Ocula_temi
Occhio semiotico sui media | Semiotic eye on media
www.ocula.it • ISSN 1724-7810 • Vol 24 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

30 Visualità della scrittura nelle arti e nella comunicazione visiva
a cura di Valentina Manchia e Salvatore Zingale

Alternative graphemics

Aztec writing system as a case study towards an integrated, digitalised model of non-typographic graphemics

Antonio Perri
Dipartimento di Scienze umanistiche, Università degli studi Suor Orsola Benincasa, Napoli, IT
antonio.perri@unisob.na.it

Luciano Perondi
Dipartimento Culture del progetto, Università Iuav di Venezia, IT
lperondi@iuav.it

Daniele Capo
CAST Società Cooperativa
daniele@c-a-s-t.com

Roberto Arista
Typotheque
roberto@typotheque.com

Giampiero Dalai
Alpaca Società Cooperativa
giampiero.dalai@alpacaprojects.com

Abstract
Aztec pictographic writing conveys semantic (and linguistic) contents through conventional patterns of units. We argue that relevant visual values are ordered in subsets (topological, spatial and logical). Each unit is a "character" (in the Unicode vocabulary), either as a positive unit (a graphic sign) or as an "empty" spatial, topo-graphical relation between signs. It is sketched, then, a digital method processing sets of basic graphic features as non-linear combinations of variables; they display an overall structure similar to Unicode encoding of emoji. Finally, more coded variables are set to account for broader compositions, since coding systems available so far are not flexible enough to accommodate the encoding of Aztec units.

Keywords
Aztec writing; Entaxis; Synsemia; Unicode standard; Non-linear digital typography

Contents

Ocula_temi
Occhio semiotico sui media | Semiotic eye on media
Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

1. Aztec writing eludes the constraints of linguistic linearity

Aztec writing shows an internal structure which is specifically designed to convey semantic (and linguistic) contents via conventional (but sometimes also figurative-analogic) non-linear patterns of units which can be subsumed to a digital encoding.

Indeed, there are several visual variables structured in the textual space, combining logo-syllabic with non-linear and analogic graphic units in an optimal blending of two visual systems or strategies: while names, dates and numbers are conveyed by logograms and syllabograms, further linguistic information is encoded by systematic use of layout and spatial disposition. Writing, unlike speech, does not need to represent phonetically an utterance in order to provide a complete encoding of linguistic content (Zamora Corona 2022). On the contrary, visual elements spatially arranged in a non-linear toposyntax (Klinkenberg & Polis 2018) do indeed signify in an often more effective way than a chrono-sequence of letters.

We used the seven variables by Bertin (1967) to single out entactic units (Vaillant 1999), and synsemic composition. These minimal units, in turn, can be combined in a multirank system giving rise to complex and hierarchically structured graphic units. The arrangement of elements is obtained through (1) associative (visual sets) and dissociative (visual hierarchies) variables; and through (2) topological and/or spatial patterns or set of rules (frames). (1) and (2) are also at stake when broader visual (or textual) arrangements of graphic units are dealt with.

The purpose of the article is to describe in detail the actual working of these principles. We argue that it is possible to show sets of relevant, basic graphic elements in Aztec writing showing a systematic correspondence both with specific visual variables, and with linguistic contents. Each graphic unit, indeed, is thought as a combination of basic elements. The system so conceived displays an overall pattern which has some similarities to the Unicode encoding of Emoji (Unicode Inc. n.d. a), but the latter is not flexible and multifaceted enough to accommodate a proper encoding of Aztec semantic units.

Moreover, some variables can be manipulated in a discrete and finite range of possibilities while others display a continuous (and thus non-finite) range of variability; specific variables – such as Shape – seem to have a virtually infinite variability and a hardly discrete set of variations. Our hypothesis is that at least some of these visual variants can be organized in more or less abstract subsets (topological, spatial, and logical).

We suggest that every semantic unit can be seen as a “character” (as in the Unicode vocabulary), regardless of its being related straight to a glyph (i.e. a graphic sign) or to a (formal, “empty”) spatial or topo-graphical relation between graphic signs. Indeed each character can be composed (and combined) entactically, i.e. non-linearly; as a consequence, the difference between entaxis and synsemia (Perondi 2012) will depend on the levels-ranks of combination between units under analysis, while the overall composition model stays the same.

According to our hypothesis, the internal structure of Aztec writing is specifically devised in order to go further and elude in many ways the constraints

Ocula_temi
Occhio semiotico sui media | Semiotic eye on media
Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

of linguistic linearity (i.e., that of vocal signifiers). While undoubtedly representing Aztec language, indeed, i.e. Classical Nahuatl, it is supposed to convey semantic (and linguistic) contents through codified patterns of units (Fedorova & Perri, in press; Zamora Corona 2022). In this article, we argue that those units and framed assemblages are clearly detectable, while often iconically grounded; moreover, we will show that Aztec pictorial characters or basic units, usually named *glyphs* by Mesoamerican scholars[1] – in spite of their having been often referred to as integrating a dense system (in the sense of Goodman 1968), therefore not possessing the basic quality of linguistic symbols – articulate framed patterns which can be made discrete and, ultimately, digitalized.

