Archeologia e Calcolatori
31.1, 2020, 7-32
doi 10.19282/ac.31.1.2020.01

IMAGE STUDY OF MESOPOTAMIAN CYLINDER SEALS THROUGH TEXTS: THE PROCRUSTES TRANSFORMATION APPLIED TO CORRESPONDENCE ANALYSIS RESULTS[1]

1. Introduction

This paper reports an experiment performed on a corpus of texts describing presentation scenes engraved on late 3rd millennium BCE Mesopotamian cylinder seals. The aim is to define a methodology of exploratory textual analysis suitable for the study of the iconographies of ancient findings, once these images are described through a text. Indeed, the translation of visual content into formalized textual encoding may lead to highly complex formal systems, which allow in-depth investigations. However a potential drawback is the introduction of too many elements, which may confuse interpretation of the results. To reduce this risk, the effects of reducing further the forms considered in the analysis process have been studied. In this paper we report on the results of this experimentation, together with their interpretation.

The analysis of the iconographic structure of images engraved on archaeological finds having a cyclical nature using textual data was first addressed by Camiz and Rova (1991) and subsequently continued by them (Camiz, Rova 1996), and by Di Ludovico and Camiz (2014). All of this research deals with corpora of Mesopotamian glyptic iconography (protohistoric in the first research project, produced by a historical complex state in the second) as described through a formalized text. Indeed, texts may describe at best all iconographic elements, when properly formalized.

The issue of textual coding of archaeological artefacts was first introduced by Gardin (1958), considering the utility of having uniform, standardized descriptions of the material at hand. Camiz and Rova (1991) followed the idea of studying the visual language of seals from an iconographical point of view, and Camiz (2004) clarified its use in the framework of other codings currently adopted in archaeology. Thus, Textual Correspondence Analysis (TCA: Lebart, Salem 1988) was the main method they adopted, owing to its ability to represent graphically similarities and dissimilarities between texts or objects, accompanied by the interpretation provided by an analogous representation of the lexical forms. Nevertheless, some drawbacks exist in the use of TCA, which are intrinsic to classical Correspondence Analysis (CA: Benzécri 1973-82) applied to sparse tables, in particular the overweighting

[1] This article is – in all its parts – the product of a joint effort of both authors. Only for academic purposes, we attribute here parts 1 (Introduction) and 5 (Conclusion) to both, parts 2 (Materials) and 4 (Results) to Alessandro Di Ludovico, and part 3 (Methods) to Sergio Camiz.

of low-frequency items, whose high inertia influences the extracted factors (Lebart, Salem 1988; Cibois 1997; Lebart *et al.* 2006). In this case, the reduction of the data table to the forms with highest frequencies is currently adopted, to provide more stable and more safely explainable factors.

Unlike classic CA, in which low frequency levels can be aggregated, when dealing with texts, the elimination of rare forms would result in a loss of meaning, hence the problem of finding a reasonable compromise between such loss and the stability of the results is raised. If the text is a special coding for iconographic elements the problem is heightened, because one may lose relevant (albeit rare) elements, which may represent significant aspects to be taken into account. The situation concerning repeated segments is even worse – sequences of forms which occur in exactly the same way within the corpus – because their number is much larger and low-frequency ones become prevalent.

In the study described in this paper, we try to identify a suitable reduction of the forms contained in 354 formalized texts describing the iconographies of seals already analyzed by Camiz and Di Ludovico (2014). We considered two double contingency tables cross-referencing the origins and dates of the seals with both textual forms and repeated segments contained in the texts, and we submitted them to several TCAs. These were applied after removing increasing percentages of low-frequency forms or segments. The results were then interpreted through the so-called characteristic forms and segments, i.e. those whose frequency at a site or during a period is significantly higher than expected. In this paper, we will only deal with forms: the results concerning repeated segments will be reported in a following paper.

2. Materials

2.1 *Glyptic compositions*

The corpus at hand corresponds to that used in previous quantitative investigations of scenes depicted on glyptic artefacts dating back to the 21st and the early 20th century BCE (Di Ludovico 2005; 2008; Di Ludovico, Camiz 2014). The whole corpus is fairly homogeneous from both compositional and iconographic points of view: it is made up of 354 specimens, including original seal impressions and cylinder seals. These scenes belong to the so-called *presentation* theme, a general subject – not always clearly defined – found in Lower Mesopotamian glyptic from the late Early Dynastic period (Di Ludovico 2005, 60-61). The production of seal representing *presentation* scenes is not exclusive to the late third and early second millennia, but there are strong clues suggesting that this subject was prevalent during the Third Dynasty of Ur (21st century BCE). This has led to hypotheses that it acquired an important official meaning and reached an unprecedented formal and compositional standardization in that period (Collon 1982, 129; Winter 1987; Di Ludovico 2005).

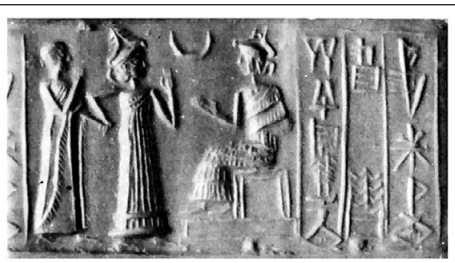
The basic scheme of a presentation scene is made of a receiving figure – which can always be clearly and unambiguously detected – represented together with at least another character, and a portion of space on the surface of the seal, marking both ends of the scene: an area not actually involved in the figurative narration. In the Ur III period, this is very often the place where the legend of the seal is carved. On the other hand, in most cases the representation includes between two and four anthropomorphic figures, which have been formalized in the data set in single descriptions expressing all their attributes, their position, and their features. In addition to this, in most cases the field also shows other symbols and objects. Here, these elements have been named *integrating motifs*.

2.2 *Textual encoding*

The presentation scenes of the corpus were described with a highly formalized text reproducing, at best, both spatial and compositional relations existing among the iconographic elements, as well as their attributes (Fig. 1). Thus, each character represented in the scene is described according to the following order: nature and gender, clothing, headgear, hairstyle, orientation, posture, position of each arm, physical attributes, and objects-attributes. Any integrating element is described with a category of belonging. This is a code indicating its specific type or shape and its position in the field, such as a scorpion placed in the bottom part of the field between two characters, while the seal legend is described for its shape and type of content, for example as three framed lines reporting the name, the patronymic and the profession of the owner of the seal. In the texts, parentheses enclose the forms which make up a substructure. Note that the encoding was drastically shortened in order to simplify the graphical representations. In the discussion of results, it will be paralleled with a clear description.

Such a text may be seen according to two points of view: i) the elements appearing in the image, such as human beings, deities, animals, objects, or symbols, and their attitudes, such as sitting, standing, holding objects, etc., which correspond to textual *forms* composing the text; ii) the small, sometimes repeated, sub-patterns – i.e. subsets of forms – which compose the image and may occur identically on different seals, such as a 'goddess standing, wearing a pleated robe and a crown with a single pair of horns, having a hairstyle with double-curl, facing right, raising her right hand before her face'. Such subsets may have major importance in the study of images traditionally composed of repeated sub-patterns, whose combination results in different images. They are a subset of the so-called *repeated segments* (Lebart, Salem 1988), sequences of forms which are repeated identically throughout the corpus, and which are usually extracted automatically from the textual corpus itself.

In textual coding, special care must be devoted to keep the 1-1 correspondence between descriptors and described objects: this may be obtained



(((Un U m_fr ac_c dx m_dx d_vt m_sx inb) p_m (una D v_a_p ti_c5 ac_n1 dx m_sx d_vt)) (at_l inh) ((una D v_a_b ti_c1 ac_n1 sx m_sx a_w m_dx inf) as_su (tr_Q1b pod_1)) (I_3_ks kt_c2))

((Una D v_a_b ti_c5 ac_n1 dx m_dx d_vt m_sx d_vt ha g3 g4) (un U m_fr ac_c dx m_dx a_w m_sx a_w) (at_l2 inh) ((un G v_a_b ti_c5 ac_b1 sx m_sx a_w m_dx inf ha bb1 ha g1) as_su (tr_Q8f pod_1)) (I_cr 4_ks kt_c3))

Fig. 1 – Two scenes with the compacted encoding adopted: BM 8909 (Collon 1982, n. 377), left, and MLC 2658 (Buchanan 1981, n. 626), right. The original texts, with the parentheses enclosing the sub-structures, are the following:

BM 8909 – (a man in fringed mantle, clean-shaven, facing right, holding his right hand before his face, his left hand downward) hand in hand (a goddess in pleated dress, wearing a headgear with multiple pairs of horns, hairstyle with a double curl behind her neck, facing right, holding her left hand before her face) (a moon crescent in the upper part of the field) (a goddess in flounced robe, wearing a headgear with a single pair of horns, hairstyle with a double curl behind her neck, facing left, holding her left hand at her waist, her right hand in front of her) sits on (square seat with a frame placed on a platform) (three cases framed legend containing the formula "PN, profession, son of PN2").

