Arqueología y *Téchne*

Métodos formales, nuevos enfoques
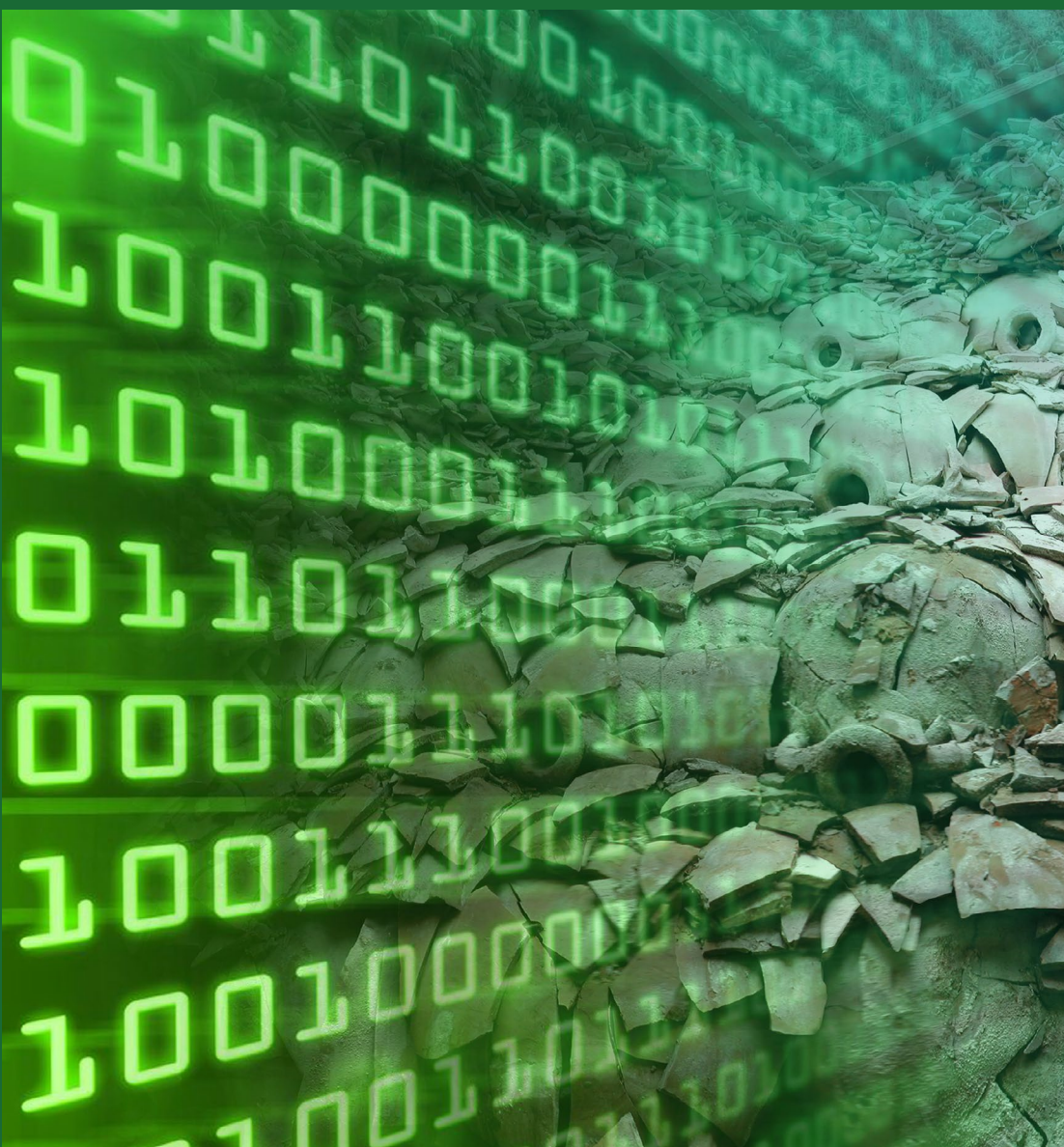
Archaeology and *Techne*

Formal methods, new approaches

Editado por
José Remesal Rodríguez
Jordi Pérez González

About Access Archaeology

Access Archaeology offers a different publishing model for specialist academic material that might traditionally prove commercially unviable, perhaps due to its sheer extent or volume of colour content, or simply due to its relatively niche field of interest. This could apply, for example, to a PhD dissertation or a catalogue of archaeological data.

All *Access Archaeology* publications are available as a free-to-download pdf eBook and in print format. The free pdf download model supports dissemination in areas of the world where budgets are more severely limited, and also allows individual academics from all over the world the opportunity to access the material privately, rather than relying solely on their university or public library. Print copies, nevertheless, remain available to individuals and institutions who need or prefer them.

The material is refereed and/or peer reviewed. Copy-editing takes place prior to submission of the work for publication and is the responsibility of the author. Academics who are able to supply print-ready material are not charged any fee to publish (including making the material available as a free-to-download pdf). In some instances the material is type-set in-house and in these cases a small charge is passed on for layout work.

Our principal effort goes into promoting the material, both the free-to-download pdf and print edition, where *Access Archaeology* books get the same level of attention as all of our publications which are marketed through e-alerts, print catalogues, displays at academic conferences, and are supported by professional distribution worldwide.

The free pdf download allows for greater dissemination of academic work than traditional print models could ever hope to support. It is common for a free-to-download pdf to be downloaded hundreds or sometimes thousands of times when it first appears on our website. Print sales of such specialist material would take years to match this figure, if indeed they ever would.

This model may well evolve over time, but its ambition will always remain to publish archaeological material that would prove commercially unviable in traditional publishing models, without passing the expense on to the academic (author or reader).

Arqueología y *Téchne*

Métodos formales, nuevos enfoques

Archaeology and *Techne*

Formal methods, new approaches

Editado por

José Remesal Rodríguez

Jordi Pérez González

Access Archaeology

This book is available direct from Archaeopress or from our website www.archaeopress.com

Contents

Similarity Analysis in Epigraphy. Syntactic Clustering of *Tituli Picti* on the PO8 Amphoras

Daniel J. Martín-Arroyo Sánchez

Iza Romanowska

1. Introduction[1]

The Pompeii 8 is a type of Roman amphora from Crete. It is known as the Schoene-Mau VIII type depicted in the *Vasorum formae* tables, published within volume IV of the *Corpus Inscriptionum Latinarum* (*CIL*) (Fig. 1). This volume is dedicated to the inscriptions from *Pompeii*, *Herculaneum*, and *Stabiae*, in the surroundings of *Mons Vesuvius*. Here we refer to this amphora type as PO08 to account for the diversity of amphora-shapes that could have been classified under this type by the authors of *CIL* IV.[2] Currently, the original design fits with the amphora-type Crétoise 2 (AC2), which was a wine trade container produced from the end of the Augustan period to the mid- or late-2nd century A.D.[3] A total of 535 PO08 carrying *tituli picti* were recorded in the *CIL* IV. These inscriptions were made in Greek and Latin, using different inks. At first sight, it is difficult to evaluate whether they were applied in a standardised way. Within the EPNet Project,[4] the Codex methodology[5] has been proposed to investigate this epigraphic set beyond its general appearance. As a research tool, the Codex methodology has been developed to simplify data management and analysis. It divides every inscription into meaningful basic units taking into account the content of the text and the confidence level of their decipherability. These units and their spatial and formal aspects are transcribed in a series of symbols. These symbols can be interpreted in a similar way to any string of characters, for example, numbers in an equation system.

Here we present the results of an epigraphic investigation and their historically recontextualised interpretation. Primarily, the study has an epigraphic objective: to identify the basic units in every inscription, to determine their meaning and to detect structural patterns. For example, the inscription ME (*CIL*, IV, 6474; Fig. 1) can be identified as a basic unit because of its isolated position, near the handle, regarding the position of the main *tituli* ΛΥΤΤΙΟϹ / Μ · Π · ΤΕΥΠωΝΟϹ, probably on the neck or shoulder of the amphora. It can be identified as a Greek or Latin word abbreviation or as the Greek number 45. By comparing its epigraphic features and context with other similar inscriptions, we can support its identification and possible meaning, for example, the weight or price of the vessel's contents. At the same time, similarities in the inscription syntax may indicate close trade contexts, in which information is broadcast using specific conventions. Such epigraphic structures would indicate the existence of a limited group of traders acting within similar mechanisms and places of exchange. Thus, the second goal of this research is to detect any patterns indicative of larger trade structures from the *tituli picti* epigraphic data. Even the same product, such as the Cretan wine, could arrive in *Pompeii* through different ports or trade routes involving different mechanisms of exchange. This research aims to establish whether these can be reconstructed based on the inscriptions on amphoras.

