Archeologia e Calcolatori
33.2, 2022, 13-32
doi 10.19282/ac.33.2.2022.02

EXPERIMENTING GENERALIZED PROCRUSTES ANALYSIS ON A CORPUS OF SCENES CARVED ON MESOPOTAMIAN CYLINDER SEALS AND ENCODED IN A TEXTUAL DATA SET

1. Introduction

In a recent paper, the Procrustes Analysis was carried out on the results of Correspondence Analyses (CA) applied on presentation scenes carved on late third and early second millennium BCE cylinder seals. In particular, the influence of rare lexical forms to the extracted factors (Di Ludovico, Camiz 2020) was investigated, by progressively withdrawing them and comparing the resulting CAs. This was a further attempt to improve the study of the iconographical structure of images through a formalized textual coding (Camiz, Rova 1996, 2001; Camiz 2004). The idea of using textual coding in iconographic studies is relevant, since a text may describe an image in a much better detail than the usual coding of iconographic elements and their attributes: this, on condition that all coding ambiguities be avoided.

Indeed, the classical coding may not take into account the relations among elements, unless explicitly coded aside. The method has advantages with respect to the pattern recognition too, because uncontrolled differences due only to the drawing/carving style are prevented. In addition, from this coding the presence/absence of selected items may be easily identified, by simply extracting the corresponding keywords: e.g., a different animal, the different positions of an individual's arms, or a kind of seat, as well as extended to include recurring iconographic substructures, such as *goddess sitting on a bench with her right arm held forward and her left one kept at her waist*. It is our opinion that textual coding, with its ability to easily be transformed, might be taken as a reference for the coding of archaeological findings (Camiz 2004).

Still very current and helpful are the pioneering methodological observations and critical thoughts developed by Gardin (1958, 1966, 1967; see also Di Ludovico, Camiz 2015a, 29-30) about the textual coding for the quantitative investigation of archaeological artefacts. Yet, the method's use is difficult, mainly depending on the very high number of forms that usually are arranged in a sparse contingency table: this gets critical the statistical tests for significance, hence the correct dimensions to take into account: moreover, the interpretation of the extracted factors may be difficult, sometimes only outlining groups of forms and corresponding groups of items where such forms appear. Indeed, as the method belongs to the exploratory multidimensional analysis framework, it must be considered an investigation tool rather than a statistical method to identify once and for all a data table characteristics: therefore, these drawbacks are not so

relevant. Nevertheless, new interpretation aids may ease the researcher's work. It is with this purpose that we present here the results of an exploration carried out on the same corpus and the same methods of Di Ludovico and Camiz (2020). This time, instead of the forms, Generalized Procrustes Analysis (GPA) is used to select the *repeated segments* extracted from the corpus.

2. Textual Correspondence Analysis and Procrustes Analysis

Textual Correspondence Analysis (TCA: Lebart, Salem 1994) is the basic exploratory tool for textual analysis, since it may show the pattern of the studied objects, corresponding to the texts, pattern based on the structure of association of the textual forms, as revealed in hierarchical order of relevance. Associated to TCA is often a classification of the objects, based either directly on the texts or on the factors deemed relevant by the study. The interpretation of the classes thus formed is subsequently developed, usually by identifying the *characteristic forms* of the classes, i.e., the lexical forms whose frequencies in a class are significantly higher or lower with respect to those in the whole corpus. Indeed, they may address to particular features common to the majority of the objects of the class. Therefore, such a procedure – also called *tandem analysis* – is not particularly different from the same applied to other kind of data, unless for the statistics involved for the characteristic forms.

On the opposite, it is relevant the availability of computer programs able to build the basic data table, necessary to deal with all the manipulations needed prior the analysis, which require specific tools to be handled properly. Among them, the possibility of collecting the so-called *repeated segments* raised our interest, given the connection with some iconographic structures, that are repeated within the corpus at hand. Indeed, to describe them properly, a formal sentence, i.e., a sequence of forms, is needed, which is supposed to repeat identically throughout the corpus. This might correspond to some repeated segments, which are built automatically by a computer program by aggregating progressively adjacent forms, until the resulting segments are repeated or – if so instructed – a punctuation sign occurs. This process issues a very large number of repeated segments, only some of them having a complete meaning – i.e., fully describing an iconographic substructure – whereas most may suggest some kind of common substructure developing toward progressively enlarging ones.

While the exploration of a text through the forms is cumbersome, given the high number of forms involved in a corpus, even worse is its study through repeated segments: first, because they are much more and fill quickly the graphical space of representation; second, because one must follow the progressive extension with new forms sometimes in different directions; third, because the contingency table built on them is even more sparse than that built on the forms and may not carry statistically relevant information. The use of repeated

segments in archaeology was already attempted on a different corpus of seals images (Camiz, Rova 2001) and the interpretation problems rose quickly. Eventually an attempt to deal with *quasi-segments* (carried out with Monique Bécue, unpublished), i.e., repeated segments with one or two different inner forms, did not succeed: in fact, a too high number of items resulted to deal with a very scarce significance of the built table. Therefore, we believe that a study addressing the selection of both the forms and the repeated segments to submit to TCA, once solved, could ease studies otherwise too difficult to face.

The Procrustes Analysis (PA: Gower 1975; Gower, Dijksterhuis 2004) aims to identify the best linear transformation of a data set into another, based on the same units, minimizing the sum of squared distances between the units in the two representations. Associated statistics measure the quality of the adjustment. Applied to analyses of factor kind, it allows a direct comparison of the clouds of points represented in the corresponding factor spaces. The GPA extends the quest to a multiple data set, providing a *compromise* representation of the different clouds in a common representation space.

Thus, our idea is to use PA to ascertain how the withdrawal of some percentage of rare elements – which usually highly affect the identification of the factors, let them be either form or repeated segments – influences the analysis. In other words, we are trying to understand to what extent the table may be reduced without losing relevant information and consequently changing the interpretation of the factors. For this task, we took into account the contingency tables crossing the seals geographic origin with either forms or repeated segments. Then, we withdrew progressively some percentage of rare items and applied the same TCA. Eventually, through PA, we examined the change that occurs on the factors' significance and possible interpretation and the reciprocal position of the considered items while the tables were reduced.