MLC 2658 – (a goddess in flounced robe, wearing a headgear with multiple pairs of horns, hairstyle with a double curl behind her neck, facing right, holding her right hand before her face, her left hand before her face, has a multiple necklace and double bracelets) (a man in fringed mantle, clean-shaven, facing right, holding his right hand at his waist, his left hand at his waist) (moon crescent with sun disk in the upper part of the field) (a god in flounced robe, wearing a headgear with multiple pairs of horns, hairstyle with hair lock behind his head, facing left, holding his left hand at his waist, his right hand in front of him, is bearded and has a bracelet) sits on (seat with double niche and multiple frames placed on a platform) (short four cases framed legend containing the formula "PN, profession, son of PN2, profession").

by using the same form for the description of the same element or the same attitude, avoiding inflections due to grammatical rules, ambiguities in the meanings of words, and the literary style used for the coding. In Fig. 1 two examples of coding are reported under the corresponding scenes.

3. Methods

3.1 *Correspondence Analysis*

TCA (Lebart, Salem 1988) is the most appropriate exploratory method which may be used to identify the most relevant sources of variation within a corpus of texts, in this case the iconographic differences found in the studied

images. Indeed, TCA is but a CA applied to the texts, once the different textual forms are automatically recognized and counted in each statistical unit (in our case, an image described by the formalized text). This way, a contingency table of images/texts-by-forms is created, in which each entry represents the number of times a particular form appears in a particular text, and thus the number of times an element appears in a scene.

CA (Benzécri 1973-1982; Greenacre 1983; Lebart, Piron, Morineau 2006) is a well-known exploratory analysis method for contingency tables, producing orthogonal factors, where to represent at the best both rows and columns of a data table – in our case, images and forms, respectively. Assuming independence between characters, the factors describe, in decreasing order of relevance, the deviations of the observed data with respect to independence. The number of relevant dimensions is usually checked on chi-squared tests based on the residuals (Malinvaud 1987), a test often impossible to carry out when dealing with large sparse tables, due to the highest number of zero frequencies. As CA is a canonical analysis (Gittins 1985), the squared root of the inertia of the factors (the eigenvalues) represents the correlation between their projections on rows and columns. Indeed, when these correlations are too low, the factor deserves no interest, since its projections on rows and columns are independent. Due to the CA metrics, which weights the items as inversely proportional to their frequency, its results are strongly affected by the rarest elements (Rao 1995, 45; Legendre 2001, 271; but see also Greenacre's 2006 discussion), that is – in our case – the poorest texts and/or the forms with least frequency in the corpus. This is a known drawback of CA, because relatively rare elements can exert a decisive influence on the structure of the axes, and should instead be excluded from the analysis: this raises the problem of identifying a suitable frequency threshold, allowing to withdraw lower frequency forms which may influence the axes formation too heavily, without the risk of losing too much relevant information.

3.2 *Characteristic forms and specific scenes*

A confirmatory complement to TCA, when dealing with groups of texts, is the identification of the forms which are typical of each group, i.e. they are either the only ones present, or they have a higher frequency in a group than others. In our case, this corresponds to the search for the iconographic elements which may distinguish the images of seals of a specific period or origin from the others. To this end, *characteristic forms* (Lebart, Salem 1988) may be looked for. These are forms appearing with significantly higher frequency at some sites or in some periods than in the whole corpus. Since, under random distribution, the frequency of forms follows a hypergeometric density law, one may consider forms characteristic of a site or a period to be those whose frequency probability of occuring at random, according to the

hypergeometric law, is lower with respect to a significance threshold (usually 5%). Identical considerations may be taken concerning items with significantly low frequencies. This will allow us to characterize the seals of each origin or period through the iconographic elements particularly present (or absent) in the images of that group.

To get a better idea of the kind of seal scenes identified as belonging to a certain group, some *specific specimens* may be extracted too, according either to their proximity to the group centroid, the average of the coordinates on the TCA factors of the scenes belonging to the group, or to the ranks of the characteristic forms which appear in it (Lebart, Salem 1988). It may be expected that they represent a kind of "iconographic centroid" of the group, i.e., they share some iconographic features with all other scenes of the group.

3.3 *Outcomes of the previous studies*

The above-mentioned studies concerning the iconographic structure of the seals were carried out in an exploratory manner using both TCA and the characteristic forms. The investigations of the corpus of presentation scenes (Di Ludovico, Camiz, Pieri 2013, 497; Di Ludovico, Camiz 2014, 15-29; 2015b, 490-491) led to highlight the weight of the contribution given to the scene by the gender of the main characters: male figures (gods and men) appear in the results as opposite to female ones. Both main types of character were consistently associated with their main relevant features. Everything related to the representation of goddesses or women seemed also to be related to the basic type of presentation scene, while *royal presentations* showed a peculiar, opposite status. Moreover, further opposition results between some types of concentrations of integrating motifs – which may reveal a reworking of the seal – and forms most peculiar to the presentations which appear in the original impressions, and so probably typical of seals still used for administrative purposes. The analysis which involved the external features (above all places of origin) confirmed this distinction as a fundamental one, as well as suggesting a clear geographic separation between southern and northern sites.

The results of the performed analyses were quite clear and consistent, but partly dispersed and perhaps somehow altered by the presence of too low-frequency elements. However, it is also true that low-frequent iconographic features could be typical and representative of specific geographic origin, historical (sub)period, or personal status. It is thus of great importance to evaluate how less frequent forms can influence the outcomes of analysis.

3.4 *Procrustes Analysis*

In ancient Greek mythology, the bandit Procrustes (Προκρούστης, "the stretcher") was a son of Poseidon with a stronghold on Mount Korydallos, situated on the sacred way between Athens and Eleusis. There he had an

iron bed in which he invited every passer-by to spend the night. According to the guest's size in respect to the bed, he used either to stretch him with his blacksmith's hammer to fit the bed, or to amputate the excess length. Indeed, nobody ever fitted the bed exactly, because Procrustes took advantage of two beds of different sizes. Procrustes continued his reign of terror until Theseus, travelling to Athens along the sacred way, "fitted" him to his own bed (Plut. *Thes.* 11a).

In data analysis, Procrustes Analysis (PA: Gower, Dijksterhuis 2004) is a method seeking the best adjustment of a test cloud, i.e. a set of points represented in the geometrical real p-dimensional space, to a given target cloud, another set of corresponding points in the same space, according to transformations composed only by translation, scaling, and rotation/reflection. Originally proposed by Mosier (1939), its name is due to Hurley and Cattell (1962). Further developments, in particular the Generalized Procrustes Analysis (GPA: Gower 1975), in which the best adjustment is sought between several clouds followed. Indeed, Borg and Groenen (2005) argue that the terminologies "Procrustes problem", "Procrustes analysis", etc., now standard, are generally misleading, «inasmuch we do not want to mutilate or distort the test cloud».

The method is largely used both in pattern recognition and in the so-called shape analysis (Dryden, Mardia 2006) as a first adjustment of more complex transformations, but it may be applied to all situations in which direct comparisons among configurations of the same objects under different representations are sought (for details on the method, see Gower, Dijksterhuis 2004; Borg, Groenen 2005, or Camiz, Denimal 2011). The method was considered by Kendall (1989) for shape analysis, quoting the geometry of Stonehenge as a possible application, and more recently by Davis *et al.* (2015) to study flaked lithic artefacts.

As a by-product of Procrustes transformation, the pairwise Procrustes correlation coefficient results. This is an index of multidimensional matching of two clouds ranging within the interval [0,1] (Legendre, Legendre 2012). It has the same meaning of the common squared correlation between variables and allows a numerical comparison of the patterns of the points issued by different analyses.

3.5 *The pathway*

After the construction of formalized texts describing the iconography of seals, and consequently producing the data table crossing the scenes with the forms, their selection was made. This was carried out first by removing the *hapax* – i.e. the forms appearing only once, and thus not informative about the multidimensional structure. Then, we applied progressive cuts to the rarest forms, in order to understand what could be considered *noise* – those

		Forms	
Analysis	*Percentage of occurrences*	*Number of different forms*	*Total occurrences of forms*
0	Original text	224	26369
1	100% without hapax	165	26310
2	95%	62	25106
3	90%	42	23852
4	85%	32	22516
5	80%	26	21226
6	75%	22	19974
7	70%	19	18610
8	60%	15	16178
9	50%	11	13062

Tab. 1 – Characteristics of the reduction of the contingency tables crossing seals with forms and repeated segments submitted to TCA: percentage of used items, corresponding number and total occurrences of used forms.

elements safely removable without losing relevant information – and what deserved to be kept – because of significantly contributing to the interpretation of the results.

