ISSN: 0544-3733

Further statistical analysis of the personal names used on Crete during the late Bronze Age[1]

Richard J. FIRTH
Bristol

This paper continues the analysis of the Greekness of personal names that was begun by the author in an earlier paper (Firth 1993). In that paper, it was demonstrated that it was possible to use a statistical approach to assess the Greekness of groups of names taken from the Knossos Linear B tablets. Furthermore, it was shown that there was a surprisingly good level of agreement between the results of that statistical assessment and assessments of Greekness by independent experts.

The aim first of this paper is to develop the techniques and the analysis put forward in the earlier paper. This will be done by extending the range of comparisons with the assessments of independent experts and reinforcing the demonstration of validity of the basic methodology.

The paper will then go on to focus on the non-Greek names of shepherds on the Knossos Linear B tablets. It will be shown that there appears to be a variation in the frequency of usage of free vowels between different regions of Crete. It will be demonstrated that this regional variation in the non-Greek names in Linear B is matched by a corresponding regional variation in the usage of free vowels in Linear A words. It is suggested that this variation is due to the extent of usage of the prothetic *a-* across the island (and in different chronological periods). At Haghia Triada in LM IB, the usage of free vowels appears to have been reduced to a minimum and the spelling rules adopted there largely prevailed when non-Greek names were written in Linear B.

On the basis of this finding, it is judged to be necessary to modify the statistical analysis. In the original analysis, the frequency of the usage of consonants and free

[1] The author wishes to gratefully acknowledge the help of José Melena and Stephie Nikoloudis who provided comments on an earlier draft of this paper.

vowels in non-Greek names was represented by the frequencies found in Linear A. However, since the usage in Linear A is shown to be variable, then it is more appropriate to use a set of frequencies based on the non-Greek shepherds' names from the Linear B tablets. In practice, this does not result in major changes to the assessed level of Greekness of groups of names although it often reduces the measure of uncertainty that is used in the assessment.

Finally it is suggested that the Linear A words *MA-KA-I-TA* can be interpreted as a Greek word that has been borrowed into the Minoan vocabulary.

1. Brief description of the methodology developed in the earlier paper

The earlier statistical paper (Firth 1993) was based on the marked difference between the frequency of usage of consonants and free vowels (expressed as percentages) within Greek names of men written on the Linear B tablets and within Minoan words written in Linear A (see Table 1).

Table 1. Percentage frequencies of consonants and free vowels

Frequency of usage (%)	Pylos Greek names (G_i)	Linear A (A_i)
vowel	17.67	10.64
d	2.96	9.67
j	7.31	4.90
k	11.42	10.80
m	6.49	7.12
n	5.26	9.06
p	6.66	5.83
q	2.71	2.61
r	15.12	13.00
s	3.12	10.35
t	12.24	12.19
w	8.79	2.61
z	0.25	1.16

The frequencies for Pylos Greek names were derived from the Pylos men's names identified as Greek within the Glossary of *Documents* (2nd edition). The frequencies for Linear A words were taken directly from the table of distributions given by Duhoux (1982 p. 254).

It was demonstrated that it was possible to estimate the proportion of Greekness within a group of men's names by using a statistical technique based on the usage of consonants and free vowels within the group. This distribution was specified in an

array for each group of names being considered. The array for each test group, T_i, was compared to the arrays for Pylos Greek names, G_i, and for Linear A usage, A_i.

$$\text{Residual Sum} = \text{minimum} \sum [\, y\, G_i + (1 - y)\, A_i - T_i \,]^2$$

[summing over all locations in the array (i), where $0 \leq y \leq 1$]. The proportion, y, was varied until the residual sum was minimised. The residual sum is a measure of uncertainty in the accuracy of the fit so that, if the residual sum is very large, then the assessment is less reliable.

As an example, when the test group was all men's names at Knossos, then y = 0.566, with a residual sum of 13, i.e. 56.6% Greekness.

In mathematical terms, this is a "least squares fit" and fits of this kind are the basis for a large amount of analysis of scientific data. The novel suggestion here is the application of this standard method of scientific analysis to data derived from the writings of Bronze Age Crete.

This approach has been found to be surprisingly robust, so long as the analysis does not result in a residual sum which is too large. Its success derives primarily from the fortuitous finding that the frequency of usage of consonants and free vowels within large groups of Greek names (written in Linear B) does not vary greatly.

This statistical approach has the advantage over the conventional method (which considers each name separately) because it sets a consistent standard across all groups of names that are considered. By contrast, the conventional approach tends to be dependent upon whether the particular expert doing the analysis is "bold" or "conservative" in judging whether particular names are Greek or non-Greek.

2. Comparison with independent assessments

In section 7.3 of the earlier paper (Firth 1993), the results of the statistical method were compared with those obtained by a number of independent methods. Table 2 includes the four sets of results already given in the earlier papers (marked with an asterisk) and adds several others.[2]

Table 2

	Independent assessment of % Greekness	Predicted % Greekness	No. of signs	Residual Sum
* Homeric names	100	100	270	49
* KN Sc + Vc(1)-series	84	89	286	24
Mycenae	72	58	364	27

[2] A commentary on the sources used for the new assessments is given in Appendix A.