3. The studied corpus

Preliminary to this work are the important experiences in the quantitative study of Mesopotamian glyptics carried out by Sergio Camiz and Elena Rova (Camiz, Rova 1991). As far as presentation scenes are concerned, the first actual quantitative experiment has been carried out on a relatively large corpus with the use of algorithms referable to the category of Artificial Neural Networks (Di Ludovico, Ramazzotti 2008). The corpus at hand here corresponds to that used in the previous paper (Di Ludovico, Camiz 2020): it is made up of 354 specimens, attested on both original impressions and seals, which belong to the category of *presentations*. This is a general subject – not always clearly defined – attested in lower Mesopotamian glyptic from the late Early Dynastic period (Di Ludovico 2005, 60-61). A presentation may be shortly defined as a scene showing a receiving anthropomorphic

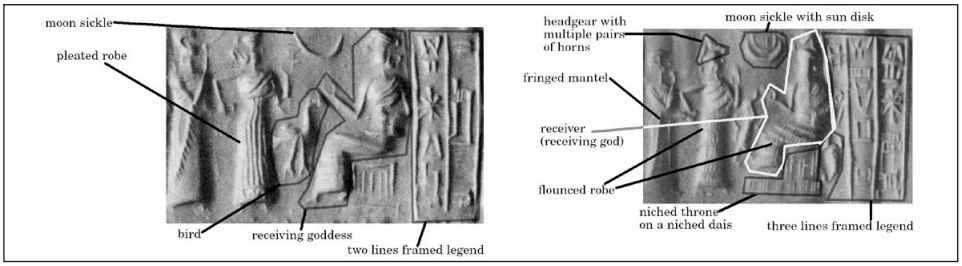

Fig. 1 – Two presentation scenes engraved on Third Dynasty (Ur III) seals, with indication of the most relevant features.

figure, facing one or several characters: within this period, the former is usually a deity, a royal, or another human figure of authority, whereas the characters have either human or divine nature. A seemingly neutral space on the cylinder's surface often hosts a short inscription, usually called *legend* in literature, while in other free parts of the surface various symbols, animals, objects, and other elements, called *integrating motifs* (Di Ludovico 2005, 60-68), can find place. In Fig. 1 are represented two presentation scenes from seals kept at the British Museum (Collon 1982, n. 92, 137), showing two figures presenting to a sitting receiver, with integrating motifs and a legend on the right, together with the indication of some typical features.

The presentation scenes considered in this paper date back to a historical age largely corresponding to the Third Dynasty (usually called Ur III) period, which followed the late Early Dynastic for few centuries. In that period, this theme tended to become homogeneous and to follow stable compositional schemes. Strong clues (Collon 1982, 129) suggest that it became the main theme of official glyptic and acquired an increasing importance for the state administration and its officials, because of the symbolic implications of its iconographies and compositions (Winter 1986, 1987). Then, after the end of Ur III period, the presentation scenes became progressively scarce (Haussperger 1991).

The textual corpus has been created by describing in full detail the images, starting from the area of the field which is most distant from the *receiving* figure, namely the figure that is always clearly recognizable in the scene, and ending with the neutral part of the field which usually hosts the legend (in former works this area was called *origin of the scene*). This was done through a formalized text composed of minimal entities called *forms*, each one representing an element (such as a *man*, a *deity*, a *lion*, a *bench)*, an element attribute (*with_long_hair, with_open_arms, holding a cup*), a feature of the body posture (*holding_a_hand_at_waist*, *standing*, *kneeling*) or a relation with another element (*on*, as in *sitting on a bench*, *placing_the_right_foot_on_a_dais*).

Such a code may be easily used for several purposes: i) identify either the presence/absence or the frequency of a selection of elements, attributes, relations, to analyse through classical methods, such as Multiple Correspondence Analysis; ii) use the text as it is and study it through TCA of forms, this way giving relevance to all forms composing the text, unless previously selected; iii) extract the repeated segments/quasi-segments and apply TCA to them, a way to broaden the study to the substructures; and iv) the restriction to the connections between forms/segments, as resulting from the text *skeleton*, obtained by synthesizing all the elements of the same kind (*main person standing, secondary person on object, man in front of a deity*) to study the kind of relations between elements, substructures, and attributes (Camiz, Rova, Tulli 1998).

An essential feature of textual coding is the strict rule to follow to describe an image (top to bottom, left to right, etc.) and its unambiguity: only one term is allowed, neither conjugated nor declined, to mean a given item, attribute, action, or relation. This results in a 1-1 correspondence between iconographic elements and describing forms. The repeated segments have been created automatically by a program, in two ways: i) only limiting their increase to 20 forms, and ii) stopping their increase should a punctuation mark occur.

The quoted previous work on this corpus (Di Ludovico, Camiz 2020) focused on the search for the most suitable reduction of the least frequent forms that did not affect excessively the factors' interpretation and avoided their disproportionate influence on the results. It resulted that the corpus of scenes under analysis is too much homogeneous to avoid, after the first cuts, a huge loss of data that are of basic importance to distinguish the groups of scenes. A 5% reduction of the tokens (from 100% to 95%, cutting the less frequent relevant forms) involves already a meaningful loss of features which allowed distinguishing classes and sub-classes of scenes. Further reductions of the data-set imply a too huge loss of information. Now the same experimentation is applied to the repeated segments of the same corpus: once again, to simplify the study, we focused on the sites of origin of the seals, although also their relative chronological frame was taken into account, this time crossed with the segments. The reason for this choice derives from the results of the quoted experiment on the forms, which showed a very weak factor in relation to the chronological sub-periods and their very limited explaining role (Di Ludovico, Camiz 2020, 29). Within this frame, in order to have more explicit results, the interpretation of the analyses focused mainly on the scenes having a known geographic origin. Thus, the total amount of scenes on which the attention has been specifically focused is 212. Of them, 75 come from Umma, 45 from Ur, 43 from Telloh, 27 from Susa, 11 from North-Bab, 6 from South and 5 from North-East.