We will detail how the system is grounded on several visual variables, which in turn frame the textual space combining logo-syllabic with non-linear and analogic graphic units in an optimal blending of two strategies of linguistic information display – embedded in the same artifact: while names, dates and numbers are conveyed by logograms and syllabograms – articulated mainly in the form of non-linear or entactic emblems (Vaillant 1999; Fedorova 2009), but rooted in segmental linguistic units such as morphemes and syllables – further linguistic information is flexibly encoded by systematic use of layout and spatial disposition of characters. Such a framing practice, in the spatial toposyntax of pictographic texts, is supposed to provide readers with a definite order of information-content processing; but, more important, it makes readers strive to locate a coherent hierarchy of processable contents. Finally, and of the utmost importance in our view, we suggest that layout and non-linear arrangement also encode glottic and textual cues for interpretation and reading of linguistic content: as Roy Harris (1998; 2003) and more recently Zamora Corona aptly noted writing, unlike speech, does not need to represent phonetically an utterance in order to provide a complete encoding of a given linguistic content (Zamora Corona 2022). On the contrary, in this reversal of the longstanding old paradigm seeing writing as visible speech (see e.g. DeFrancis 1989) visual elements spatially arranged in non-linear patterns (Klinkemberg & Polis 2018) can signify in an often more effective way than a linear chrono-sequence of (glottic) letters.

2. A new encoding system for Aztec writing

As stated above, it is possible to feature sets of relevant, basic graphic elements in pictorial writing which show a systematic correspondence both with specific visual variables and with linguistic contents. Each graphic unit, indeed, is thought and articulated as a combination of basic elements (lines,

[1] In this article we will not use, however, such a notion in the technical sense of Americanists and specialists of Mesoamerican iconography: indeed, there is a concrete risk of conflating this term with "glyph" as meant in digital typography, i.e. in Unicode terminology. According to Unicode Consortium (Unicode Inc. n.d. a), a glyph is any token-instantiation (via a definite font) of a coded abstract character. When mentioning the type-units of Aztec pictorial writing and variables as coded in our model, therefore, we will label them as characters (in the sense of Unicode standard terminology), while of course they are still named glyphs in the jargon of Americanists. It is obviously not relevant here to state if we could suppose, in the digital domain, that different Aztec characters-as-types will be materially instantiated by different scripts or notational forms, actually corresponding to one and the same character.

surfaces, colors). The system so conceived displays an overall pattern which has some similarities to the Unicode encoding of Emoji (Unicode Inc. n.d. a); but the latter is not flexible and multifaceted enough in order to accommodate a proper encoding of Aztec pictorial and semantic units.

Moreover, we suggest that while some variables can be segmented in a discrete and finite range of values (hence, of possible manifestations in texts), others – more similar to dense structures in Goodmanian terms, as stated above – display a continuous (thus in principle non-finite) range of variability: therefore specific variables – such as shape – seem to have a virtually non-finite variability, and a hardly discrete set of positive values. Despite this we argue, as stated above, that it is possible to account for at least some of the visual variants involved in terms of more or less abstract and discrete subsets (topological, spatial, and logical).

We posit, then, that any semantic unit is to be seen as a "character" (following the Unicode terminology, see footnote 1), regardless of its being related to a possible, perceptually visual item (i.e. a graphic sign, manifested by a positive token) or to a (formal, thus "empty", not visually manifested by an individual material feature) spatial or topo-graphical relation *between* graphic signs-units. Indeed, each character can be composed (and combined) entactically, i.e. through a non-linear assemblage of units; as a consequence, the difference between the *entaxis* of a bundle of agglutinated units and the *synsemic display* seen in larger portions of the text will depend on the levels-ranks of combination between specific units under analysis, without affecting the overall compositional pattern.

The coding method we propose is articulated in two main levels (albeit strictly related, and ultimately merging one into the other): elsewhere we have called them *entactic* and *synsemic*. The former, low and bottom-up level is to be seen as coding agglutinated units in order to form logograms, simple phrases or formulas; while the latter, high and top-down level encodes the overall organization of units. It therefore puts in significant relation two, or more, entactic compositions of units – or multiple basic units.

Of course, if we limit our approach to a "classical" typographic strategy – thus coding written entactic units, but only insofar they can be sequentially arranged in clusters of growing complexity, as in the Unicode standard for glottic scripts – then the complete meaning (and reading) of the text will be lost since, as we stated before, characters and groups of agglutinated units are always spatially arranged according to a non-linear toposyntax (Klinkenberg & Polis 2018) whose role is intrinsically linguistic.

In order to provide a solid theoretical background to the choice of visual variables, manipulated to single out both entactic units (Vaillant 1999) and synsemic composition, we resorted to the seven variables devised by Bertin (1967). These variables, in turn, are combined in a multirank system giving rise to complex and hierarchically structured graphic units.

The arrangement of elements is obtained through (1) associative (visual sets) and dissociative (visual hierarchies) variables; and through (2) topological and/or spatial patterns or sets of rules (frames). (1) and (2) are also at stake when any broader visual (or textual) arrangement of graphic units is dealt with.

Ocula_temi
Occhio semiotico sui media | Semiotic eye on media
Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

More specifically, we propose that Bertin's seven variables are related to both the entactic and synsemic layers of articulation of any text.

Variables and layers set up in order to properly encode Aztec pictorial and semantic units are summarized in Table 1.

Layer	Sub-layer	Operationalization
a. Entaxis		
	a.1 – Shape	*Characters meant to create entactic compositions. E.g. "MAN", "WARRIOR HAIRSTYLE", "HOUSE"*
	a.2 – Color	*Discrete sets of chromas and saturation combinations*
	a.3 – Texture	*Set of samples of textures*
	a.4 – Amount	*Quantifications of countable or uncountable items (the latter enumerable via their containers)*
	a.5 – Position	*Set of relative positionings between items*
	a.6 – Orientation	*Set of relative orientation of items*
	a.7 – Value	*Discrete set of lightness values*
b. Synsemia		
	b.1 – Orientation	*Set of relative orientation in entactic compositions of units*
	b.2 – Size	*Set of hierarchical ranks resulting from relative differences in sizes between entactic composition of units*
	b.3 – Position	*Set of relative positioning of clusters in layouts with respect to other entactic compositions of units*

Table 1. List of visual variables related to both entactic and synsemic layers of articulation in order to code Aztec pictorial writing.