	Independent assessment of % Greekness	Predicted % Greekness	No. of signs	Residual Sum
KN B-series	71	78	298	8
KN D-series	44	35	1095	18
* KN As-series	35	45	436	35
KN C-series	31	59	203	140
KN Ap-series	18	30	169	162
* KN 2-sign names	7	4	200	136
KN names ending in -i	0	6	118	158
Lists of non-Greek names				
Astour (Semitic names)	0	48	257	116
Billigmeier	0	1	228	133
Ilievski (KN D-series)	0	3	599	63
Killen	0	0	163	102
Landau	0	14	200	107

The main conclusion is that the further analyses support the validity of the statistical analysis as a method for indicating the level of Greekness within a group of names.

The exception is the attempt by Astour (1965, pp. 336-344) to isolate of group of Linear B names which are based on Semitic words. In this case, the statistical analysis shows that the group of names proposed by Astour has a level of Greekness of 48%, which is comparable to that of any randomly selected group of names at Knossos (57%). This finding supports the *communis opinio* of Linear B experts, that Astour's claims are not valid in this respect.

Another conclusion is that the residual sums are substantially higher if the level of non-Greekness is high. In practice, this arises because the frequency of usage of consonants and free vowels is repeatable for Mycenaean Greek names but much more variable for non-Greek names.

The above table includes an assessment of the Greekness of personal names from the Mycenae tablets. In order to give some overall perspective, Table 3 gives the predicted percentage of Greekness of the names from the four main sites from which Linear B tablets have been found.[3]

[3] The results for Knossos and Pylos were given in the earlier paper (Firth 1993). The results for Thebes are based on men's names. They include the newly published names (Aravantinos *et al.* 2001) but exclude the names found on the Inscribed Stirrup Jars (because the latter originated from Crete).

Table 3

	Predicted % Greekness	No. of signs	Residual Sum
Pylos (males only)	88	2624	8
Knossos (males only)	57	3094	13
Mycenae	58	364	27
Thebes (males only)	91	366	18

3. Sensitivity study using an alternative frequency array for non-Greeks

I received the criticism that the original statistics paper used Duhoux's array of Linear A signs and, in the absence of a decipherment of Linear A, the transliteration of the Linear A signs is not proven. It would be possible to respond to this by setting out references to the detailed arguments which have been put forward to justify the readings of the individual Linear A signs which are now widely accepted. However, it is more convincing to demonstrate that the results of the analysis can be achieved using an alternative statistical approach that does not rely on the transliteration of Linear A signs.

The alternative statistical approach is based on the large group of non-Greek shepherds' names given by Ilievski (1992) from the D-series of the Knossos Linear B tablets. This group has already been considered in Table 2 where it was assessed, using the original statistical method to have a very low level of Greekness. We can develop an alternative statistical method using the same basic approach but replacing the frequency array of consonants and free vowels based on the Linear A signs with one based on Ilievski's non-Greek shepherds' names. It is convenient to refer to the original method (using Duhoux's Linear A array) as Method A and that using an array derived from Ilievski's non-Greek shepherds as Method B. For completeness, a comparison of the two frequency arrays is given in Table 4.

Table 4

	Duhoux' Linear A (Method A)	Ilievski (Method B)
vowel	10.64	7.32
d	9.67	7.49
j	4.90	4.99
k	10.80	8.32
m	7.12	5.99
n	9.06	8.82
p	5.83	2.66

	Duhoux' Linear A (Method A)	Ilievski (Method B)
q	2.61	4.33
r	13.00	15.14
s	10.35	11.15
t	12.19	14.64
w	2.61	5.99
z	1.16	3.16

Table 5 compares the predicted levels of Greekness of the names from the four main Linear B sites using Methods A and B.

Table 5

Predicted % Greekness	Method A	Method B
Pylos (males only)	88	84
Knossos (males only)	57	56
Mycenae	58	62
Thebes (males only)	91	86

This shows that the effect of changing from Method A to Method B is small and within the accuracy that might reasonably be expected of these statistical methods.

We might suppose that the difference between the two methods would be greater when the groups of names contain low levels of Greekness, as the use of a different non-Greek array should be more significant. However, again, we generally find a high level of agreement (see Table 6). The greatest variation is for the list of non-Greek names extracted from Killen (2004). This is because those names are not ideally suited to this type of analysis as they were chosen because they included particular signs, i.e. *me*, *mo*, *ne*, *pi-ja-*, *pu*, *re*, *se*, *te* and so there is an implicit bias in the frequency distribution of Linear B signs. Therefore we should not place undue weight on the analysis of that list of names here.

Table 6

Predicted % Greekness	Method A	Method B
KN 2-sign names	4	5
KN names ending in *-i*	6	6
Billigmeier's list of non-Greek names	1	7
Killen's list of non-Greek names	0	17
Landau's list of non-Greek names	14	11

The conclusion is that the statistical method is not sensitive to which of the two arrays (given in Table 4) is used to represent non-Greek names. This finding arises because the effectiveness of the statistical method is based on the repeatability of the frequency of consonants and free vowels amongst different groups of Greek names. By contrast, the non-Greekness is effectively defined by the lack of alignment to the Greek frequency array. Thus the validity of the statistical method is not dependent on the transliteration of Linear A signs.

