Table 8.6 *The ten Cypriot Syllabic signs in Table* 8.1 *form sixty different word-internal pairs in the Cypriot Syllabic target text*



From the Linear B benchmark text as defined in Step 2a, I tabulated the total number of unique word-internal pairs consisting solely of the 10 Linear B signs in Table 8.1 resulting in a list of seventy-eight unique word-internal pairs, as shown in Table 8.7:

[17] Masson 1983. This concatenated body of texts consists of inscriptions no. 1 up through the sequence [illegible] in Line 26 of no. 217 (Masson 1983: 95–237). In counting word-internal pairs, pairs containing any sign transcription in non-italic text, denoting an uncertain reading, were ignored.

Table 8.7 *The ten Linear B signs in Table 8.1 form seventy-eight different word-internal pairs in the Linear B benchmark text*

Ͳ├	ͲȲ	Ͳǂ	Ͳㄣ	Ͳᓚ	Ͳᛉ	Ͳᒥ	Ͳᗰ	ͲŦ	├├	├Ȳ	├ǂ
├ㄣ	├ᓚ	├ᛉ	├ᒥ	├Ŧ	ȲͲ	Ȳㄣ	Ȳᓚ	Ȳᛉ	Ȳᒥ	Ȳᗰ	ȲŦ
ǂ├	ǂȲ	ǂǂ	ǂㄣ	ǂᓚ	ǂᛉ	ǂᒥ	ǂᗰ	ǂŦ	ㄣ├	ㄣȲ	ㄣㄣ
ㄣᓚ	ㄣᒥ	ㄣᗰ	ㄣŦ	ᓚͲ	ᓚ├	ᓚȲ	ᓚǂ	ᓚㄣ	ᓚᛉ	ᓚᒥ	ᓚᗰ
ᓚŦ	ᛉ├	ᛉȲ	ᛉǂ	ᛉㄣ	ᛉᓚ	ᛉᛉ	ᛉᗰ	ᛉŦ	ᒥ├	ᒥᓚ	ᒥᗰ
ᒥŦ	ᗰͲ	ᗰ├	ᗰȲ	ᗰǂ	ᗰᓚ	ᗰᛉ	ᗰᒥ	ᗰŦ	ŦͲ	ŦȲ	Ŧǂ
Ŧㄣ	Ŧᓚ	Ŧᛉ	Ŧᒥ	Ŧᗰ	ŦŦ						

Finally: from the Linear A benchmark text as defined in Step 2a, I tabulated the total number of unique word-internal pairs consisting solely of the ten Linear A signs in Table 8.1, resulting in a list of fifty-two unique word-internal pairs, as shown in Table 8.8:

Table 8.8 *The ten Linear A signs in Table 8.1 form fifty-two different word-internal pairs in the Linear A benchmark text*

Ͳ├	Ͳī	Ͳǂ	Ͳᓚ	ͲY	ͲШ	ͲΛ	ͲŦ	├Ͳ	├├	├ī	├ᓚ
├Ш	├Λ	├Ŧ	īͲ	ī├	īī	īǂ	īᓚ	īY	īΛ	ǂ├	ǂī
ǂᓚ	ǂY	ǂШ	ʕͲ	ʕŦ	ᓚͲ	ᓚī	ᓚᓚ	ᓚΛ	Y├	Yī	Yʕ
Yᓚ	YY	YŦ	Шī	Шᓚ	ШY	ШŦ	ΛͲ	Λ├	Λī	Λǂ	Λᓚ
ΛY	ΛШ	ΛΛ	ŦY								

8.3 Steps 3a and 4a: Completing the Control Experiment

Step 3a *Scoring the Cypriot Syllabic Target Text for its Syllabotactic Similarity to the Linear B and Linear A Benchmark Texts*

In this step, the Cypriot Syllabic target text is scored for its syllabotactic similarity to the Linear B and Linear A benchmark texts. The procedure is straightforward: first, determine the number of word-internal pairs in the Cypriot Syllabic target text (Table 8.6) whose *Linear B homomorphs* appear in the Linear B benchmark text (Table 8.7). (The Linear B homomorphs of the pairs in the Cypriot Syllabic target text are arrived at through consulting the list of Linear B/Cypriot Syllabic homomorphs in Table 8.1.) The results are shown in Table 8.9:

Table 8.9 *Fifty-two pairs in the Cypriot Syllabic target text (Table 8.6) whose Linear B homomorphs appear in the Linear B benchmark text (Table 8.7)*

✳⊦	✳₸	✳ǂ	✳ʕ	✳⩡	✳ᗰ	✳↑	✳F	⊦⊦	⊦₸	⊦ǂ	⊦ʕ
⊦⩡	⊦V	⊦ᗰ	₸V	₸ᗰ	₸↑	₸F	ǂ⊦	ǂʕ	ǂ⩡	ǂV	ǂᗰ
ǂ↑	ǂF	ʕᗰ	ʕF	⩡✳	⩡⊦	⩡ᗰ	⩡F	V⊦	Vǂ	VV	VF
ᗰ⊦	ᗰ⩡	ᗰ↑	ᗰF	↑ǂ	↑⩡	↑V	↑F	F✳	F₸	Fǂ	Fʕ
FV	Fᗰ	F↑	FF								

As there are sixty pairs in the Cypriot Syllabic target text, the syllabotactic similarity score for the Cypriot Syllabic target text vs the Linear B benchmark text is thus **52/60**. Finally, determine the number of word-internal pairs in the Cypriot Syllabic target text (Table 8.6) whose *Linear A homomorphs* appear in the Linear A benchmark text (Table 8.8). (The Linear A homomorphs of the pairs in the Cypriot Syllabic target text are arrived at through consulting the list of Linear A/Cypriot Syllabic homomorphs in Table 8.1.) The results are shown in Table 8.10:

Table 8.10 *Thirty pairs in the Cypriot Syllabic target text (Table 8.6) whose Linear A homomorphs appear in the Linear A benchmark text (Table 8.8)*

✳⊦	✳₸	✳ǂ	✳⩡	✳ᗰ	✳↑	✳F	⊦✳	⊦⊦	⊦₸	⊦⩡	⊦ᗰ
₸V	₸↑	ǂ⊦	ǂ⩡	ǂV	ǂᗰ	ʕF	⩡✳	V⊦	VV	VF	ᗰ⩡
ᗰV	ᗰF	↑ǂ	↑⩡	↑V	FV						

As there are (again) sixty pairs in the Cypriot Syllabic target text, the syllabotactic similarity score for the Cypriot Syllabic target text vs the Linear A benchmark text is thus **30/60**.

