The density of types and the dignity of the fragment A website approach to archaeological typology <sup>1</sup>

For

Between variability and singularity: crossing theoretical, qualitative and computer-based approaches to types and typologies in archaeology

Giorgio Buccellati and Marilyn Kelly-Buccellati

Typology hinges on categorization, and the two main axes of categorization are the roster and the lexicon: the first defines elements from an -emic, and the second from an (e)-tic point of view, i. e., as a closed or an open system, respectively. Archaeology poses a special problem with regard to the -emic dimension, since it deals essentially with broken traditions, i. e., with cultural wholes for which there are no living carriers. Dealing with this problem entails that one should engage in a very intensive type of distributional analysis, where density plays a major role. – Having so defined the universe of data, we deal with the question of how best to articulate and communicate the results. We propose in this regard the use of the website as a privileged epistemic system, one that allows to preserve the dignity of the fragment in its singularity while at the same time offering a full-fledged representation of the larger typological classes. The notion of digital discourse makes it possible to integrate the data within the flow of a coherent multi-planar argument. – The paper explains these principles using concrete evidence from Urkesh, one of the earliest known urban sites in Syro-Mesopotamia, with an excavated ceramic record of close to one million items.

This article is part of the publication program of the project *Cybernetica Mesopotamica*, undertaken under the aegis of the Balzan Foundation. See the website <u>cyb-mes.net</u>. We are grateful to Federico Buccedllati for his comments on an earlier version of this paper.

## Table of Contents

28	1 Introductory	3
29	I. DATA ORGANIZATION	3
30 31 32 33 34 35	2 Categorization 2.1 (e)-tic and -emic 2.2 Open and closed systems 2.3 Scalarity and congruence 2.4 Procedures and criteria 2.5 Density	3 4 5 5
36 37 38 39 40	3.1 Broken traditions	7 7 7
41	INTERLUDE	10
42	4 Typology in action	10
43	5 Data processing	10
44	II. THE PROCESS	11
45 46 47 48 49 50	6 The website as epistemic system. 6.1 The concept	11 11 12
51 52 53 54 55 56	7 Digital discourse	14 14 15
57	CONCLUSION	18
58 59 60 61	8 Variability and singularity	18 19 19
62	References	
63 64	Figures	23

- 67 Pensar es olvidar diferencias, es generalizar, abstraer.
- En el abarrotado mundo de Funes no había sino detalles, casi inmediatos.
- 69 Borges, Funes

## 70 1 Introductory

- 71 One can "forget differences" (*olvidar diferencias*) only once they have been identified.
- And one can go beyond a mass of purely contiguous facts (*detalles, casi inmediatos*) only if one gains a sense of the whole.
- That is the message we read in the Borges' text cited as epigram. And we will seek here an answer to the implicit question along two main lines of inquiry.
- The first is the nature of the categorization system, the one through which we establish the differences. It refers to the way in which the data are organized.
- The second issue concerns the process through which we can articulate and convey knowledge relying on typological classification..
- In the conclusion, we will see how all of this pertains to the central topic of the volume, i. e., to the relationship between variability and singularity.

#### I. DATA ORGANIZATION

## 83 2 Categorization

- 84 2.1 (e)-tic and -emic
- The terms (e)-tic and -emic have come to be in common use to refer to whether a given cultural phenomenon is being studied from without or from within the culture to which it belongs.
  - The term was borrowed from linguistics, and it is in recognition of this original use of the terms that here we write (e)-tic instead of -etic: the (e) is in fact not part of the suffix, but is rather derived from the word to which it is attached, in this case *phōnē*, while it is (o)-tic in "antibiotic" or, with a consonant, (s)-tic in "fantastic." In this article we keep the form (e)-tic throughout, though in point of fact one may well ignore this fine point of etymology and follow the current use which has entrenched "etic" and "emic" in the dictionary, thus allowing also for the (useful) adjectives "etical" and "emical."
- What really matters, beyond terminology, are the concepts behind the two terms (Buccellati 2006 "(e)tic" and Buccellati 2017 *Critique* 3.4.). They are powerful heuristic concepts, which need to be better defined: it will then appear how useful they can be in approaching the subject of typology, particularly with regard to variability and singularity.

82

88

89

90 91

92

93

94

#### 2.2 Open and closed systems

Structurally, an (e)tic system is open and an -emic system is closed. To illustrate the concepts we will use the cases of phonemics and of a street stoplight.

An (e)-tic system is open in the sense that there is no intrinsic limit on the number of types that are possible within the system. An acoustic inventory of phones articulated by different speakers is open ended because the "phonetic" variations that can be measured acoustically are practically unlimited, and there is always room for new ones to be added. Analogously, the measuring of chromatic differences among different embodiment of the three basic colors of green, yellow and red in different street stoplights can yield great variations, and still leave room for more.<sup>2</sup>

An -emic system, on the other hand, is closed in the sense that, within a given system, only a fixed number of possible realizations is found. There is a fixed set of phonemes in any given language, and a fixed set of three colors in any given stoplight. The variable acoustic realizations of a given phoneme must fit within a range of possibilities for communication to be possible, just as the variable chromatic realizations of a given color in a stoplight must fit within a range of possibilities for the message to be understood (one will be alerted to stop whatever shade of yellow or even orange the intermediate signal may have).

The elements of a closed system are mutually exclusive in the sense that their function depends on this exclusivity: the red color of a stoplight entails the stopping of traffic not because of an intrinsic quality of the color red, but because of its contrast with the yellow and green colors.

The elements of an open system, instead, are mutually exclusive because of an intrinsic quality of the element: thus the chromatic value of a given shade of red is different not only from that of yellow and green, but also from a multitude of other shades of red.

Several additional factors play a role in defining a closed system, for instance duration. The length of a phone (i. e., how long the articulation of a phone lasts) may be -emic in some languages and not in others. Analogously, in a stoplight in the US or Italy the red and green signals appear for the same amount of time, whereas yellow appears for a shorter period, the reason being that green and red are linked to the time when vehicular traffic may or may not flow, whereas yellow is intended only as a brief warning – thus enhancing the -emic status of the contrast among the three colors.

<sup>12 2</sup> 

We use the terms "speakers" and "phones" as they are in linguistics: a speaker is a native speaker, and a phone is a sound as defined either by the way it is articulated in the human vocal apparatus or by the way it can be defined acoustically. The stoplight, in turn, refers to the three color type that is most commonly in use, without regard for possible variations in the nature of the display; also, variations in the shade or intensity of color may be attributed to incrustations on the face of the fixtures, to different manufacturers, to some alteration in the electrical connections, or the like.

It must be stressed that a closed system is closed not in the sense that it is unchangeable, but in the sense that if a change occurs (deletion of an element or addition of another), then the whole system has to be re-calibrated. That is because the system is endowed with an in-ternal integrity that cannot be altered without the system changing its identity. For instance, one might add a blue signal to a stoplight to indicate heavy traffic ahead, but this would change the nature of the system, in this case simply because it would reduce the immediacy of the messages conveyed by the signal, an immediacy that is especially important precisely with heavy traffic.