4. Further comparisons with independent assessments of Greekness

Thus far it has been demonstrated that statistical Method A gives results which are in agreement with independent assessments. It has also been shown that its results are not sensitive to its use of a different frequency array for non-Greek names (i.e. using Methods A or B). The next step is to demonstrate the use of Methods A & B as a basis for examining more detailed analyses of the Greekness of personal names by different experts.

The conventional approach to assessing the Greekness of a group of Linear B names is to consider the Greekness of each name separately and then to draw up separate lists of the Greek and the non-Greek names. Some experts require a higher level of confidence before assigning a name to be Greek. In other cases, names which could be interpreted as both Greek or non-Greek are listed separately as "probably Greek" or "possibly Greek" and so on. Experts frequently identify a name as Greek although it is an ethnic name based on a non-Greek toponym. Because of such differences, different experts can produce different lists of Greek and non-Greek names even though they are considering the same data.

A number of the comparisons with independent assessments given in Section 2 were based on mixed groups of Greek and non-Greek names. It is possible to repeat some of these comparisons where the groups have already been divided into Greek and non-Greek lists. The key benefit of using the statistical method to examine the results of these previous analyses is that the comparisons are done against a consistent standard.

4.1 *Ilievski's analysis of the shepherds in the D-series*

Ilievski (1992) assesses the Greekness of shepherds' names by considering each name individually. Of the ~350 shepherds' names (excluding collectors), he suggests that:

40 are Greek with a high degree of certainty [Ilievski's group (a)]
44 are probably Greek [Ilievski's group (b)]

71 are possibly Greek [Ilievski's group (c)]
197 are non-Greek [which we shall name group (d)]

If we apply the statistical methodology of Methods A & B to each of Ilievski's groups of names then we obtain the results given in Tables 7a & 7b.[4]

Table 7a

Ilievski's Groups of Shepherds	Predicted % Greekness (Method A)	Residual Sum
(a) high degree of certainty	91	66
(b) probable	73	25
(c) possible	60	84
(d) non-Greek[5]	3	63

Table 7b

Ilievski's Groups of Shepherds	Predicted % Greekness (Method B)	Residual Sum
(a) high degree of certainty	100	68
(b) probable	74	23
(c) possible	70	95
(d) non-Greek[6]	0	0

The percentages of Greekness are in reasonable agreement between Method A and Method B. Furthermore, both variations of the statistical method are in good agreement with Ilievski's analysis, in that,

- group (a) names have a high degree of Greekness,
- groups (b) and (c) have progressively reduced levels of Greekness, but are nevertheless more Greek than non-Greek,
- group (d) names have a very low level of Greekness.

This is clearly very encouraging.

4.2 *Baumbach's analysis of the Ap, As & C-series*

Baumbach lists the Greek names within each of the Ap, As and C-series. However, in view of the smallness of the numbers of Greek names in the Ap and C-series, it

[4] There were 143 signs in the analysis of group (a); 142 signs for group (b); 211 for group (c); and 599 for group (d).
[5] This result has already been given in Table 2.
[6] This result gives a Greekness which is identically equal to zero precisely because Method B uses Ilievski's non-Greek shepherds' names as the basis for establishing non-Greekness.

is convenient to combine the lists of Greek names for the Ap, As and C-series (and similarly for the non-Greek names). If we analyse arrays of signs from Baumbach's Greek and non-Greek groups of names then we get the results given in Table 8.[7]

TABLE 8

Ap, As & C-series Baumbach's categories	Predicted % Greekness (Method A)	Predicted % Greekness (Method B)
Greek names	93	82
Non-Greek names	23	24

The analyses both agree that Baumbach has succeeded in isolating most of the Greek names. However, the results imply that there remains a proportion of Greek names amongst the list of names that Baumbach has judged to be non-Greek. This is consistent with Ilievski's suggestion that Baumbach had been too cautious in identifying names which are Greek, erring towards the assumption that "doubtful names" are more likely to be non-Greek than Greek.[8]

4.3 *Varias' analysis of the personal names from the KN B-series and from Mycenae*

Varias separated the personal names from the KN B-series and from Mycenae into three groups: Greek names; doubtful Greek names; and non-Greek names. If we take these six groups individually, they are relatively small and so, for the purposes of this paper, we will combine the names from the KN B-series and Mycenae to form three groups: Greek names; doubtful Greek names; and non-Greek names. Table 9 gives the results of the statistical analysis using these three groups using both Methods A and B.[9]

TABLE 9

Varias' groups	Predicted % Greekness (Method A)	Predicted % Greekness (Method B)
Greek names	66	72
Doubtful Greek names	85	86
Non-Greek names	50	49

[7] The residual sums for Method A are 55 and 51 for Baumbach's groups of Greek and non-Greek names respectively. There are 255 signs in the analysis of Greek names and 553 signs for non-Greek names.

[8] Ilievski 1992, p. 323, footnote 8. For example, Baumbach implied that the following names are non-Greek, *i-ne-u*, *ke-u-po-da*, *mi-ru-ro*, *o-ku*, *qe-ro*, *qo-te-ro*, *wi-ja-ma-ro*, whereas Ilievski (1992) suggests that they are Greek. Similarly, *a-qi-ti-ta*, *i-ta-mo*, *wa-ra-ti* are given as Greek or Greek? by Landenius-Enegren (1995) although Baumbach has implied that they are non-Greek.