Thus, from the whole table without *hapax,* we progressively built reduced tables, by keeping a percentage of forms corresponding to 95%, 90%, 85%, 80%, 75%, 70%, 60%, and 50% of the most frequent ones. In Table 1 these percentages are reported, together with the corresponding number of distinct forms, and their total occurrences, respectively.

Dealing with the whole corpus – namely the table crossing lexical forms with scenes – for this study would be too difficult, owing both to its size and the overly large and sparse resultant table, which would likely issue non-significant chi-square tests. Thus, we focussed on the eight toponyms referring to the origins of the seals – namely *North-Bab, North-East, South, Susa, Telloh, Umma, Ur*, and *Unknown,* gathering the seals with unidentified origin – and three phases within the period of the Third Dynasty of Ur: *Early-Ur-III, Core-Ur-III*, and *Late-Ur-III*.

TCA was run on the nine tables, joining the two built on sites and periods, and six factors were taken into account, regardless of their relevance and/or significance. In particular, the square roots of the eigenvalues were considered as a measure of the relationship between sites/periods and forms and their percentage of explained inertia as a measure of the explaining power of the factors within TCA. This suggested limiting to two the factors to be studied in depth, considering the eigenvalues and the significance of their associated canonical correlations.

Then, the first six coordinates of sites and periods issued by the various TCAs were pairwise compared using Procrustes analysis, to get pairwise

graphics and Procrustes correlations. This allowed us to check the variation at each step owing to the reduction of forms and retaining the Procrustes correlation between the two tables and between the current and the total. This led us to identify the dataset resulting from the first cut of frequencies, say limited to 95% of the occurrences, as the most suitable to take into account.

Eventually, the interpretation of each group was carried out through the identification of characteristic forms, paying particular attention to the effects of their withdrawal. Indeed, this check was carried out on the sets of characteristic forms identified after the reduction of forms, thus observing the loss in the descriptive ability of the progressively reduced sets.

For the purposes of this study, all textual manipulations and analyses have been carried out with the SPAD package (Lebart *et al.* 1999), whereas for the Procrustes methods the two *R* (*R* Core Team 2019) packages Shapes (Dryden, Mardia 2006) and FactoMineR (Lê, Josse, Husson 2008) have been used.

4. Results

4.1 *CA of external features*

In Table 2 the contingency table cross-referencing the dates and origins of scenes is reported, together with the representation of both sites and periods on the plane spanned by the first two factors of CA. In the graphic of Fig. 2 the following pattern is visible: *Early-UrIII* is isolated on the first axis, as opposed to the other two periods, which are separated along the second. Note that, whereas *Late/URIII* is at its extreme, *Core-UrIII* is much closer to the origin, because of its higher frequency with respect to the others.

The toponyms *Umma*, *Ur*, and *North-East* repeat a similar pattern, owing to some agreement with the corresponding datation. Indeed, the central position of most sites is due to the much higher frequency of the scenes found in these places dated *Core-UrIII*.

	Early-UrIII	*Core-UrIII*	*Late-UrIII*	*Total*
North-East	0	3	2	5
North-Bab	2	8	1	11
Susa	6	17	4	27
Ur	6	36	3	45
Tello	6	34	3	43
South	1	4	1	6
Umma	27	43	5	75
Unknown	21	106	15	142
Total	69	251	34	354

Tab. 2 – Contingency table crossing toponym and period of the scenes.

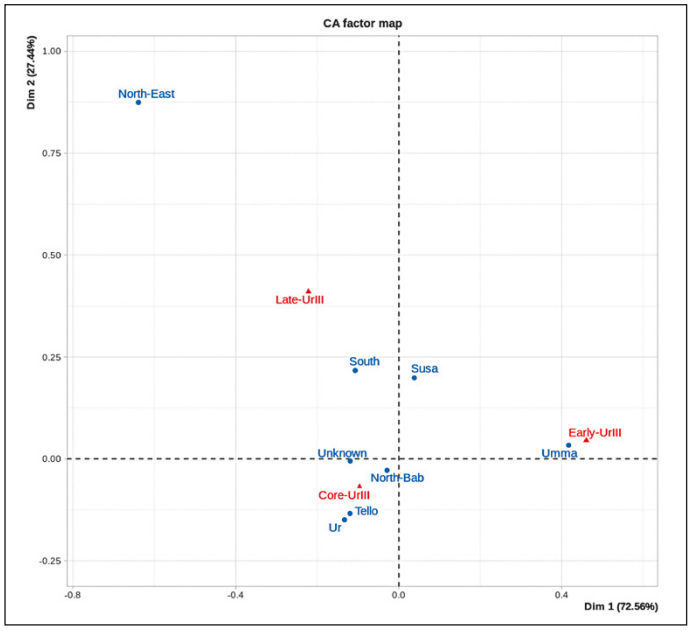


Fig. 2 – The items on the first factor plane issued by the CA of the contingency table crossing toponyms and periods of the scenes.

4.2 *Procrustes analysis*

In Table 3 the main results of the 9 TCAs are reported, together with their relations. For every analysis the first two canonical correlations are given, which inform about the relationship between table rows and columns, with the corresponding percentage of total inertia which gives information regarding the amount of data provided. This is followed by the Procrustes correlation, based on 6 TCA factors, with the analysis performed on the whole set of forms (excluding *hapax*), and the correlation with the previous one, both of which inform us about the loss of meaning of the analysis due to the reduction of the forms.

As a result, the canonical correlations are rather weak in each analysis, but, according to the threshold level for significance of the correlation (which, for 354 units and the probability level of 0.05, equals 0.1042556), those corresponding to the first component are all significant, whereas those corresponding to the second one are only significant up to 95%. Thus, it may not be advisable to cut over 5% of the forms, considering the loss of one dimension for larger cuts; on the other hand, should one decide to ignore the second dimension, a larger cut seems possible without losing meaning. This is

Forms Analysis	*Canonical Correlations*	*% Inertia*	*Procrustes correlation with 100% analysis*	*Procrustes correlation with former analysis*
100%	1 – 0.204	33.15	1.000	1.000
	2 – 0.152	18.41		
95%	1 – 0.189	51.88	0.817	0.817
	2 – 0.117	19.94		
90%	1 – 0.180	67.57	0.808	0.942
	2 – 0.079	13.18		
85%	1 – 0.162	71.47	0.827	0.991
	2 – 0.067	12.23		
80%	1 – 0.153	78.86	0.800	0.992
	2 – 0.055	10.07		
75%	1 – 0.128	80.69	0.814	0.977
	2 – 0.039	7.43		
70%	1 – 0.111	77.99	0.831	0.996
	2 – 0.039	9.43		
60%	1 – 0.117	84.15	0.792	0.995
	2 – 0.035	7.32		
50%	1 – 0.129	87.43	0.784	0.998
	2 – 0.036	6.81		

Tab. 3 – Procrustes transformations on the results of TCAs on the forms. In the columns: the percentage of forms in TCA, the first two canonical correlations and the corresponding explained inertia, the Procrustes correlation of the analysis with that on the whole table without *hapax*, and its correlation with the former one. Note that the correlation significance threshold, for n = 354 and p = 0.05, is 0.104.

also reflected by the following columns: observing the Procrustes correlations between analyses, it results that all of them maintain their correlation with the total one within the interval (0.830, 0.784) without a relevant decrease after the reduction of forms. At the first step, which produces the 95% table, this means that the structure of the following tables, concerning the position in the graphics of both sites and periods, does not change very much. Indeed, the Procrustes correlation between two adjacent analyses after the total one varies within the interval (0.942, 0.998), thus very little.

Therefore, one could imagine that, after the first screening of forms, which appears to affect the pattern of the periods first and then also of the sites, the following analyses would appear alike enough, suggesting the choice of reducing the number of forms significantly. On the other hand, the significance of the correlations would set a relevant limit to 95%, if one wishes to interpret both the first two factors, and to 50%, limiting attention to the first factor only. Indeed, the following results will show that the first cut does not significantly reduce the interpretation, whereas the subsequent ones will do.