---

[1] Many thanks to Nick Bennet-Britton for the English revision.

[2] Martín-Arroyo *et al.* 2017: 181-182; Martín-Arroyo 2018: 316-317.

[3] Marangou-Lerat 1995: 81.

[4] Remesal *et al.* 2014.

[5] The original idea of Codex, based on the *tituli* delta from the Dressel 20 amphora type, in Remesal and Aguilera 1999: 102-103; Aguilera 2000: 1232. A previous PO08 study case in Martín-Arroyo 2019.
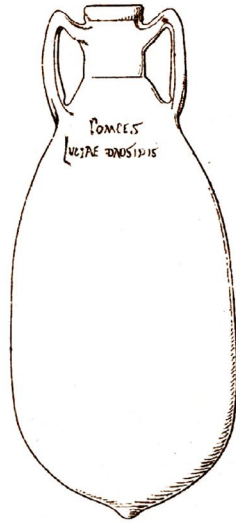
Figure 1. PO08 amphora type and related inscriptions: Schoene-Mau VIII type (PO08 prototype); *tituli picti CIL*, IV, 6474 (with partial tracing by Marangou-Lerat 1995: n. P7, fig. 88) and 6483.

In this paper, we will present the research methodology used in the case study: the *tituli picti* from the PO08 amphora found at *Pompeii*, including the introduction to the Edit Distance algorithm. The results will be depicted as a series of graphical representations and interpreted within their historical context. We will evaluate the usefulness of formal data modelling techniques with regards to the reduction of the uncertainty and the detection of epigraphic structures. Finally, we will discuss the challenges and potential of this type of computational tools to analyse data sources common in humanities research, especially in the light of their qualitative complexity and size.

2. Source management and resulting data set

The analysed dataset consists of 533 PO08 amphoras carrying *tituli picti* from the site of *Pompeii* and therefore dated before the year 79 AD. At first glance, almost all *tituli picti* are different: they record information regarding particular individuals and figures. Here we assume that similar commercial histories have a higher than random probability of sharing similarities in the general structure of the inscriptions. Although this is far from being an absolute distinction, we argue that the similarity of administrative documentation as recorded in *tituli* can be treated, with a degree of caution, as a proxy for commercial history. Thus, by grouping the sets of containers with the same epigraphic structure of inscriptions we may be able to identify specific administrative processes, a similar transport route or even specific cargo loads. However, quantifying the similarity of the *tituli* contents is highly problematic. Consider two *tituli*: one bearing the name of *Marcus Iulius* and a quantity of 30, and one with the name *Marcus Valerius,* and the quantity is in Greek (Λ = 30). It is clear that although the literal content of inscriptions carries a wealth of information, it also makes it difficult to compare the inscriptions, and it masks the underlying commercial process. To overcome this issue of data heterogeneity we apply the Codex methodology to the collected *tituli picti*. The Codex methodology codifies each epigraphic unit following simple semantic rules. Thus, the actual values of the epigraphic units, such as the names *Marcus Iulius* and *Marcus Valerius*, the figure *30* (Λ), or the content of an amphora, such as *wine*, are replaced with their semantic values ("Latin name", "Latin/Greek number", and "Latin product-noun" respectively). As a result, we only record the structure of the inscription: the presence/absence of given information and its form. Below we provide a detailed description of the Codex methodology.

2.1. The Codex methodology

The Codex methodology has been developed by the EPNet Project to investigate epigraphic data beyond its lexical layer, which is noisy and notoriously difficult to interpret. To the best of the authors' knowledge, there is no other epigraphic standard to classify accurately the syntactic structure of ancient inscriptions that would be applicable for data management and analysis of the PO08 *tituli*. Previous classification systems similar to the Codex methodology have been applied to the more regular sequence of data recorded in *tituli* from Dressel 20,[6] and to amphoras carrying fish-based products.[7] Here, these systems were adapted to the more heterogeneous epigraphic of PO08 by expanding the set of *concepta* used to classify the inscriptions and by including coding of the inscription language.

The Codex methodology consists of the following steps. First, every inscription is divided into meaningful basic units on the basis of their syntactic and lexical meaning. Second, these units, their location and formal aspects are transcribed in a set of characters. Finally, the confidence level given to the decryption is also recorded.

The identification of the basic units of every inscription is based on:

- The location in respect to other inscriptions on the amphora;
- Lexical meaning;
- Distribution in different lines of text;

---

[6] Aguilera 2012.
[7] Martínez 1998; Lagóstena 2002-2003; Lagóstena 2004.

- Different inks;
- Different languages (Greek or Latin).

For example, the *titulus* ΜΕ appeared beside the *tituli* ΛΥΤΤΙΟC / Μ · Π · ΤΕΥΠωΝΟC in the inscription *CIL*, IV, 6474. A similar case is NT in 6483 (see both cases in Fig. 1). ΜΕ is identified as a basic meaningful unit because of its isolated position on the amphora's body near the handle. The meaning of "ΜΕ" might have been a Greek or Latin word abbreviation or the Greek number 45. This example shows the epigraphic value of using formal methods to compare inscriptions. If a similar set of syntactic units repeatedly appears in inscriptions, it may give some weight to one interpretation over the other. For example, if there is a good sample of amphoras where the weight of the contents is noted in a similar context to the amphora 6483's *titulus* ME, then it supports its interpretation as a Greek number 45 rather than a Latin abbreviation. Such epigraphic structures could have been used by a limited group of traders acting within similar mechanisms and areas of commerce or by   harbour administration.

Every basic unit identified in the PO08 epigraphic record has been classified as one of the *concepta* defined by the Codex methodology (Fig. 2). Every *conceptum* expresses an isolable semantic unit, such as a determined word, number or idiom. The *conceptum* "Productus" refers to the commodity contained in the amphora or some related adjective. "Nomen" is the name of one or more individuals. If it is a plural name or a group of names the Latin conjunction *et* is included in the Nomen class. Abbreviations appear as the first part of a word or the initials of a name. The latter was always assumed if full stops followed the letters. In cases where the *titulus* could be read as either an abbreviation or a number, it was provisionally classified as the latter. The conceptual "Numerus" refers to numbers and "Consules" to consular datings. "Signum" is a mark, i.e., a nor-alphabetic symbol. The *conceptum* "Reliquiae" was applied to vestiges of inscriptions that could not be deciphered because of their poor preservation. "Incertum" was used if it was not possible to establish the type of *concepta*. All identified units which had more than one potential meaning were included in the class of *concepta incerta* and additionally marked with a question mark.

Each inscription was transcribed to a script. To specify the alphabet (*Litterae*) of the inscription, we used uppercase (Latin) and lowercase (Greek) letters. The resulting set of symbols is named *Regestum*. We added spatial and ink characteristics by a complementary set of symbols (Fig. 3). In that way, we obtain simple formulae or *codices*, enabling us to compare the syntactic structure of inscriptions. For example, the inscription *CIL*, IV, 6483 (Fig. 1) can be read as:

ΛΥΤ(ΤΙΟC) / Μ. · Π(ΟΠΥΛ)(---) · /= ΤΕΥΠΩΝ[ΟC] // Ν^Τ Λ ΙΙC S(---)

The *codex* of the inscription is:

p/n/=//x/\X

p represents ΛΥΤΤΙΟC, which was a variety of wine from the Cretan city of *Lyttus*.[8]
n stands for the name: Μ. ΠΟΠΥΛ. ΤΕΥΠΩΝΟC.[9]
/= indicates that the text continues in a subsequent line.

---

[8] Martín-Arroyo 2020.
[9] For example, *CIL*, IV, 10455 and 10460.

x is the NT - the Greek number 350.[10]

\ shows that the colour of the ink has changed.

X is capitalized because IICS must be Latin number 98 since the type "S" is exclusively Latin. The last *conceptum* was written with white ink, different to the rest of the inscription, making it quite clear that it is a Latin expression added to Greek *tituli*. In this case the interpretation is relatively straight forward, but in most of the researched cases, the reading of a *titulus* is more controversial and alternative readings can be proposed. To account for this uncertainty, we recorded alternative readings of an inscription separately under different *codices* and indicated their uncertainty with a '?'. Thus, a *regestum* "p?" could be "p" or "i" in different *codices*.

CONCEPTUM	REGESTUM
Productus	P/p
Nomen	N/n
Numerus	X/x
Consules	C/c
Signum	S
Reliquiae	...
Incertum	I/i
Productus incertus	P?/p?
Nomen incertum	N?/n?
Numerus incertus	X?/x?
Consules incerti	C?/c?