The textual analyses have been carried out through SPAD.T program (Lebart, Salem 1994); the Procrustes analysis through the FactoMineR package (Lê, Josse, Husson 2008) of R environment (R Core Team 2017).

4. Tables organization

The quantitative experiments carried out in the past on this corpus of presentation scenes gave results which stressed the contribution that the gender of the characters acquires in the scene, affecting deeply its interpretation. Male and female characters and their typical attributes were clearly opposed to each other, when they play similar roles in the scene, since they were basically opposed toward the two ends of the first factor (Di Ludovico, Camiz, Pieri 2013, 497; Di Ludovico, Camiz 2014, 15-29; 2015b, 490-491). Another remarkable opposition which emerged along both first two factors is that between the features typical of scenes subject to secondary work interventions and the traits related to the presence of a male human (royal or similar) receiver. The elements more easily detectable on the original seal impressions (i.e., belonging to seals still clearly active in the administration) were placed in the resulting graphics in a position similar to that of the latter (that is, related to the male human receiver).

In the previous analyses, as in the case of the forms, the results concerning the repeated segments presented many low-frequency features that could bias the interpretation. Therefore, we progressively reduced the corpus of segments, removing the least frequent, this way obtaining a series of contingency tables, all crossing the set of sites of origin with a progressively reduced number of the lowest frequency segments. The PA was then applied both pairwise to the adjacent tables of the sequence, to examine their similarity and globally to get a general image of the variation. Indeed, the main target was to identify which table could provide most of the interpretable information with the minimum number of segments, that is try to identify which segments could be considered as informative and which simply noise.

The original segments contained in the data table – 934 – were reduced to 875, once the 59 hapax forms had been withdrawn. Then, eight subsequent tables have been built progressively: the first by reducing 5% of occurrences each, and the last two by withdrawing 10%. These last larger reductions were done to cope with the small number of remaining segments: indeed, a 5% reduction would have changed too little the table structure. To choose among encapsulated segments – that is, one a prolongation of the other – we preferred the longest if it carried actual further information, otherwise the most frequent. As a consequence, starting from the complete table, while the progressive reduction of the total occurrences of the segments was regular, in each reduced table a small, irregular, amount of segments was withdrawn too. To these tables, a tenth was added, where only segments especially expressive of meaningful iconographic parts of the scenes and/or related to specific toponyms were selected. This table contains a relatively high number of segments, corresponding to around 93.5% of the total occurrences. In Tab. 1, the percentage of occurrences and the number of different segments of each table are reported.

5. Results

Just like in the quoted analysis on the forms, here a great change is observed passing from 100% to 95% analysis. Comparing the figures corresponding to those of that analysis, the relative differences in the number of remaining segments are smaller. Following the 80% cut, the relative tendency of the number of persisting segments parallels that of the forms. This seems to confirm the strong inner homogeneity and consistency of the whole data set, since a (comparatively) large number of segments is an evidence of the homogeneous compositional structure of the records. Indeed, the number and significance of the dimensions – represented by the axes of the graphics – reduces dramatically after the first reductions (Tab. 1). The manual selection, essentially based on qualitative features, has been conceived as a compromise between the representativeness of the retained segments and the table expressiveness: the analysis results showed the information more spread through the factors, but one may wonder to what extent this dispersion may depend on the high number of zeros in the table.

Indeed, looking at the pattern of the segments on the factorial planes, one may notice that while the number of involved segments decreases, their spreading reduces to a more compact representation. This might depend on the progressive reduction of rare segments typical of only one site that spread according to the sites positions.

Table	Percentage of occurrences	Number of occurrences	Number of different segments	Number of meaningful axes	Main contribution to the formation of the first axis (first two toponyms)
1	100%	24727	875	4	Umma (57.8%), Telloh (9.2%)
2	95%	23492	338	3	Umma (58.5%), Telloh (9.4%)
3	90%	22441	231	2	Umma (58.2%), Telloh (9.2%)
4	85%	21067	159	2	Umma (57.8%), Telloh (8%)
5	80%	19795	169	1	Umma (53.1%), Telloh (11.3%)
6	75%	18562	98	1	Umma (54.8%), Telloh (6.8%)
7	70%	17489	82	1	Umma (56.4%), Telloh (6.8%)
8	60%	16107	58	1 (weak)	Umma (51.2%), Telloh (7.5%)
9	50%	13540	43	1 (very weak)	Umma (44.7%), Telloh (9.7%)
10	~93.5% (manual selection)	23109	509	5	Umma (49.9%), Telloh (25.6%)

Tab. 1 – Percentage, number of occurrences and number of different segments for each of the contingency tables crossing site origin with repeated segments built from the first through progressive reduction.