In Fig. 3 the clouds of items paired by the Procrustes analyses of subsequent tables are represented. Note that the two axes may not correspond to those issued by TCA, due to the rotation intrinsic to the method. Looking at

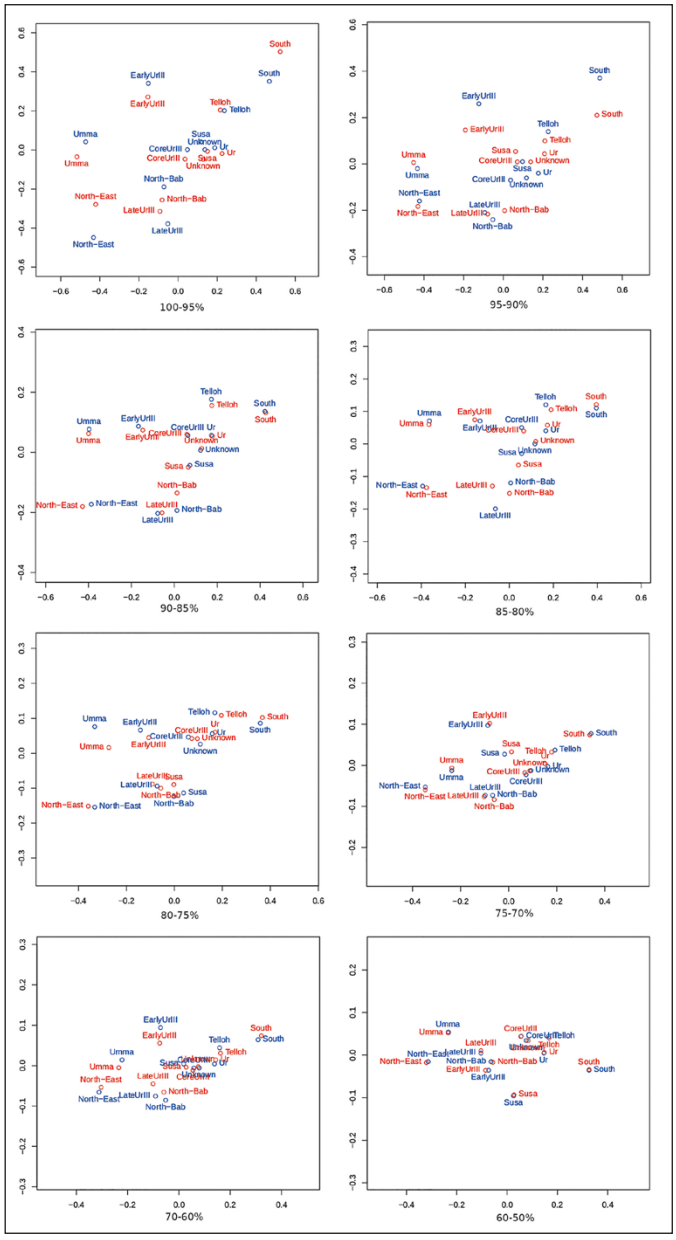


Fig. 3 – Analysis of forms: results of Procrustes transformations between pairs of graphical representation of both toponyms and dates. Each graphic represents the items according to two subsequent reductions of forms prior to TCA, adjusted to best fit according to Procrustes transformation. Below each graphic the percentages of forms present in the two analyses are set.

the series of graphics, the progressive shrinking of the cloud of items along the second axis is evident, whereas no such strong reduction occurs along the first one, even if in the last comparisons the items appear closer and closer. On the opposite, the last three comparisons appear somehow different from the previous ones. This allows us to imagine that in this case the Procrustes rotation issued, as main axes, directions different from those of the previous ones. Thus, in the last analyses the relevant information extracted in the first one was practically lost, due to the cut of the forms that characterized the first analyses.

The figures related to the outcomes of all TCAs revealed that only the first axis was actually meaningful and representative of the relations among the forms. A fairly constant opposition which could be observed in all analyses is that between forms related to scenes with features – like the presence of a character with material attributes (e.g. a weapon, jewels, a cup, etc.) – that are mostly typical of (probably) recent scenes or scenes known from original seal impressions, on the one hand, and features – like the divine headgear with one pair of horns or the hairstyle with a double curl – which are typical of female characters or of scenes known from (probably earlier) cylinders, on the other hand. The toponyms *Umma* and *North-East* are related specifically to the first ones, while *Ur*, *Telloh*, and *South* may be associated with the second group of forms, i.e., with the opposite half of the same axis. *North-Bab*, *Susa*, and *Unknown* (provenance) are positioned close to the centre of the axis (to its coordinate 0), while *North-Bab* is tendentially closer to *Umma* and *North-East*, and the other two rather to *Ur*, *Telloh*, and *South*.

This general frame resulted evident even from the few forms remaining in the analysis performed on 50% of active occurrences, but this is mainly because the previous ones had already made clear the meaning of the axes, as explained by the withdrawn forms.

In the extremely minimal analysis – represented here by that of the 50% data set – only part of the traits typical for the positive side of Axis 1 are actually evident. They are meaningful as far as they express the presence in the scene of a female deity wearing a headgear with a single pair of horns, but this does not say, by itself, very much about the structure of the scene and the possible features which would help in distinguishing the specific types of represented elements.

Only with the 70% analysis is there the chance of singling out some meaningful forms, insofar as they hint at compositional features of the scene (forms like *p_m*, referring to the couple of figures represented hand in hand, or *inh*, related to the presence in the scene of one or more integrating motifs placed in the upper part of the field), but they are very few, and limited to the positive side of the axis.

Few more emerges from the forms remaining after the 80% and 85% cuts. In the first case there are aspects possibly bound to chronological differences, like

ti_c5 (divine headgear with multiple pairs of horns) and *g3* (multiple necklaces, a typical attribute of goddesses) – both giving quite an important contribution to the formation of the negative side of the axis and hinting at a relatively late date – on one hand, and *v_a_p* (pleated robe), which, on the positive side, seems to express a quite early date, on the other. Only in the 85% cut can one find forms relating to the type of legend, observing a contrast which clearly emerges between the 3-line framed legend (*I_3_ks*), on the negative side, and the 2-line framed legend (*I_2_ks*), on the positive one. This could also have chronological reasons but is, by itself, not clear evidence. More expressive is the presence of the forms *bb1* (bearded character) on the negative side and *at_l2* (moon crescent with sun-disk – negative side), *at_l* (moon crescent), *inb* (integrating element placed in the lower part of the field) and *s_sp* (integrating elements placed one above the other in the field), these latter three relate to the positive part of the axis. This means that in the 85% analysis a gender opposition, probably to be referred to receiving characters, emerges between the two parts of the axis (*bb1*, on the negative side, versus *D* and – possibly – *una*, on the positive one).

An actual distinction of the tendencies represented by the two parts of the axis can be observed in the analyses on 90% and 95% of the corpus of forms. Compared to them, the investigation of the full corpus gives details about some specific elements and features shown in the scenes, but no further information on how the possible structures or compositions of the scenes distinguish from one another.

The 90% analysis outlines many basic traits of the scenes which can be connected to the two sides of the first axis fairly well. On the negative side, there are elements which clearly point at scenes of the "royal presentation" type (i.e., showing a royal male figure in the role of the receiver) – *tr_H4*, the typical padded stool, and *ti_cl*, the skull cap – and presentation before a male god (*G*, god; *ac_b1*, hairstyle – type b1), in addition to elements which are kept after the larger cuts, and which, in general, can be both related to scenes known from original impressions and scenes with a relatively late date. Examples of this are the forms *ti_c5* (headgear with multiple pairs of horns), *g3* (multiple necklace), and *at_l2* (sun disc with moon crescent in the field). On the positive side of the axis the tendency is, in contrast to earlier features, like *tr_Q1b* (square seat with a frame), *at_l* (moon crescent in the field), *v_a_p* (pleated robe), a concentration of integrating elements in the field (suggested by forms like *inh* and *inb*, respectively elements placed in the upper and in the lower part of the field, which remain in the successive cuts, and *inx*, elements placed in the middle of the field), which can be a symptom of secondary cut works on the scenes, besides the presence of female receivers, which continues to be clear following cuts.

In the 95% analysis this framework is emphasised by further elements which express more precisely the tendencies of the two sides of the axis, giving

more detailed traits about the main iconographic features which compose the scenes of different types. In the 100% analysis the extension of the first axis is enlarged, especially on its negative side, due to the wide number of rare forms, which find peripheral positions in the graphic. In the following we shall discuss the results of TCA applied to these two tables, also considering secondary factors, which seem in both cases to carry interesting information.

4.3 *TCA of the first two tables*

In Fig. 4 both periods and toponyms and the most relevant forms are represented on the planes spanned by the first two TCAs, run to the whole corpus without *hapax* and on the table limited to 95% of the forms. Observing them side by side, the relevant reduction in the number of forms is evident.