Step 4a *Evaluating the Scores*

Step 3a has now produced the following two syllabotactic similarity scores for the control experiment:

The Cypriot Syllabic target text vs the Linear B benchmark text = 52/60
The Cypriot Syllabic target text vs the Linear A benchmark text = 30/60

In this final step of the control experiment, each of these scores is evaluated against a representative average score produced by chance alone; thus, for each of the two comparisons listed above, the average score produced by chance alone must be calculated first. The method of calculating this average score is straightforward (if somewhat tedious). I will begin by calculating a representative average score produced by chance alone for the first of the comparisons listed above: the Cypriot Syllabic target text vs the Linear B benchmark text.

In Table 8.11, row LB1 contains the ten Linear B signs from Table 8.1, while row LB2 contains the *Linear B homomorphs* of the ten Cypriot Syllabic signs present in Table 8.6:

Table 8.11 *Row LB1: the ten Linear B signs from Table 8.1; row LB2: the Linear B homomorphs of the ten Cypriot Syllabic signs present in Table 8.6 (the pairs in the Cypriot Syllabic target text)*

	1	2	3	4	5	6	7	8	9	10
LB1	𐀫	𐀙	𐀭	𐀞	𐀮	𐀨	𐀆	𐀅	𐀴	𐀵
LB2	𐀫	𐀙	𐀭	𐀞	𐀮	𐀨	𐀆	𐀅	𐀴	𐀵

In Table 8.12, the ten signs in row LB2 of Table 8.11 have been randomly rearranged:

Table 8.12 *Table 8.11, with the ten signs in row LB2 randomly rearranged*

	1	2	3	4	5	6	7	8	9	10
LB1	𐀫	𐀙	𐀭	𐀞	𐀮	𐀨	𐀆	𐀅	𐀴	𐀵
LB2	𐀵	𐀅	𐀆	𐀨	𐀞	𐀮	𐀙	𐀴	𐀭	𐀫

In Table 8.13, row LB3 contains the *Linear B homomorphs* of the pairs in the first row of Table 8.6 (i.e. the first twelve pairs in the Cypriot Syllabic target text), while row LB4 contains those same pairs with each sign replaced by the corresponding (random) sign from row LB2 of Table 8.12:

Table 8.13 *Row LB3: the Linear B homomorphs of the Cypriot Syllabic pairs in the first row of Table 8.6; row LB4: those same pairs retranscribed according to the random permutation of signs in row LB2 of Table 8.12*

LB3	𐀫𐀅	𐀫𐀭	𐀫𐀞	𐀫𐀮	𐀫𐀙	𐀫𐀆	𐀫𐀴	𐀫𐀵	𐀅𐀫	𐀅𐀅	𐀅𐀭	𐀅𐀞
LB4	𐀵𐀴	𐀵𐀆	𐀵𐀨	𐀵𐀞	𐀵𐀅	𐀵𐀙	𐀵𐀭	𐀵𐀫	𐀴𐀵	𐀴𐀴	𐀴𐀆	𐀴𐀨

If we continue in this way to retranscribe the *Linear B homomorphs* of all sixty Cypriot Syllabic pairs in Table 8.6 according to the random permutation of signs in row LB2 of Table 8.12, then count the number of randomly retranscribed pairs in row LB4 that also appear in the Linear B benchmark text, the result is forty-five – that is, the syllabotactic similarity score for this randomly retranscribed set of pairs from the Cypriot Syllabic target text vs the Linear B benchmark text is **45/60**.

Thus this score of 45/60 is one representative of a score produced by chance alone; yet this score is the product of the single permutation of signs in row LB2 of Table 8.12, when in fact those ten signs can be rearranged in a very large number of ways.[18] To generate a much more representative average score produced by chance alone, I therefore rearranged the signs in row LB2 of Table 8.12 in 1,000,000 different random ways, retranscribed the *Linear B homomorphs* of the sixty Cypriot Syllabic pairs in Table 8.6 according to each random permutation of signs, and scored each set of retranscribed pairs by counting how many of them also appear in the Linear B benchmark text, just as was done in the preceding example. The results are shown in Table 8.14:[19]

Table 8.14 *Syllabotactic similarity scores for 1,000,000 different sets of randomly retranscribed Linear B homomorphs of the Cypriot Syllabic pairs in Table 8.6 (the pairs in the Cypriot Syllabic target text) vs the Linear B benchmark text*

Score out of 60:	**Permutations with that score:**	**% of permutations:**	**Score out of 60:**	**Permutations with that score:**	**% of permutations:**
40	40	0.0040%	51	42,619	4.2619%
41	586	0.0586%	52	16,597	1.6597%
42	3,794	0.3794%	53	4,872	0.4872%
43	15,973	1.5973%	54	1,069	0.1069%
44	47,728	4.7728%	55	142	0.0142%
45	101,940	10.1940%	56	12	0.0012%
46	160,380	16.0380%	**Tot. permutations:**	1,000,000	100%
47	193,008	19.3008%	**Average score:**	47.422 / 60	
48	183,939	18.3939%	**Std deviation (σ):**	2.013 / 60	
49	141,057	14.1057%	**Avg. score + 2σ:**	51.448 / 60	
50	86,244	8.6244%	**Score of 52/60:**	Avg. + 2.270σ	$p = 0.0227$

[18] The total number of possible permutations is equal to 10! (‘10 factorial’) = the product of integers 1 through 10 = 3,628,800.