[9] The residual sums for Method A are, respectively, 27, 31 and 50 for Varias' Greek, doubtful Greek and non-Greek groups and the number of signs in the analysis of each group is 366, 161 and 135, respectively.

The results of Methods A and B are in good agreement. However, in view of the success that the statistical method has had in interpreting the results by Ilievski and Baumbach, the present results seem surprising, at first sight, because Varias' group of "Greek names" would appear to be less Greek than the "Doubtful Greek names".

On closer inspection, we can see that Varias has included ethnics in his list of Greek names, even though many Greek place-names are pre-Greek in origin. This highlights a difference between the statistical analysis and the conventional approach. It is conventional to regard all names ending in *-ios* as Greek whatever the base to which the suffix is applied, since the *-ios* ending is a Greek form. However, the statistical assessment is based on the frequency of signs and so if the base of a name is pre-Greek then including it in a group of names being tested is likely to tilt the assessment towards a lower level of Greekness.

We can demonstrate this as follows. According to the statistical analyses, Varias' "Doubtful Greek names" are probably Greek. Therefore, we will combine Varias' Greek and Doubtful Greek names but remove the ethnics from this list. Instead, the ethnics will be included as non-Greek. This gives the following results (using Method A).

Table 10

Varias' groups	% Greekness	Residual Sum
Greeks + Doubtful Greeks (exc. ethnics)	76	16
Non-Greeks plus ethnics	45	96

This gives some improvement. However, the statistical analysis would tend to imply that there is still the possibility that Varias might have been too cautious in identifying Greek names, leading to the relatively high percentage of Greekness amongst names which he has listed as non-Greek.[10]

5. Considering regional variation of Cretan names

In section 4.1, we included a discussion of Ilievski's analysis of the Knossos shepherds' names into lists: Greek (a); Greek (b); Greek (c); and (d) non-Greek. In this section, we are going to extend that discussion further by separating the shepherds' names into regional groups according to their associated toponyms. We shall use the regional groups of toponyms set out by Wilson (1977), given in Table 11.

[10] For example,
a-wo-ro: *Documents*[2] interprets this as *Aworos*, implying Greekness.
]e-ri-ko[: if we take this to be *]e-ri-ko[-wo?*, then *Documents*[1] reads this as *Erikowos* or *Erigowos*.
o-to-wo-wi-je is compared by Ruijgh (1967, p. 294 FN 22) to *Orthwowes*, though he notes that the *-i-je* ending is obscure.

Table 11

Group I	Group II	Group III	Group IV	Group V
da-wo	pu-na-so	a-ka	o-du-ru-wo	a-mi-ni-so
da-*22-to	ra-ja	pu-so	*56-ko-we	ko-no-so
e-ko-so	ra-su-to	qa-mo	si-ra-ro	se-to-i-ja
e-ra	ra-to	qa-ra	wa-to	
pa-i-to	ri-jo-no	ru-ki-to		
ku-ta-to	tu-ni-ja	su-ri-mo		
	do-ti-ja	ti-ri-to		
		tu-ri-so		
		u-ta-no		

In practical terms, we shall be looking at Groups I, II and III. (We do not have lists of names associated with the western towns given in Group IV. The Group V places are listed together because they are important centres, rather than because they are a geographical grouping.)

We will begin with the groups of names divided between "Greek" and "Non-Greek" (as set out by Ilievski 1992) and divided between the regions (as set out by Wilson). The results using statistical Method A are given in Table 12. This table also includes the results for the combined group of Greek and Non-Greek names.

Table 12

Statistical assessment	Group I		Group II		Group III	
	% Gkness	resid. sum	% Gkness	resid. sum	% Gkness	resid. sum
Combined group	34	52	44	34	33	34
"Greek"	79	43	91	152	71	48
"Non-Greek"	0	137	8	77	5	83

At first sight, this appears to be reasonably successful, as the groups containing names which Ilievski has assessed to be Greek are predicted to have relatively high levels of Greekness (using the statistical method) and similarly for the groups assessed to be non-Greek. However, there is an underlying problem which is indicated by the high residual sum associated with the Greeks of Group II.[11]

We can see, from Table 12, that the Greekness of the combined groups of names from each of Groups I, II and III is roughly the same (varying between 33-44%). So

[11] We have already noted above that the residual sums are typically lower for groups with high levels of Greekness, compared to groups with low levels of Greekness.

we might suppose that these groups of combined names have similar frequencies of consonants and free vowels. Table 13a shows the frequencies from the three groups. Table 13b gives the squares of the differences between pairs of groups. The sum of these squares of differences is a measure of the level of agreement between these sets of frequencies.

Table 13a

	Group I	Group II	Group III
free vowels	10.2	15.9	10.7
d	5.9	5.7	5.4
j	7.1	4.9	7.0
k	9.4	11.8	10.7
m	6.1	7.3	4.4
n	8.1	9.3	7.4
p	2.5	3.3	4.4
q	4.6	1.6	4.0
r	17.0	12.2	14.1
s	7.4	9.3	9.4
t	15.0	10.6	14.8
w	4.8	6.9	5.7
z	1.8	1.2	2.0
	100	100	100

Table 13b

	Gps I & II	Gps II & III	Gps I & III
free vowels	32.21	26.17	0.31
d	0.03	0.10	0.23
j	5.05	4.70	0.01
k	5.64	1.10	1.75
m	1.46	8.73	3.04
n	1.46	3.87	0.58
p	0.50	1.23	3.30
q	8.73	5.76	0.31
r	23.55	3.61	8.73
s	3.88	0.00	4.07
t	19.75	17.61	0.06
w	4.31	1.45	0.76
z	0.32	0.63	0.05
Sum of Squares:	106.88	74.97	23.20

It can be seen that Groups I & III show similar frequencies with a relatively low sum of squares (i.e. 23.2). However, Group II differs markedly from Group I and III, particularly in respect of the frequency of free vowels, which is substantially higher in Group II (see Table 13a).