#### 2.3 Scalarity and congruence

An open system is scalar in the sense that any given element is seen in and of itself, without a necessary relationship to other neighboring elements. Using our two examples, the acoustic or the chromatic definition of a given sound or a given color does not depend on a relationship of that element to any other elements in the system. Or rather, it depends only in the sense that both are defined by the same external categorization system, not because of the way in which the two elements interact with each other.

A closed system, on the other hand, is congruent in the sense that any given element depends on its relationship to one or more of its neighbors. Thus, in the phonemic system of a given language, a voiced dental phoneme includes a range of phones that, in turn, exclude the neighboring voiceless dental phoneme, while in another phonemic system (one that does not recognize the distinction between voiced and voiceless), the range includes both voiced and voiceless. Analogously, in a stoplight the import of the red signal is congruent with that of the yellow and green signals: there is an equilibrium or balance among them so that each depends on the other for its meaning to be fully perceived.

A more concrete example might be that of the organizational chart for a supermarket display. It is scalar in the sense that any new product (e. g., a new brand of mustard), or any new category (e.g., a curry condiment) may be added at will. It is simply a matter of making room on the shelves, in the appropriate place. It is the model of the website seen as a container, not as a system. Or again, the collection of photos on our cell phones is also scalar, being generally organized by date, and thus remaining open for any possible future addition.

#### 2.4 Procedures and criteria

The procedures to categorize elements within the two systems differ depending on the system.

The categorization of elements within an open system operates on the basis of criteria that are *external to the system*. Neither acoustics nor chromatic analysis are concerned with

the communicative function of language or with vehicular traffic. They use wholly independent standards, and for this reason they are seen as more "objective."

It appears then that, in effect, an open system is also closed – except that it is closed at a higher level than that of its constitutive elements. Acoustic phonetics is a closed system, but one that allows for an immensely larger set of possible realizations than the phonemic system. Analogously, a chromatic scale includes a seemingly infinite set of possible realizations of red, yellow or green, whereas the -emic version of a stoplight accepts only a set of three "colors," each of which can include an immense variety of actual chromatic realizations.

An -emic closed system, in fact, operates on the basis of *internal contrasts* that have to be identified through a sophisticated use of distributional analysis – especially where there is no carrier of a living tradition to provide a verification of assumed contrasts. When such an analysis is undertaken, the result is no less "objective" than with an open system.

#### 2.5 Density

With the term "density" we mean the quantity of items documented for any given type in its relationship to the total number of items in the inventory. The term is thus not used in the sense it has in statistics: it rather refers to a scale of values for the population attested within a type. Alternatively, we may say that density is a measure of percentage seen explicitly in function of the total to which it refers.

We may express density by means of a superscript referring to the total, added to the percentage of the item in question in relationship to the same total. For our current purposes, the superscript corresponds to a group of 100 items: thus the superscript <sup>2</sup> stands for a group of between 200 and 299 items, the superscript <sup>20</sup> for a group of between 2,000 and 2,099 items, the superscript <sup>200</sup> for a group of between 20,000 and 20,099 items, and so forthgiven

Let us consider two examples.

- (1) If we have an inventory of 220 items, of which a given type includes 1 item, we may express the density of this type with the formula  $0\%^2$ . The figure 0% gives the percentage of 1 item vis-à-vis the total of 220; it is a rounded figure for the longer 0.45454545454545. The superscript  $^2$  indicates that there are between 200 and 299 items in the inventory.
- (2) If we have an inventory of 20,020 items, of which a given type includes again only 1 item, we would express the density of this type with the formula  $0\%^{200}$ . The percentage given is the rounded figure for 0.004995004995005%, and  $^{200}$  is the abbreviated figure for the total 20,020.

This formalization is useful in giving a quick visible expression to the actual import of percentages:  $0\%^2$  expresses a much lower density than  $0\%^{200}$ . The deeper import of this notation is that it indicates that a case with 0% occurrences within a total of between 200 and

- 201 299 items carries much less meaning than a case of 0% occurrences within a total of between
- 202 20,000 and 20,099 items. We may say the  $0\%^2$  exhibits a thin, and  $0\%^{200}$  exhibits a thick den-
- 203 sity.

## 204 3 Implementation

#### 205 3.1 Broken traditions

The definition of an -emic system is particularly difficult in the case of broken traditions, i. e., cultures for which there are no living carriers capable of articulating their awareness for the contrastive nature of the elements of the system. We can illustrate this with reference to our two examples: the phonemic system of a language and the stoplight.

In the case of phonemics as applicable to "dead" languages, i. e., languages for which there are no living speakers, we cannot ask questions as to whether a given difference in, e. g., articulation is felt to be contrastive or not. We must deal with the additional filter of the written medium, and in this case "decipherment" means establishing distributional classes of the mute evidence that is assumed to match the spoken reality forever lost to us.

Or, from the point of view of an extraterrestrial who can watch the flow of traffic in one of our cities (Buccellati 2006 "Mars"), the distributional class would be one that takes into account the way in which the traffic flows, stops or begins to stop in correlation to the changing of colors: a mere chromatic analysis of the colors would not yield any meaning.

### 3.2 Native digitality of the archaeological record

Archaeology deals by definition with broken traditions: there are no informants who can share their awareness of an underlying typology. This forces us to define categorization systems, or "grammars," solely on the basis of the identification of distributional classes from within in the data, and, naturally, the larger the inventory, the higher is the probability of success in this endeavor.

There is in fact one more difficulty that is specific to the archaeological record. At the source, we only have disconnected fragments, as they come out of the ground: data are scattered in ways that do not match any typological ordering. The most distinctive and unique aspect of archaeology is, in fact, not the complete object that is brought to light – a building, a statue, a text. What is distinctive and unique is what happens at the moment of excavation, and at that moment we only have disconnected fragments, or rather: fragments that are connected only by virtue of how they are situated in the ground, as the result of a largely unplanned depositional process. This we call "emplacement." That is the only degree of association we can objectively record.

It is in this sense that the fragments are natively digital: they are atoms that do not co-here, emplacement-wise, into a typological whole. For instance, a text may be one of these atoms, and then a whole new expertise is then called into question, one that deals with oth-er wholes received as such, in fact, several such types of expertise: epigraphy, philology, lin-guistics, literary analysis. But the atoms in their aggregation are a different matter: that par-ticular text would have been found with sherds, bones, lithics, other texts, all in a matrix of soil that has its own distinctive nature, and connected only through the vagaries of a depo-sitional process that is, for the most part, unplanned.

#### 3.3 The Urkesh Global Record

The Urkesh Global Record (UGR) is a recording system that is meant to address explicitly these problems. For a general overview of the excavations we will refer to <a href="https://www.urkesh.org/glance">urkesh.org/glance</a>.