In the first analysis (100% of the forms), both the first two factors may be considered meaningful and representative of the corpus (see the figures in Table 3). Concerning the formation of the first factor, one finds, on the positive side, elements that may refer to scenes carved on seals from the core region of Southern Mesopotamia and possibly dating to different periods: the short legend containing names or dedicatory formulas (*I_ln*, *I_ks*, *kt_b*, *kt_d*, *kt_d2*, *kt_a2*), the standing lion (*bv_u1b*), the bull (*an_r2*, *an_r2b*), the

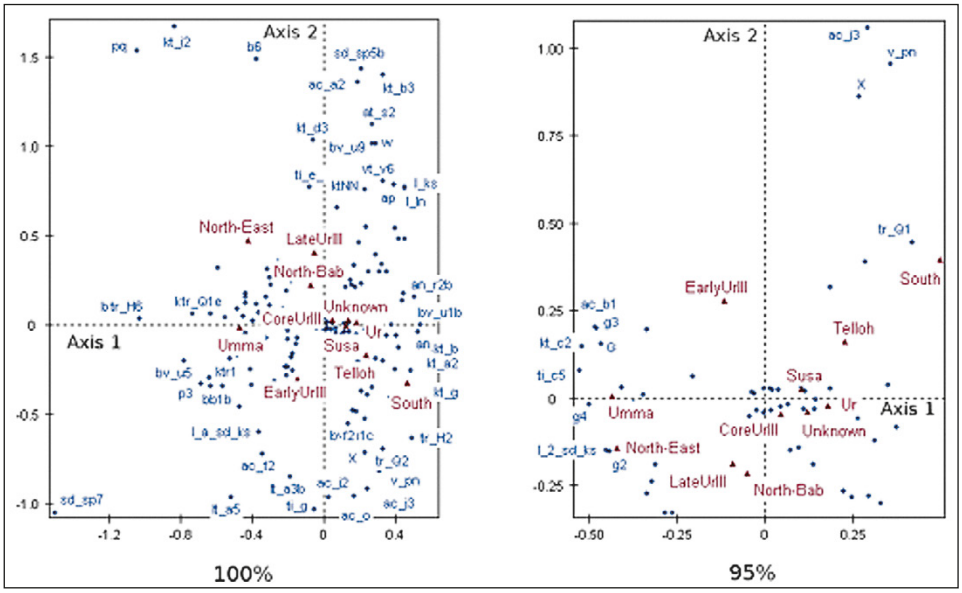


Fig. 4 – TCA of the forms describing the scenes of the corpus with respect to periods and toponyms. Representation of the items on the first factor plane of the whole corpus without *hapax* (left) and of the table reduced to 95% of the forms.

simple niched throne (*tr_H2*) and the rectangular one (*tr_Q1*, also with the net of lines on its side, *tr_Q2*), the duck-throne (*tr_HB*), the pine-like tree (*vt_v6*), the "ball-and-staff" (*ap*), the scorpion (*rt_u4b*), the bird (*an_u6*), the dwarf (*w*), the monkey (*u9*), the plain robes (*v_pn*). On the negative side, a contribution to the structuring of the same factor is given by a number of elements together with longer and more complex legends. An important role is played by elements like the standard topped by a bull (*sd_sp7*), the presence of a smaller deity in the field (*pq*), the throne with a high backrest (*tr_H6*), the lion (sitting, *bv_u5c*, or striding, *bv_u5*), the legends including formulas related to the royal person and titles (*kt_i2*, *ktr4*, *ktr1*, *ktr2*), the throne with the striding lion (*tr_Q1e*), the altar with the palm tree (*lt_a5*), the weapons appearing on the shoulders of the sitting deity (*e3*), the "rod-and-ring" (*b6*), etc. It seems that this part of the first factor is in general related to official sealings or scenes which have been poorly or never reworked.

The second factor seems more connected with chronology. On the positive side, one finds features that date to relatively late times, like the abundance of integrating elements, and especially the presence of standard with griffins (*sd_sp5b*), the "rod-and-ring" (*b6*), the star (*at_s2*), the monkey (*bv_u9*), the dwarf (*w*), the "ball-and-staff" (*ap*), the pine-like tree (*vt_v6*). Further iconographic elements related to the positive part of factor 2 are the small goddess (*pq*), long legends mentioning the seal as a gift from the king or short ones with divine names (*kt_i2*, *kt_b3*, *kt_d3*), and physical features, such as the striped simple hairstyle or the striped skull-cap without brim (*ac_a2*, *ti_e_*). Contributing to the negative part of Factor 2 are much earlier elements and those which are not unusual in original impressions, like the standard topped by the bull (*sd_sp7*), the table with offerings or the altar with the date palm (*lt_a3b*, *lt_a5*), the lizard, the lion-griffin (*rt_u4d*, *bv_u1c*), variations of the square throne (*tr_Q1*, *tr_Q2*) and the niched one (*tr_H2*), but also few attributes of the characters, like the weapons on the shoulders of the goddess or the triangular cup (*e3*, *f2*), and a great variety of hairstyles (*ac_j2*, *ac_j3*, *ac_o*, *ac_f2*, all female ones, and relatively early in date) and plain robes (*v_pn*, *m_piano*).

The second analysis (95% of the forms) shows a first factor which is still strong and meaningful, but a much less significant second one, almost useless for interpretation purposes. Meaningful forms for developing the structure of the first factor are now much less expressive of large parts of the content of the scenes, but they can indirectly suggest the presence of further meaningful elements. Forms related to the positive side of the first factor express scenes with short legends, basically mentioning name, patronymic or profession (*kt_d*, *I_2_ks*), the receiver sitting on a relatively simple throne (*tr_Q1*), and the possible presence of female human figures and few integrating elements, among which the moon crescent (*at_l*) and the scorpion (*rt_u4b*) or a bird (*an_u6*). These are all symptoms typical of rather basic compositions, quite

frequent in cylinders from the South that have not been re-carved. The opposite side of the same factor is clearly developed through the forms concerning scenes more largely used in official administration, such as the so-called "royal presentations". The legends are large and complex (*kt_c2*, *I_2_sd_ks*) and the presence of divine or human male characters (including royal ones) in the scene is especially frequent: they are suggested by physical or object attributes like, for instance, some hairstyles (*ac_b1*), the beard (*bb1*) or jewels (*g1*, *g2*, *g3*, *g4*).

Few forms taking part to the formation of the second factor emerge. On the positive side, an essential role is played by the hairstyle with a lock of hair on the top-back of the head (*ac_j3*), the plain robe (*v_pn*), the presence of a woman in the scene (*X*), and the simple square throne (*tr_Q1*). On the negative side of factor 2, one may just distinguish a very weak contribution of few forms, namely those related to the cup (as attribute of a character holding it: *f6*), to the throne with footstool (*pod_1+1b*), and to the goose-like bird and the scorpion as integrating elements (*an_u6*, *rt_u4b*).

What emerges from the third analysis (90% of the forms) is a similar picture, but the few forms producing the structure of the first factor are altogether not enough meaningful to suggest a proper interpretation of the factor itself. Their distribution in the graphic may be instead interpreted with some precision only through the results of the former analyses. In the following analyses (85% to 50%) the situation is much worse, as one would expect.

It seems, in general, that the first factor mainly explains the geographic differences, while the second one is related to features depending on diachronic variations: the latter changes are not as strong as the previous ones, limited as they are to a few rare forms, which are the first to be withdrawn. This interpretation could at least be valid for the first analysis, while the results of the second one (on 95% of the forms) show an overly weak factor 2. The loss of information owing to the first cut of the forms affected its structure so much that the remaining forms can no longer explain its main features and inner structure. With regard to the explicit contribution given by the forms, one can state that the data sets resulting from subsequent cuts of the 95% corpus cannot be used, and even this cut must be considered with a certain prudence.

4.4 *Typical forms and compositions*

The characteristic forms associated with the three periods are very few after the first cut, and their number is drastically reduced after the third: this is consistent with the weakness of the second factor of TCA, and its loss of significance due to the reduction of forms.

1) For the *Early Ur III* phase, present forms concern the lock of hair on the top-back of the head (*ac_j3*) and other quite unusual (for the main part of the Ur III period's iconography) hairstyles (*ac_o*, *ac_j2*, *ac_f2*); the plain robe

(*v_pn*); the skirt (*gn*); a female human character; elements like the standard with a bull (*sd_sp7*) or an altar (*lt_a5*, *lt_a3b*). On the opposite, the goddess' hairstyle with double curl behind the neck (*ac_n1*), the cup as attribute of a character (*f6*), the moon crescent and the sun disc inscribed in the moon crescent (*at_l*, *at_l2*), the goose-like animal (*an_u6*), and the square seat with a frame (*tr_Q1b*) are typically absent.

2) The *Core Ur III* period shows some of those absent in the former phase, like *ac_n1*, *tr_Q1b*, and *ac_n1* as typically present forms, but also those related to the pleated robe (*v_a_p*, frequently worn by standing goddesses) and the character with a clean-shaven head (*a_c*). Absent forms are here *ac_j3* (lock of hair on the top-back of the head), *gn* (skirt), *X* (female human character), *v_pn* (plain robe), *sd_sp7* (standard with a bull), *ac_j2* (a lock of hair on the head), *ac_a2* (simple striped hairstyle), *ac_o* (braid), *lt_a5* (altar).

3) In the scenes of the *Late Ur III* period one finds the monkey (*bv_u9*), the standard topped by griffins (*sd_sp5b*), the asterisk (*at_s2*), the seats of the types square with frame and vertical support and padded stool (*tr_Q8e*, *tr_H4*), the legend containing a divine name (*kt_b3*) or a royal dedication (*kt_i2*, *ktr2*), the dwarf (*w*). Typically absent are *v_pn*, *v_a_p*, *D* (presence of a goddess in the scene), *tr_Q1* (simple square seat), and *kt_d* (legend containing the formula "PN1, son of PN2").