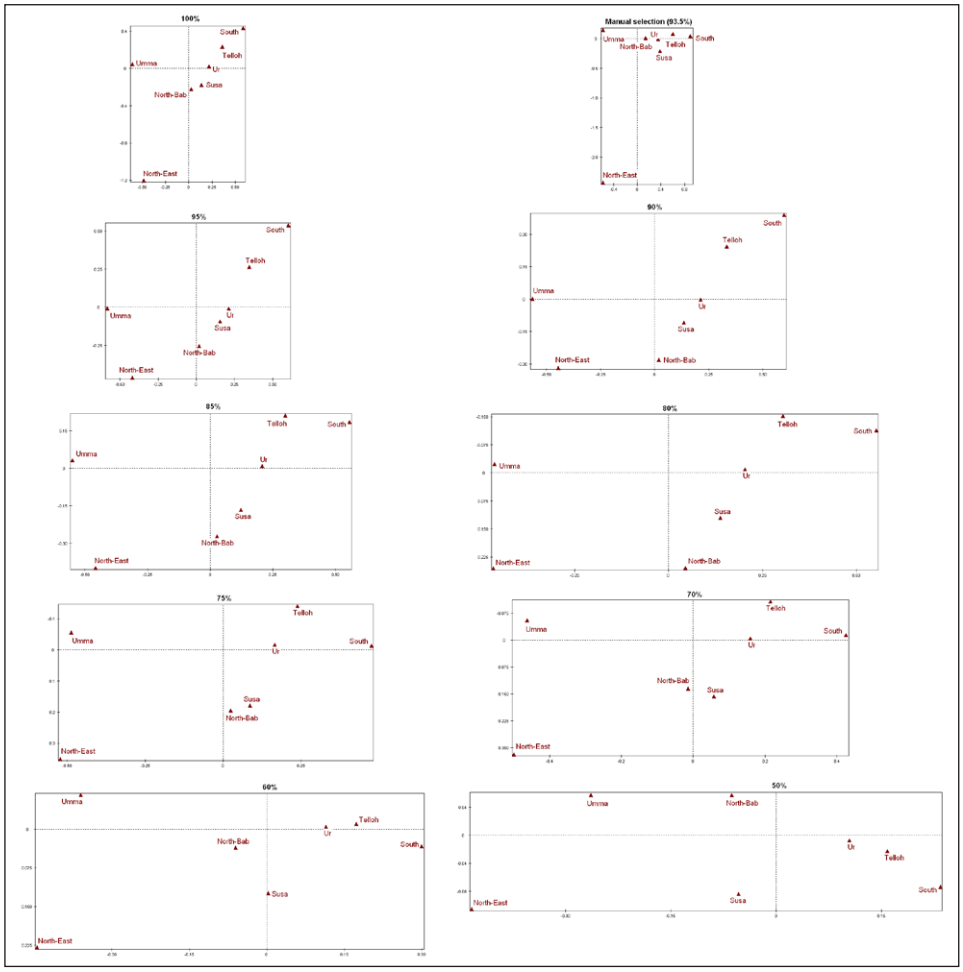

Fig. 2 – Representation of the geographic origins on the first factor planes issued from the TCA of the progressively reduced data sets.

In Fig. 2 the sites of origin are represented on the planes spanned by the first two factors issued by the ten TCAs carried out on the progressively reduced data sets. Considering the first axes, the opposition of both toponyms, Umma and Telloh, along them keeps about the same in all analyses, albeit much more evident in that of the full table, progressively decreasing its evidence as the data set is reduced. Therefore, the meaning of this axis keeps the same in all analyses, with minor changes, like the position of other toponyms associated to the quoted, respectively: North-East close to Umma and South and Ur,

close to Telloh. Note the higher contribution of North-East to the second axis (90.2%) together with some segments nearly exclusive, like those referring to specific *male receivers with niched throne* or to a *goddess*. The position and the weight of Umma in the graphics is remarkable, especially with reference to the first axis: this toponym, placed at its negative extreme, mainly opposed to Telloh, South and Ur, gives a crucial contribution to the formation of the axis. Analogous results are issued by the analysis of the manually selected corpus: here, on the first axis Umma – its main contributor – is opposed to Telloh and South. North-East is very close to Umma on the first axis, but it seems better explained by the second one, almost exclusively (90.2%) based on this toponym and on some segments, distributed in analogous position and less related to the other toponyms: e.g., segments referring to specific *male receivers with niched throne* or to a *goddess* represented in particularly small size. The latter, visible in few presentation scenes, is difficult to interpret: it may either represent a *tiny goddess* – whose meaning is unknown – or result of a further inclusion in a composition, where the little space had been a constraint.

Considering the manual selection analysis, the same opposition occurs on the first axis, but with a larger contribution of Telloh too: this is the only analysis where both Telloh and its associated segments play a more relevant role. Note also the high number of significant dimensions, 5, compared to the 4 of the full table and the 3 and less of the reduced ones: this may mean that a *reasoned selection* implied an *a priori* identification of both forms and segments discriminating among toponyms, hence better associated to dimensions that become more relevant. On the opposite, the reduction of segments consequently reduces the relevance of the dimensions following the first: indeed a sign of loss of information, once ascertained their importance in the larger tables. As for the position of the segments in this analysis, main contributions come from segments referring to the presence of a *female divine receiver with a double-curl hairstyle and a headgear showing a single pair of horns*, on the positive side, as Telloh is, and to a *bearded male god as receiver*, on the negative side where Umma is situated. Indeed, other segments give an important contribution to the axis, still expression of the presence of female deities and further attributes of the male receiver, respectively. Such opposition had been already observed in other analyses and experiments (Di Ludovico, Camiz 2014, 14-18; Di Ludovico 2018, 99-104), not mainly related to the gender, but rather referred to certain types of presentation scenes. So, these with female divine receivers are the standard and earlier scenes opposed to those with male god: these were probably developed later and had a longer official administrative use, because they are especially attested from original seal impressions on administrative documents. This is further confirmed by the coordinates on the same axis of other segments: on the negative side, describing either the goddess with the headgear with multiple pairs of horns

or the male receiver with royal attributes; and on the positive side describing a goddess or a woman with the plain robe.

The second axis of this analysis is mainly structured by North-East and its main features, which are typical of the more recent – or very late – official scenes depicting male receivers. What emerges from the first two axes is also that the scenes corresponding to a stable official use include integrating motives, albeit in low quantity, and mostly composed of animals like *lion* or *long legged bird*. The third axis appears defined by the opposition between Susa on the positive side and Telloh on the negative one: both are related to many features typical of the earlier phases, albeit Susa shows some peculiarities, like presence of *characters wearing plain garments*, *men having the simple hairstyle,* or *men and goddesses wearing skirts,* that may have a regional origin. On the fourth axis, a similar opposition between Susa on the positive side and both North-East and Ur on the negative one results: this seems due to the opposition between *original impressions* and *presentations before a male figure*. Note also that North-East seems related to features appearing in later scenes, like the presence of a *little monkey* in the field. This is further evident in the structure of the fifth axis, where Ur and North-East are clearly opposed because of the opposition between original impressions and presentations before a male figure, both contributing meaningfully to the formation of the axis.