Figure 2. Codex keys of *concepta*.

---

[10] Any name with NT as initial letters in *Lexicon of Greek Personal Names* (*LGPN*) (http://clas-lgpn2.classics.ox.ac.uk/name).

&	In the same line of text
/	In a following line of text
/=	The *conceptum* continues in the following line of text; e.g. N/=
//	*Ab altera parte*; elsewhere
!	Doubtful position
\	Written with a different ink

Figure 3. Codification of spatial and ink characteristics.

2.2. Summary statistics

The summary statistics of the resulting data set are described in tables (Fig. 4-7). The length of inscriptions ranges from 1 to 9 *concepta* (see Fig. 4). On average, there are 1,89 *concepta* for each amphora. Most of *concepta* come in pairs. They occur less frequently singly or in groups of three or four units. Cases of five or more *concepta* are exceptional. In general, 80% of amphoras carry inscriptions of one or two *concepta*.

Nº *concepta*	Total *concepta*	Nº amphoras	% *concepta*	% amphoras
1	213	213	21,1%	40,0%
2	416	208	41,1%	39,0%
3	249	83	24,6%	15,6%
4	84	21	8,3%	3,9%
5	15	3	1,5%	0,6%
6	18	3	1,8%	0,6%
7	7	1	0,7%	0,2%
8	0	0	0,0%	0,0%
9	9	1	0,9%	0,2%
TOTAL	1011	533	100,0%	100,0%

Figure 4. Breakdown of the inscriptions based on the number of *concepta*.

60% of inscriptions are written in Greek (GRA) and the remaining ones in the Latin (LAT) alphabet (Fig. 5). There does not seem to be a difference between the languages in respect to the length of the inscription, i.e., both Greek and Latin *tituli* have a similar proportion of one, two and three *concepta*. It should be noted that the language could not be identified for *concepta* recorded as *signa* or *reliquiae*.

About half of *concepta* have uncertain decryption (*concepta incerta*) (Fig. 6). The most numerous category is "Nomen" in both Greek (23% of deciphered *concepta*) and Latin (15% of deciphered *concepta*). This is followed by inscriptions containing a number ('Numerus'; 7% of deciphered *concepta*) and amphora content ('Productus'; 4.5% of deciphered *concepta*). Regarding the proportionality of the *concepta certa* mentioned above, there is a significant number of items in the 'Numerus incertus' category (11%).

Nº *concepta*	GRA	LAT	Total	% GRA	% LAT	% Total
1	120	92	213	11,9%	9,1%	21,1%
2	246	167	416	24,3%	16,5%	41,1%
3	155	91	249	15,3%	9,0%	24,6%
4	42	41	84	4,2%	4,1%	8,3%
5	12	1	15	1,2%	0,1%	1,5%
6	12	5	18	1,2%	0,5%	1,8%
7	3	4	7	0,3%	0,4%	0,7%
8	0	0	0	0,0%	0,0%	0,0%
9	0	9	9	0,0%	0,9%	0,9%
TOTAL	590	410	1011	58,4%	40,6%	100,0%

Figure 5. Use of Greek (GRA) and Latin (LAT).

CONCEPTUM	GRA	LAT	Total	% GRA	% LAT	% Total
Productus	34	12	46	3,4%	1,2%	4,5%
Nomen	234	156	390	23,1%	15,4%	38,6%
Numerus	19	53	72	1,9%	5,2%	7,1%
Consules	2	3	5	0,2%	0,3%	0,5%
Signum			5			0,5%
Reliquiae			6			0,6%
Incertum	116	86	202	11,5%	8,5%	20,0%
Productus incertus	21	5	26	2,1%	0,5%	2,6%
Nomen incertum	66	76	142	6,5%	7,5%	14,0%
Numerus incertus	98	19	117	9,7%	1,9%	11,6%
Consules incerti	0	0	0	0,0%	0,0%	0,0%
Total of *concepta certa*	289	224	524	28,6%	22,2%	51,8%
Total of *concepta incerta*	301	186	487	29,8%	18,4%	48,2%
Total of records	590	410	1011	58,4%	40,6%	100,0%

Figure 6. Quantification of *concepta* based on the language (Graecae and Latinae) and uncertainty.

Finally, Figure 7 shows the distribution of the *regesta* (i.e., *concepta* including the uncertainty factor). What is most striking is the high number of *numeri incerti Graeci* (x?). These, in many cases, could be read as *nomina Graeca* (n), that is, Greek name abbreviations. In all contested cases they have been classified as *numerus incertus*. Nevertheless, recent approaches on the PO08 amphora epigraphy[11] suggest that at least some *tituli* written in well-delimited contexts should be identified as Greek names of people who were dealing with the amphoras. This line of research should be   pursued further.

Figure 7 also shows that in cases of amphoras with only one *conceptum*, 71% of them could be read with a high degree of certainty (*concepta certa*), which is significantly higher than the average of 50% certain readings. This is due to the completeness of written Greek names and the presence of recognisable initials of *tria nomina* in amphoras with just one *conceptum*. In that sense, the presence of Greek uncertain names and numbers is more reduced than in the following groups. This difference could indicate a specific pattern or context in which names are written in a more complete manner, in order to satisfy the functionality of these *tituli*. In that sense, the structural patterns of groups of *tituli* could be indicative of the commercial process, in which the identification of an individual seems to play an important role.

Nº concepta	p	P	n	N	x	X	c	C	S	...
1	1	2	90	52	1	2	2	1	1	
2	18	4	82	56	6	26		1		3
3	11	3	41	31	8	10			2	1
4	2	1	12	13	3	9				1
5	1		5						1	1
6	1		4	2					1	
7				1	1	3				
8										
9		2		1		3		1		
TOTAL	34	12	234	156	19	53	2	3	5	6

Nº concepta	p?	P?	n?	N?	x?	X?	c?	C?	i	I	Certa	Incerta
1		2	3	16	7	2			16	15	152	61
2	6	2	25	34	52	8			57	36	196	220
3	11	1	27	18	22	6			35	22	107	142
4	3		6	6	13	3			3	9	41	43
5			2	1	3				1		8	7
6	1		3	1	1				2	2	8	10
7									2		5	2
8											0	0
9										2	7	2
TOTAL	21	5	66	76	98	19	0	0	116	86	524	487

Figure 7. Quantification of *regesta* regarding to the Number of *concepta* in each amphora.

---

[11] Martín-Arroyo 2019.

3. Edit distance

Edit distance (also known as Levenshtein distance) is a well known algorithm used in computational linguistics and other disciplines for quantifying the difference between two sets of characters (strings). The algorithm counts the minimum number of operations a string has to undergo in order to transform it into another string. Permitted operations include: insertion, deletion and substitution and each is scored equally. For example, the minimum distance between 'past' and 'present' is 4 (substitution of 'a' to 'r' and insertion of 'e', 'e' and 'n') (Fig. 8).

The edit distance algorithm is much more robust in the case of incomplete strings than the standard pairwise comparison methods since it strives for an optimal alignment. For example, the edit distance between strings 'banana' and 'anana' is only 1 (insertion of 'b') rather than 6 achieved using pairwise comparison ('b' compared with 'a', 'a' compared with 'n'...). Thus, it is particularly well suited for analysis of incomplete strings, such as ancient inscriptions.

Performing the edit distance analysis on a set of strings results in a matrix of distances from each string to every other string. The resulting matrix can be represented as a heatmap and used in a cluster analysis to compute dendrograms of similarity (Fig. 9).