In order to understand this finding we must consider the variation of signs used in the Linear A inscriptions. Packard (1974) showed that the frequency of signs on the Haghia Triada tablets differed from that on the remaining inscriptions.[12] At that stage, this might have been attributed to the difference between the economic records from Haghia Triada and the religious inscriptions which made up much of the remaining inscriptions that were available to Packard. However, there have been a significant number of new inscriptions published since Packard's analysis. For the purposes of this paper, the frequencies of Linear A signs have been derived using the transliterations proposed on Younger's web-site.[13] Since this paper is primarily concerned with personal names, Linear A words that have been tentatively identified by Younger as being transaction terms were ignored, similarly, religious inscriptions have not been included.[14] On this basis, Table 14 contains the frequency distributions from the Haghia Triada inscriptions and the remaining (non-Haghia Triada) inscriptions.

Table 14. Percentage frequencies of consonants and free vowels

	Haghia Triada (HT_i)	Non-Haghia Triada ($nonHT_i$)
vowel	7.84	15.63
d	11.65	8.04
j	4.85	5.78
k	12.27	8.49
m	7.22	8.95
n	8.45	7.81
p	4.74	5.44

[12] Packard 1974, Appendix E

[13] http://www.people.ku.edu/~jyounger/LinearA/ (downloaded on 9 October 2004). By way of verification of the contents of his web-site, Younger states that "The transcribed texts are based on a transnumeration and phonetic normalization finished 22 March 1994 by John G. Younger; Jean-Pierre Olivier checked this document against *GORILA* vols. I-V and a ms. of VI, and put in tabular form in January and February 1997. Since then, there have been continual updates". It could be argued that Younger's web-site should not be considered as a source for transliteration of Linear A signs. However, it should be emphasised that this is a statistical study and, for present purposes, it is not essential that every single sign is correctly interpreted.

[14] In order to avoid double accounting, each word was only counted once and if an incomplete word could potentially be the same as a longer word then it was not included in the count.

	Haghia Triada (HT_i)	Non-Haghia Triada ($nonHT_i$)
q	3.81	2.15
r	14.85	10.31
s	9.69	10.53
t	12.68	13.25
w	1.55	1.81
z	0.41	1.81

The most substantial difference between these two frequency distributions is the frequency of free-vowels, where the frequency for the non-Haghia Triada inscriptions is double that at Haghia Triada.

At this stage, we will introduce statistical Method C. This is very similar to the previous methods, however, Method C is intended not only to estimate the proportion of Greekness within a group of men's names but also to determine the proportions of "Haghia Triada" and "non-Haghia Triada" non-Greekness. Therefore, in this case, the array for each test group, T_i , was compared to the arrays for Pylos Greek names, G_i , and for Haghia Triada usage (HT_i) and non-Haghia Triada usage ($nonHT_i$).

$$\text{Residual Sum} = \text{minimum} \sum [(y + z) G_i + (1 - y) HT_i + (1- z) nonHT_i - T_i]^2$$

[summing over all locations in the array (i), where $0 \le y \le 1$, $0 \le z \le 1$, $0 \le (y+z) \le 1$]. The proportions, y and z, were varied until the residual sum was minimised.

Table 15 presents the results of the analysis of the three regional groups of shepherds' names using Method C. It also includes the analysis for the complete set of men's names found at Knossos.

Table 15

	Greek (%)	Haghia Triada (%)	non-Haghia Triada (%)	Residual Sum
Group I	40	60	-	41
Group II	40	13	47	34
Group III	39	51	10	35
All KN men's names	62	35	3	14

The statistical analysis suggests that the non-Greek names in Group I are all characterised by the Haghia Triada frequency of consonants and free vowels[15]. This is

[15] I.e. in Table 15, the percentage for non-Haghia Triada is zero.

also largely true of the Group III names. However, the non-Greek names in Group II are predominantly characterised by the non-Haghia Triada frequency of consonants and free vowels.

We should now recall that Group I names are from a geographical region including Phaistos, which would, of course, include Haghia Triada. Group III includes Tylisos and probably represents a geographical region near to Tylisos and Knossos.[16] Group II toponyms are not so easy to locate because they do not contain a toponym which can be securely identified with a physical location.

In other words, it is possible to use the variation in the frequency of consonants and free vowels from the Linear A inscriptions and predict which of the three groups of shepherds from the Linear B tablets is most likely to represent the region that includes Haghia Triada. Furthermore, it is also possible to use them to predict which of the three groups is least influenced by Haghia Triada.

This is clearly an important result because it is suggesting, firstly, that there was a variation in the frequency of usage of consonants and free vowels in Linear A words between Haghia Triada and the other inscriptions and, secondly, that this variation persisted and was still present in the non-Greek Linear B names of the shepherds.