The analysis carried out on 95% of the forms gives the same results for the first three axes – the only truly meaningful here – of the manual selection: here, the prevailing segments are shorter, thus not as rich in information as the others. Moreover, the weight of North-East is not as evident in the formation of an axis. The analyses on the further reduced corpora produced results that may be interpreted the same way, but this thanks to the comparison with

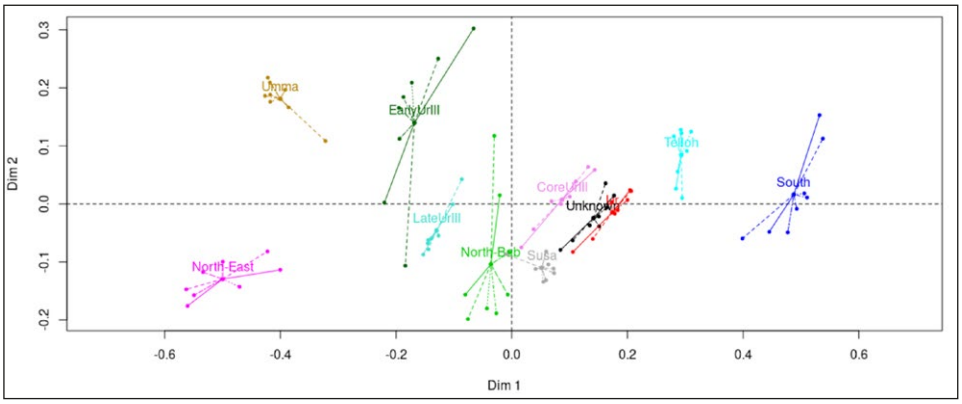

Fig. 3 – Generalized Procrustes Analysis results. Positions of the sites on the first factorial planes, as resulting from the original TCAs, on the plane of issued common representation.

those already interpreted of both the whole and the manually reduced tables. Without them, the analyses would be most difficult to interpret, since the segments most contributing to the meaningful axes tend to become shorter while reducing the table, thus less informative. As an example, in the case of the tables with less than 70% the main contributions to the axes come from very general segments only indicating that the figures in the scene have some personal attributes. In these cases it is also difficult to recognize features opposed to each other along the axes.

In Fig. 3 the sites positions on the first factorial planes of each TCA are represented on a common plane around their compromise position, as issued by GPA. It is evident the major importance of the first dimension, given that the second one shows a minimum variation.

6. Characteristic segments of the sites

In order to identify the *characteristic segments* associated to each toponym, a first session of analyses has been carried out by selecting the segments corresponding to complete substructures only. This was done by keeping in the coding the brackets isolating them, since the segments may not exceed the brackets. A second session followed, this time without brackets, thus allowing substructures incomplete. The latter results proved clearer and better interpretable. Therefore, it seems that longer segments spanning beyond substructures limits may result more informative: indeed, an unexpected result.

In Tab. 2 the amount of characteristic segments, both significantly present and absent in a site, are reported for the nine progressively reducing corpora. The figures show the progressive reduction in their number, while reducing the tables. This becomes dramatic in the last three reductions. Indeed, considering the quality and the degree of expressiveness of the characteristic segments, two meaningful reduction steps emerge: one corresponding to the reduction from 90% to 85% and one from 75% to 70%: actually, the latter cut withdraws too many segments with elements that may discriminate the specimens according to their places of origin. The reduction to 85% could seem quite reasonable, since it keeps enough segments to suggest sufficiently well the types of scenes and the general features characterizing the site; nevertheless, compared to the 90% data set, too many elements – especially the integrating ones – as well as their position in the scene, are lost, leaving uncertain the characterization of some toponyms. This drives the attention to the 90% cut, although it highly reduces the characterization of Susa with respect to the 95% one.

The amount of segments typically associated with the toponyms is huge for Ur, Telloh, Susa, and Umma, both on the positively and negatively, but a thorough interpretation seems crystal clear only in the analysis carried out on the full corpus.

Toponyms	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	30	30	30	30	30	30	30	30	30	30	30	30	30	26	30	9	30	9
Ur	30	30	30	30	29	30	25	30	25	30	20	26	8	16	3	7	3	7
North-Bab	12	2	12	2	12	2	10	2	6	2	5	2	1	1	1	0	1	0
South	19	2	19	2	19	2	13	2	10	2	10	2	4	2	2	2	2	2
Telloh	30	30	30	30	30	30	30	30	30	30	28	30	25	27	22	7	22	7
North-East	30	15	28	13	23	13	14	13	14	12	11	12	6	11	3	10	3	10
Susa	30	26	22	23	15	23	9	23	6	23	5	23	2	13	1	3	1	3
Umma	30	30	30	30	30	30	30	30	30	30	30	30	30	30	13	30	13	30

Tab. 2 – Number of characteristic segments for all toponyms, according to the nine progressively reduced table, subdivided in positive (significantly abundant in the site with respect to the whole corpus) and negative (significantly poor). In the heading, the percentage of kept segments in the table.

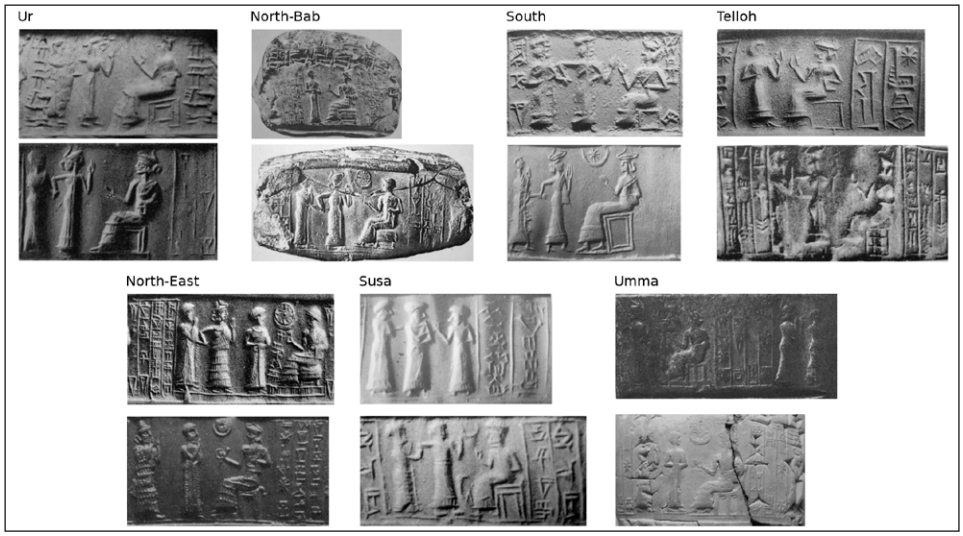

Fig. 4 – The two most typical scenes of each seals' site of origin.