2.1. Description of the layers in the proposed model

Having previously explained the nature of the two topmost layers of the model (a. Entaxis and b. Synsemia), we will sketch in the following pages an illustration of each sub-layer, together with its operationalization.

Layer a – Entaxis

Sub-layer a.1 – Shape

It takes a single character out of a set, purposefully meant to create an entactic composition of units. E.g. MAN, WOMAN, WARRIOR HAIRSTYLE, MINISTER HAIRSTYLE, HILL, HOUSE, HAND, SILO, etc.[2] Notice that those kinds of basic characters often assume a "default" variable as far as colour (a.2), texture (a.3) and other features are concerned. For example, the standard entactic composition unit for "hill" is coloured in green, but the word corresponding to it in Nahuatl reading of the unit is just TEPETL (i.e. "hill", not "green hill"). This feature, indeed, is similar to the way emojis are currently depicted.

[2] Capitalization is used for character names according to Unicode conventions (see Unicode Standard Version 15.0 – Core Specification, Chapter 4.8, page 180 <https://www.unicode.org/versions/Unicode15.0.0/ch04.pdf>).

Sub-layer a.2 – Color

This variable is defined as a discrete set of combinations of chromas and saturation values (assuming an LCH type colour space, but it can be any) attested in aztec manuscripts. The value of lightness is expressed by the Sub-layer a.7 – Value.

Sub-layer a.3 – Texture

This variable is defined as a set of samples of custom drawn textures, attested in aztec manuscripts.

Sub-layer a.4 – Amount/Number of items

This variable is defined as quantification by number. Even in the case of uncountable amounts of any given item, the variable is still described as a quantitative value corresponding to the amount of containers, taken as units of measure. E.g. “two *quauhcuexcomatl* [silos] of X”, “one *comitl* [vase] of Y”, “one *caxitl* [bowl] of Z” and so on. This is not considered as an individual variable by Bertin.

Sub-layer a.5 – Position

It is a variable whose values are only defined in relative terms. E.g. OVER, “ON THE SIDE”, ON/ABOVE, “BELOW, BEHIND, IN FRONT of a(nother) unit.

Sub-layer a.6 – Orientation

Such a variable is to be thought up not only as measure of absolute directions in orientation, but rather as a character out of a predefinite set deliberately devised to produce entactic compositions of units (e.g. STANDING, SITTING, LEANING, CROUCHING, LAYING). As a practical example, in the complex unit read in nahuatl *tepehualiztli*, “defeat”, the character ROOF of the temple is leaning and collapsing; or, to obtain the complex, reduplicated character labeled by scholars *othli*, “road”, a series of single footprints is oriented from the starting unit to the end unit in the written space.

Sub-layer a.7 – Value

This variable refers to the amount of lightness, defined as a discrete set of values (depending on the value of colour space, according to Sub-layer a.2). E.g. Grey 20%, Grey 30% etc. This variable, however, is not attested in Aztec writing as such and possibly should be excluded – at least if the purpose is encoding only Aztec pictorial writing. However, we choose to maintain it because it can prove useful for the coding of other notations or scripts.

Layer b – Synsemia

Sub-layer b.1 – Orientation

Each entactic composition of units is given a value in terms of orientation, defined by its correlation to other entactic compositions of units or to any other basic unit. E.g. LEADING, TRAILING, FOLLOWING ALONG, SIDING, STANDING IN FRONT.

Sub-layer b.2 – Size

Each entactic composition of units is assigned a different rank (higher, same or lower) in the visual and linguistic hierarchy defined by its relative size, compared with other entactic compositions of units or any other basic unit. E.g. MORE PROMINENT, LESS IMPORTANT, EQUALLY IMPORTANT.

Sub-layer b.3 – Position

Each entactic composition of units is labeled by a specific value of relative positioning with respect to other entactic compositions of units or any other basic unit. E.g. ABOVE, BELOW, INSIDE, SURROUNDING. Consider, to illustrate this point, the synsemic frame in fol. 2r of Codex Mendoza where the warriors are INSIDE a specific portion of the landscape sectioned by the water, and they in turn are SURROUNDING Tenochtzin.

2.2. Encoding principles

Even if the complete encoding provided by the model is not yet fully explored and, of course, transposable to a formal Unicode standard in terms of programming, in the following lines we will sketch a tentative proposal whose aim is to logically structure the set of composition variables described in § 2.1.

The starting point for a possible encoding of the model is the idea that each layer includes either a set of *n* basic characters, e.g. B1, B2, B3 … B*n* (e.g. from the layer a.1 – Shape, the character corresponding to a Nahuatl reading as *tepetl*, "hill", or *chimalli*, "shield") and a predefined set of *n* variants working as modifiers for those basic characters (which is suited to collect only specific modifications attested), e.g. C1, C2, C3 … C*m* in the sub-layer a.2 – Color). The combinatorial matrix of these B*n*×C*m* characters will form sequences which will "depict" or process the manifested unit of Aztec script by selecting the chosen variables and values. An example of a B1 C2 combination could be the "mountain coloured in red" (instead of maintaining the default value of "green"). We posit, however, that the whole combinatorial matrix of variables isn't completely manifested, due to the absence (in Nahuatl language, or in agglutinative rules combining units) of some theoretically possible values: this means, then, that the typeset/font will not include the glyph resulting from those combinations – i.e. a fallback method should be determined. As we mentioned in § 2.1., Sub-layer a.1 – Shape, it is possible that specific characters already include multiple variables emerging from different relevant layers by default; in this case it is impossible to find – both in visual notation and in Aztec language, if the proper reading is concerned – any example or depiction of a given character where those intrinsic or default features are missing. A fallback method could be used, then, to display the basic glyphs in a way that they clearly show that the proper visualization can't be achieved (e.g. instead of showing a red mountain, the text editor will show a HILL – *tepetl,* which by default is green – and a red square). The fallback method requires to be investigated further.