The edit distance analysis was run on the *regesta* taking into account the following factors: *Conceptum*, *Littera* and uncertainty. Although interesting, the ink colour factor had to be removed from the analysis because not all the entries in *CIL* have recorded this characteristic. Similarly, the spatial relationship of the *regesta* was omitted in the light of a recent study[12] which showed that the inscription position was not a relevant factor even in similar epigraphic structures. The high diversity and degree of abbreviation of the researched *tituli* are considered as evidence of their informal and limited contexts, where the information was easily understood by a small number of users.
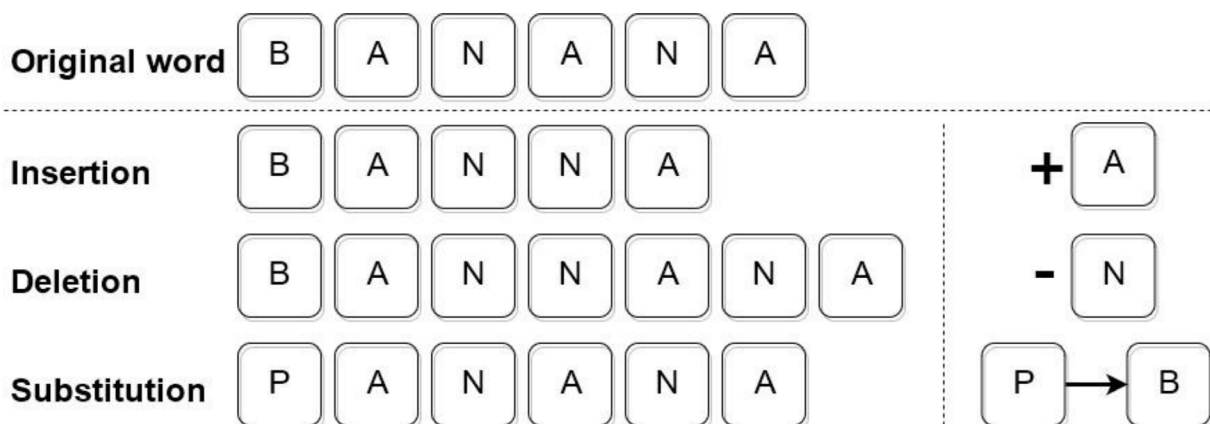

Figure 8. Edit distance operations needed to achieve the desired word: insertion, deletion and substitution. Each operation has the same 'cost'.

---

[12] See groups 5A and 5B in Martín-Arroyo 2019: 447, tab. 2, where common or parallel elements were located in different ways in the group and the subgroups respectively.
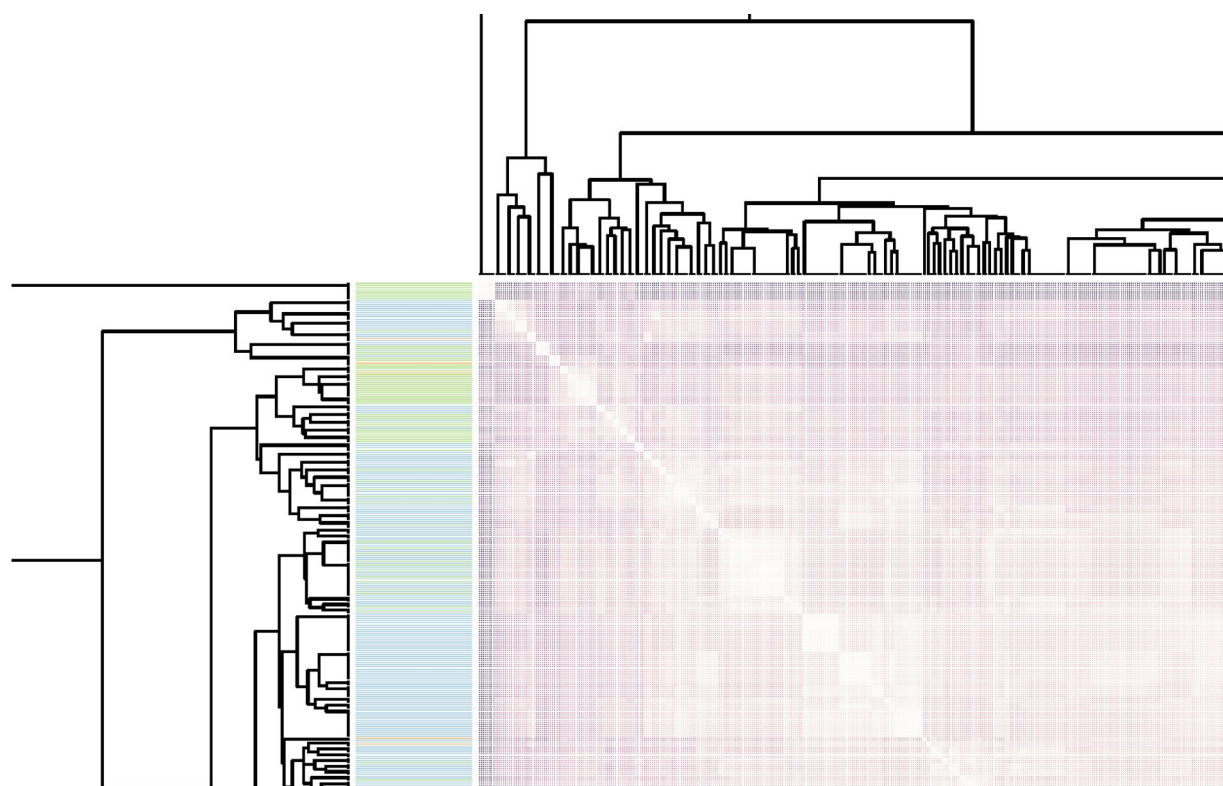
Figure 9. Detail of a heatmap with an associated dendrogram. The left-hand side column shows an independent variable – the language (*Littera*), with green stripes indicating Latin and blue ones Greek (orange stripes stand for 'uncertain').

The Codex methodology simplifies data management and analysis. The *codices* synthesise a remarkable amount of information within just a few symbols so that even elaborate inscriptions can be easily compared and contrasted. Second, we can carry out a more sophisticated analysis through different strategies of filtering and grouping of data. In this analysis, three versions of Codex were developed based on the level of uncertainty. In a progressive classification, Codex I maximises uncertainty, Codex II preserves it, and Codex III eliminates it. For example, a *regestum* "n?" from Codex II will be transformed in "i" for Codex I or in "n" for Codex III.

In order to extract maximum information contained in every *regestum*, we additionally treated the data. To give an example, without this treatment the *codices* "n" and "N" would be considered equally distant as "n" and "P" since just one substitution is required to move from one to the other. However, conceptually "n" and "N" are more similar than "n" and "P", because both of them are a name, *nomina*. Thus piece of the information (language, *conceptum*, *regestum*) was spelt out separately by adding additional characters to the Codex, marking the type of *littera* and the type of *conceptum* used in the inscription. For example, the *codex* "n" is transcribed as "Nng", where "N…" is the *conceptum* Nomen, "…n…" is the *regestum* Nomen Graecum, and "…g" is the *littera* Graeca. After this transformation, the comparison between the syntactic structure of the inscription was based on all available information. For example, "Nng" and "NNl" are more similar than "Nng" and "PPl", with distances of 2 and 3 respectively. Then *codices* with similar *concepta* and *litterae* are thus clustered closer together, but even more when the combination of them offers common *regesta*.

In the *regesta* from Codex V every *conceptum incertum* includes the character "?", so the presence and quantity of *concepta incerta* in each *codex* results in an added factor of similarity.

3.1. Cluster diagrams

The clusters have been calculated using the average algorithm. The results have been checked using the Cophenetic Correlation Coefficient test, which demonstrated that the original structure of the data had been well preserved. The resulting clusters are visualised in a cluster map and an associated dendrogram. Additionally, to facilitate visual evaluation the dendrogram was visualised with different levels of thresholding (Fig. 10).

3.2. Statistical analysis

The resulting clusters have been analysed in terms of independent variable association. In simple words, we looked at whether the detected clusters at different threshold levels correlate with any one of the independently recorded variables: the language of the *tituli* (the *litterae*), the *conceptum*, the *regestum* or the location within *Pompeii*, recorded at the scale of the neighbourhood (*Regio*), block (*Insula*) or particular building (*Aedificium*). In addition, a quantitative analysis of the identified clustered was undertaken with an epigraphic approach.

3.3. Software used

The analysis has been done in Python 3.6, including the following libraries: editdistance 0.3.1,[13] matplotlib 2.0.2,[14] numpy 1.12.1,[15] pandas 0.20.1,[16] and seaborn 0.7.1.[17]

4. Case study and results

The result of the analysis is several combinations of *codices* from PO08 amphoras displayed graphically in the form of heatmaps and dendrograms. These graphics show the number of combinations of *regesta* and the number of amphoras involved in each one of them.

To interpret the resulting graphics and to evaluate the significance of the resulting clusters we have used a combination of quantitative and qualitative methods. The former was used to compare sets with a considerable number of amphoras and *regesta*, in order to investigate the general patterns of sequences of *regesta*. This was followed by more traditional evaluation of the epigraphic contents highlighted by the analysis. For example, the analytical methods identified a structural group of *codices* with pairs of Greek names, but to interpret them we appraised whether some of these names corresponded to the

---

[13] Hyyrö 2002; Hyyrö 2005.
[14] Hunter 2007. The code is available here: https://www.scipy.org/citing.html 10.5281/zenodo.592845. https://github.com/aflc/editdistance
[15] Oliphant 2006.
[16] McKinney 2010.
[17] Waskom *et al.* 2017. The code is available here: https://github.com/izaromanowska/edit_distance.

same or different people and whether some of these characters were also present in other structural groups. In that way we can research phenomena like the contextual or sequential configuration of inscriptions, as well as the possible participation of one or more people in specific contexts.