The foundation of the system is the Grammar of the Archaeological Record, which is found in the digital book <u>urkesh.org/grammar</u>. As of this writing, the book is still unfinished and will be completed during the tenure of the Balzan project mentioned above in note 1; it is, however, sufficiently indicative of the approach and of the goals we have set.

Some of the central concerns of the system may be given as follows.

- (1) While categorization systems are at the basis of all archaeological projects, what is specific to the UGR is the explicit concern with *emplacement* and a theoretical framework that *i*ntegrates emplacement with all other aspects of fieldwork and subsequent analysis. Emplacement is seen as the only factor that is exclusively and properly archaeological in nature, meaning that it presents a set of problems that no other discipline faces.
- (2) With regard to emplacement in particular, the record aims at being *global* in the sense that it includes in perpetuity every single observation ever made during the excavation, in the belief that this is the only objective data in view of the inability to repeat the experiment.
- (3) The *website* is adopted as the primary epistemic system, meaning that the data are recorded from the start in a format that is natively digital so as to be ideally suited for integration in the format of a browser edition.

The <u>urkesh.org</u> website is in effect a cluster of websites, called digital books, each one devoted to either individual excavation units or topical books. In the exemplification below we will use the ceramic topical book (<u>urkesh.org/ceramics</u>; this website is a work in progress, access is currently accessible with <cer> for both ID and PW) and some of the digital books dedicated to individual excavation units.

#### 3.4 Rosters and lexica

An essential aspect of the UGR, and one that is of interest to us here, is the categorization system that is used on the excavation for the analysis of the data with regard to their emplacement in the first place, and then their general morphological classification. The two major components of this system are the roster and the lexicon. There are different rosters and lexica for each class of data, with regard to both emplacement and typology; in this article we will use the ceramic data as an example. The whole ceramic inventory studied so far applies to different chronological horizons, spanning over some two millennia, and it applies to an inventory of some 300,000 between whole vessels and sherds that have been fully analyzed (out of a grand total of about one million that have been excavated and recorded as to emplacement, with only a preliminary typological analysis).

The roster is an -emic, closed system. Each slot records the typological categories that are seen as being mutually exclusive and congruent within the overall inventory of types. As an example, Fig. 1 lists the roster categories divided into various sub-categories. Thus, the twelve main shapes include vessels that are mutually exclusive in terms of how they can be handled and used: a bowl is primarily used for holding and can be held with two hands, a bottle for pouring with one hand, a jar for pouring with two hands, a cup for drinking, and so on.

The lexicon is an (e)-tic, open system. It includes an open ended list of attributes or definitions that are mutually exclusive in terms not of their neighbors, but of an extrinsic hierarchy that determines the distinctive traits of each element. Thus a jar may be high necked, with handles, with a spout, or with none of these traits, and in addition it would be of one ware or another, with decoration or without – but it would always remain a jar.

Rosters and lexica may be seen respectively as a list of variables and a list of variants. Thus the stoplight, as an -emic system, includes three variables, meaning that each of the three colors may vary within itself: "red" can include a variety of hues from pink to purple, "yellow" can include lemon to orange, "green" can include jade or emerald. These alternatives are the variants that would be listed in a lexicon, should one wish to have a lexicon for such color variants.

In other words: *variability* implies the possibility of *variation*. This possibility is limited, from an -emic point of view, by the reciprocal contrast among variables, and from an (e)-tic point of view by the (much wider) range of possible variants. Which in turn speaks to *singularity*. A variable retains its status whether it includes one or more variants, and whether it includes one or more specimens.

To summarize, we may list as follows the relevant terms:

306

307

308

309

311

312

313

314

315

317

318

319

320

322

323

324

325

326

327

328

329

330

331

332

variables	variants
categories	attributes

#### **INTERLUDE** 303

## 304 4 Typology in action

Typology is not an end in itself. The classification effort that produces it assigns categories that can be reconfigured in different ways. The common way in which this happens is through spreadsheet programs, such as Microsoft EXCEL, that produce tabulations in which the entries are sorted according to a variety of different categories. This is the normal process through which typology is put to use.

310 What we propose in the next section is the implementation of the website as a parallel, and more powerful, process for using typology. This approach offers not only a more dynamic use of the data, but also the possibility of integrating data and argument in ways that cannot be done with the printed medium. Not that databases are jettisoned, quite the contrary; they are fully maintained, but integrated in the broader framework of a browser edition.

316 We will discuss (6) the nature of the website as an epistemic system with regard in particular to the notion of typology, and will then elaborate (7) on the notion of digital discourse, which concerns specifically the way in which different arguments interact with each other. In each case we will use examples from the ceramic inventory of the Urkesh excavations.

#### Data processing 321

The categorization system that undergirds typology is essential for the data to be coherently processed. And such processing is universally understood today as being digital in nature. We can no longer think of a paper spreadsheet (even though the word retains a reference to the paper format), it is second nature for everyone to use a digital version – and this is inevitably based on a typology, however rudimentary and unexpressed it may be.

For our purposes, a dedicated programming platform is needed, one that adheres closely to the grammar and its rosters and lexica, and one that produces outputs suited for the kind of interweaving of planes which we will discuss. Such programming has been at the forefront of our efforts since the late seventies, undergoing a variety of transformations, and in this phase we are indebted to Bernardo Forni for overseeing the entire current suite of programs that underlie the system. The effort requires an in depth understanding of the

337

343

344

345

346

347

348

349

350

351352

353

354

355

356

357

358359

360

361

362

32

33

34

333 goals and the potential of the system, and Bernardo has proven to be as brilliant in this as

334 he is sophisticated in writing and documenting the programs.

#### II. THE PROCESS

## 336 6 The website as epistemic system

#### 6.1 The concept

The browser format, as exemplified in a website, is ideally suited to serve as an epistemic system, i. e., a system for the articulation and communication of knowledge. Essentially, we see websites as serving to integrate fragments into a new whole, and to develop a new type of argumentation, in ways that are not possible with printed publications such as books or articles.

We will mention here only two aspects of this system that are relevant for our current concerns: the concept of inter-planarity (6.2) and the way in which a dynamic inter-planar universe makes it possible to protect singularity (6.3). We will then give some examples from the Urkesh record (6.4 and 6.5)

#### 6.2 Multiplanarity

Critical thought is based on the concurrence of parallel planes of inquiry. With a printed publication this functions in the sense that, while one follows the argument developed by the author, one also develops parallel planes in one's own mind, adducing either additional facts or additional arguments, which may support or contradict the author's reasoning. This falls, in broad terms, within the notion of inter-planarity.<sup>3</sup>

Digital inter-planarity implies that these parallel planes are not only in one's own mind, but are accessible directly within one and the same structural system, the website. We must distinguish between multiplanarity and inter-planarity.