Fig. 5 shows the typical scenes of the three periods, identified according to their proximity to the group centroid in the analysis. All specimens are based on the compositional scheme in which the couple of characters hand in hand is represented before a third sitting figure. The most central scene of the *Early Ur III* period is a basic presentation with three female characters in plain robes and few attributes, while that of the *Core Ur III* period shows a presentation before a sitting goddess, with attributes and robes that are more varied. In the third scene (*Late Ur III* period), the composition is more similar to that of the *Core Ur III*, but the receiver has "royal features" and the field contains many integrating motifs, some of which take the place of the legend: these are all symptomatic of secondary engraving on the seals.

The outcomes of the analysis on the forms which are most typically associated with toponyms are not homogeneous from the quantitative point of view. In Table 4 the number of significant forms for each toponym is reported, considering both positive (highest frequencies) and negative (lowest frequencies) ones, for each table's reduction. After the first cuts the decrease in the number of forms playing an important role in defining external features is evident, while the number of toponyms lacking characteristic forms increases. More specifically, after the first and third cuts there are general and significant reductions in the number of forms which could be useful to describe distinctive elements of the scenes, while after the second cut – and

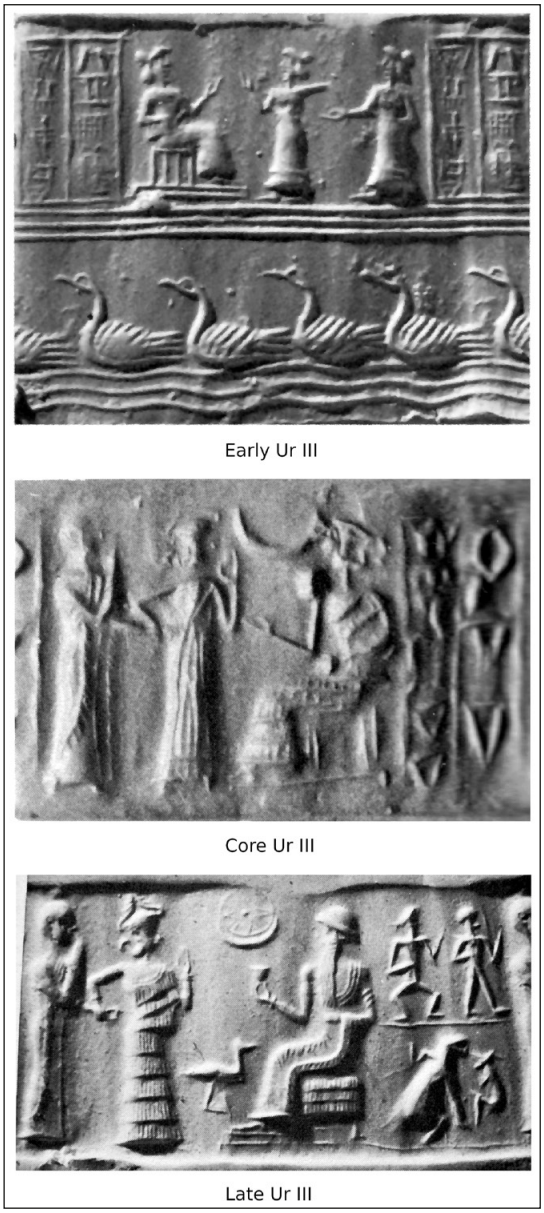


Fig. 5 – The most representative scene of each period, according to the results of the first analysis (100% of the whole corpus): *Early Ur III* (Buchanan 1981, n. 561); *Core Ur III* (Collon 1982, n. 368); *Late Ur III* (Buchanan 1981, n. 629).

Toponym	*100*		*95*		*90*		*85*		*80*		*75*		*70*		*60*		*50*	
	pos	*neg*	*pos*	*neg*	*pos*	*neg*	*pos*	*neg*	*pos*	*neg*	*pos*	*neg*	*pos*	*neg*	*pos*	*neg*	*pos*	*neg*
Unknown	13	11	9	10	8	9	6	6	3	4	2	2	2	1	2	1	2	1
Ur	10	9	4	9	2	7	1	4	0	3	0	2	0	1	1	1	0	1
North-Bab	3	2	2	2	2	1	0	1	0	0	0	0	0	0	0	0	0	0
South	4	2	4	2	1	2	0	2	0	2	0	2	0	1	0	1	0	1
Telloh	12	13	9	13	3	10	2	5	2	3	2	2	2	1	2	1	2	1
North-East	8	1	4	1	3	1	4	1	3	1	3	1	2	1	2	0	1	0
Susa	11	6	4	5	2	3	2	2	0	2	0	1	0	1	0	1	0	1
Umma	29	26	19	24	13	16	8	12	5	9	4	5	3	4	3	4	3	4

Tab. 4 – Amount of characteristic forms of the toponyms, both positive and negative, resulting from the progressive cuts.

even more so after the third – such meaningful forms are drastically reduced or disappear for many toponyms.

In the following, the most significant typical forms, both positive and negative, are reported for each of the seven known toponyms, together with the description of the most typical seal: the selection does not have any significance for the *Unknown* toponym and is not reported. The forms are reported in order of relative significance (the most significant are the first):

1) *Ur.* Typically present in *Ur* are many forms, like plain robe, legend containing a divine name, one line-legend, headgear with multiple pairs of horns, presence of a female character, square seat with a frame, single-case legend, and legend containing a personal name. On the opposite, the absent forms are skull cap, padded stool, double bracelet, throne with footstool, 3-cases legend, legend of the kind "name, profession, son of [other name]", presence of a male character, and multiple necklaces.

2) *North-Bab*. In *North-Bab* little amphora, skull cap, and padded stool are present. Absent forms are headgear with multiple pairs of horns, and 3-cases legend.

3) *South*. Here typically present are plain robe, square seat with a frame, woman, and footstool before the throne, whereas typically absent are presence of a male character and presence of a character with concrete attributes.

4) *Telloh*. Many forms are typical of *Telloh*, specifically those present are: plain robe, headgear with one pair of horns, square seat, female character, simple hairstyle, legend mentioning name and profession, niched throne, dragon-like heads sprouting from the shoulders, and bird of prey with spread wings. On the other hand, typically absent are: square seat with a frame, female character at the end, moon sickle with sun disk, beard, skull cap, cup-attribute, bracelet, padded stool, male character, and headgear with multiple pairs of horns.

5) *North-East*. Here, the significantly present forms are goddess of small dimensions, presence of a character with concrete attributes, striped headgear,

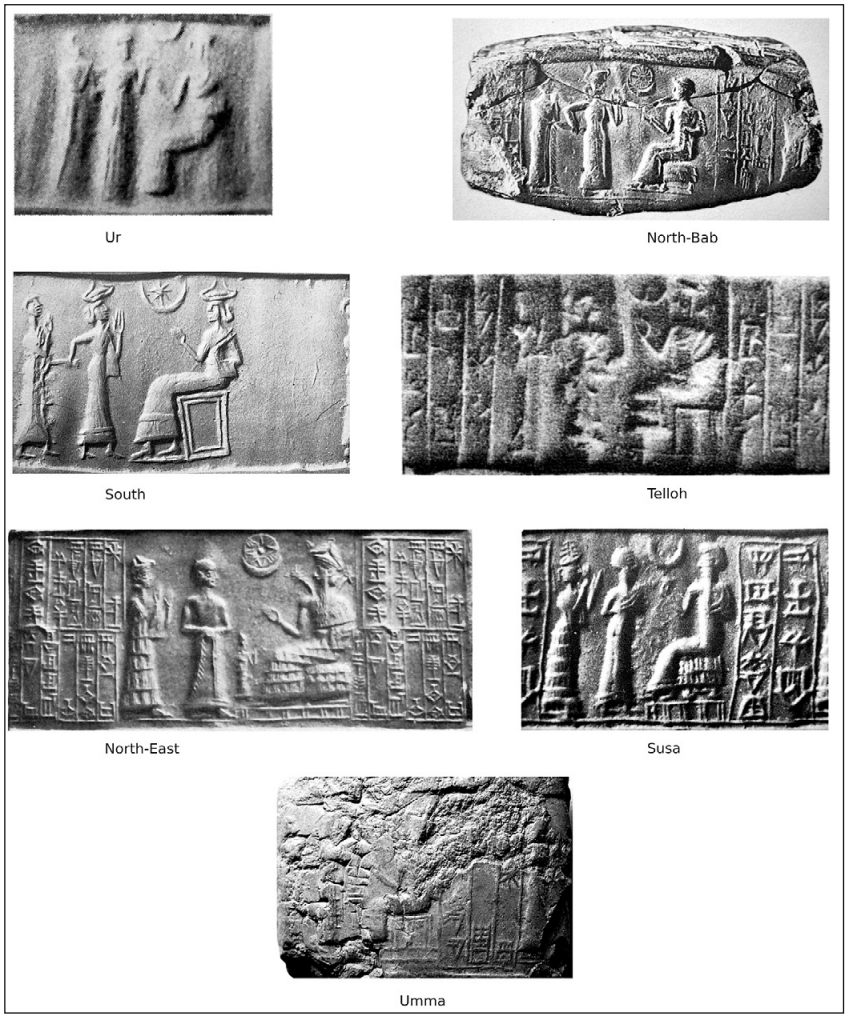


Fig. 6 – The most representative scene of each toponym, according to the results of the first analysis (100% of the whole corpus): *Ur* (Legrain 1951, n. 360); *North-Bab* (Legrain 1925, n. 287); *South* (Kjaerum 1983, n. 369); *Telloh* (Parrot 1954, n. 133); North-East (Frankfort 1955, n. 709); *Susa* (Amiet 1972, n. 1693); *Umma* (Ybc 1630).

moon sickle with sun disk, 4-cases legend, necklace, and headgear with multiple pairs of horns, whereas the only significantly absent form is the couple of characters hand in hand.