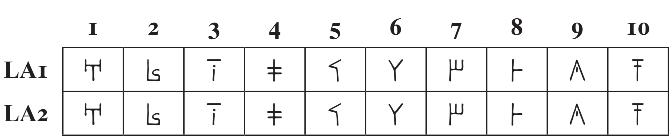
[19] Note to statisticians: as each of the four statistical analyses in this paper is based on a random sample of 1,000,000 permutations (out of populations of 10! in the first two analyses, and 21! in the other two – thus never the whole population), σ^2 is always calculated as a sample variance – that is: $\frac{\sum(x-\bar{x})^2}{(n-1)}$.

As Table 8.14 shows, the 1,000,000 random permutations produce an average score of 47.422/60, with a standard deviation (σ) of 2.013/60, such that the region of 'statistical significance' (2σ or more above the average) begins at 51.448/60 (i.e. scores of **52/60** or greater). Meanwhile, the original syllabotactic similarity score for the Cypriot Syllabic target text vs the Linear B benchmark text (as shown in Table 8.9) is **52/60**, which is **2.27σ above** the average score produced by chance alone (i.e. within the region of statistical significance); and the *p*-value (produced by adding the percentages in the right column for all scores of 52/60 or more) is 0.0227, meaning that of the 1,000,000 random permutations, just 2.27% of them scored 52/60 or higher. The potential meanings of this score will be discussed shortly.

The Cypriot Syllabic Target Text vs the Linear A Benchmark Text

Next, a representative average score produced by chance alone can be calculated for the Cypriot Syllabic target text vs the Linear A benchmark text in just the same way, as follows. In Table 8.15, row LA1 contains the ten Linear A signs from Table 8.1, while row LA2 contains the *Linear A homomorphs* of the ten Cypriot Syllabic signs present in Table 8.6 (the pairs in the Cypriot Syllabic target text):

Table 8.15 *Row LA1: the ten Linear A signs from Table 8.1; row LA2: the Linear A homomorphs of the ten Cypriot Syllabic signs present in Table 8.6 (the pairs in the Cypriot Syllabic target text)*



Rearranging the ten signs in row LA2 of Table 8.15 in 1,000,000 different random ways;[20] retranscribing the *Linear A homomorphs* of the sixty Cypriot Syllabic pairs in Table 8.6 according to each random permutation of signs; and scoring each set of retranscribed pairs by counting how many of them also appear in the Linear A benchmark text produces the results shown in Table 8.16:

[20] The total number of possible permutations is 3,628,800; see note 18.

Table 8.16 *Syllabotactic similarity scores for 1,000,000 different sets of randomly retranscribed Linear A homomorphs of the Cypriot Syllabic pairs in Table 8.6 (the pairs in the Cypriot Syllabic target text) vs the Linear A benchmark text*

Score out of 60:	Permutations with that score:	% of permutations:	Score out of 60:	Permutations with that score:	% of permutations:
19	13	0.0013%	35	64,527	6.4527%
20	39	0.0039%	36	38,994	3.8994%
21	197	0.0197%	37	20,069	2.0069%
22	826	0.0826%	38	9,151	0.9151%
23	2,758	0.2758%	39	3,385	0.3385%
24	7,317	0.7317%	40	1,030	0.1030%
25	16,305	1.6305%	41	205	0.0205%
26	30,416	3.0416%	42	44	0.0044%
27	49,971	4.9971%	43	4	0.0004%
28	73,679	7.3679%	44	1	0.0001%
29	98,143	9.8143%	**Tot. permutations:**	1,000,000	100%
30	118,721	11.8721%	**Average score:**	31.239/60	
31	129,553	12.9553%	**Std deviation (σ):**	2.967/60	
32	128,968	12.8968%	**Avg. score + 2σ:**	37.174/60	
33	114,389	11.4389%	**Score of 30/60:**	Avg. – 0.42σ	$p = 0.7203$
34	91,295	9.1295%			

As Table 8.16 shows, the 1,000,000 random permutations produce an average score of 31.239/60, with a standard deviation (σ) of 2.967/60, such that the region of 'statistical significance' (2σ or more above the average) begins at 37.174/60 (i.e. scores of **38/60** or greater). Meanwhile, the original syllabotactic similarity score for the Cypriot Syllabic target text vs the Linear A benchmark text (as shown in Table 8.10) is **30/60**, which is **0.42σ below** the average score produced by chance alone (i.e. nowhere near the region of statistical significance); and the *p*-value (produced by adding the percentages in the right column for all scores of 30/60 or more) is 0.7203, meaning that of the 1,000,000 random permutations, more than 72% of them scored 30/60 or higher.