(1) The type in use in curent websites may be called *incidental multiplanarity*. The planes are there, but are independent of each other. In this case a hyperlink in a given text invokes a target in what exists indeed as a parallel plane, but the "invocation" is exhausted the moment the target is reached. It is the kind of epistemic activity we engage in when looking up a word in a dictionary: we are satisfied with the immediate answer the target gives, without concerns for the structural whole, the dictionary, and what else it may tell us (from a list of synonyms to matters of semantics or semiotics,).

<sup>&</sup>lt;sup>3</sup>We distinguish between multiplanarity as implemented in standard websites and inter-planarity as implemented in the digital discourse type of website. For the original notoin of multiplanarity see Buccellati 2022 "Transformative"; Paolini 2022 "Modeling". Paolini ws the first to call attention to the concept of planarity.

(2) The type we are arguing for is instead a *systemic inter-planarity*. Here the interaction among planes presupposes the whole plane as a target, and not only the detail with which we are incidentally concerned. There is a segment of the linear argument that is appropriated, but it is a segment that can be understood only within the larger context of its own plane. We will give examples below (7.2 and 7.4).

A website has the potential of articulating these planes concurrently. It has to be "written" with this in mind, and "read" in the same way. Such an approach to websites is not in current use. The current best, and extremely valuable, use of scholarly websites is one that sees them as *containers*: they are data bases of the known, from which one can draw at will individual data through aimed queries that extract information. When and if, in current use, a website develops an argument, it is in the form of texts that are analogical transpositions of printed texts, generally in a .PDF format. But these are essentially static, and so are the hyperlinks it contains: they direct us to a detail, but this detail is not incorporated in a narrative that is presupposed as such in the originating text.

#### 6.3 Protecting singularity: the dignity of the fragment

The notion of inter-planarity is relevant to the topic of our volume because it allows us to give full recognition to even the most minute and seemingly unimportant element while seeing it, at the same time, integrated in a web of related, parallel universes within which this small piece may otherwise lose its identity and disappear. A website, conceived as a new epistemic system, preserves the dignity of the fragment while at the same time allowing for the whole, in fact, for multiple wholes, to be identified in their fullness.

The way in which this happens is by placing the single fragment, regardless of how seemingly unimportant it may be, within a track or plane that is written exclusively for it, and which is invoked, precisely as a plane, by other planes without the fragment losing its identity. This creates a proper "digital discourse" among planes that gives pride of place to each plane in its own integrity, one that includes and enhances the singularity and integrity of the fragment.

#### 6.4 Sherds, wares and shapes

We will show here how this happens with regard to Urkesh ceramics, in terms of wares and shapes. We will illustrate this with screen shots taken from <a href="urkesh.org">urkesh.org</a>, which one may want to access for a full dynamic "reading" of the results. Also, the figures give the full view of a page as one scrolls down to the end: the purpose of the figures is to show the overall structure of the page, even if one cannot read the content in this printed version; to do so, one has to go to the website and actually scroll down to the end.

Fig. 2 shows a full page dedicated to a simple sherd from a cup (accessible at urkesh.org/A16q779-p1). This is the biography, so to speak, of this sherd, with full details about the emplacement and the typology.

Fig. 3 shows the frequency of the total inventory of sherds for unit A16 (accessible at <u>urkesh.org/A16-freq-ceram</u>). This inventory includes 59,834 vessels and sherds. Clicking on the "Fine Red-Orange Calcite Tempered" ware yields the total of 6,958 sherds of this ware, each sherd being accessible individually with its own page, just as for A16q779-p1.

Fig. 4 takes us out of the digital book for unit A16 and into the topical book on Ceramics (<u>urkesh.org/cer</u>) and shows the full description of the "Fine Red-Orange Calcite Tempered" ware (accessible at <u>urkesh.org/cer-RC1</u>).

Fig. 5 shows the detailed analysis of the ware in the Ur III-Isin Larsa period with the illustration of some sherds, including our sherd A16q779-p1 (both are accessible by scrolling down on the page <u>urkesh.org/cer-RC1</u>).

Fig. 6 remains within the topical book of ceramics, and shows a synopsis of all the bowl shape types from the urban contraction horizon at Urkesh, which corresponds to the Ur III / Isin-Larsa period in Mesopotamia (accessible at <a href="urkesh.org/ceram-horiz-IL">urkesh.org/ceram-horiz-IL</a>).

#### 6.5 Densities

Tabulations are the universal method through which typology is used: even the simplest spreadsheet is based on a categorization system, and tabulations correlate the data, which have been so categorized, with sorts and computations that provide a variety of clustering representations. It is only because of this underlying categorization that it is possible to arrive at sophisticated methods of statistical analysis such as are now current in archaeology.

Statistics speaks to density (2.5). What a website approach offers is the opportunity to have these densities speak to each other and to descend dynamically from the totals to the individuals, and to have them interact with arguments that are properly digital and embedded as such in the website.

For an example, we may look at Fig. 3. We see that in our total population of 59,834 ceramic vessels and sherds, from the excavation unit A16 there are only 10 for the Early Trans-Caucasian ware, for an effective percentage value of 0%, in contrast with the highest density of 57% for the Chaff Tempered ware. And here, too, we can have a full page for each one of the ten sherds, as in the case of <a href="https://www.urkesh.org/A16q314-p5">urkesh.org/A16q314-p5</a>. Using the formula given above (2.5) the two figures for Early Trans-Caucasian ware and Chaff Tempered ware would be rendered respectively as  $0\%^{590}$  and as  $57\%^{590}$ .

These figures must be assessed with regard the total volume of the excavated area in the unit in question, A16 (see <u>urkesh.org/A16-overview</u>). It consists of eight 5x5 squares, for a

total of approximately 350 cubic meters, with a large number of elements recorded (see 433 <u>urkesh.org/A16-dataset</u>).

Obviously, the value of any density calculation has to be assessed in relationship to the universe of data to which it applies (see 2.5). In this regard, it is especially meaningful to consider the value of a statement of non-occurrence. For example, in the Urkesh ceramic roster, based, as we said, on some 300,000 specimens, we have no dishes. Should even just one single dish be found, we would have to create a roster category for it, and we would then also create a single lexical definition of this one lonely example. On the other hand, the lack of dishes implies a type of functional use of foodstuff that it correlative to the kind of flatware that might have been used, if any, and which certainly did not include any knives.

## 442 7 Digital discourse

#### 7.1 The concept: the fragments and the whole

With the term "digital discourse" we refer to the way in which a website goes beyond serving as a mere container, however well articulated, and develops mechanisms for the articulation of a narrative flow among the data and their interpretation (<u>d-discourse.net</u>; this website is a work in progress; access is currently possible with <discourse> for both ID and PW). Digital discourse presupposes distinct but congruent categorization systems (2.3), and it provides avenues for them to "discourse" with each other.