The obvious next question is what is the implication of this difference between Haghia Triada words and those from other inscriptions. One's initial thoughts might be that it could represent a difference of language or of personal names arising from different linguistic influences. However, the simplest and most likely explanation is that the Haghia Triada scribes were less likely to have included prothetic initial vowels.[17]

If we exclude the religious inscriptions, transaction terms, repeated words and words where there is not confidence that we know the first sign, then only 15% of the words in texts from Haghia Triada begin with a vowel, compared to 36% of words from the other Linear A sites.[18] Table 16 presents a more detailed statement of this result.

Table 16. Frequency of words beginning with a vowel

	A	E	I	O	U
Haghia Triada	7.2	–	3.8	0.6	3.1
non-Haghia Triada	23.7	2.1	5.9	1.7	2.1

[16] See Melena 1975 (p. 126) for the suggestion that su-ri-mo is probably Sylamos, just south of Knossos.

[17] An explanation of prothetic vowels is given by Beekes (2003). "The definition is 'initial vowel that is present or absent in (nearly) identical forms'; for we cannot say whether the vowel disappeared or was added under certain circumstances; still another possibility is that it represents a kind of laryngeal sound, that was sometimes heard as a vowel and sometimes not. The vowel is in most cases an α. The numbers by Furnée are as follows: α ~90, ο 10, ε 5, ι 3, υ 0, η 6, αι 2".

[18] This is based on a total of 319 words for Haghia Triada and 236 words for the other Linear A sites.

Thus, it is clear that the major difference between the usage of initial vowels between Haghia Triada and the other sites is for the initial *-a* and, as already noted in the description of prothetic vowels (in footnote 17), this vowel predominates. However, Table 17 shows that the frequency distribution of consonants is similar between Haghia Triada and the other sites. [The frequency distribution is also included for Pylos Greek names as an example of the level of difference when there is a difference of language.[19] Note in particular the large differences in the percentages for d, s and w between the Pylos Greek names and the two Linear A frequencies].

Table 17. Percentage frequencies of consonants

	Pylos Gk names	Haghia Triada	non-Haghia Triada
d	3.60	12.64	9.53
j	8.88	5.26	6.85
k	13.87	13.31	10.07
m	7.88	7.83	10.60
n	6.39	9.17	9.26
p	8.09	5.15	6.44
q	3.29	4.14	2.55
r	18.37	16.11	12.21
s	3.79	10.51	12.48
t	14.87	13.76	15.70
w	10.68	1.68	2.15
z	0.30	0.45	2.15

Thus, it is suggested that one key difference between the Haghia Triada inscriptions and the other Linear A inscriptions is that the LM IB Haghia Triada scribes have a greater tendency not to use the initial vowel *a-*. We have not sufficient material to state with confidence whether this resulted from a difference in dialect or a difference in the way that the language was annotated using Linear A signs. However, one might tentatively suggest that it was due to the way in which the language was written down because it allowed words to be written in a more concise form and it could be seen as a late development in the writing of Linear A that appears to emanate from LM IB Haghia Triada.

We will now return to the discussion on the development of a statistical method for assessing Greekness in groups of Linear B names. We have demonstrated that

[19] A numerical indication of the differences between these frequency distributions is given by the sum of the squares of differences. These are 245 for Greeks and Haghia Triada, 263 for Greeks and non-Haghia Triada, but only 61 between Haghia Triada and non-Haghia Triada.

there is a variability of the usage of free vowels within the Linear A inscriptions. By using a frequency array based on all Linear A inscriptions we are implicitly adopting a particular balance between the Haghia Triada usage of free vowels and that found at other sites that is governed by the inscriptions that have been found. However, in Table 15, we have shown that the non-Greek men's names at Knossos are biased towards the Haghia Triada usage of free vowels. One possible way around this, within the statistical method, would be to adopt a version of the method which was based only on the frequency of consonants, however, this was found to be inefficient precisely because it neglected the information provided by the frequency of usage of free vowels. The alternative is to use a frequency array which specifically represented the non-Greek names of Knossos and the obvious choice is to adopt statistical Method B, based on Ilievski's non-Greek shepherds (in preference to statistical Method A). We have already given numerous results using Method B, however, for completeness, Appendix B summarises the results of the analyses using Method B.

6. Discussion

There are two aspects to this final discussion. The first is a set of general points arising from the present discussion. The second is consideration of the interpretation of two Linear A words which emerged during the course of this work.

It is worthwhile starting by making a few general points:

- It is possible to find many examples where different experts will describe the same name as Greek or non-Greek. In other words, it is often a matter of judgement whether a name is Greek or not Greek and experts will differ in their views. Ilievski (1992, p. 331) wrote, "Some Mycenaeologists used to classify all the uninterpreted Mycenaean names as non-Greek. But it would be too bold to claim a name is non-Greek simply because we cannot explain it as Greek". Ilievski (1992, p. 323) also notes "not all uninterpreted names are non-Greek, just as not all interpreted names are only Greek".
- One could have a discussion on whether a name which is a Mycenaean name formed from a pre-Hellenic base results in a Greek or a non-Greek name. It is conventional in Mycenaean studies to regard such names as Greek because the morphology is Greek (even though the base is non-Greek). However, in terms of the present statistical method, which considers the frequency of usage of consonants and free vowels, then the weighting between Greek and non-Greek is divided but the greater emphasis is on the non-Greek base, because it gives rise to the larger number of syllables.
- This point becomes even more focussed when the personal name is simply a toponym, such as *qa-ra-i-so* or *u-ta-no*. Ilievski (1992) has listed these names

as Greek although they are toponyms which are very probably non-Greek in origin (see, for example, Brown 1985, pp. 196-7 on *qa-ra-i-so*). In these cases, the only contribution of Greek morphology is likely to have been the change of the final vowel to an *-o*. In terms of the statistical method, names based on non-Greek toponyms would probably have a non-Greek weighting.