4.1 General description of results

The edit distance analysis has generated three main dendrograms or cluster diagrams (see examples in Fig. 10), each of which depends on the versions Codex IV, V, and VI, respectively. Their horizontal axis shows the inventory numbers of inscriptions. Their vertical axis shows the thresholds of similarity. We tested the clustering strength with the Index of Similarity P, which indicates that all clusters under the threshold of 70 are valid for our research, i.e., it demonstrates that these clusters are not the result of analysing a random sample.

The number of clusters varies from 10 for the threshold of ninety to over 100 for the threshold ten going up to 144 in case of Codex V, which includes uncertainty. If we take the threshold 50 as an example from the middle of the displayed range, we observe that approximately ⅓ of the clusters contain over ten inscriptions for Codex IV and VI, but only approximately ¼ for Codex V. In conclusion, different strategies of research can be proposed regarding the distribution of clusters resulting from the different treatments of the uncertainty. On the one hand, the selection of thresholds can be modified according to the convenient size of the related clusters of inscriptions. On the other hand, we can contrast different ways of clustering to obtain a more complete understanding of evidence and method.

4.2. Selection of samples to analyse the epigraphic contents

4.2.1. Pairs of Greek names

Codices nn (two Greek *nomina*) are an interesting epigraphic structure.[18] Here, we examine them by comparing the dendrograms from Codex IV, V and VI (C.IV, V and VI) using the threshold 50. Three groups appear clearly. The first one contains all the *codices* with just one *regestum*. The second and third ones contain the two-*regesta codices*, with a majority of *regesta Graeca* and *regesta Latina* respectively.

Within the group of two-*regesta-Graeca codices* from C.IV and V, a main division distinguishes two new subgroups. In the first subgroup, all types of *regesta* are combined with *nomina Graeca* (n+). In the other subgroup, all types of *regesta* are combined with *incerta* (i+). However, the subgroup n+ from C.VI differs from another subgroup composed of i+ and x+ (*incerta* or *numeri Graeci* with another *regestum*). In such subgroup, the presence of x+ is quite relevant because of the number of items.

Curiously, there is a low ratio of x and x? compared to n and n? within the one-*regestum codices*, as seen in the dendrogram C.V. The higher number of complete *nomina Graeca* is an explanation for this difference. This completeness of the names makes it easier for their identification as *nomina Graeca* and reduces the uncertainty. On the other hand, the degree of abbreviation is higher within the two-*regesta-Graeca codices.* As a result, there is a higher degree of uncertainty and potential for alternative codifications. This difference could be considered an additional criterion for distinguishing epigraphic structures.

---

[18] Martín-Arroyo 2019: 466.

dendrogram of the CODEX 4 data por con_fit_sorted Threshold: 80

dendrogram of the CODEX 4 data por con_fit_sorted Threshold: 50
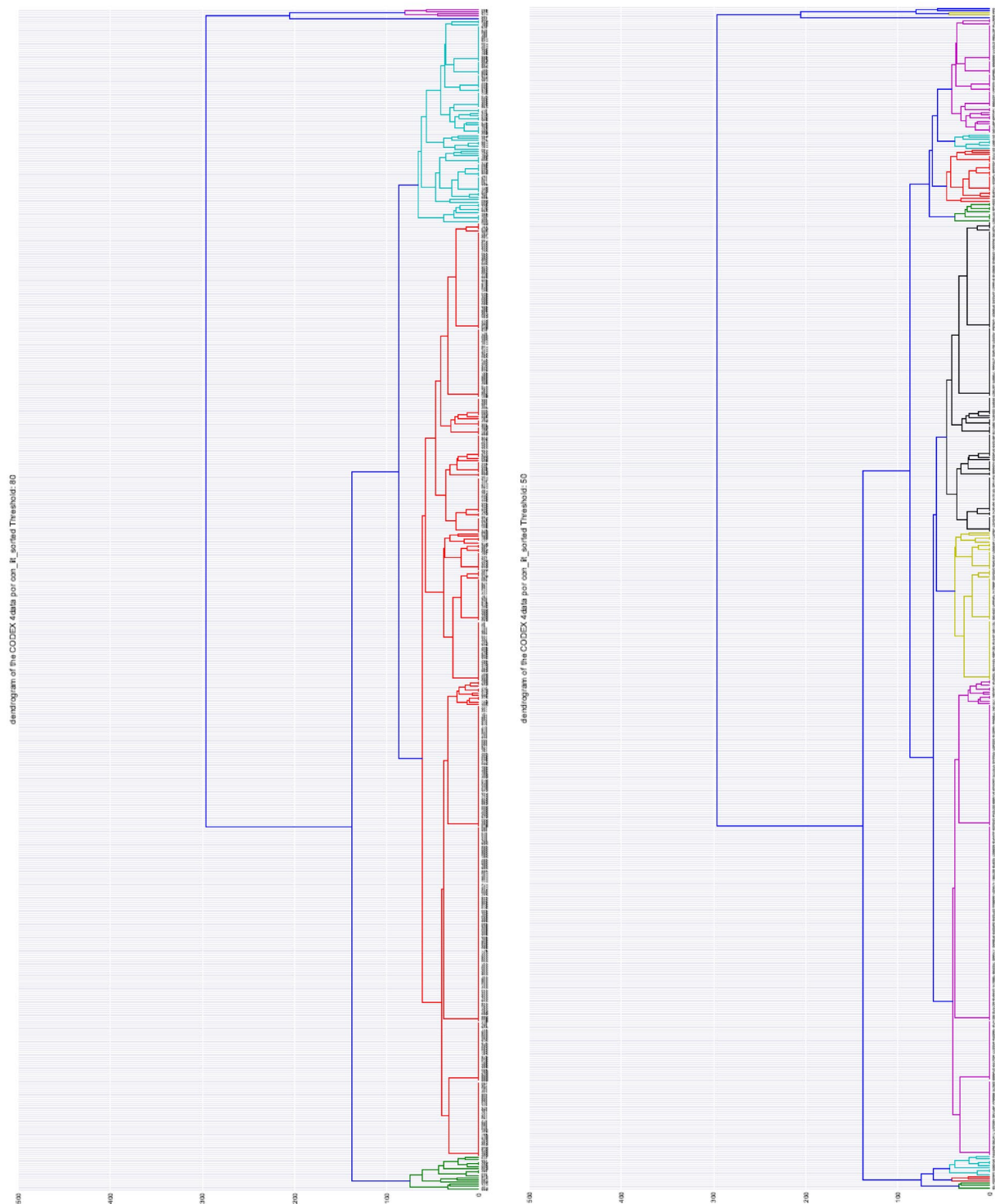
Figure 10. Example dendrograms with clusters coloured, based on the value of the threshold.

Most of the inscriptions with two *regesta* would be contemporary rather than resulting from ascribing an additional name at a later time. On the one hand, there are not many *codices* nn (7 items in C.V), even when the uncertainty has been reduced in C.VI (15 items). Thus, pairs of names were used, but they were frequently written in an abbreviated way. On the other hand, the commented ratio of x and x? to n and n? in C.V supports this general interpretation. In other words, the ratio of possible abbreviated names in one-*regestum codices* is lower than within two-*regesta codices*. In fact, n has a relevant position in all dendrograms (88, 88 and 92 items respectively). The weight of i within the dendrogram C.IV (27 items) is moderately distributed between the other *regesta* in dendrograms C.V (14 items) and VI (14 items), without affecting n too much. Consequently, few isolated *incerta Graeca* (i) can be interpreted as *nomina Graeca*. It is not the case of the two-*regesta* n+. *Codices* ii from C.IV (44 items) increases the quantities of other combinations within C.V and VI (eight remaining "ii" in each one), including n?n? (5), n?i (5), in? (2), x?x? (10), x?i (5) and ix? (6 items in C.V). These *codices* could contain abbreviated names within nn structures in alternative interpretations.

It is noted that two different structural criteria are observed for *codices* n and nn, with more abbreviations within the later ones. *Codices* nx (1) and xn (1 item) are rare and there is no xx in C.V. Few *numeri Graecae* are clearly identified. Furthermore, in *codices* np (1) and pn (15), the *productus* ΛΥΤΤΙΟC (Fig. 1) involves the majority of the evidence (15 items), as an exception to the rule of lack of references to the transported wine.[19] Numbers and products seem to be exceptions. In summary, the *codex* nn is the most likely interpretation in cases of uncertainty in the two-*regesta-Graeca* inscriptions. Differences in the writing of names between *codices* n and nn could involve different exchange contexts or procedures and, consequently, different functionalities of the *tituli picti* and the mentioned individuals.