Typological analysis aims at distinguishing and breaking up a given universe into categories and attributes. Statistics (6.5) is the dominant way in which these fragments are recomposed into a whole. Digital discourse, as implemented in a website, adds a major new dimension: it allows the integration of statistics into a higher level argument that is made possible through inter-planarity (6.2), and in this fashion one can obtain a different understanding of the underlying relationships. It is in this sense that digital discourse shows the full epistemic value of the website model.

We will first stress the importance of dynamically integrating argument and data into a unified system, or "discourse," rather than just aiming for a static juxtaposition as is the case with current scholarly websites (7.2). This entails an understanding of non-linearity as being in effect *multi*-linear, in ways that are complementary with inter-planarity.

Examples from the Urkesh Global Record are given, to show how we may obtain a narrative flow that integrates data and interpretation (7.3) and how we may create conditions through which a dynamic interaction among different planes is possible (7.4).

#### 7.2 Multilinearity

In current practice, when a website develops an argument it does so in ways that are not properly digital. It is essentially a static text, whether as an analog of a printed text (such as .PDF) or as formatted text in a browser format (such as .HTML). They do certainly contain hyperlinks to other texts or to data (this is formally embedded in the HTML acronym). But these links are anecdotal, they are not systemic.

A systemic link means that the target is written concurrently with the invoking text, and vice versa. Such concurrence is functional: the invoking argument expects the target argument to unfold as an argument, and expects therefore its premises and its conclusions. This applies also to the data, in the measure in which a single piece of information is expected to be based on a given system which is shared by the invoking argument.

An anecdotal link, on the other hand, invokes only the detail. It does, certainly, assume that this detail is embedded in a rational overarching system, but this system is not shared, much less has it been written in function of the invoking text. The target is not the system, but only the detail. It is not inter-planar.

Now, an argument is intrinsically linear, in the sense that it leads linearly from a premise to a conclusion, adducing facts and references to other arguments in the process. Multilinearity is what makes it possible for inter-planarity to develop into a proper discourse. Inter-planarity implies in fact the existence of parallel linear tracks, written in view of each other, where each expects a given linear development in the others. It is in this sense that we can speak of multilinearity. As one "reads" a website, the argument flows integrating linear segments from parallel tracks into a single unitary linear (or "multi"-linear) sequence (Fig. 7).

One normally speaks of "non-linearity" as a positive aspect of website use. It is indeed what we are all accustomed to doing, and the term "browsing" describes it well: one "nibbles" at this or that, without an explicit sequential flow. We do not in fact "read" websites in common use: we "surf" on the surface, led by the casualness of what carries us. With scholarly websites it ought to be different, and that is what our websites are doing. We will see now, with some concrete examples, what this means for a study of typology.

The difference between incidental multiplanarity and systemic inter-planarity (6.2) depends precisely on the application of multilinearity. Planes exist, *in nuce*, in any form of critical thought; they also exist, digitally, in any website, as soon as there is a link. For a systemic approach to inter-planarity, however, it is necessary to expect multilinearity linking the planes, For an example see below, (7.4).

#### 7.3 A narrative flow – staccato and legato

The UGR categorization system based on the (e)-tic and -emic systems is the presupposition for digital discourse, and this particularly in an archaeological context. By applying rigorously this categorization, especially at the moment of excavation (see above, 3.2), the data are not only available as such; they also enter into a narrative flow that distinguishes in essential ways this presentation from what happens with a database (Buccellati 2020 "Degrees"; Buccellati and Kelly-Buccellati 2020 "Narratives; forthc. "Clustering"). The difference lies in the fact that the individual data are not given as independent cells, where one cell relates to the other only in function of the slot in which it fits within the overarching architecture of the database; rather, the individual data are linked to each other through a sequential logic that highlights the interdependence of the data.

In the Urkesh Global Record we have two types of such narrative, which we define with terms derived from musical notation.

The staccato type consists of segments that are created individually, at different moments in time and by different people. It is by virtue of the categorization system, which defines each segment at the moment it is created, that these segments cohere into a sequential flow. We may refer to Fig. 2 for an example of this type (and one would have to look at the website page <a href="urkesh.org/A16q879-p1">urkesh.org/A16q879-p1</a> for the details). There is a logical sequence that rests on a variety of different inputs. After the first two sections that give a general overview and details of identification, there are:

- a section on stratigraphy that gives first the date of discovery (2002-9-23) and the feature within which it was found (the feature can be accessed for full detail), and then the assignment to a stratum, which was finalized at a much later date (2015-5-20)
- a section on typology that gives details of measurement and morphology, written up in the ceramic lab some three weeks after discovery (2002-10-12)
- a section on reference which gives photos (taken in 2006 and 2009) and a drawing (done in 2003).

The legato type is instead produced as a normal text file that corresponds to a single display page. It is written as a whole, generally by a single author; there may be revisions and additions, but the page as such is always seen as a complete whole, not as an aggregate of segments coming from different sources. What distinguishes it is the fact that the target is seen as an integral part of the invoking argument. Thus, for example, Fig. 8 (accessible at <a href="urkesh.org/J5-surfaces">urkesh.org/J5-surfaces</a>) gives a page from unit J5 which describes one type of emplacement (horizontal surfaces), with links to each of the features in the unit that fit in that category. Thus the link to f288 (Fig. 9; accessible at <a href="urkesh.org/J5f288">urkesh.org/J5f288</a>) expects a full knowledge of that feature, as if it were included in the legato argument, while conversely, a link in the f288

page expects full knowledge of the other surfaces defined by discrete components as presented in the pertinent page (Fig. 8).

The relevance for typology will be clear from the example just cited. The notion of pavement as a "surface defined by discrete components" is made clear not just by a verbal description, but by examples of the same type as exposed in the same excavation unit (and, at a different referential level, for the entire site). A "pavement" is not just a label accompanied by a number giving the total of actual occurrences; it is an assemblage of concrete individuals, seen in their individuality and identity (8.2).

#### 7.4 Interaction among planes – inter-planar and multilinear

The multilinear approach constructs a full argument that expects portions of the argument to come from parallel planes. Only if it is multilinear does inter-planarity serve the purposes of digital discourse (7.2; Buccellati and Kelly-Buccellati forthc. "Clustering").

Fig. 10 shows diagrammatically how the examples we have seen earlier fit into this picture. The "biography" of a simple sherd (A16q979-1), given as plane A, is placed in the much larger context of the RC1 ware, both in terms of the density within the entire ceramic inventory for unit A16 (plane B) and in terms of the precise definition of the ware itself (plane C, from the distinct website dedicated to ceramic typology). Planes B and C are invoked from within plane A, and so are a number of other planes, in particular the one about the emplacement of a given feature and the one about the stratum and phase to which the feature (hence the sherd) belongs. In other words, a variety of different planes, from different digital books or websites, are all tightly interlaced and "written" with each other in mind.