6) *Susa*. Forms significantly present in *Susa*: fringed robe, striped headgear, flat headgear, simple hairstyle, 2-cases legend, hairstyle with double curl, legend

of the kind "servant of"; significantly absent are: male hairstyle with hair lock behind the head, bracelet, 2-cases legend, and necklace.

7) *Umma*. At *Umma* many forms are significant. Meaningfully present are: headgear with multiple pairs of horns, necklace, legend of the kind "name, profession, son of [other name]", beard, male hairstyle with hair lock behind the head, god, moon sickle with sun disk, 3-cases legend, legend made of 2 rows of cases, bracelet, legend of the kind with royal titles and "his servant" formula, standard with a bull, skull cap, padded stool, sitting lion, long axe, square seat with a vertical support, throne with footstool, cup-attribute, and seat with striding lion. Many forms are also seen to be absent: couple of characters hand in hand, headgear with pairs of horns, presence of elements in the field, pleated robe, character bringing an arm forth, fringed robe, female hairstyle with a double curl behind the neck, presence of a male character, square seat, hairstyle with a lock of hair on the top-back of the head, scorpion in the field, plain robe, goose-like bird, moon sickle, and 2-cases legend.

The most typical scenes for each toponym, according to the significance of the contained features, are illustrated in Fig. 6. Looking at it, most central specimens of *Ur, Telloh*, and *South* appear quite similar to one another: all of them show presentations with the man and the goddess hand in hand before a sitting female deity. In all cases there is an astral symbol placed in the upper part of the field, just before the receiver, and a quite large space designed to host the legend. Many features of the typical scene for *South* (more specifically, this seal was excavated on Failaka island), like the plain robes, the seat and physical traits of the characters, distinguish it from the other two scenes.

The case of the three most central scenes of *North-Bab, North-East*, and *Susa* is similar: all of them belong to the category of the royal presentation scenes, but meaningful differences may be located here as well. The scene of *North-Bab* is known from an original seal impression and shows the couple hand in hand before a sitting male figure with the typical attributes and clothes of the ruler. This scene may be logically interpreted as an interface between the traditional presentations before a goddess and the most typically formalised administrative version of the royal presentation. Very different is the most central specimen of *North-East*, which is a well-known reworked royal presentation scene, with a large legend and a receiver changed into a god having unusual features, but still maintaining some attributes of the ruler's iconography. Furthermore, in the reworking, a miniaturised goddess has been introduced before the receiver. The scene from *Susa* has typical traits of the royal presentation, although it is clearly an unusual version.

The most central example of *Umma* has also been recorded using an original impression, and it stands out for representing a presentation before a god. This is not unusual, but relatively infrequent.

5. Conclusion

Textual analysis applied to the description of iconographic corpora stimulates detailed studies for its best application. The Procrustean analysis proved successful in showing the influence of the reduction of forms on the interpretability of the results. The corpus of glyptic compositions investigated in this research was produced in a short span of time and may be attributed to a very centralised political and administrative machinery, that of the Ur III state. Besides this, it has the peculiarity of being quite homogeneous, since all specimens share many general and basic structural features.

Indeed, the study shows very few differences attributable to the engraving period. This explains the weakness of the second factor and its sudden loss of interpretability after reducing the forms, whereas the distinction between sites seems more consistent, thus justifying the more robust meaning of the first factor, which is less susceptible to the reduction of forms reduction. Nevertheless, the huge loss of information after the first cut is evident, and those which follow progressively cause an impoverishment of meaningful forms, which prevents proper interpretation. This appears even worse when considering characteristic forms. Their dramatic reduction owing to the cuts makes a clear distinction between different sealing groups impossible. Thus, in this case, a relevant reduction of forms prior to analysis is not advisable, whereas studies concerning more heterogeneous corpora may lead to different conclusions.

Alessandro Di Ludovico
Dipartimento di Scienze dell'Antichità
Sapienza Università di Roma
alediludo@gmail.com

Sergio Camiz
Istituto di Scienze del Patrimonio Culturale – CNR
sergio@camiz.net

Acknowledgements

We are most grateful to Paola Moscati and Elena Rova for their precious encouragement, suggestions, and help in the carrying out of this work.

References

Benzécri J.P. 1973-1982, *L'analyse des données*, 2, Paris, Dunod.

Borg I., Groenen P.J. 2005, *Modern Multidimensional Scaling: Theory and Applications*, New York, Springer.

Buchanan B. 1981, *Early Near Eastern Seals in the Yale Babylonian Collection*, New Haven-London, Yale University Press.

Camiz S. 2004, *On the coding of archaeological data*, «Archeologia e Calcolatori», 15, 201-218.

Camiz S., Denimal J.J. 2011, *Procrustes analysis and stock markets*, «Case Studies in Business, Industry and Government Statistics», 4, 2, 93-100.

Camiz S., Rova E. 1991, *Analysis of ancient Near-Eastern cylinder seals (late Fourth Millennium B.C.)*, in E. Diday, Y. Lechevallier (eds.), *Symbolic-Numeric Data Analysis and Learning*, New York, Nova Science, 131-142.

Camiz S., Rova E. 1996, *Metodi di analisi per lo studio di un gruppo di sigilli cilindrici vicino-orientali e di altre immagini strutturate*, «Archeologia e Calcolatori», 7, 647-659.

Camiz S., Rova E. 2001, *Exploratory analyses of structured images: A test on different coding procedures and analysis methods*, «Archeologia e Calcolatori», 12, 7-46.

Camiz S., Rova E. 2003, *Quantitative study of images in archaeology: I. Textual coding*, in M. Schader, W. Gaul, M. Vichi (eds.), *Between Data Science and Applied Data Analysis*, Berlin, Springer, 624-632.

Cibois P. 1997, *Les pièges de l'analyse des correspondances*, «Histoire et Mesure», 12, 3/4, 299-320.

Collon D. 1982, *Catalogue of the Western Asiatic Seals in the British Museum. Cylinder Seals 2, Akkadian – Post-Akkadian – Ur III Periods*, London, The British Museum.

Davis L.G., Bean D.W., Nyers A.J., Brauner D.R. 2015, *GLIMR: A GIS-based method for the geometric morphometric analysis of artifacts*, «Lithic Technology», 40, 3, 199-217.

Di Ludovico A. 2005, *Scene-in-frammenti: una proposta di analisi delle "scene di presentazione" dei sigilli a cilindro mesopotamici orientata all'elaborazione statistica ed informatica dei dati*, in A. Di Ludovico, D. Nadali (eds.), *Studi in onore di Paolo Matthiae presentati in occasione del suo sessantacinquesimo compleanno*, Contributi e Materiali di Archeologia Orientale, 10, Special Issue, Roma, Sapienza, 57-95.

Di Ludovico A. 2008, *Between Akkad and Ur III: Observations on a "short century" from the point of view of glyptic*, in H. Kühne, R.M. Czichon, F.J. Kreppner (eds.), *Proceedings of the 4th International Congress on the Archaeology of the Ancient Near East (Berlin 2004)*, 1, Wiesbaden, Harrassowitz, 321-341.

Di Ludovico A. 2010, *La glittica della fine del Terzo Millennio come strumento di controllo e di consolidamento del potere in Mesopotamia*, in R. Dolce (ed.), *Quale Oriente? Omaggio a un Maestro. Studi di Arte e di Archeologia del Vicino Oriente in memoria di A. Moortgat a trenta anni dalla sua morte*, Palermo, Flaccovio, 241-261.