The model for composition and assemblage of units described so far is used primarily to encode basic units in the Layer a – Entaxis. It is conceptually easy to mix and match one modifier variable to one basic unit, in an analogous way to what emoji already accomplish in the Unicode specification. On the other hand, things become more complex (and problems arising are not solved at present in the current Unicode standard, as mentioned in § 1) when it comes to provide a specific list of the composition rules involved in accounting for Layer b – Synsemia. In order to do this, indeed, we need to add new types of (non-positive or not displayed-as-signs) characters which will function as markers within the sentence, in order to precisely single out and locate each entactic composition of units interested by sinsemic combinations at stake: we suggest to label those special characters an *initial marker of agglutination*, an *end marker of agglutination* and a *counter*.

As a practical example, in Aztec language, we can refer to units-lexemes such as the already mentioned *othli* ("road"): the displayed glyph is seen in most of the cases, when a predicative reading is involved, as a series of footprints going from (a) starting unit to (an) end unit. These, in turn, will be coded as AZTEC CONNECTING CHARACTERS, insofar they require an *initial marker of agglutination* and an *end marker of agglutination* to be properly coded and drawn. Therefore, in order to define the generic read in Nahuatl *othli* ("road") as C1, the encoding will require a sequence of at least B1 C1 B2, defining B1 and B2 as possible locations connected by the character C1.

A theoretical representation of the overall encoding sequence, which contains every descriptive layer of the entactic composition between units, could work as follows:

> | initial marker of agglutination [number/counter] | Unit 1 [layer a.1 – Shape, layer a.2 – Color, layer a.3 – Texture, layer a.4 – Amount, layer a.5 – Position, layer a.6 – Orientation, layer a.7 – Value]; Unit 2 [layer a.1 – Shape, layer a.2 – Color, layer a.3 – Texture, layer a.4 – Amount, layer a.5 – Position, layer a.6 – Orientation, layer a.7 – Value]; ... ; Unit *n-1* [layer a.1 – Shape, layer a.2 – Color, layer a.3 – Texture, layer a.4 – Amount, layer a.5 – Position, layer a.6 – Orientation, layer a.7 – Value]; Unit *n* [layer a.1 – Shape (OTHLI, AZTEC CONNECTING CHARACTER), layer a.2 – Color (-), layer a.3 – Texture (-), layer a.4 – Amount (-), layer a.5 – Position (INBETWEEN), layer a.6 – Orientation (-), layer a.7 – Value (100%)] | end marker of agglutination | layer b.1 – Orientation (counter for start unit; counter for end unit; counter for connector unit), layer b.2 – Size (counter for start unit; counter for end unit), layer b.3 – Position (counter for start unit; counter for end unit)

3. Some examples, and discussion of the model

A basic prerequisite for encoding Aztec writing is a knowledge of rules of composition concerning any text; this, however, necessarily involves a linguistic understanding: it is therefore nonsensical to imagine a "typographic" setting of characters-units completely machine-guided, syntactic-grounded but completely non semantic, such as in Searle's Chinese room argument (Searle 1980). In this paragraph, then, we provide some examples suggesting a rationale for interpreting pictorial units as linguistic elements; also, a coding procedure is devised according to the model of § 2.

The overall premise to the use of an encoding as the one proposed in this article is that the data entry of Unicode characters should consist of a series of basic units composed by a series of sub-elements encoded and linked together (in layer a of the matrix, see Table 1). These units are then arranged in space according to the constraints brought by the variables at the second level of the matrix (Layer b). Each choice of layer/variable is done directly by the writer, through a graphic composition system.

As it happens with writing systems such as the Arabic, the so-called "logical order" is to be distinguished and carefully separated from "visual order". In the case of Aztec script encoding, several different logical sequences can produce the same visual output; to achieve this, it is therefore necessary to identify classes into which individual characters can be placed and which describe how logical order is to be transformed into visual ordering. A software devised to visualize Aztec script must implement these rules by employing special typefaces which, in turn, must be designed in order to allow patterns of entactic composition between units. The typefaces should enable – and the standard should allow – a kind of organizing principles similar to those involved in composing accented letters of the Latin alphabet, or to choosing correctly between initial, middle and final forms in writing systems such as Arabic (see Chapter 9 of the Unicode Standard; Unicode Inc. 2022).

If we take the example of the glyph HILL, usually read as "tepetl" in Nahuatl, we can argue that it has at least five ways of being combined with other basic units in the script. In terms of encoding, individual elements will be provided as a sequence, and Aztec characters will have to contain anchor-type instructions for instantiating a "correct" graphic composition. The method, indeed, is similar to the one used for diacritics in glottic writings, e.g. classifying characters into five groups (Top, Right, Bottom and Left and Over Joining).

Writing softwares "interpreting" a text encoded in such a way will succeed in producing an appropriate re-arrangement of characters, in order to display their attested agglutinated form.