In Fig. 4 two most *typical* scenes of every origin are represented. We mean as typical those scenes with minimum distances from the centroid of the group of seals with the same origin. As in TCA the centroids of the classes are defined based on the weighed (chi-square) distance between the segments, it is expected that such scenes present a similar balance of the most relevant features characterizing the origins.

In the figure, the scenes from Ur show earlier features, like those from Telloh, which have primarily female characters. Those representative of North-Bab are the two most relevant compositions of presentation before

a royal receiver, both coming from original impressions. Those from South have basic compositions and characters with plain robes, with a tendency to an earlier style, while those from North-East have rather late features. The scenes from Susa show own arrangements of characters and their attributes, with the most representative one having a standing receiver. Lastly, scenes from Umma are known form original impressions and are the only ones, among all most typical scenes, showing a bearded god in the role of receiver.

In the following, the sites are characterized according to a selection of distinctive segments – built without brackets, considering the most frequent and most absent, respectively. Indeed, some segments resulting typical for a toponym may resemble very much to each other, sometimes one being an extension of another, or varying in length or clarity. Therefore, in the lists below, the clearest and most expressive segments have been selected, whereas those too similar to them and forming incomplete structures have been withdrawn. This way, repetitions have been avoided and a clear picture of typical sub-structures could be recognized. Those most present are represented in the accompanying figures prior each list.

Ur (Fig. 5)

Fig. 5 – Most characteristic sub-images of seals from Ur.

Present:

1) *Goddess with flounced robe and hairstyle with double curl, with headgear having a single pair of horns looking left with her left hand at waist and right hand brought forward sitting on a square seat with a frame standing on a simple dais*; 2) *goddess with hairstyle with double curl looking left with her left hand at waist and right hand brought forward sitting on a square seat with a frame standing on a simple dais*; 3) *goddess with flounced robe and a lock of hair on the top-back of the head*; 4) *moon sickle with sun disk above, before a seated goddess*; 5) *goddess with a lock of hair on the top-back of the head stands before a seated goddess with flounced robe and a lock of hair on the top-back of the head with her left hand at waist and right hand brought forward*; 6) *a bird in the field placed above with a scorpion thereunder*.

Absent:
left hand before the face with bracelets; male figure; male receiving figure on a padded stool; man in fringed mantle; receiving bearded god in flounced robe, having a headgear with multiple pairs of horns, hair-lock behind his head, left hand by the waist, right hand brought forward; goddess with hairstyle with double curl turned right and holding hand by hand a man; simple dais by the receiver and three lines legend in a frame with content of the like "x, profession, son of y".

North-Bab (Fig. 6)

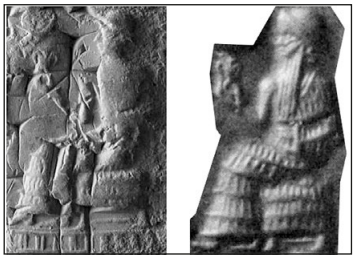

Fig. 6 – Most characteristic sub-images of seals from North-Bab.

Present:
1) *royal receiving figure with skull-cap, a bracelet, and possibly holding a little amphora*; 2) *royal receiving figure with skull-cap, flounced robe and attributes.*

South (Fig. 7)

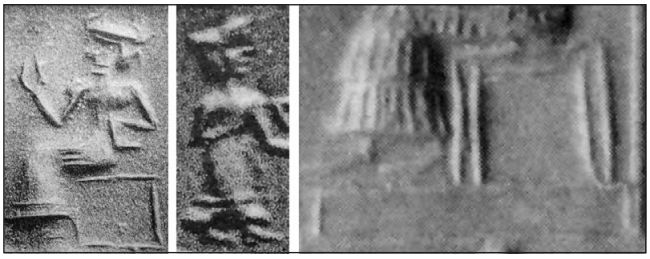

Fig. 7 – Most characteristic sub-images of seals from South.

Present:
1) *goddess with plain robe, double curl behind her neck, and headgear with single pair of horns*; 2) *woman with plain robe and a lock of hair on the top-back of her head*; 3) *square throne with a frame, simple dais, and a footstool.*

Telloh (Fig. 8)

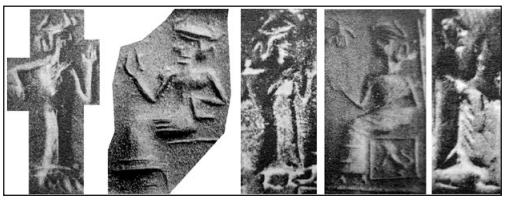

Fig. 8 – Most characteristic sub-images of seals from Telloh.

Present:

1) *goddess in plain robe with headgear with a single pair of horns and double curl behind the neck, having her left hand before her face*; 2) *goddess in plain robe and lock of hair on the top-back of her head*; 3) *goddess in fringed robe with headgear with a single pair of horns*; 4) *receiving goddess in flounced robe with headgear with a single pair of horns, double curl behind the neck, left hand at her waist, right hand forward, and multiple necklace*; 5) *received man with fringed mantle and simple hairstyle.*

Absent:

receiving man with fringed mantle, skull-cap, with left hand at the waist and right hand forward, and having some attributes; goddess in flounced robe with headgear with multiple pair of horns and double curl behind the neck; character with both hands before the face; bearded receiving character with left hand at the waist, right hand forward and some attributes.

North-East (Fig. 9)

Fig. 9 – Most characteristic sub-images of seals from North-East.