Once again, the relevance for typology will be clear from the example just cited. The assignment of this sherd's ware to category RC1 is much more substantial than a simple code in a spreadsheet. The interaction between staccato and legato types of text is particularly indicative of the way in which typology emerges not just as a list, but as an argument, one that is construed from a variety of different inputs that cohere into a single whole.

A website conceived as a digital discourse, then, presents multiple parallel tracks which are "written" as arguments that are linked *qua* arguments, and not just mined for individual, isolated words or concepts. The difficulty lies in developing a disciplined approach to "writing" and "reading," which we put here in quotation marks precisely because the notion is alien to our current mindset regarding websites. Currently one "constructs" a website, meaning that one prepares a container which can be altered at will and within which one can place data at will; one does not "write" it as a cogent narrative that develops an argument from a premise to a conclusion. And one "uses" a website by "surfing," "browsing,"

577578

579

580

581582

583

584

585

586

587588

589 590

- 570 "querying," but without ever gaining a sense of the whole without ever even thinking of 571 "reading" it "from cover to cover."
- 572 7.5 Big data
- The notions of inter-planarity and multilinearity add one important dimension to what is known as big data.
  - On the face of it, the term may be taken to refer simply to a large inventory of items. But that would be misleading, because what is really meant is a very rich and complex categorization system, based on rosters and lexica, so that, for any single item there is a highly differentiated set of attributes that fit into an equally highly differentiated set of categories, exponentially multiplying the amount of detail that can be set in correlation with each other.

Digital discourse adds one more degree of complexity, and a major one at that: the correlation among planes, a correlation that is made explicit because of the flow among planes that multilinearity makes possible. In other words, multilinearity affords a higher degree of control on an even greater amount of detail than either the sheer size of the inventory or the complexity of the categorization system imply. The Urkesh Global Record may rightfully be seen as a good example of such inter-planar and multilinear big data system in action: the interaction among planes within a single digital book, or website, and then among parallel digital books, happens on account of the inter-planar and multilinear design with which they have been written, always allowing for a fully dynamic interaction among the overall narrative on the one hand and, on the other, the most minute of details that is found in the inventory.

### 591 **CONCLUSION**

## 592 8 Variability and singularity

- 593 8.1 An assessment
- By way of conclusion, we may now consider how all of this affects a discussion about variability and singularity. In this article, we have emphasized two major aspects.
- The first about the data (sections 2 and 3) is the relationship between (e)-tic and -emic systems. The distinction is fundamental for an appreciation of what typology is, and while it stresses the difference between the two, it also shows how inextricably linked they are: there cannot be one without the other.
- The second about the process (sections 6 and 7) deals with the role that the website, as a re-imagined epistemic system, must have in allowing for this reciprocal distinctiveness and interaction to be fully implemented. The dynamics intrinsic in such a system is structurally

different from other static systems, in essence from the printed page or its analogous electronic equivalent. It is different because it allows variability and singularity to talk to each other in real time.

When applied to typology, this approach yields a much richer picture, which we may here view from two particular perspectives (8.2 and 8.3), both of them speaking to the dignity of the fragment (see above, 6.3).

#### 8.2 Anonymity and identity

The first perspective is the one through which we can best appreciate the identity of the fragments in their singularity and distinctiveness. Multilinearity in the sense described above (7.2) means that every single fragment, even the smallest sherd, is present at all times in its full identity. A detail of Fig. 3 deserves special attention in this regard: every single sherd that is included in the total of 59,834 for unit A16 is accessible through the list in the lower portion of the right hand sidebar.

The interplay between roster and lexicon (3.4) can best be appreciated in this light. The insistence on typological classification does contribute to the loss of individuality. We certainly need to reduce a real item to the status of a type in order to control how it is is distributed in a given universe. But we must be aware that we are in fact "reducing" it. On the one hand we are contributing to a better understanding of its properties by seeing it in its similarity to others. And yet, on the other hand, we must not make of the reduction the ultimate tool for understanding. Variability tells us where the fragment fits, but singularity can still only be appreciated outside of a controlled variability. With some exaggeration, we may say that even the most non-distinctive of sherds has a dignity of its own that should not be jettisoned and lost in the anonymity of group membership. We must remain aware of the fact that the statistical whirl, while pivoting around individuals and thriving on variability, inevitably erases their very singularity and individuality.

To go back to our two examples of the phonemic system and the stoplight (2.2), we may say that the timber or tone of the voice of a friend will remain open to our perception regardless of how clearly we may define the phonemic classes or the acoustic parameters. And even a modest stoplight may acquire as if a personality once we notice that, at a familiar intersection, the welcoming "Go!" sign has, for some unknown reason, gained greater brilliance and turned into a brighter shade of green...

### 8.3 Density and globality

The second perspective concerns the notion of density (2.5). Here the dignity of the fragment emerges in a different light: precisely because every fragment counts, nothing is discarded without being recorded, so that density is not based on a sample, but on the totality

	roster		lovicon
main category	secondary c	lexicon (entries)	
		bowls	286
		cups	51
		stands	14
		pitchers	3
		jars	258
	type	workdisk	1
		platters	28
		other	12
chana		pots	23
shape		strainers	11
		bottles	18
	family	•	17
	sub-family		21
		rims	60
	rima hasa handla	handles	18
	rims,base,handle	bases	28
	other		2
	forming techniques		
	main ware	25	
	main temper	1	
	other temper	1	
	surface (interior)		8
	surface (exterior)	7	
ware	inclusion	7	
	inclusion frequenc	3	
	firing	4	
	fracture	8	
	subfamily	1	
	technique		5
	applique	9	
	bitumen	8	
decoration	incised	10	
	Khabur painted	14	
	Mittani painted	32	
	stamped	3	
measuremer	1		
color			10
function			8
assignment			5
7	36		1025



#### Back to top: A16q879-p1

#### 2. IDENTIFICATION

Labeling

Roster	Date	Author	Record
Reference to a-lo	t !!	!!	a879

#### Designation

Roster	Date	Author	Record
Definition	2002-10-12	hh	cup [Input ZA823S.j]
Quantity of components	2002-10-12	hh	[Input: ZA823S.j]

Back to top: A16q879-p1

#### 3. STRATIGRAPHY

#### Contact Association

Roster	Date	Author	Record
Type of contact: contemporary events/movable items	2002-09-23	!!	q879-p1 (cup) sits in f325 (^ly) [Imput: MX12LR.j]

Time Sequencing

Roster	Date	Author	Record	
Stratum (to which element belongs)	2015-05-20	!!	S380AAH [Input: ZA520CJC.j]	
Phase (to which element belongs)	2015-05-20	!!	h5cAAH [Input: ZA520CJC.j]	

Back to top: A16q879-p1

#### 4. TYPOLOGY

#### Measurements

Roster	Date	Author	Record
Width 1 or diameter of rim	2002-10-12	hh	11 [Input: ZA8238.j]