3.1. Layer a – Entaxis, minimal units and their possible clustering

In Aztec writing, as in every notation, it is possible to detect minimal pictorial units (elsewhere we called them *picto-graphemes*, but the term is somehow misleading since it evokes the linear and segmental structure of purely glottic notations, cfr. Perri 2010). These units occur only in a few cases as isolated items, since they are most of the time assembled with other units in plastic and agglutinating visual clusters, often corresponding to agglutinative morphology of spoken Nahuatl. Of course pictorial minimal units, while identifying a small inventory of words or morphemes, sharply differ from full-fledged autonomous logographic characters (such as the ones of Chinese script, coded in a one-to-one ratio with the corresponding word by the Unicode standard), insofar they engage in multidimensional meaningful arrangements with other units in order to form more complex entactic combinations and visual clusters.

Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

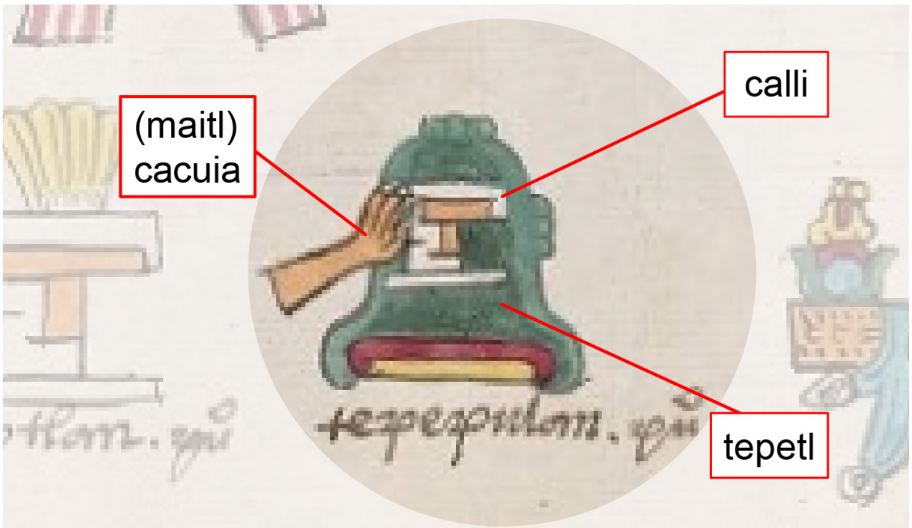

Figure 1. Detail of f. 20r of the Codex Mendoza depicting the Aztec glyph for *Tepecacuilco*, with description of the components of the glyph.

The case of place names (toponyms) best represents the process we alluded to. In the complex place name which we argue should be read as *Tepecacuilco*,[3] then, we can see an entactic combination of pictographic units-characters such as *maitl*, "hand", which actually codes a predicate (*cacuia*, "to press") through its relation with two other units: the character to be read *calli* ("house") in *Nahuatl* and the character we already know is read as *tepetl* ("hill"). The Nahuatl locative suffix -*co* ("place of") is not graphically marked in writing, since it is inferred from its relative position in the textual space: indeed it is detected (and coded) at the higher Level b (i.e. synsemic overall articulation of entactic groups).

In order to encode the agglutinative process described so far, it is important to establish in which ways a given sequence of units has to be clustered in a single entatic unit – therefore the order and principles of agglutinative composition, which of course are meaningful for linguistic reading.

In the place name at stake, then, we can appreciate a superposition of layers: the HAND (layer 1) press the HOUSE (layer 2) against the HILL (layer 3), thus this entactic order should be clearly coded in order to properly read the "full" Nahuatl active sentence *in maitl quicacuia in calli tepeticpac,* "the hand press the house against the hill", then "reduced" to the passive form *in te-*

[3] In Codex Mendoza, the source we used to pick up pictographic texts under analysis, reading in Nahuatl of pictorial places names is also provided by alphabetic glosses written down by an unknown Spanish "interpreter" before the manuscript left Mexico; indeed, it was sent to Spain as a special gift to Emperor Charles V (however the manuscript didn't reach the intended destination, and nowadays is preserved in Oxford, at the Bodleian Library). While those alphabetic glosses are at times appropriate, and represent a useful cue in the segmentation and reading of pictographic minimal units, they are often incomplete, misleading or totally wrong: in the case at stake, e.g., the gloss reads as *Tepepulan*, whose suggested interpretation in Nahuatl during XIX century was "site where [houses] are done", with a verb *(tzoqui)poloa* ("prepare mud") not expressed by the pictographic assemblage. We rejected the suggestion provided by the gloss, trying to infer the correct place name reading straight from the agglutinative cluster of pictorial minimal units.

petl (cal)cacuilo, "the hill is pressed by a house" which is, in turn, read as the place name *Tepecacuilco*, "in the place where hills are pressed". Notice that in this kind of notation (but the same can be said of visual-gestural systems such as Sign Languages) diathesis of vocal languages is virtually absent, since the assignment of syntactic roles entirely depends on the layer sequencing (or, in other cases, on the path of processing into a single layer): in the place name discussed, ordering minimal units from layer 3 to 1 would automatically change the active process expressed into a passive one.

A tentative encoding in order to account for such a complex unit could read as follows.

Tepecacuilco, "in the place where hills are pressed" (Fig. 1)

> | initial marker of agglutination [1] | Unit 1 (HAND, *cacuia*) [layer a.1 – Shape (A HAND PRESSING SOMETHING), layer a.2 – Color (YELLOW), layer a.3 – Texture (-), layer a.4 – Amount (1), layer a.5 – Position (ON THE SIDE-LEFT), layer a.6 – Orientation (-), layer a.7 – Value (100%)]; Unit 2 (HOUSE, *calli*) [layer a.1 – Shape (A HOUSE), layer a.2 – Color (YELLOW), layer a.3 – Texture (-), layer a.4 – Amount (1), layer a.5 – Position (INSIDE), layer a.6 – Orientation (-), layer a.7 – Value (100%)]; Unit 3 (HILL, *tepetl*) [layer a.1 – Shape (A HILL), layer a.2 – Color (GREEN), layer a.3 – Texture (-), layer a.4 – Amount (1), layer a.5 – Position (SURROUNDING), layer a.6 – Orientation (-), layer a.7 – Value (100%)] | end marker of agglutination [1] | layer b not present, entaxis only.