It was inevitable that, during a general statistical study of Greekness and non-Greekness of personal names, there should b detailed findings. This last short section considers the possible interpretation of a Linear word.

MA-KA-I-TA (PK 1 and ZA 5): It is suggested this could be read as μαχαίτας (or plural μαχαῖται) 'fighter, warrior(s)', cf. *ma-ka-ta* on PY Jn 725.[20]

The Mycenaean language was used for a long period on the Mainland before it was adopted on Crete and, therefore, it is not surprising that we should find a few Greek names and Greek words amongst the Linear A texts (c.f. *I-JA-TE* on a pithos sherd, PH Zb 4 and *i-ja-te/* Hom. ιητήϱ 'physician' on PY Eq 146).

We should also note an earlier paper by Jan Driessen (1984) suggesting that there were mercenaries on Crete that were listed on the Linear B tablets. This puts the present suggestion that there were μαχαῖται, Greek warriors, at Zakros and Palaikastro during LM IB into some context.[21]

Bibliography

Aravantinos, V. L.; Godart, L. & Sacconi, A., 2001, *Thèbes. Fouilles de la Cadmée*, vol. I: *Les Tablettes en Linéaire B de la Odos Pelopidou,* Rome.

Astour, M. C., 1965, *Hellenosemitica,* Leiden.

Baumbach, L., 1983, "An examination of the personal names in the Knossos tablets as evidence for the social structure of Crete in the Late Minoan II period", in *Minoan Society: proceedings of the Cambridge Colloquium,* edited by O. Krzyszkowska & L. Nixon, pp. 3-10.

Baumbach, L., 1986, "The Personal Names on the Knossos Ap tablets", in *O-O-PE-RO-SI: Festschrift für Ernst Risch*, edited by A. Etter, pp. 273-278.

Baumbach, L. 1992, "The People of Knossos: Further thoughts on some of the personal names", in *Mykenaïka*, edited by J.-P. Olivier, pp. 57-63, Paris.

Beekes, R. S. P., 2003, "Pre-Greek. A language reconstructed" (second version, November 2003).

Billigmeier, J. C., 1969, "An inquiry into the non-Geek names on the Linear B tablets from Knossos and their relationship to languages of Asia Minor", *Minos* 10, pp. 177-183.

[20] The word, *MA-KA-I-SE* also appears at Zakro (ZA 8), and we might speculate that this could be a "Minoanised" form of *MA-KA-I-TA*.

[21] The LM IB dating for these tablets is based on Raison & Pope 1994.

Brown, R. A., 1985, *Evidence for Pre-Greek speech on Crete from Greek alphabetic sources,* Amsterdam.

Documents = Ventris, M. & Chadwick, J., 1973, *Documents in Mycenaean Greek,* 2nd edition, Cambridge University Press.

Driessen, J., 1984, "Some Military Aspects of the Aegean in the Late Fifteenth and Early Fourteenth Centuries B.C.: Part I. Mercenaries at Knossos?", *BSA* 79, pp. 49-56.

Duhoux, Y., 1982, *L'Étéocrétois: Les texts – la langue,* Amsterdam.

Firth, R. J., 1993, "A Statistical Analysis of the Greekness of Men's Names on the Knossos Linear B Tablets", *Minos* 27-28, pp. 83-100.

GORILA = Godart, L. & Olivier, J.-P., 1976-1985, *Recueil des Inscriptions en Linéaire A,* Libraire Orientaliste Paul Geuthner, Paris.

Ilievski, P. H., 1992, "Observations on the Personal Names from the Knossos D Tablets", Mykenaïka, *BCH Suppl.* 25, Paris, pp. 321-348.

Killen, J. T., 2004, "Names in *-e* and *-e-u* in Mycenaean Greek", in *Indo-European Perspectives: studies in honour of Anna Morpurgo Davies*, Oxford, pp. 217-235.

Landau, O., 1958, *Mykenisch-Griechische Personennamen*, Göteborg.

Landenius-Enegren, H., 1995, "A Prosopographical Study of Scribal Hand 103, Methods, Aims and Problems", in *Politeia: Society and State in the Aegean Bronze Age*, edited by R. Laffineur and W.-D. Niemeier, pp. 115-130.

Melena, J. L., 1975, *Studies on some Mycenaean inscriptions from Knossos dealing with Textiles, Supl. Minos* 5.

Morpurgo Davies, A., 1999, "The Morphology of Personal Names in Mycenaean and Greek: Some Observations", in *Floreant Studia Mycenaea*, pp. 389-405.

Packard, D. W., 1974, *Minoan Linear A*, University of California Press.