Morphology

Roster	Date	Author	Record	
Ware or Material, species	2002-10-12	hh	Fine Red Orange Calcite Temper	Ware [Input: ZA8238.j]
Shape, form	2002-10-12		cup	[Input: ZA8238.j]
Zca.MEASUREMENTS: Percentage of rim	2002-10-12	hh	17.5%	[Input: ZA8238.j]
Zca.SHAPE: Family	2002-10-12	hh	conical (cup)	[Input: ZA823S.j]
Zca.SHAPE: Type	2002-10-12	hh	2	[Input: ZA823S.j]



General Intro
home: unit A16
preface
shortcuts
Urkesh website
Mozan sitewide
area AA
table of contents
page layout

## A16 excavations (contract/expand)

INTRODUCTION

STRATIGRAPHY
TYPOLOGY
BUILT ENVIRONMEN
OBJECTS
overview
ceramics
chaption

glyptics
epigraphics
figurines
lithics
metals
other

## A16 tabulations introduction

indices
indices
frequencies
data set
logs
special tabulations
correlations
strat.summaries

sitewide

## A16 Frequency computations

Introduction

Installations and deposits definition

definition phase stratum

#### Objects and samples

ware or material
phase
feature

#### Ceramic vessels

ware
overall shape
shape family
shape sub-family
shape type
details of shape

stratum feature body sherds

#### Glyptics

stratum/feature aggregate

### A16 Frequency Computations Ceramic vessels and sherds in order of frequency by ware

#### Processed on 10-10-2016

Note. Included in this category are ceramic vessels, whether complete objects or sherds, and whether <u>items</u> or <u>q-items</u>.



Grammar
overview
the Grammar
the Manual
shortcuts

A16 constituer introduction -incidentals a aggregates b assemblages c composites f features

b assemblages
c composites
f features
h phases
i items
k loci
p plots
q lots
qb bones
qi q-items
qp pottery
r relays

Finer Red-Orange Calcite Tempered

sequential q3-p92 q3-p93 q3-p94 q3-p95 q4-p77 q4-p78 q4-p79 q4-p80 q6-p82 q6-p83 q6-p84 n7-n78 q11-p87 q11-p88 q11-p89 q11-p90 q11-p92 q12-p80 q12-p81 q12-p82 q15-p76 q15-p77 q15-p78

<u>q20-p90</u>

q17-p81

q17-p82 q20-p89

950-p31 950-p33 950-p34 950-p40 950-p45 950-p47 9951-p8 9951-p9 9951-p10 9951-p11 9951-p12 9951-p13 9951-p14 9951-p14 9951-p15 Elements: 1932

Components: 5958

## Ceramic Typology shortcuts home Urkesh website THE BOOK preface authorship page layout left and right hand side relationship to UGR THE DATA the Urkesh corpus history of the project Categorization THE SYSTEM introduction roster and lexicon ROSTER types and attributes roster LEXICON introduction lexicon: overall wares shapes decoration color measurement function time assignment

## Ceramic wares

### 1. Introduction

overview principles synchrony/diachrony

### 2. 4th mill.

introduction LCH: Chaff Temper Ware LGR: Grav Ware LCF: Fine Ware LCT: Tan Slipped Ware LCGb: Green-

## 3. 3rd-2nd mill.

Buff Slipped Ware

introduction NIN: Ninevite 5 Ware CH: Chaff Temper Ware FC: Fine Chaff Temper Ware

RC: Calcite Temper Ware

RC1: Fine Calcite Temper W

R: Rough Ware

P: Pebble Temper Ware

FP: Fine Pebble Temper War

S: Simple Ware

WS: Wet Smooth Ware

G: Grav Ware

M. Metallic Ware

IM: Imitation Metallic Ware ETC: Early Transcaucasian V

BC: Bi-Color Ware

DBR: Dark Brick Red Ware

## **Attributes: Wares** 3rd and 2nd Millennium

## Fine Red Orange Calcite Temper Ware (RC1)

April 2006 - Marilyn Kelly-Buccellati Updated May 2016 - L. Recht

General Description Characteristics Phase 4 - Akkadian Sherd Illustrations Phase 5 - UrIII/Isin-Larsa Sherd Illustrations Phase 6 - Khabur Sherd Illustrations Phase 7 - Mittani Sherd Illustrations

## **General Description**

Fine Red Organge Calcite Temper ware is very similar to the Red Orange Calcite Temper ware with a red-orange color and much calcite temper. The finer RC1 type can be distinguished from the RC by the finer temper. Additionally, the shapes are also smaller and finer than the RC types.

## Horizons Introduction Index 2. Temple Terrace (LC3) 3. Urban growth 1 (EDII) 2 (EDIII) 4.Great alliances (Akk) 5. Urban contraction (Url Shapes

# INTRODUCTION

DATABASE INDICES AND SYNOPSES phase stratum feature ware shape: overall

## Whole vessels

sub-family

INDICES ONLY

family

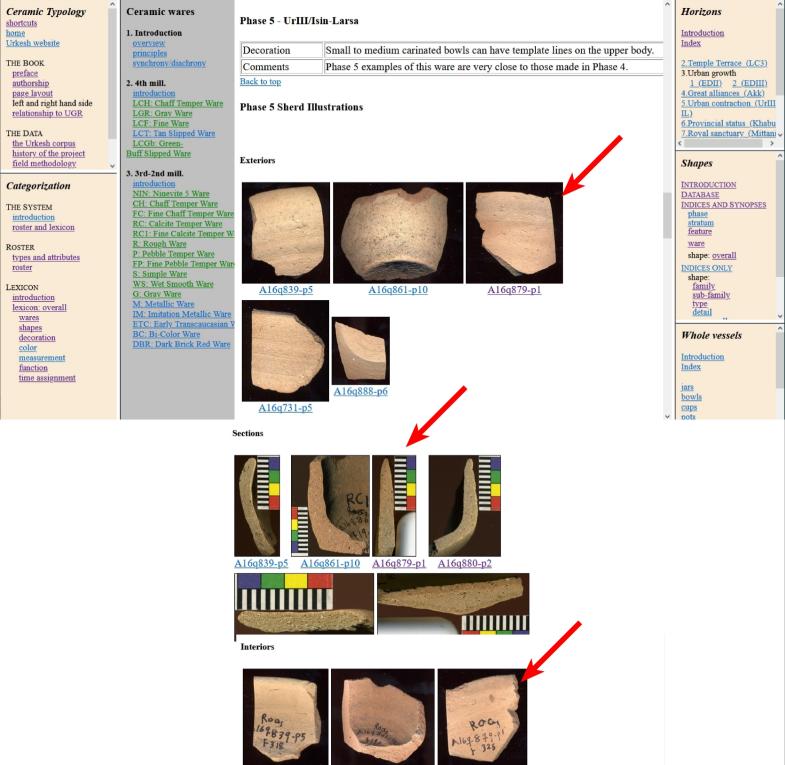
shape:

Introduction Index

jars

bowls cups

pots



A16q861-p10

A16q888-p6

A16q879-p1

A16q839-p5

A16q731-p5

131-F



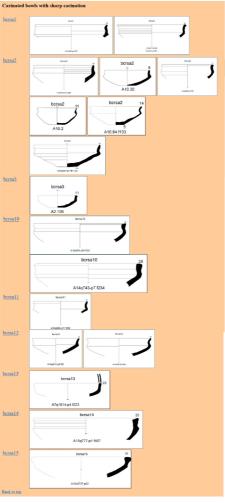
Back to top

Carinated Bowls

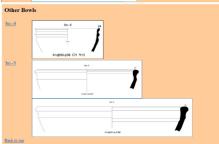
berrl

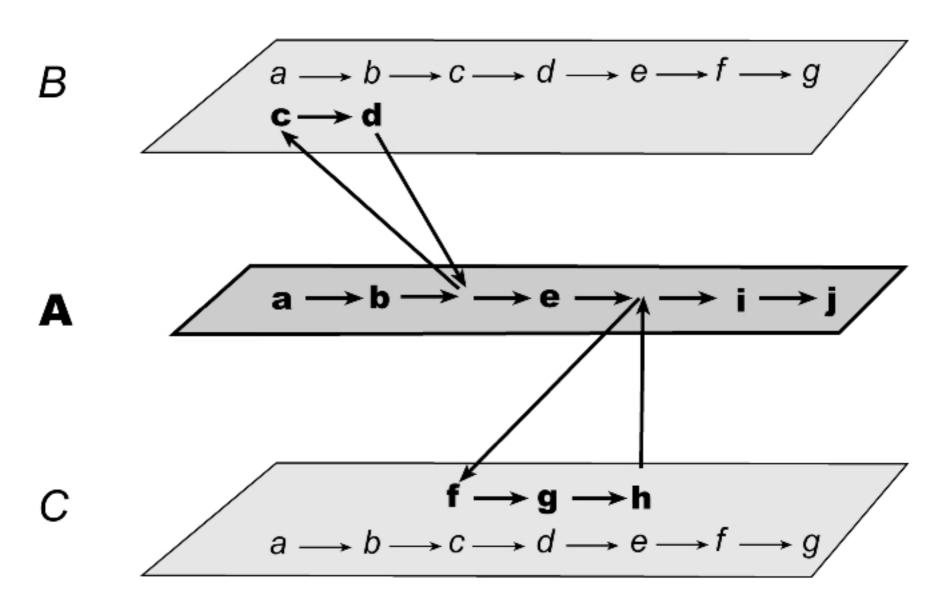
berr2

berr3









**Emplacement** General Intro introduction Urkesh website Mozan sitewide ELEMENTS. statistics surfaces table of contents layering amorphous ordered **FEATURES** IN CONTACT statistics J5 excavations 1.cut (contract/expand) 2.1ean 3.abut INTRODUCTION 4.cover STRATIGRAPHY 5.cap emplacement ITEMS IN CONTACT statistics BUILT ENVIRONMENT Conservation PRESENTATION J5 tabulations strat.summaries

preface

shortcuts

area AA

page layout

topical index

principles

volumetry

deposition

TYPOLOGY

OBJECTS

SAMPLES

INTEGRATIVE

introduction

indices frequencies

data set logs

phases/strata

unit J5

## Emplacement for unit J5 Horizontal surfaces

September 2011 - James Walker The home for this page is J5

Surfaces defined by discrete components (pavements) Surfaces defined by soil compaction (floors)

In the Main Lexicon there are 19 defined horizontal surfaces reflexing alignment in a single plane. There are 5 pavements defined by their composition and 14 floors generally defined by the degree of compaction.

Few floor surfaces and pavements have been excavated in this unit, despite a large amount of construction over several millennia. There were two pebble pavements, £288 and £247, associated with the Early Dynastic period first escarpment, ^esc1. In the Mitanni period there were a number of flows - most were associated with a monumental staircase contructed during the late Mitanni extension of the sacral complex to the west. Floor f45 is one example Back to top

#### Surfaces defined by discrete components (pavements)

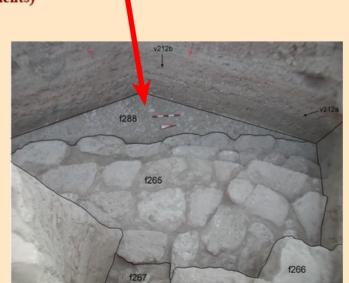
The most significant pavement, £288, was located to the southwest of the revetment wall and its protective escarpment. Further to the east there was a similar pavement, £249 at a similar, but slightly higher elevation. There was insufficient exposure of either to determine what purpose they may have served or how

they may have been related. Another pavement, £246, was associated with the early Mittani staircase, "strc1 that provided access to the revetment wall top. One other Mittani pavement f247, directly covered Early Dynastic pavement f249 (described above) with no evidence of intermediate occupation.

These five pavements excavated were classified as pavement type C. A brick surface, f196 is problematical. It may be a pavement type B, or the top of a substantial wall. Unfortuantely excavation was paused before the true nature of this feature could be determined.

Back to top

Surfaces defined by soil compaction (floors)



Grammar

the Grammar

J5 Constituents

the Manual

introduction

a aggregates

incidentals

f features

h phases

items

k loci

p plots

q q-lots

r relays

s strata

z zeros

qi q-items

qp pottery

t templates

v view photos

specific labels

typological

data base

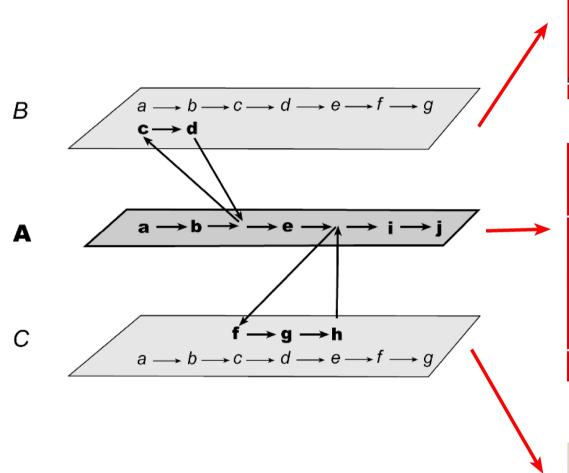
w view drawings

shortcuts

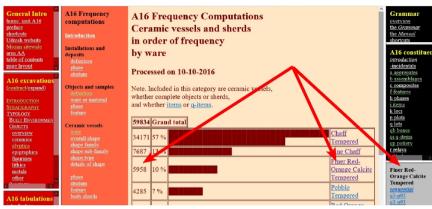
overview

Please click on one of constituents above





#### ware frequencies



#### sherd "biography"



#### ware description for a given horizon