The encoding procedure sketched so far, therefore, needs to account for different patterns of composition between units: while in the above example we managed a multiple layering of the overall synsemic space, it is also possible to find other ways of clustering, not necessarily involving superposition of ordered layers: thus the basic unit for HOUSE is placed on the

Figure 2. Detail of f. 16r of the Codex Mendoza depicting the Aztec glyph for *Caltepec*.

Ocula_temi
Occhio semiotico sui media | Semiotic eye on media
Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

top of the minimal unit for HILL in the place name *Caltepec* ("At the hill of the house", CM fol. 16r), and we should legitimately wonder if any external (i.e. non-linguistic) perceptual constraint would prevent from imagining a reverse combination – in which the HILL over an HOUSE would express the (indeed unattested) Aztec toponym *Tepecalco*, "At the House of the hill" (Perri 2006).

Caltepec, "At the hill of the house" (Fig. 2)

> | initial marker of agglutination [1] | Unit 1 (HOUSE, *calli*) [layer a.1 – Shape (A HOUSE), layer a.2 – Color (YELLOW), layer a.3 – Texture (-), layer a.4 – Amount (1), layer a.5 – Position (ABOVE), layer a.6 – Orientation (-), layer a.7 – Value (100%)]; Unit 2 (HILL, *tepetl*) [layer a.1 – Shape (A HILL), layer a.2 – Color (GREEN), layer a.3 – Texture (-), layer a.4 – Amount (1), layer a.5 – Position (BELOW), layer a.6 – Orientation (-), layer a.7 – Value (100%)] | end marker of agglutination [1] | layer b not present, entaxis only.

But there is more. Any minimal unit could change its meaning and linguistic content according to potentially admissible values of relevant variables listed above: thus the form of a HILL unit is curved (variable Shape) to express the nahuatl sequence *Colhuâ-* (a possessive form interpreted as "having", *huâ*, "a curve", *coltic*) attested in the glyph *Colhuacan*; or color and/or texture can change as in *Iztacxaltepec*, "At the white (*iztac*) sandy (*xalli*) hill".

Iztacxaltepec, "At the white (*iztac*) sandy (*xalli*) hill" (Fig.3)

> | initial marker of agglutination [1] | Unit 1 (HILL, *tepetl*) [layer a.1 – Shape (A HILL), layer a.2 – Color (WHITE, *iztac*), layer a.3 – Texture (SANDY, *xalli*), layer a.4 – Amount (1), layer a.5 – Position (-), layer a.6 – Orientation (-), layer a.7 – Value (100%)] | end marker of agglutination [1] | layer b – not present, entaxis only.

Figure 3. Detail of f. 20r (left) and 13r (right) of the Codex Mendoza depicting the glyph *Colhuacan* (left) and *Iztacxaltepec* (right).

Ocula_temi
Occhio semiotico sui media | Semiotic eye on media
Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

Figure 4. Detail of f. 2r of the Codex Mendoza showcasing a variety of combinations of layers, while retaining a certain degree of regularity in the composition.

From this analysis of Aztec writing, it comes out that agglutinative processes of entactic combinations of units allow for a series of graphic (and linguistic) features, which are supported by combinations of layers in the proposed model of encoding. The features are:

– *superposition* of one or more units over another (o more than a single one), suggesting a "sequence" of ordered layers "compressed/flattened" on the planar surface of text (e.g. *Tepecacuilco*). Feature supported by a.1 – Shape, a.5 – Position, a.6 – Orientation;
– *relative position* of a pictographic unit on the top, the bottom, to the left or to the right of another identifiable unit(s), thus producing a graphic compound (e.g. *Caltepec*). This is obtained through visual contact between the units, but without a fusion (partial or total)[4]. Feature supported by a.1 – Shape, a.5 – Position, a.6 – Orientation;
– *relevant change* in the *basic form* of the unit-character. Feature supported by a.1 – Shape, a.4 – Amount, a.6 – Orientation;
– *relevant change* in internal *texture* and/or *color* (e.g. *Iztacxaltepec*). Feature supported by a.2 – Color, a.3 – Texture, a.7 – Value.

Any entactic composition of units resulting from those features, however, is always seen as a unique blending of all relevant relationships between layers (see, for example, Fig. 4).

3.2. Layer a – Entaxis, discrete variants and connectors

Assuming that shape, color, texture and number are fundamental variables for the encoding model, it is important to assess if the variants of such variables are discrete or continuous, and if the current available set of variants is enough to cover all possible modifications of the basic units.

4 Since Nahuatl is an agglutinative language, we often find complex words formed with two (or even more) lexical roots involved in various semantic relations between them. When pictographic entaxis suggests a sort of "blending" or "hybridization" between two different characters, however, instead of a simple juxtaposition with visual contact this is usually a case in point of a dependency relation between roots, over coordinative or simple relations. Typically, this is expressed by changes affecting the texture.