Present:

1) *both hands before the face, multiple necklace and bracelets*; 2) *small goddess in flounced robe*; 3) *man with fringed mantle and striped headgear*; 4) *goddess in flounced robe wearing a headgear with multiple pairs of horns and having a double curl behind her neck*; 5) *man in fringed robe with shaved head and right hand before his face*; 6) *man in fringed mantle with shaved head and both hands at his waist; 7) royal padded stool with niched dais.*

Absent:
goddess with headgear with single pair of horns ad double curl behind her neck; received man in fringed mantle with shaved head and right hand before his face.

Susa (Fig. 10)

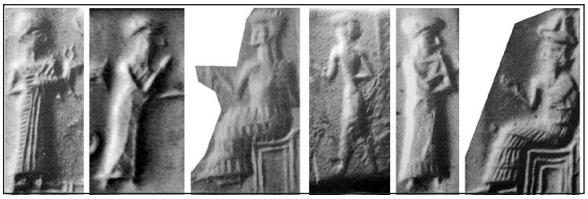

Fig. 10 – Most characteristic sub-images of seals from Susa.

Present:
1) *man in fringed robe*; 2) *man in fringed mantle and striped headgear with his right hand before his face*; 3) *male receiver with simple hairstyle, left hand at the waist and right hand forward, having beard and bracelets*; 4) *man in plain robe and simple hairstyle*; 5) *received man in fringed robe, with flat headgear, right hand before his face, left hand at his waist*; 6) *bearded god in flounced robe with headgear with multiple pairs of horns, hairlock behind his head, left hand at his waist, right hand forward*.

Absent:
goddess in flounced robe with headgear with multiple pairs of horns, double curl behind her neck, multiple necklace and left hand before her face; god receiver, bearded, with left hand at his waist and right hand forward.

Umma (Fig. 11)
Present:

Fig. 11 – Most characteristic sub-images of seals from Umma.

1) *goddess in flounced robe, headgear with multiple pairs of horns, double curl behind her neck and left hand before her face, bracelets and multiple necklace*; 2) *goddess in pleated robe, headgear with multiple pairs of horns,*

double curl behind her neck and both hands before her face; 3) *bearded god in flounced robe with hair lock behind his head, headgear with multiple pairs of horns, left hand at his waist and right hand forward*; 4) *three-lines framed legend with content type "xxx, profession, son of yyy"*.

Absent:

goddess in pleated robe, headgear with single pair of horns, double curl behind her neck; goddess in flounced robe, headgear with single pair of horns, double curl behind her neck, her left hand at her waist, her right hand forward; two-lines framed legend with content type "xxx, son of yyy".

The segments typically present in the specimens from Ur are related to types of presentation scenes that are still quite various and show some comparatively old motifs. Their compositional structure is not really uniform. The segments typically absent are usually related with more standardized scenes and presentations before a seated male figure. These features place Ur quite close to South and in clear opposition to North-Bab, in which particularly present are the segments of the presentations before a royal figure. The typically present and absent segments of Telloh place it in a seemingly isolated position, but rather in affinity with Ur and South. Similarly, North-East segments are generally quite close to those of North-Bab, but in a peculiar position. The typical and non-typical segments of Susa show a picture very close to that of North-East, but with a much higher differentiation of features. In the case of Umma the situation is quite peculiar as well, since scenes of standardized composition are found, which are more bound to their administrative use, in opposition to older features.

7. Conclusion

Considering the general structure of the presentation scenes of the Ur III glyptic, the analysis on the repeated segments seems especially recommended, because it reveals, more explicitly than the analysis on the forms, the relative affinities and differences among compositional structures and substructures. Actually, the deep interpretation of these scenes heavily depends on the spatial relations of the iconographic elements in the field. Both peculiarities and relevance of the scenes known from original impressions emerge clearly in the graphics and especially influence the relative position of the toponym *Umma*. In particular, the opposition of *Umma* to both *Telloh* and *Ur* shows a specific development, in the Ur III period, of the way seals were conceived in the state administration after the reforms of the first two kings of the Dynasty. On the other hand, the regional features typical of the scenes from northern *Babylonia,* as well as those from *Susa*, may be considered slight variations of the themes officially adopted in the core of the kingdom.

Compared to its use on the forms, the PA on the repeated segments allowed a better fine-tuning of the processed data, based on both their frequency

and iconographical meaning. This way, the quality of the outcomes improved significantly as well as the comparison between different selections. Indeed, this is due to the specific nature of the repeated segments and their intrinsic quality of being able to represent parts of the scenes of various size. In fact, a more nuanced selection becomes possible, aiming at keeping a larger variety of possible sub-structures of the scenes, while the analysis on the forms involved a more radical choice of keeping or excluding items. Such choice may not be subject to an *a priori* evaluation, since hypotheses about their values should be based rather on the outcomes of the analyses.

As far as it emerges from the investigation described here, the experiences with the use of PA on the repeated segments allow a clear definition of its possible limits and its global impact on TCA results. The exam of the outputs corresponding to regular reduction of the data sets, while decreasing the solutions' dimension, gets clearer the most relevant aspects of the differences among the seals origins. Indeed, the origins are well separated, but what eventually emerges as segments is a common core of features, whereas the most specific attributes – precisely because of their specificity – tend to be withdrawn. For this reason, when the data set has a considerable amount of inner unevenness – or a large quantity of zeros – then a *manual* selection of segments is advisable: this way, the automatic reduction based on the frequency may be smoothed, to keep some particularly descriptive and/or specific segments. On one hand, this should keep the basic information expressed by rare but meaningful segments, while on the other, it should exclude segments that – though frequent – replicate information provided by slightly more frequent and clearer ones. Differently from the forms, this cannot be done by adjusting the original coding, because segments are aggregation of already defined forms. Therefore, the proposed mixed strategy (automatic and manual) may lead to an ultimate improvement and a good compromise in terms of meaning and precision of the results.

The inspection of the graphics representing the scattering of the sites on the first factorial plane suggests a curvilinear pattern; this may be interpreted as a Guttman effect (Camiz 2005), indicating a seriation in the data. A specific study, aiming at identifying some kind of continuous variation of the scenes' features according to the sites of origin, might be the subject of a further paper. Indeed, this pattern corresponds to the linear one provided by the GPA.