ΔΙΔΥ(---) OR ΔΙΟΚΛΕ(---) OR NIKIA(---)

ΔΙ(---)

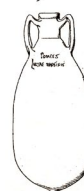**Ti.C.O.**

X 5

KO(---) OR M(---) OR [M]CY(---)

ΦΗΛΙΚΟC

C.A.R.

X 6

ΤΡΥ(---)

ΙΟΥΔΑΙΚΟC

L.E.S.
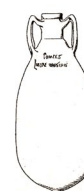
X 4

Figure 11. Comparison of epigraphic contents. Proposal of parallel cases.

---

[19] Martín-Arroyo 2020.

4.2.2. *Search of parallel inscriptions starting from a preceding structure case*

In a previous paper,[20] the probable Latin initials *C. A. R.* and *Ti. C. O.* were highlighted (Figs. 12-13) because they were written with yellow ink (*color gilvus*), which is not a common usage. Both appeared alone or with other *tituli*. Six structural groups of *codices* were defined to classify the inscriptions. Group 5 contains remarkable parallel cases. *C. A. R.* appeared on six amphoras PO08 (Fig. 11, nº 1-6). *Ti. C. O.* appeared on five Cilician Pompeii XIII (PO13) (Fig. 11, nº 7-11). These two probable initials were combined with pairs of complete or abbreviated Greek names written in black ink (*atramentum* or *color nigrum*).

Nº	INV.	*CIL*, IV	YEAR / *DOMUS*	TYPE	INK	INITIAL	*TITULI*	C. I	C. II	C. III
1	1305	6583	1876 VI 14, 38	PO08	GIL	C. A. R.	KO / ΦΗΛΙΚΟC // CAR	i/n//\I	n?/n/\N?	n/n/\N
2	1306	6584	1876 VI 14, 38	PO08	GIL	C. A. R.	M / ΦΗΛΙΚΟC // CAR	i/n//\I	n?/n/\N?	n/n/\N
3	1307	6584	1876 VI 14, 38	PO08	GIL	C. A. R.	M / ΦΗΛΙΚΟC // CAR	i/n//\I	n?/n/\N?	n/n/\N
4	1308	6584	1876 VI 14, 38	PO08	GIL	C. A. R.	M / ΦΗΛΙΚΟC // CAR	i/n//\I	n?/n/\N?	n/n/\N
5	1310	6586	1876 VI 14, 38	PO08	GIL	C. A. R.	[M]CY / ΦΗΝΥΚΟC / CAR	i/i/\I	n?/n?/\N?	n/n/\N
6	1311	6586	1876 VI 14, 38	PO08	GIL	C. A. R.	[M]CY / ΦΗΝΥΚΟC / CAR	i/i/\I	n?/ n?/\N?	n/n/\N
7		6335	1876 VI 13, 16	PO13	GIL	Ti. C. O.	Δ[I] / ΔΙΔΥ^ΚΟ… / TICO	i/n/\N	n?/n/\N	n/n/\N
8		6336	1876 VI 13, 16	PO13	GIL	T. C. O.	[Δ]I / ΔΙΔΥM(---) / TCO	i/n/\N	n?/n/\N	n/n/\N
9		6341	1876 VI 13, 16	PO13	GIL	T. C. O.	ΔI / ΔΙΟΚΛΕ[ΟΥC] / TCO	i/n/\N	n?/n/\N	n/n/\N
10		6341	1876 VI 13, 16	PO13	GIL	T. C. O.	ΔI / ΔΙΟΚΛΕ[ΟΥC] / TCO	i/n/\N	n?/n/\N	n/n/\N
11		6500	1876 VI 13, 16	PO13	GIL	Ti. C. O.	ΔI / NIKIA(--?) // TICO	i/n/\N	n?/n/\N	n/n/\N
12	1503	9757	1939 I 8, 8-9	PO08	NIG	L. E. S.	TPY / LES / IOYΔAIKOC	i/I/i	n?/N?/n?	n/N/n
13	1504	9757	1940 I 8, 8-9	PO08	NIG	L. E. S.	TPY / LES / IOYΔAIKOC	i/I/i	n?/N?/n?	n/N/n
14	1505	9757	1941 I 8, 8-9	PO08	NIG	L. E. S.	TPY / LES / IOYΔAIKOC	i/I/i	n?/N?/n?	n/N/n
15	1506	9757	1942 I 8, 8-9	PO08	NIG	L. E. S.	TPY / LES / IOYΔAIKOC	i/I/i	n?/N?/n?	n/N/n
16	2260	10293	1952 I 9, 13	PO08	NIG	Q. L. K.	[Q]LK / ANNI·[M]ω // AKA	N/n//n	N/n//n	N/n//N

Figure 12. Comparative table showing the Cluster 69 / Threshold 10 / Codex VI (nº 1-6, 12-16) and other parallel inscriptions on PO13 amphoras (nº 7-11).

Looking at the dendrograms, all the inscriptions with the *tituli* CAR (nº 1-6) are included in cluster 69 from threshold 10 in C.VI. This group contains five other amphoras PO08 with 3 *regesta* in the inscriptions of each one. The resulting 11 *codices* in C.VI have these structures: nnN, nNn, and Nnn. In Figure 12 these structures appear in the column C.III. Codex I, II, and III are equivalent to Codex IV, V, and VI respectively, but including spatial and ink characters (Figs. 2-3). Details about the coding of the inscriptions studied previously have already been published.[21] Now the *regesta* from the remaining inscriptions will be examined to test the level of similarity.

---

[20] Martín-Arroyo 2019.
[21] Martín-Arroyo 2019.

Firstly, similarity is rather doubtful in the case of the inscription nº 16. Its first line of text contains the probable initials of a *tria nomina*. The letter L indicates that it is a Latin *titulus*. The first letter is incomplete. The CIL proposes it is "O", but "Q" is suggested here with reference to the *praenomen Quintus*. No *cognomen* starts with K, and therefore it is probably a mistake. No Latin word or name begins with OLK. The three-character structure of the *titulus* reinforces the *tria nomina* hypothesis.

The *CIL* proposal for the reading of the second line is *Anni Mo*(*desti?*), because a similar abbreviated name appeared in a *signaculum*. The Greek letter ω and the very poor preservation of the character M make such an interpretation difficult. It is probably a Greek name with two words linked by a dot. It could be the nominative form Ἄννι[22] or an abbreviated name. The second word can be an abbreviated or partially erased patronym. Or, contrarily, it is possible to interpret it as a transliteration of the genitive form of the *nomen Annius*, as this type of case has been previously observed.[23] However, these cases are rare, making its interpretation more likely as a Greek personal name. In a similar way, the last *titulus* can be the nominative Ἄκα[24] or an abbreviated form.

The inscription 16 is written completely in black, without any yellow component. The distribution of its components makes it difficult to associate the Greek names. The first one was on the shoulder or back of the amphora (*in umero*) and the second one was under the handle (*sub ansa*). Furthermore, both of them can be read as nominatives, not as genitives. Because of all these differences, the parallelism between the inscription 16 and the CAR inscriptions is likely to be coincidental.

The inscriptions 12-15 have a common transcription in *CIL*. In that sense, no alternative reading of their characters is possible, as limited by the typography used. Another PO08 with a similar inscription[25] was found in 1904, at the Pompeian building (*domus*) VI 16, 10. Here, the proposed transcription is ΤΥ / ΙΛΙΙΧ / ΙΟΥΔΑΙΚΟ[Υ?]. There is the possibility of reading a genitive of Ιουδαικος, but this name is not recorded in *LGPN*. The most similar form is Ἰουδαῖος[26] and the most frequent is Ἰούδας, with 33 entries.

The *titulus* ΤΡΥ is interpreted as the Greek abbreviated name of a product in *CIL*, though this hypothesis is controversial.[27] The characters in ΤΡΥ can be read as Latin letters. However, they are probably not the initials of a *tria nomina* because there is no male *cognomen* beginning with Y in the *Onomasticon Provinciarum Europae Latinarum* (*OPEL*), other than one female case. However, these characters fit the first letters of 41 entries in *LGPN*. The 438 records of the name Τρύφων must be highlighted. As Greek names, ΤΡΥ(---) and ΙΟΥΔΑΙΚΟΣ are similarly expressed as the aforementioned pairs of names. Furthermore, the possible variety of the cases ΤΡΥ and ΤΥ (if this one is not a writing or translation mistake) points to the diversity of abbreviated names that are linked to the common names within the considered groups of amphoras, as it is ΙΟΥΔΑΙΚΟΣ on this occasion.