Ocula_temi
Occhio semiotico sui media | Semiotic eye on media
Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

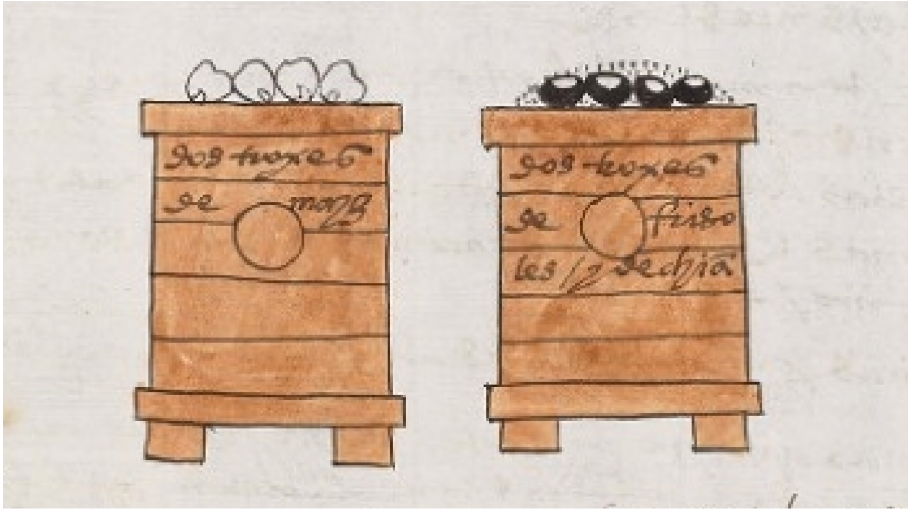

Figure 5. Detail of f. 44r of the Codex Mendoza showcasing a quantifier objects on top of their containers.

Let us consider for example the relevant strategies for numbering of items. In Fig. 5 we see BEANS (read in Nahuatl as *etl*) and CORN grains (read in Nahuatl as *tlaolli*) lying at the top of a SILO, (*quauhcuezcomatl*, "wooden silo"); but their function is to number the amount of units-containers, as well as their content (e.g. 4 silos full of corn and 4 silos full of beans). Beans and corn grains, indeed, are discrete numeric indicators and act as figures. This method of numbering occurs whenever specific quantities are detailed, as was usually the case in a tribute register. But the script could also have recourse to specific "abstract" numerals: for example, in Fig. 6 the FLAGS above (and visually connected to) different containers or products-items are discrete numeric indicators, too (the reading *pantli*, "flag", corresponds in Nahuatl to a row of "20" units).

This special feature of Aztec writing, indeed, demands a "numerator", whenever we find entactic compositions involving units used (and read) as figures/quantities.

As far as we know, no entactic compositions of units in Aztec writing denote continuous numbering or quantity; however, in principle it would be possible to provide a specific encoding as well for the latter in our model.

In the model we have sketched, the coding for 4 SILOS CONTAINING BEANS would be as follows:

> | initial marker of agglutination [1] | Unit 1 (SILO) [layer a.1 – Shape (A SILO), layer a.2 – Color (BROWN), layer a.3 – Texture (HORIZONTAL LINES STACKED), layer a.4 – Amount (1), layer a.5 – Position (BELOW), layer a.6 – Orientation (-), layer a.7 – Value (100%)]; Unit 2 (BEANS) [layer a.1 – Shape (A BEAN), layer a.2 – Color (BLACK), layer a.3 – Texture (-), layer a.4 – Amount (4), layer a.5 – Position (ABOVE), layer a.6 – Orientation (-), layer a.7 – Value (100%)] | end marker of agglutination [1] | layer b not present, entaxis only.

Ocula_temi
Occhio semiotico sui media | Semiotic eye on media
Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