We now have final evaluations of the two syllabotactic similarity scores produced by the control experiment, with respect to representative averages produced by chance alone:

The Cypriot Syllabic target text vs the Linear B benchmark text = 52/60 = 2.27σ above the average
The Cypriot Syllabic target text vs the Linear A benchmark text = 30/60 = 0.42σ below the average

Figure 8.1 depicts these two evaluations, with the region of statistical significance shaded light grey:

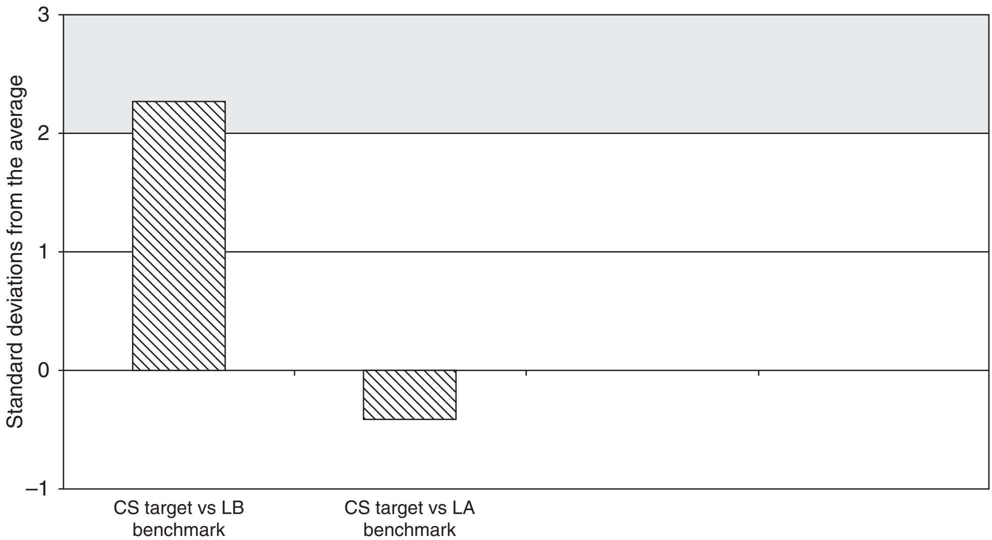


Figure 8.1 Evaluation of results: syllabotactic similarity of Cypriot Syllabic target text vs Linear B and Linear A benchmark texts

8.4 Results of the Control Experiment: Discussion

In the first half of the control experiment (evaluating the syllabotactic similarity score for the Cypriot Syllabic target text vs the Linear B benchmark text), Table 8.14 is an example of a *frequentist statistical analysis*, in which the syllabotactic similarity score of 52/60 for the Cypriot Syllabic target text vs the Linear B benchmark text is essentially being evaluated against the notion that the score is accidental – that is, due to chance alone. In frequentist statistics, this notion that the result being evaluated is accidental is called the 'null hypothesis' (abbreviated as H_0); thus, for Table 8.14, H_0 can be stated as follows:

- H_0 = The syllabotactic similarity score of 52/60 for the Cypriot Syllabic target text vs the Linear B benchmark text is due to chance alone.

What Table 8.14 tells us, though, is that to a statistically significant degree, the score of 16/16 is 'statistically incompatible'[21] with

[21] Wasserstein and Lazar 2016: 131.

H_o, and that we should therefore reject H_o in favour of the 'alternative hypothesis' (H_A):

- H_A = The syllabotactic similarity score of 52/60 for the Cypriot Syllabic target text vs the Linear B benchmark text is *not* due to chance alone.

Importantly, rejecting H_o in favour of H_A results in a crucial corollary of H_A regarding the phonetic values of the Cypriot Syllabic and Linear B signs:[22] by preferring H_A over H_o, we are also implicitly preferring the corollary that the Cypriot Syllabic and Linear B homomorphs in Table 8.1 are homophones, or nearly so ... for if the Cypriot Syllabic signs have markedly different phonetic values than their Linear B homomorphs, then the score of 52/60 (indicating that the two sets of homomorphs form word-internal pairs in remarkably similar ways) must clearly be due to chance alone – i.e. the opposite of what H_A asserts. In short: if we are to favour the hypothesis that the score is *not* due to chance alone (i.e. H_A), as the statistics tell us we should, then we must, to the same degree, favour the corollary that the Cypriot Syllabic and Linear B homomorphs are also homophones (or nearly so), as H_A and this corollary are inextricably linked. H_A can therefore be augmented with its corollary as follows:

- H_A = The syllabotactic similarity score of 52/60 for the Cypriot Syllabic target text vs the Linear B benchmark text is *not* due to chance alone. *Corollary:* The Cypriot Syllabic and Linear B homomorphs in Table 8.1 are thus at least closely homophonous.

What the statistics in the control experiment cannot tell us, though, is precisely *why* the Cypriot Syllabic and Linear B texts form pairs in such similar ways; if we imagine Cypriot Syllabic and/or Linear B to be undeciphered, this would have to be a matter for interpretation. However, the simplest explanation for this phenomenon involves a single assumption: that in H_A above, the target and benchmark texts encode the same language (or perhaps two very closely related languages, or two dialects of the same language, or two chronological stages of the same language), as any explanation for this phenomenon based on the notion that the two scripts encode two *unrelated* languages invariably requires more than one assumption.[23] Thus Occam's

[22] Remember that, for the purposes of this control experiment, we are for the moment treating Cypriot Syllabic as an undeciphered script, and the phonetic values of the Cypriot Syllabic signs as unknown.

[23] Two examples: (1) The target and benchmark texts encode unrelated languages, and the languages share a large number of loanwords (two assumptions); (2) The target and benchmark texts encode unrelated languages, and the two languages are phonemically very similar, and the two languages happen to construct words using very similar syllabotactic constraints (three assumptions).

Razor suggests that, based on the data, we should adopt Hypothesis 1 (and its corollary) as the most preferable one for the first half of the control experiment:

Hypothesis 1:

In the control experiment, the Cypriot Syllabic and the Linear B texts both encode the same language (or perhaps two very closely related languages, or two dialects of the same language, or two chronological stages of the same language).

Corollary: The Cypriot Syllabic and Linear B homomorphs in Table 8.1 are thus at least closely homophonous, such that the phonetic values of the Cypriot Syllabic signs can tentatively be assigned to their Linear B homomorphs, and vice versa.