The *titulus* LES contains the Latin characters L and S. It was written between the Greek names at a later date. Greek inscriptions would precede Latin *tituli* if the amphoras were transported from the Eastern to the Western Mediterranean, i.e., from a Greek to a Latin speaking context. Furthermore,

---

[22] *LGPN*, V4-1823-1826 and V5a 4331-4332.
[23] Martín-Arroyo 2020.
[24] *LGPN*, V3a-37206, V5a-4137 and V5a-40987.
[25] *CIL*, IV, 6990.
[26] *LGPN*, V3b-42421.
[27] Martín-Arroyo 2020.

LES appears alone on another PO08[28] from the same building as inscriptions 12-15. This fact points to the cases of CAR and TICO, which appeared alone sometimes. They were written without dots and, occasionally, in *carbo* (charcoal) or *colore nigrum* (in black), which is the case of LES. CAR and TICO were written in different parts of the respective of the amphoras, which reaffirm their independence from the Greek *tituli*. In summary, epigraphic context, punctuation, ink, and position of CAR, TICO and LES are comparable. In a similar position to LES, the *titulus* ILIIX could be a Latin number that was written in a non-standard way and could be interpreted as the number 47, or 49 and 8, or some other combination. However, its epigraphic context suggests a mistake in the transcription of LES. The reason could be poor preservation, similar to the missing last part of ΙΟΥΔΑΙΚΟ[Υ?]. Equally, it could be a wrong transcription of any other word with the same function as LES.

The *titulus* ΙΟΥΔΑΙΚΟC has been related to the wine as well as to the trader of these amphoras.[29] It translates as "the Jew" and its male gender fits οἶνος, which means "wine". However, it could be considered as the genitive form of a personal name. The case ΛΥΤΤΙΟC as appellation of origin seems exceptional. It is mostly linked to M. ΠΟΠΥΛ. ΤΕΥΠΩΝΟC, a quite complete *tria nomina*. In a different epigraphic structure, ΙΟΥΔΑΙΚΟC appears linked to abbreviated Greek names. It is similar to the case of ΦΗΛΙΚΟC (Fig. 11, nº 1-4), the genitive form of Φῆλιξ.[30] Equally, some possible abbreviated Greek names are linked to ΠΑΙΟΝΟC, genitive form of Παίων.[31] This genitive form appears on Cretan,[32] Cilician,[33] and PO12 amphoras.[34] Furthermore, this *titulus* is linked to the characters AN written with green ink. This type of ink is as similarly uncommon as the yellow ink. All these characteristics are similar to those from the inscriptions 1-11 in Figure 12, written on Cretan and Cilician amphoras. Contrary to ΦΗΛΙΚΟC-Φῆλιξ and ΠΑΙΟΝΟC-Παίων, ΙΟΥΔΑΙΚΟC and ΤΕΥΠΩΝΟC do not have clear references in *LGPN*. As presumably happened in the case ΦΗΛΙΚΟC-ΦΗΝΥΚΟC (Fig. 11, nº 1-4 and 5-6 respectively), mistakes in the writing of names and their genitive forms may offer an explanation. Consequently, ΙΟΥΔΑΙΚΟC could be an appellation of origin or the genitive form of a personal name.

The hypothetical framework concerning all these *tituli* can be summarized as follows. ΛΥΤΤΙΟC and ΙΟΥΔΑΙΚΟC can be appellations of origin. ΛΥΤΤΙΟC forms its own structural epigraphic pattern. On the other hand, ΙΟΥΔΑΙΚΟC is related to the pattern concerning ΦΗΛΙΚΟC and ΠΑΙΟΝΟC, which are the genitive forms of personal names. Three main hypotheses are proposed:

1. The epigraphic structure is not the result of the same trade context or pattern. The similarity of the *tituli* is completely or mostly a coincidence. ΙΟΥΔΑΙΚΟC was an adjective relating to wine. Φῆλιξ and Παίων were traders.

2. The epigraphic structure is the result of the same trade context or pattern. ΙΟΥΔΑΙΚΟC, ΦΗΛΙΚΟC, and ΠΑΙΟΝΟC are adjectives relating to wine.

3. Similarly, the epigraphic structure is the result of the same trade context or pattern, but ΙΟΥΔΑΙΚΟC, ΦΗΛΙΚΟC, and ΠΑΙΟΝΟC refer to names of traders.

---

[28] *CIL*, IV, 9758.

[29] ROCCA 2006: 294-295, note 28; MARTÍN-ARROYO 2020.

[30] In *LGPN*, 48 entries, hits from 100 BC to 999 AD.

[31] In *LGPN*, 28 entries, from 600 BC to 300 AD.

[32] PO08 and 10; *CIL*, IV, 5961 and 5963 respectively.

[33] PO28; *CIL*, IV, 5964.

[34] Probable Eastern Dressel 2-4 amphoras; *CIL*, IV, 5960 and 5962.

As the explanation of these trade contexts or patterns, the hypothesis number 2 highlights the geographical feature of ΙΟΥΔΑΙΚΟC. Other adjectives related to wine appear in the amphoras epigraphy from *Pompeii*.[35] Latin geographical adjectives are clearly identified. Some of them allude to Eastern islands (Cos, Chios, Crete and Rhodes) or cities (Cnidus). The inscription *passum Lycium*[36] alludes to the Roman province of *Lycia* in a similar way that ΙΟΥΔΑΙΚΟC would refer to Ἰουδαία. However, only the cities of *Kantanos* and *Lyttus* would have provided such a type of adjective in the Greek language. The exceptionality of these cases and the refutation of preceding similar hypotheses reinforce the possible relationship of ΙΟΥΔΑΙΚΟC with a personal name. Latin adjectives with the suffix *–anum* are usually related to wine. They derive from names of rural properties,[37] which in turn derive from personal names. Beyond the hypothetical use of the Greek genitive form in this way, no clear parallel words of this type have been found. Hence ΦΗΛΙΚΟC and ΠΑΙΟΝΟC are more convincingly related to names of traders.

Hypothesis number 3 involves three types of traders. Firstly, a variety of Greek providers of wine: ΚΟ(---), Μ(---), [Μ]CΥ(---), ΔΙΔΥΜ(---), ΔΙΟΚΛΕ(ΟΥC), ΝΙΚΙΑ(--?), ΤΡΥ(---), and maybe ΤΥ(---). Secondly, Greek gatherers of goods, wine at least: Φῆλιξ, Δι(---), and Ἰουδαῖος (or another similar name). The case of Παίων requires further examination in order to include it in this group. Thirdly, Roman citizens in the Latin speaking context: *C. A. R.*, *Ti. C. O.* and *L. E. S.* The predominance of abbreviated forms point to limited contexts, where a simple mark was enough to differentiate between amphoras owners. These amphoras were brought together to be traded or transported and the similarity of inscriptions can indicate a common origin in a specific warehouse, marketplace, or ship's hold. The *tituli* point to at least two transactions. One of them, presumably the first, was carried out in a Greek context followed by another one carried out by a Latin speaker.

The diverse provenance of the specified amphoras must be discussed too. If hypothesis 3 is accepted and all these amphoras are regarded as primarily to be used as containers, then the activity of the providers can be located along the coasts of Crete and Cilicia. In this way, the diversity of amphoras with the *titulus* ΠΑΙΟΝΟC can be explained. However, it is also possible that providers from the same region reused amphoras of different provenance. Either way, the similarity of the discussed epigraphic structures can be regarded as evidence of a trade network with multiple providers and a limited number of ports and traders.

These results were reached only thanks to the combination of traditional epigraphic critique with computational analysis and filtering of a large dataset. It is thanks to this systematisation of the evidence, from a highly diverse and challenging dataset, that the confirmation of an epigraphic pattern and a new parallel case were found, reinforcing the preceding theoretical framework of research.

5.0. Discussion

In this paper, we have presented a methodology for evaluating the syntactic structure of ancient inscriptions, quantifying the similarity between them and visualising them in several ways. We showed how it could be applied to large datasets and aid detailed epigraphic critique of the sources. A review of contents in clustered inscriptions yielded some relevant parallels, as attested in the case of group 69

---

[35] Martín-Arroyo 2020.

[36] *CIL*, IV, 5594.

[37] Panella and Fano 1977: 158, note 41.

from threshold 10 in Codex VI. Through this case study we have identified a likely commercial process involving Cretan and Cilician wine arriving in *Pompeii* before the eruption of the *Mons Vesuvius.* We also found that the display of the size of the groups and their clustering in dendrograms was helpful for the understanding of the uncertainty in the *codices* nn.