Raison, J. & Pope, M., 1994, *Corpus Transnuméré du Linéaire A*, 2nd edition, Peeters, Louvain-la Neuve.

Ruijgh, C. J., 1967, *Études sur la Grammaire et le Vocabulaire du Grec Mycénien,* Amsterdam.

Varias, C., 2002, "The Personal Names from the Knossos B-Tablets and from the Mycenae Tablets", in *A-NA-QO-TA: Studies presented to J. T. Killen, Minos* 33-34, pp. 349-370.

Wilson, A. L., 1977, "The place-names in the Linear B tablets from Knossos: Some preliminary considerations", *Minos* 16, pp. 67-125.

Appendix A: Commentary on the new assessments used in section 2

Mycenae and B-series (Varias 2002): Varias divides the names of both the personal names from Mycenae and the KN B-series into three lists: Greek, Doubtful Greek and Non-Greek (Varias 2002, Tables 2 and 1, respectively). The percentages attributed to an independent assessment of Greekness given in Table 2 of this paper are based on a combination of the Greek and Doubtful Greek names.

D-series (Ilievski 1992): Ilievski divides his Greek names into three categories: (a) high degree of certainty; (b) probable; (c) possible. The percentage quoted in Table 2 of this paper is a summation of all of these three categories. All the figures quoted exclude the collectors.

Ap, As and C-series (Baumbach 1986, 1983, 1992, respectively): According to the statistical analysis, Baumbach systematically underestimates the amount of Greekness in each of these three groups of names. Ilievski (1992, FN 8) comments on Baumbach's analysis of the KN D-series personal names, "Applying strict criteria for identification of the names in the KN D-series she found 90 names which can be interpreted as Greek, and a proportion of one Greek to four non-Greek names. But, as we shall see below, not all the uninterpreted names are non-Greek, just as not all the interpreted names are only Greek". In other words, Baumbach suggested that only 20% of the D-series names were Greek whereas Ilievski suggested a considerably higher percentage.

KN names ending in *-i* (Morpurgo Davies 1999): That work concluded that, "The *i*-stem names of the Mycenaean texts morphologically share few if any of the categories of the alphabetic period. If so, it is difficult, if not impossible, to understand why in Knossos, as contrasted with Pylos, there are so many *i*-stem names; the analysis of these forms (already proposed by Killen) confirms their non Greek origin."

Astour (1965, pp. 336-344) provides a list of Linear B names which he claims may reasonably be taken to have a Semitic or Hurrian origin. This suggestion has not been accepted by the Linear B community. The present analysis supports that view. The statistical analysis suggests that 57% of the men's names at Knossos are Greek (Table 3). By comparison, the level of Greekness found for Astour's list is 48%. In other words, the level of non-Greekness is little better than would be achieved by a random selection from the men's names of Knossos.

Billigmeier (1969) provides a list of non-Greek names by suggesting a series of non-Greek prefixes and suffixes which are used to form compound names. Whether or not this approach is correct in all its detail, the present analysis suggests that Billigmeier has been successful in isolating a group of non-Greek names from Knossos.

Killen (2004) gives lists of names that have a non-Greek appearance. However, it is emphasised that the non-Greek names given in that paper are not ideally suited to this type of analysis since they were chosen because they include particular signs, i.e. *me*, *mo*, *ne*, *pi-ja-*, *pu*, *re*, *se*, *te* and so there is an implicit bias in the frequency distribution of Linear B signs.

Landau gives a list of non-Greek Linear B names on pp. 268-273 (Landau 1958).

Appendix B: The results of the assessment of Greekness using the recommended method: Method B

Table B1. Percentage frequencies of consonants and free vowels

Frequency of usage (%)	Pylos Greek names (G_i)	Ilievski
vowel	17.67	7.32
d	2.96	7.49
j	7.31	4.99
k	11.42	8.32
m	6.49	5.99
n	5.26	8.82
p	6.66	2.66
q	2.71	4.33
r	15.12	15.14
s	3.12	11.15
t	12.24	14.64
w	8.79	5.99
z	0.25	3.16

Table B2

	Independent assessment of % Greekness	Predicted % Greekness	No. of signs	Residual Sum
KN Sc + Vc(1)-series	84	90	286	24
Mycenae	72	62	364	26
KN B-series	71	81	298	10
KN D-series	44	37	1095	5
KN As-series	35	45	436	20
KN C-series	31	42	203	88
KN Ap-series	18	37	169	164
KN 2-sign names	7	5	200	97
KN names ending in *-i*	0	6	118	127
Lists of non-Greek names				
Astour (Semitic names)	0	70	257	150
Billigmeier	0	8	228	113
Killen	0	17	163	190
Landau	0	11	200	55
Cretan place names on KN tablets		0.5	138	24

Table B3

	Predicted % Greekness	No. of signs	Residual Sum
Pylos (males only)	84	2624	4
Knossos (males only)	56	3094	4
Mycenae	62	364	26
Thebes (males only)	86	366	14

The analyses for:
· Ilievski's D-series names of shepherds are given in Table 7b.
· Baumabach's Ap, As and C-series names are given in Table 8.

Table B4

Varias' groups	% Greekness	Residual Sum
Greeks + Doubtful Greeks (exc. ethnics)	82	20
Non-Greeks plus ethnics	42	72