Di Ludovico A. 2011, *Experimental approaches to glyptic art using artificial neural networks. An investigation into the UR III iconological context*, in E. Jerem, F. Redő, V. Szeverényi (eds.), *On the Road to Reconstructing the Past. Proceedings of the 36th International Conference on Computer Applications and Quantitative Methods in Archaeology (Budapest 2008)*, Budapest, Archaeolingua, 135-146.

Di Ludovico A. 2012, *The uses of the cylinder seal as clues of mental structuring processes inside Ur III state machinery*, in G. Wilhelm (ed.), *Organization, Representation, and Symbols of Power in the Ancient Near East. Proceedings of the 54th Rencontre Assyriologique Internationale (Würzburg 2008)*, Winona Lake, In, Eisenbrauns, 275-289.

Di Ludovico A. 2013, *Symbols and bureaucratic performances in Ur III administrative sphere. An interpretation through data mining*, in S. Garfinkle, M. Molina (eds.), *From the 21st Century B.C. to the 21st Century A.D. Proceedings of the International Conference on Sumerian Studies (Madrid 2010)*, Winona Lake, In, Eisenbrauns, 125-151.

Di Ludovico A., Camiz S. 2014, *A quantitative approach to UR III Mesopotamian figurative languages: Reflections, results, and new proposals*, «Archeologia e Calcolatori», 25, 7-32.

Di Ludovico A., Camiz S. 2015a, *Art history of the ancient Near East and mathematical models. An overview*, in F. Giligny, F. Djindjian, L. Costa, P. Moscati, S. Robert (eds.), *21st Century Archaeology. Concepts, Methods and Tools. Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology (Paris 2014)*, Oxford, Archaeopress, 29-34.

Di Ludovico A., Camiz S. 2015b, *Ancient Mesopotamian glyptic products, statistics and*

data mining: A research proposal, in F. Giligny, F. Djindjian, L. Costa, P. Moscati, S. Robert (eds.), *21st Century Archaeology. Concepts, Methods and Tools. Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology (Paris 2014)*, Oxford, Archaeopress, 489-496.

Di Ludovico A., Camiz S., Pieri G. 2013, *Comparative use of mathematical models in an investigation on Mesopotamian cylinder seals*, in F. Contreras, M. Farjas, F.J. Melero (eds.), *Fusion of Cultures. Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology (Granada 2010)*, Oxford, BAR International Series 2494, Archaeopress, 495-498.

Di Ludovico A., Pieri G. 2011a, *How to facilitate interpretation of natural computation results by converting binary codes of images back to images*, «CiiT International Journal of Artificial Intelligent Systems and Machine Learning», 3, 7, 437-446.

Di Ludovico A., Pieri G. 2011b, *Artificial neural networks and ancient artefacts: Justifications for a multiform integrated approach using PST and Auto-CM models*, «Archeologia e Calcolatori», 22, 99-128.

Di Ludovico A., Ramazzotti M. 2008, *Reconstructing lexicography in glyptic art: Structural relations between the Akkadian age and the Ur III period*, in R.D. Biggs, J. Myers, M. Roth (eds.), *Proceedings of the 51st Rencontre Assyriologique Internationale (Chicago 2005)*, Studies in Ancient Oriental Civilization 62, Chicago Oriental Institute, 263-280.

Dryden I. 2000, *Statistical Shape Analysis in Archaeology*, Paper presented at the Workshop Spatial Statistics in Archaeology (Chieti 2000) (https://www.researchgate.net/publication/247644994_Statistical_shape_analysis_in_archaeology).

Dryden I., Mardia K. 2006, *Statistical Shape Analysis with Applications in R*, New York, John Wiley.

Eckart C., Young G. 1936, *The approximation of one matrix by another of lower rank*, «Psychometrika», 1, 3, 211-218.

Frankfort H. 1955, *Stratified Cylinder Seals from the Diyala Region*, Oriental Institute Publications 72, Chicago, Oriental Institute.

Gardin J.-C. 1958, *Four codes for the description of artifacts: An essay in archeological technique and theory*, «American Anthropologist», 60, 335-357.

Gittins R. 1985, *Canonical Analysis*, Berlin, Springer.

Golub G.H., Van Loan C.F. 1996, *Matrix Computations*, 3rd ed., Baltimore, Johns Hopkins University Press.

Gower J.C. 1975, *Generalized Procrustes analysis*, «Psychometrika», 40, 1, 33-51.

Gower J.C., Dijksterhuis G.B. 2004, *Procrustes Problems*, Oxford, Oxford University Press.

Greenacre M.J. 1983, *Theory and Application of Correspondence Analysis*, London, Academic Press.

Greenacre M.J. 2006, *Tying up the loose ends in simple, multiple and joint correspondence analysis*, in A. Rizzi, M. Vichi (eds.), *Compstat 2006. Proceedings in Computational Statistics*, *17th Symposium (Rome 2006)*, Wien, Physica, 163-185.

Hurley J.R., Cattell R.B. 1962, *The Procrustes Program: Producing direct rotation to test a hypothesized factor structure*, «Computers in Behavioral Sciences», 7, 258-262.

Jolliffe I.T. 2002, *Principal Component Analysis*, 2nd ed., Berlin, Springer.

Kendall D.G. 1989, *A survey of the statistical theory of shape*, «Statistical Science», 4, 2, 87-99.

Kjaerum P. 1983, *Failaka/Dilmun. The Second Millennium Settlements*. 1,1. *The Stamp and Cylinder Seals*, Jutland Archaeological Society Publications 17, 1, Aarhus, Jysk Arkæologisk Selskab.

Lê S., Josse J., Husson F. 2008, *FactoMineR: An R package for multivariate analysis*, «Journal of Statistical Software», 25, 1, 1-18.

Lebart L., Morineau A., Lambert T., Pleuvret P. 1999, *SPAD Système pour l'analyse des données*, version 5.5, Paris, Cisia-Ceresta.

Lebart L., Piron M., Morineau A. 2006, *Statistique exploratoire multidimensionnelle: visualisations et inférences en fouille de données*, Paris, Dunod.

Lebart L., Salem A. 1988, *Analyse statistique des données textuelles: questions ouvertes et lexicométrie*, Paris, Dunod.

Legendre L., Legendre P. 2012, *Numerical Ecology*, 3rd ed., Amsterdam, Elsevier.

Legendre P. 2001, *Ecologically meaningful transformations for ordination of species data*, «Oecologia», 129, 271-280.

Legrain L. 1925, *The Culture of the Babylonians from their Seals in the Collections of the Museum*, Publications of the Babylonian Section 14, Philadelphia, University Museum.

Legrain L. 1951, *Ur Excavations X. Seal Cylinders*, Philadelphia, University Museum.

Macleod N. 2009, *Who is Procrustes and what has he done with my data*, «Palaeontological Association Newsletter», 70, 21-36.

Malinvaud E. 1987, *Data Analysis in Applied Socio-Economic Statistics with Special Consideration of Correspondence Analysis*, in *Marketing Science Conference*, Jouy en Josas (France), HEC-ISA.

Mosier C.I. 1939, *Determine a simple structure when loadings for certain tests are known*, «Psychometrika», 4, 2, 149-162.

Parrot A. 1954, *Glyptique mésopotamienne. Fouilles de Lagash (Tello) et de Larsa (Senkereh) (1931-1933)*, Paris, Paul Geuthner.

R Core Team 2019, *R: A language and environment for statistical computing*, Vienna, R Foundation for Statistical Computing (https://www.R-project.org/).

Rao C.R. 1995, *A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance*, «Qüestiió», 19, 23-63.

Tal A. 2014, *3D shape analysis for archaeology*, in M. Ioannides, E. Quak (eds.), *3D Research Challenges in Cultural Heritage*, Berlin, Springer, 50-63.

Winter I.J. 1987, *Legitimation of authority through image and legend: Seals belonging to officials in the administrative bureaucracy of the Ur III State*, in McG. Gibson, R.D. Biggs (eds.), *The Organization of Power: Aspects of Bureaucracy in the Ancient Near East*, Studies in Ancient Oriental Civilization 46, Chicago, The University of Chicago Press.

ABSTRACT

The use of Textual Correspondence Analysis to investigate a corpus of iconographic compositions carved on Mesopotamian cylinder seals proved to be very useful to understand the peculiarities of the specimens from the points of view of geographic origin, typology and inner chronology. The presence of a relatively high number of rare forms in the data set – besides the hapax – led to think, however, that the extraction of the factors – so the outcomes of the analyses – could have been influenced heavily by them. For this reason, looking for an optimal composition and for the most effective encoding of the data set, a reduction of its rarest forms was performed to find the threshold which could allow to reconcile the need for keeping the useful encoded information with the best possible reduction of elements producing high inertia. Adopting the methodology known as Procrustes, the data table was thus progressively reduced, and each time investigated: the results so obtained were then used to reach a global assessment about how much each reduced data set could correspond to the optimized one.