Importantly: note that we have been able to complete this analysis and formulate this hypothesis based solely on the *syllabotactic behaviour* of the Cypriot Syllabic and Linear B signs, not on their phonetic values (which are mentioned nowhere in this chapter) – that is, we would still have arrived at Hypothesis 1 even if Cypriot Syllabic and/or Linear B really were undeciphered.

Of course, both scripts have in fact been deciphered, and the phonetic values of the Cypriot Syllabic and Linear B homomorphs are known, so we can directly assess the accuracy of Hypothesis 1 and its corollary: they are both correct, in that Cypriot Syllabic and Linear B both encode Greek – although different dialects of Greek separated in time by a few centuries.[24] With regard to the corollary of Hypothesis 1, Table 8.17 shows the actual phonetic values of the ten Cypriot Syllabic and Linear B homomorphs in Table 8.1:

Table 8.17 *Actual phonetic values of the ten Cypriot Syllabic and Linear B homomorphs in Table 8.1*

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
CS	𐠀	𐠭	𐠙	𐠞	𐠡	𐠏	𐠨	𐠩	𐠯	𐠰
	a	ta, t^ha, da	na	pa, p^ha, ba	po, p^ho, bo	la	sa	se	ti, t^hi, di	to, t^ho, do
LB	𐀀	𐀅	𐀙	𐀞	𐀡	𐀨	𐀭	𐀮	𐀴	𐀵
	a	da	na	pa, p^ha, ba	po, p^ho, bo	ra, la	sa	se	ti, t^hi	to, t^ho

24 Linear B did not survive long past 1200 BC, while the earliest Cypriot Syllabic inscription is dated to 1050–950 BC (Duhoux 2012: 71).

Thus, the phonetic values of the Cypriot Syllabic and Linear B homomorphs are exactly the same in six cases, and partly the same in the other four. There are some other notable differences between the two scripts as well:[25]

- Linear B omits certain phonemes from its writing more often than Cypriot Syllabic does;
- The two scripts spell two-consonant clusters in different ways;
- The two scripts treat word-final consonants in different ways;
- Linear B almost always separates words with word-dividers, whereas Cypriot Syllabic is less rigorous in doing so (especially when it comes to the definite article: e.g. *to-na-ra-ku-ro-ne* /ton arguron/ 'the silver [acc.]'); and
- Linear B, especially at Knossos, contains a substantial number of non-Greek (most likely Minoan) person- and place-names, whereas Cypriot Syllabic does not.

The differences outlined in Table 8.17 and in the bulleted list above must surely be adding a certain amount of noise to the data; and yet the results in Table 8.14, together with the relative accuracy of Hypothesis 1 and its corollary, strongly suggest that this noise is simply not strong enough to prevent this method of syllabotactic analysis from detecting the signal that the same language is behind both scripts.

However, a crucial point must be made here: if Cypriot Syllabic and/or Linear B actually were undeciphered, we could not in any way claim that the analysis in the first half of the control experiment has on its own proven Hypothesis 1, which would still remain a hypothesis – though one supported by strong statistical data, such that we *could* validly claim that this hypothesis should be adopted as the prevailing one regarding the nature of the language behind the Cypriot Syllabic and Linear B texts used in the control experiments. Indeed – if either Cypriot Syllabic or Linear B were undeciphered, adopting this hypothesis as the prevailing one would clearly be a productive move for scholars of the undeciphered script, as the hypothesis correctly suggests that the script encodes Greek, and identifying the language behind an undeciphered script is a primary key to its decipherment.

As for the last half of the control experiment (evaluating the syllabotactic similarity score for the Cypriot Syllabic target text vs the Linear A benchmark text), Table 8.16 makes it clear that, in this case, we have no grounds for rejecting H_0:

[25] For a much fuller discussion of these differences, with examples, see Davis 2018: 387–8.

- H_0 = The syllabotactic similarity score of 30/60 for the Cypriot Syllabic target text vs the Linear A benchmark text is due to chance alone.

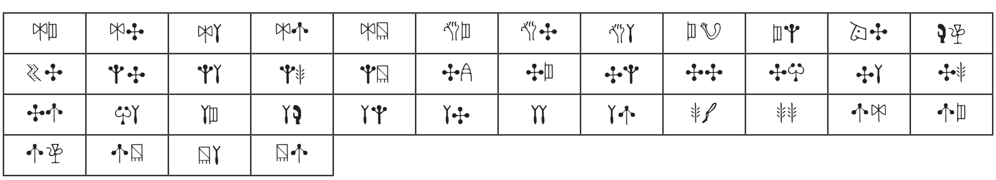
That is: the strong implication is that in the last half of the control experiment, the Cypriot Syllabic target text encodes a language *unrelated* to the language behind the Linear A benchmark text. Thus, as Greek is the language behind the Cypriot Syllabic target text, this result adds to the growing body of statistical evidence that Greek is *not* the language behind Linear A.