In the course of the investigation, we have identified certain limitations of the resources and methods employed. For example, the dataset is composed of a set of quite brief inscriptions. The general low number of *regesta* (1,89 per amphora on average) makes it more difficult to detect structural syntactic patterns. Consequently, comparative analysis has been focused on the treatment of uncertainty. In the scope of future work, the inclusion of a wider set of inscriptions would be useful to detect additional features; for example, contrasting the epigraphy on Cretan (PO08 and 10) and Cilician amphoras (PO13). There are considerable quantities of these other inscribed amphoras (305 PO10 and 207 PO13). Assuming that the same individuals traded all of them along similar routes, one would expect a sizeable part of them to carry *tituli* with a similar syntactic structure. However, the structural diversity of inscriptions would indicate particular circumstances in the wine trade across the Mediterranean.

Further development should also involve a review of the codification system concerning the range of *regesta*. Variations on the Codex configuration are useful in the study of inscriptions through texts and tables or computational analysis. A new expanded set of *regesta* could include aspects such as the ink colour. However, these changes should be done in a rigorous way. For example, a wider variety of *regesta* (regarding *conceptum*, *litterae*, uncertainty and ink) would make the comparison of structures in texts and tables more complicated. A simple logic, composed of a set of unique symbols would avoid problems such as those related to question marks in Codex V and would facilitate the computational analysis. As previously mentioned, not all the entries in *CIL* include information on the used inks. So the partiality of all new incorporated factors must be considered. Similarly, additional analysis could be carried out by regarding aspects such as the position of the *regesta* on the amphoras or their relative position with respect to each other.

Graphics in relation to general parameters, such as the composition of the epigraphic evidence by Litterae or the spatial distribution of amphoras in *Pompeii*, should be further explored. Interactive graphics in digital formats would be more useful for displaying the results, for example, showing the common *codex* of a cluster or the contents of a particular inscription by clicking on its inventory number. The development of such tools could contribute to the use of the Codex methodology in other realms, such as lapidary epigraphy or diplomatics.

6.0. Bibliography

Aguilera Martín, A. 2000. Los *tituli picti* δ del convento astigitano en el primer tercio del s. III d.C., in *Conservas, aceite y vino de la Bética en el Imperio Romano. Congreso Internacional Ex Baetica amphorae (Sevilla – Écija, 17 al 20 de diciembre de 1998).* 1231-1240. Écija: Gráficas Sol.

Aguilera Martín, A. 2012. La normalisation de l'épigraphie amphorique: les *tituli picti* des amphores Dressel 20, in M.E. Fuchs, R. Sylvestre and C.S. Heidenreich (Dirs.) *Inscriptions mineures: nouveautés et refexions. Actes du premier colloque Ductus (19-20 juin 2008, Université de Lausanne). 135-143.* Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford, Wien.

Hunter, J.D. 2007. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* 9: 90-95. DOI:10.1109/MCSE.2007.55.

Hyyrö, H. 2002. *Explaining and Extending the Bit-parallel Approximate String Matching Algorithm of Myers.* Working Paper Department of Computer and Information Sciences, University of Tampere, Finland http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.19.7158&rep=rep1&type=pdf Accessed 31/10/2019.

Hyyrö, H. 2005. Bit-parallel approximate string matching algorithms with transposition. *Journal of Discrete Algorithms* 3 (2–4): 215-229. doi.org/10.1016/j.jda.2004.08.006

Lagóstena Barrios, L. 2002-2003. Aportación al conocimiento de la sociedad de la costa de la *Ulterior* en época republicana y julio-claudia. El registro δ en los *tituli picti* de las ánforas salsarias de *Castra Praetoria. Lucentum* XXI-XXII: 227-236.

Lagóstena Barrios, L. 2004. Las ánforas salsarias de *Baetica*. Consideraciones sobre sus elementos epigráficos, in J. Remesal, J. (Ed.) *Epigrafía anfórica. Workshop. Barcelona, 9-10 mayo 2003* (Instrumenta 17). 197-219. Barcelona: Publicacions de la Universitat de Barcelona.

Marangou-Lerat, A. 1995. *Le vin et les amphores de Crète. De l'époque classique à l'époque impériale* (Études Crétoises 30). Paris.

Martín-Arroyo Sánchez, D.J. 2018. Contexto arqueológico e historiográfico de las ánforas Crétoise 2 en *Pompeii*: una crítica cuantitativa, in J. Remesal Rodríguez, V. Revilla Calvo and J.M. Bermúdez Lorenzo (eds.) *Cuantificar las economías antiguas: problemas y métodos* (Instrumenta 60). 315-340. Barcelona: Publicacions de la Universitat de Barcelona.

Martín-Arroyo Sánchez, D.J. 2019. Los *tituli picti gilvi* de las ánforas Pompeii 8 – Crétoise 2. *Epigraphica* LXXXI, 1-2: 443-469.

Martín-Arroyo Sánchez, D.J. 2020. ΛΥΤΤΙΟC: epigrafía anfórica y contexto económico de un vino de la Creta romana. *Latomus* 79 (1): 126-157.

Martín-Arroyo Sánchez *et al.* 2017 = Martín-Arroyo Sánchez, D.J., L. Prignano, I. Morer Zapata, G. Rull Fort, M. García-Sánchez, A. Díaz-Guilera and J. Remesal Rodríguez, J. 2017. The Wine Trade of Roman Crete: Construction of Onomastic and Geographical Networks, in J. Velaza (Eed.) *Insularity, Identity and Epigraphy in the Roman World*. 177-196. Cambridge: Cambridge Scholars Publishing.

Martínez Maganto, J. 1998 Inscripciones sobre ánforas de salazón: interpretación sobre la estructura y significado comercial de los *tituli picti*, in AA.VV. 1998. *Congreso internacional* Ex Baetica amphorae. *Conservas, aceite y vino de la Bética en el Imperio Romano (Sevilla-Écija, 17 al 20 de diciembre de 1998)*. 1207-1219. Écija.

McKinney, W. 2010. Data Structures for Statistical Computing in Python, in Walt, S. van der and J. Millman (eds.) *Proceedings of the 9th Python in Science Conference.* 51-56. Austin. http://conference.scipy.org/proceedings/scipy2010/pdfs/mckinney.pdf Accessed 31/10/2019.

Oliphant, T.E. 2006. *A guide to NumPy.* Trelgol. http://web.mit.edu/dvp/Public/numpybook.pdf.

Panella, C. and M. Fano 1977. Le anfore con anse bifide conservate a Pompei: contributo ad una loro classificazione, in *Méthodes classiques et méthodes formelles dans l'étude typologique des amphores. Actes du colloque de Rome, 27-29 mai 1974*. 133-177. Roma.

Remesal Rodríguez, J. and A. Aguilera Martín 1999. *Addenda et corrigenda* a los *tituli picti* y a los sellos de la campaña de 1989, in J.M. Blázquez Martínez, J. Remesal Rodríguez (eds.) *Estudios sobre el Monte Testaccio (Roma) I* (Instrumenta 6). 101-127. Barcelona: Publicacions de la Universitat de Barcelona.

Remesal *et al.* 2014 = Remesal Rodríguez, J., A. Díaz-Guilera, B. Rondelli, X. Rubio Campillo, A. Aguilera Martín, D.J. Martín-Arroyo Sánchez, A. Mosca and G. Rull Fort 2014. The EPNet Project. Production and distribution of food during the Roman Empire: Economic and Political Dynamics, in: S. Orlandi, S. Santucci, V. Casarosa and P.M. Liuzzo (eds.) *Information Technologies for Epigraphy and Cultural Heritage. Proceedings of the First EAGLE International Conference. Europeana Eagle Project* (Studi Umanistici. Serie Antichistica). 455-464 Roma, .

Rocca, S. 2006. A Jewish gladiator in *Pompeii. Materia Giudaica* 12, 1-2, 287-301.

Waskom *et al.* 2017 = Waskom, M., O. Botvinnik, drewokane, P. Hobson, Y. Halchenko, S. Lukauskas, J.B. Cole, J. Warmenhoven, J. de Ruiter, S. Hoyer, J. Vanderplas, S. Villalba, G. Kunter, E. Quintero, M. Martin, A. Miles, K. Meyer, T. Augspurger, T. Yarkoni, P. Bachant, M. Williams, C. Evans, C. Fitzgerald, D. Wehner, G. Hitz, E. Ziegler, A. Qalieh and A. Lee, A. 2017. *seaborn: v.0.7.1 (June 2016)*. Zenodo. http://doi.org/10.5281/zenodo.54844.