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

REFERENCES

Camiz S. 2004, *On the coding of archaeological data*, «Archeologia e Calcolatori», 15, 201-218 (http://www.archcalc.cnr.it/indice/PDF15/12_Camiz.pdf).

Camiz S. 2005, *The Guttman effect: Its interpretation and a new redressing method,* «Τετραδία Ανάλυσης Δεδομένων» («Data Analysis Bulletin»), 5, 7-34.

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 (http://www.archcalc.cnr.it/indice/PDF7/51_Camiz_Rova.pdf).

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 (http://www.archcalc.cnr.it/indice/PDF12/01Camiz-Rova.pdf).

Camiz S., Rova E., Tulli V. 1998, *Exploratory analysis of images engraved on ancient Near-Eastern seals based on a distance among strings*, «Statistica», 58, 4, 669-689 (https://doi.org/10.6092/issn.1973-2201/1114).

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

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, Roma, Università La Sapienza, 57-95.

Di Ludovico A. 2018, *Data description and the integrated study of ancient Near Eastern works of art: The potential of cylinder seals*, in V. Bigot Juloux, A.R. Gansell, A. Di Ludovico (eds.), *CyberResearch on the Ancient Near East and Neighboring Regions: Case Studies Archaeological Data, Objects, Texts, and Digital Archiving*, Digital Biblical Studies, 2, Leiden-Boston, Brill, 85-110 (https://doi.org/10.1163/9789004375086_005).

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 (http://www.archcalc.cnr.it/indice/PDF25/01_Camiz_Di_Ludovico.pdf).

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 (https://www.archaeopress.com/ArchaeopressShop/Public/download.asp?id={998F1DEA-319D-4D61-B3F9-13F278B5ADF2}).

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 Archeaology. Concepts, Methods and Tools. Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology (Paris 2014)*, Oxford, Archaeopress, 489-496 (https://www.archaeopress.com/ArchaeopressShop/Public/download.asp?id={998F1DEA-319D-4D61-B3F9-13F278B5ADF2}).

Di Ludovico A., Camiz S. 2020, *Image study of Mesopotamian cylinder seals through texts: The Procrustes transformation applied to Correspondence Analysis results*, «Archeologia e Calcolatori», 31.1, 7-32 (https://doi.org/10.19282/ac.31.1.2020.01).

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)*, BAR International Series 2494, Oxford, Archaeopress, 495-498.

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, Held at the Oriental Institute of the University of Chicago (Chicago 2005)*, Studies in Ancient Oriental Civilization, 62, Chicago, Oriental Institute, 283-288.

Gardin J.-C. 1958, *Four codes for the description of artifacts: An essay in archeological technique and theory*, «American Anthropologist», 60, 2, 335-357 (https://doi.org/10.1525/aa.1958.60.2.02a00090).

Gardin J.-C. 1966, *Éléments d'un modèle pour la description des lexiques documentaires*, «Bulletin des bibliothèques de France», 5, 171-182 (https://bbf.enssib.fr/consulter/bbf-1966-05-0171-001).

Gardin J.-C. 1967, *Methods for the descriptive analysis of archaeological material*, «American Antiquity», 32, 13-30 (https://doi.org/10.2307/278775).

Gower J.C. 1975, *Generalized Procrustes Analysis*, «Psychometrika», 40, 1, 33-51 (https://doi.org/10.1007/BF02291478).

Gower J.C., Dijksterhuis G.B. 2004, *Procrustes Problems*, Statistical Science Series, 30, Oxford, Oxford University Press.

Haussperger M. 1991, *Die Einführungsszene. Entwicklung eines mesopotamischen Motivs von der altakkadischen bis zum Ende der altbabylonischen Zeit*, Münchener Vorderasiatische Studien, 11, München-Wien, Profil.

Lê S., Josse J., Husson F. 2008, *FactoMineR: An R package for Multivariate Analysis*, «Journal of Statistical Software», 25, 1, 1-18 (https://doi.org/10.18637/jss.v025.i01).

Lebart L., Morineau A., Bécue M., Haeusler L. 1994, *Spad.T – Version 1.5 – Manuel de référence*, Paris, Cisia.

Lebart L., Salem A. 1988, *Analyse statistique des données textuelles*, Paris, Dunod.

Lebart L., Salem A. 1994, *Statistique textuelle*, Paris, Dunod.

R Core Team 2017, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria (https://www.R-project.org/).

Winter I.J. 1986, *The King and the cup: Iconography of the royal presentation scene on Ur III seals*, in M. Kelly-Buccellati, P. Matthiae, M.N. Van Loon (eds.), *Insight through Images. Studies in Honor of Edith Porada*, Bibliotheca Mesopotamica, 21, Malibu, Undena, 253-268.

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*, Studies in Ancient Oriental Civilization, 46, Chicago, Oriental Institute, 69-116 (https://oi.uchicago.edu/sites/oi.uchicago.edu/files/uploads/shared/docs/saoc46.pdf).

ABSTRACT

In a recently published study, a textual dataset encoding a group of scenes depicted on Mesopotamian cylinder seals from the third millennium was submitted to Procrustes Analysis, considering Correspondence Analysis variations due to progressive reductions of the forms involved. The results seemed to indicate that a slight reduction in the number of forms used to describe the raw data would improve the Correspondence Analysis results. Indeed, the actual impact of each reduction on the outcome of the analyses could not be adequately identified if the forms were not considered as elements forming segments, i.e., sub-images of the scenes. In this paper, the results of the same methodology are presented, this time applied to repeated segments within the same dataset, i.e., sequences of textual forms that may describe sub-images of the described scenes. The comparison between the progressive reductions of repeated segments showed relevant differences between small (around 10% of the rarest segments) and large reductions. Indeed, the latter may undermine a consistent interpretation of the different finding sites, yet well represented keeping 90% of segments. Moreover, a reduction performed by hand and not on the frequency of occurrences provided better results.