8.5 Steps 3b and 4b: Completing the Main Experiment

Step 3b *Scoring the Cretan Hieroglyphic Target Text for its Syllabotactic Similarity to the Linear B and Linear A Benchmark Texts*

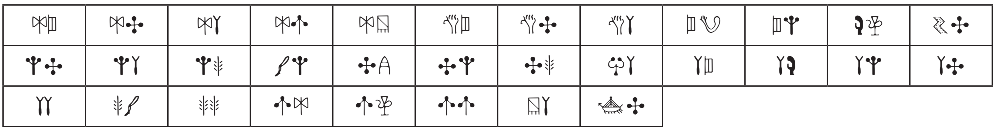
In this step, the Cretan Hieroglyphic target text is scored for its syllabotactic similarity to the Linear B and Linear A benchmark texts. As in the control experiment, the procedure is straightforward: first, determine the number of word-internal pairs in the Cretan Hieroglyphic target text (Table 8.5) whose *Linear B homomorphs* appear in the Linear B benchmark text (Table 8.3). (The Linear B homomorphs of the pairs in the Cretan Hieroglyphic target text are arrived at through consulting the list of Linear B/Cretan Hieroglyphic homomorphs in Table 8.2.) The results are shown in Table 8.18:

Table 8.18 *Forty pairs in the Cretan Hieroglyphic target text (Table 8.5) whose Linear B homomorphs appear in the Linear B benchmark text (Table 8.3)*



As there are sixty pairs in the Cretan Hieroglyphic target text, the syllabotactic similarity score for the Cretan Hieroglyphic target text vs the Linear B benchmark text is thus **40/60**. Finally, determine the number of word-internal pairs in the Cretan Hieroglyphic target text (Table 8.5) whose *Linear A homomorphs* appear in the Linear A benchmark text (Table 8.4). (The Linear A homomorphs of the pairs in the Cretan Hieroglyphic target text are arrived at through consulting the list of Linear A/Cretan Hieroglyphic homomorphs in Table 8.2.) The results are shown in Table 8.19:

Table 8.19 *Thirty-two pairs in the Cretan Hieroglyphic target text (Table 8.5) whose Linear A homomorphs appear in the Linear A benchmark text (Table 8.4)*



As there are (again) sixty pairs in the Cretan Hieroglyphic target text, the syllabotactic similarity score for the Cretan Hieroglyphic target text vs the Linear A benchmark text is thus **32/60**.

Step 4b *Evaluating the Scores*

Step 3b has now produced the following two syllabotactic similarity scores for the main experiment:

> The Cretan Hieroglyphic target text vs the Linear B benchmark text = 40/60
>
> The Cretan Hieroglyphic target text vs the Linear A benchmark text = 32/60

In this final step of the main experiment, as in the final step of the control experiment, each of these scores is evaluated against a representative average score produced by chance alone, as follows:

The Cretan Hieroglyphic Target Text vs the Linear B Benchmark Text

In Table 8.20, row LB1 contains the twenty-one Linear B signs from Table 8.2, while row LB2 contains the *Linear B homomorphs* of the twenty-one Cretan Hieroglyphic signs present in Table 8.5 (the pairs in the Cretan Hieroglyphic target text):

Table 8.20 *Row LB1: the twenty-one Linear B signs from Table 8.2; Row LB2: the Linear B homomorphs of the twenty-one Cretan Hieroglyphic signs present in Table 8.5 (the pairs in the Cretan Hieroglyphic target text)*

Rearranging the twenty-one signs in row LB2 of Table 8.20 in 1,000,000 different random ways;[26] retranscribing the *Linear B homomorphs* of the sixty Cretan Hieroglyphic pairs in Table 8.5 according to each random permutation of signs, and scoring each set of retranscribed pairs by counting how many of them also appear in the Linear B benchmark text produces the results shown in Table 8.21:

Table 8.21 *Syllabotactic similarity scores for 1,000,000 different sets of randomly retranscribed Linear B homomorphs of the Cretan Hieroglyphic pairs in Table 8.5 (the pairs in the Cretan Hieroglyphic target text) vs the Linear B benchmark text*

Score out of 60:	Permutations with that score:	% of permutations:	Score out of 60:	Permutations with that score:	% of permutations:
4	1	0.0001%	31	63,116	6.3116%
6	2	0.0002%	32	60,421	6.0421%
7	6	0.0006%	33	56,410	5.6410%
8	17	0.0017%	34	51,546	5.1546%
9	53	0.0053%	35	45,893	4.5893%
10	141	0.0141%	36	39,576	3.9576%
11	335	0.0335%	37	33,179	3.3179%
12	686	0.0686%	38	26,557	2.6557%
13	1,240	0.1240%	39	20,645	2.0645%
14	2,077	0.2077%	40	15,652	1.5652%
15	3,435	0.3435%	41	11,026	1.1026%
16	5,277	0.5277%	42	7,407	0.7407%
17	7,871	0.7871%	43	4,753	0.4753%
18	11,097	1.1097%	44	2,686	0.2686%
19	15,113	1.5113%	45	1,566	0.1566%
20	19,911	1.9911%	46	744	0.0744%
21	25,525	2.5525%	47	339	0.0339%
22	31,069	3.1069%	48	153	0.0153%
23	37,374	3.7374%	49	53	0.0053%
24	43,707	4.3707%	50	17	0.0017%
25	49,303	4.9303%	51	1	0.0001%
26	54,630	5.4630%	**Tot. permutations:**	1,000,000	100%
27	59,162	5.9162%	**Average score:**	29.531 / 60	
28	62,420	6.2420%	**Std deviation (σ):**	5.941 / 60	
29	63,486	6.3486%	**Avg. score + 2σ:**	41.413 / 60	
30	64,322	6.4322%	**Score of 40/60:**	Avg. + 1.76σ	$p = 0.0444$

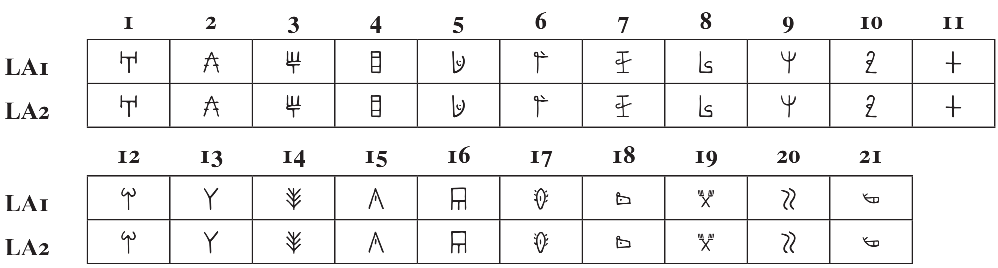
[26] The total number of possible permutations is equal to 21! ('21 factorial') = the product of integers 1 through 21 = 5.1×10^{19}.

As Table 8.21 shows, the 1,000,000 random permutations produce an average score of 29.531/60, with a standard deviation (σ) of 5.941/60, such that the region of 'statistical significance' (2σ or more above the average) begins at 41.413/60 (i.e. scores of **42/60** or greater). Meanwhile, the original syllabotactic similarity score for the Cretan Hieroglyphic target text vs the Linear B benchmark text (as shown in Table 8.18) is **40/60**, which is **1.76σ above** the average score produced by chance alone (i.e. not within the region of statistical significance); and the *p*-value (produced by adding the percentages in the right column for all scores of 40/60 or more) is 0.0444, meaning that of the 1,000,000 random permutations, 4.44% of them scored 40/60 or higher. The potential meanings of this score will be discussed shortly.

The Cretan Hieroglyphic Target Text vs the Linear A Benchmark Text

In Table 8.22, row LA1 contains the twenty-one Linear A signs from Table 8.2, while row LA2 contains the *Linear A homomorphs* of the twenty-one Cretan Hieroglyphic signs present in Table 8.5 (the pairs in the Cretan Hieroglyphic target text):

Table 8.22 *Row LA1: the twenty-one Linear A signs from Table 8.2; row LA2: the Linear A homomorphs of the twenty-one Cretan Hieroglyphic signs present in Table 8.5 (the pairs in the Cretan Hieroglyphic target text)*



Rearranging the twenty-one signs in row LA2 of Table 8.22 in 1,000,000 different random ways;[27] retranscribing the *Linear A homomorphs* of the sixty Cretan Hieroglyphic pairs in Table 8.5 according to each random permutation of signs; and scoring each set of retranscribed pairs by counting how many of them also appear in the Linear A benchmark text produces the results shown in Table 8.23:

[27] The total number of possible permutations is 5.1×10^{19}; see note 26.

Table 8.23 *Syllabotactic similarity scores for 1,000,000 different sets of randomly retranscribed Linear A homomorphs of the Cretan Hieroglyphic pairs in Table 8.5 (the pairs in the Cretan Hieroglyphic target text) vs the Linear A benchmark text*

Score out of 60:	Permutations with that score:	% of permutations:	Score out of 60:	Permutations with that score:	% of permutations:
2	4	0.0004%	24	46,370	4.6370%
3	13	0.0013%	25	36,004	3.6004%
4	72	0.0072%	26	26,509	2.6509%
5	214	0.0214%	27	18,853	1.8853%
6	697	0.0697%	28	12,570	1.2570%
7	1,689	0.1689%	29	7,876	0.7876%
8	3,506	0.3506%	30	4,709	0.4709%
9	6,878	0.6878%	31	2,653	0.2653%
10	12,215	1.2215%	32	1,490	0.1490%
11	19,362	1.9362%	33	732	0.0732%
12	28,724	2.8724%	34	355	0.0355%
13	39,537	3.9537%	35	143	0.0143%
14	51,576	5.1576%	36	58	0.0058%
15	63,740	6.3740%	37	25	0.0025%
16	73,358	7.3358%	38	10	0.0010%
17	81,344	8.1344%	39	3	0.0003%
18	85,821	8.5821%	**Tot. permutations:**	1,000,000	100%
19	87,278	8.7278%	**Average score:**	18.944/60	
20	83,192	8.3192%	**Std deviation (σ):**	4.489/60	
21	77,317	7.7317%	**Avg. score + 2σ:**	27.922/60	
22	67,602	6.7602%	**Score of 32/60:**	Avg. + 2.91σ	$p = 0.0028$
23	57,501	5.7501%			

As Table 8.23 shows, the 1,000,000 random permutations produce an average score of 18.944/60, with a standard deviation (σ) of 4.489/60, such that the region of 'statistical significance' (2σ or more above the average) begins at 27.922/60 (i.e. scores of **28/60** or greater). Meanwhile, the original syllabotactic similarity score for the Cretan Hieroglyphic target text vs the Linear A benchmark text (as shown in Table 8.19) is **21/60**, which is **2.91σ above** the average score produced by chance alone (i.e. well within the region of statistical significance); and the *p*-value (produced by adding the percentages in the right column for all scores of 32/60 or more) is 0.0028, meaning that of the 1,000,000 random permutations, just 0.28% of them scored 32/60 or higher.

We now have final evaluations of the two syllabotactic similarity scores produced by the main experiment, with respect to representative averages produced by chance alone:

The Cretan Hieroglyphic target text vs the Linear B benchmark text = 40/60 = 1.76σ above the average

The Cretan Hieroglyphic target text vs the Linear A benchmark text = 32/60 = 2.91σ above the average

The right half of Figure 8.2 depicts these two evaluations, while the left half of Figure 8.2 contains the two evaluations produced earlier by the control experiment (as shown before in Figure 8.1). The region of statistical significance is shaded light grey:

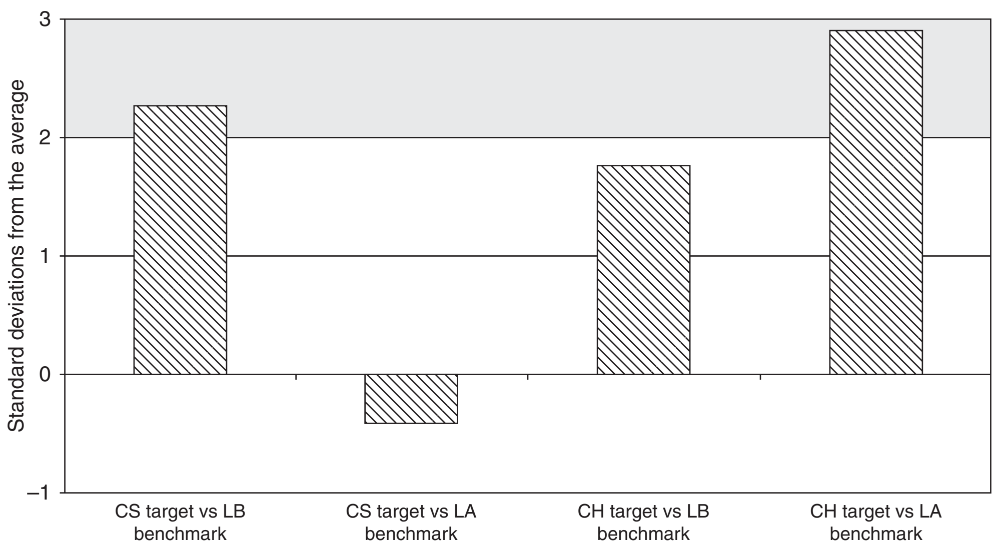


Figure 8.2 Evaluation of results: syllabotactic similarity of Cypriot Syllabic and Cretan Hieroglyphic target text vs Linear B and Linear A benchmark texts

8.6 Results of the Main Experiment: Discussion

In the first half of the main experiment (evaluating the syllabotactic similarity score for the Cretan Hieroglyphic target text vs the Linear B benchmark text), Table 8.21 makes it clear that we have no grounds for rejecting H_o:

- H_o = The syllabotactic similarity score of 40/60 for the Cretan Hieroglyphic target text vs the Linear B benchmark is due to chance alone.

That is: the strong implication is that the Cretan Hieroglyphic target text encodes a language *unrelated* to the language behind the Linear B benchmark text. Thus, as Greek is the language behind the Linear B benchmark text, this result constitutes strong statistical evidence that Greek is *not* the language behind Cretan Hieroglyphic. However, in the last half of the main experiment (evaluating the syllabotactic similarity score for the Cretan Hieroglyphic target text vs the Linear A benchmark text), Table 8.23 clearly tells us that, in this case, the syllabotactic similarity score is to a statistically significant degree 'statistically incompatible'[28] with the 'null hypothesis' that the score is due to chance alone (H_o), and that we should therefore reject H_o in favour of the 'alternative hypothesis' (H_A):

- H_A = The syllabotactic similarity score of 32/60 for the Cretan Hieroglyphic target text vs the Linear A benchmark text is *not* due to chance alone.

By the same logic outlined in the discussion of the results of the first half of the control experiment, rejecting H_o in favour of H_A results in an inextricably linked corollary of H_A regarding the phonetic values of the Cretan Hieroglyphic and Linear A signs, such that H_A can be augmented with that corollary as follows:

- H_A = The syllabotactic similarity score of 32/60 for the Cretan Hieroglyphic target text vs the Linear A benchmark text is *not* due to chance alone.

  Corollary: The Cretan Hieroglyphic and Linear A homomorphs in Table 8.2 are thus at least closely homophonous.

As in the discussion of the first half of the control experiment, the simplest explanation for the fact that the Cretan Hieroglyphic and Linear A homomorphs form pairs in such similar ways is that in H_A above, the target and benchmark texts encode the same language (or perhaps two very closely related languages, or two very similar dialects of the same language, or two chronological stages of the same language); thus Occam's Razor suggests that, based on the data, we should adopt Hypothesis 2 (and its corollary) as the most preferable one for the last half of the main experiment:

> *Hypothesis 2:*
> In the main experiment, the Cretan Hieroglyphic and Linear A texts both encode the same language (or perhaps two very closely

[28] Wasserstein and Lazar 2016: 131.

related languages, or two very similar dialects of the same language, or two chronological stages of the same language).

Corollary: The Cretan Hieroglyphic and Linear A homomorphs in Table 8.2 are thus at least closely homophonous, such that the phonetic values of the Linear A signs can tentatively be assigned to their Cretan Hieroglyphic homomorphs.

Of course, as is the case with Hypothesis 1 in the discussion of the first half of the control experiment, the analysis in the last half of the main experiment has not *proven* Hypothesis 2: it still remains a hypothesis – though one supported by strong statistical data, such that this hypothesis should be adopted as the prevailing one regarding the notion of a linguistic connection between Linear A and Cretan Hieroglyphic. Indeed, the results of the first half of the control experiment strongly suggest that adopting Hypothesis 2 as the prevailing one would be a productive move for scholars of Linear A and Cretan Hieroglyphic – because this hypothesis has at least two important implications that could very well be of substantial assistance in the process of deciphering both scripts:

(1) Assigning tentative phonetic values to Cretan Hieroglyphic signs based on the phonetic values of their Linear A homomorphs may be a much more productive and valuable method than has been previously thought; and
(2) The notion of a linguistic connection between Linear A and Cretan Hieroglyphic effectively links the decipherment of the two scripts, in that advances in the study of one of them have the potential to produce parallel advances in the study of the other, while the decipherment of one of them could very well lead to decipherment of the other.

Thus, we should employ Hypothesis 2 to underpin all future work on Linear A and Cretan Hieroglyphic, at least until we have a strong, data-supported reason for doing otherwise.

Acknowledgements. I am deeply indebted to the Institute of Classical Studies, University of London, for supporting this research through the 2019 Michael Ventris Award for Mycenaean Studies; to Daniel Stiebel, for giving me invaluable feedback on an early version of this study; and to the editors of this volume, for offering me the opportunity to publish this research in conjunction with so many eminent scholars of Aegean syllabic scripts.