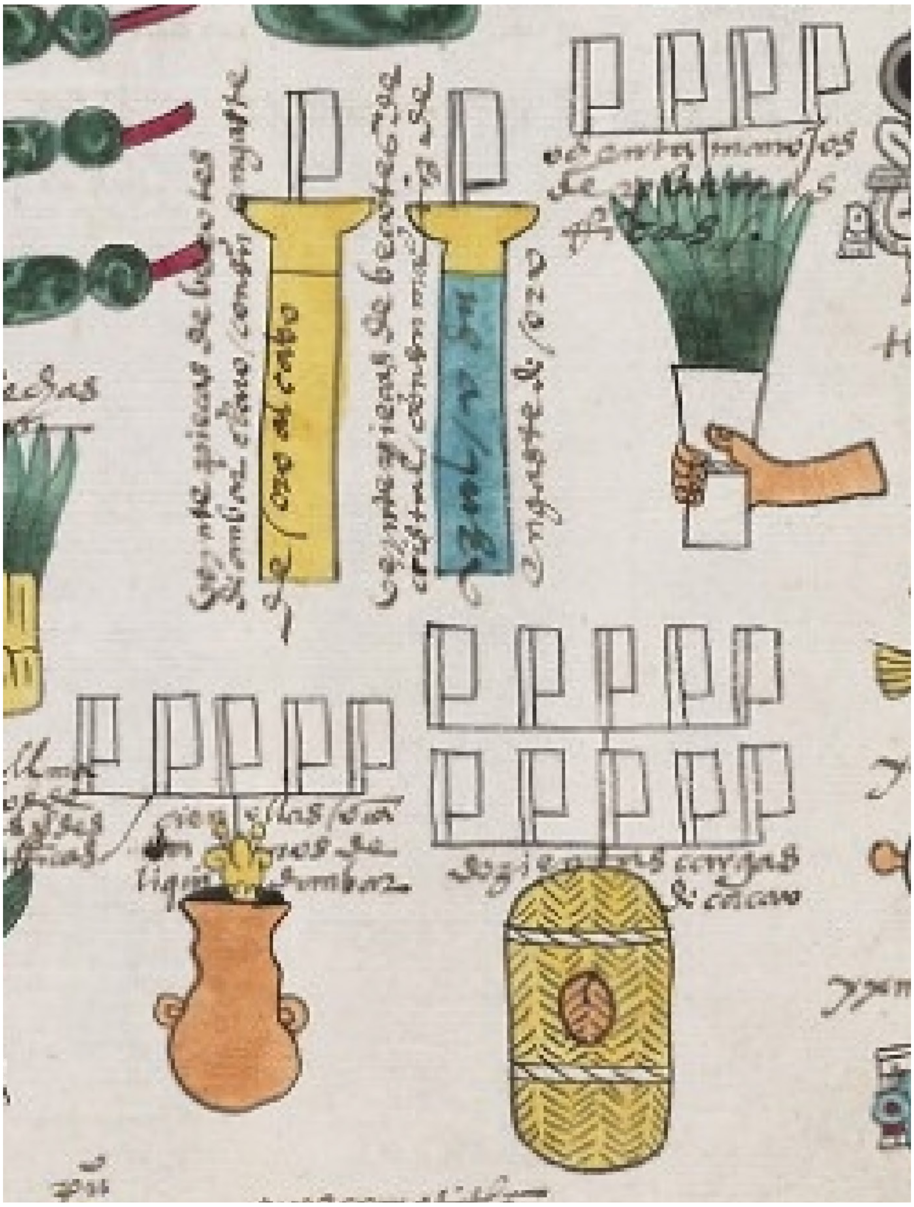

Figure 6. Detail of f. 46r of the Codex Mendoza showcasing a series of numerators connected to their containers.

The coding for a generic BASKET/VESSEL CONTAINING *N* UNITS OF PRODUCT (as in Figure 6, folio 46r of CM) is slightly different, because it needs to take into account of the (graphic) connector between the figure/number(s) and the (numbered) container.

A graphic device similar to the one used for quantities numbered with "abstract" figures occurs in the case of anthroponyms: indeed, a proper name is usually attached to the human pictorial figure bearing it (usually to his head) by a visible and cognizable stroke. In this case, the coding process is thought to

Ocula_temi
Occhio semiotico sui media | Semiotic eye on media
Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

Figure 7. Detail of f. 2r of the Codex Mendoza depicting the nobleman Ocelopantzin.

appeal to an “invisible” entactic unit of connection and function labeled NAME OF PERSON. The tentative sequence for encoding, then, could be something as BASICUNIT1 NAME-OF-PERSON BASICUNIT2, resulting in a display of BASICUNIT2 attached via a stroke to BASICUNIT1.

Figure 7 provides an example from CM f. 2r.

| initial marker of agglutination [1] | Unit 1 (*BASICUNIT1* “FLAG”) [layer a.1 – Shape (A FLAG), layer a.2 – Color (YELLOW), layer a.3 – Texture (OCELOT-SKIN), layer a.4 – Amount (1), layer a.5 – Position (ON THE SIDE-LEFT), layer a.6 – Orientation (-), layer a.7 – Value (100%)]; Unit 1 (*NAME-OF.PERSON* CONNECTOR) [layer a.1 – Shape (STROKE, AZTEC CONNECTING CHARACTER), layer a.2 – Color (-), layer a.3 – Texture (.), layer a.4 – Amount (-), layer a.5 – Position (INBETWEEN), layer a.6 – Orientation (-), layer a.7 – Value (100%)]; Unit 3 (*BASICUNIT2* WARRIOR) [layer a.1 – Shape (A WARRIOR), layer a.2 – Color (WHITE (dress)), layer a.3 – Texture (-), layer a.4 – Amount (1), layer a.5 – Position (-), layer a.6 – Orientation (SITTING), layer a.7 – Value (100%)]; Unit 4 (WICKER MAT) [layer a.1 – Shape (A WICKER MAT), layer a.2 – Color (GREEN), layer a.3 – Texture (STACKED STICKS), layer a.4 – Amount (1), layer a.5 – Position (BELOW), layer a.6 – Orientation (-), layer a.7 – Value (100%)] | end marker of agglutination [1] | layer b – not present, entaxis only.

Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

3.3. Layer b – Synsemia

Assuming the correct encoding of the entactic composition units at entactical level (see above §§ 3.1 and 3.2), in order to obtain a complete analysis of Aztec writing it is necessary to encode conventional relationships *between* those compositional entactic units. Indeed, the structure and meaning of any Aztec text depends on spatial relations between elements; synsemia (i.e. the macro-layer b) provides suitable conceptual and practical strategies for encoding this higher level of contents.

The main and principal problem we have to face trying to encode Aztec synsemia is that in any spatial synsemic system it is not possible (nor advisable) to "force" relations between entactic compositions of units in order to change them in linearly ordered sequences: every element or graphic cluster is suitable to being connected with many others; therefore it is not possible to restructure this complex network of relations in a sequential and unilinear pattern. For example, in CM f. 2r, the ruler and priest *Tenochtzin* sits in a landscape sectioned by water (rivers), surrounded by others noble Aztec warriors. He is then put in relation by the visual text both to elements-units denoting landscape and to other nobles occurring in the framed space.

There emerges, then, the need to outline an encoding system powerful enough to connect "distant" elements in a (temporalized) sequence of processing, regardless of how distant these units are from one another.

For this reason, we thought it necessary to introduce, in the encoding structure, a counter: it is useful to identify various entactic compositions of units in the text, sequencing them and enabling the reader to bring them together in a logical and sound frame. Such a feature is currently hard to imagine as an integral part of typographic-based systems such as Unicode, since it requires the introduction of specific notations whose function is to "transparently" refer (and point) to other elements-units in the linear flow of the encoded text beyond adjacent items. A possible way to achieve this aim would be to introduce a number/counter, which identifies the relative position of the "linked" character. Needless to say, this number/counter would require to be encoded as a distinct unit from numeric-figures characters.

All the sub-layers of b – Synsemia, thus, are provided of this feature: each layer will show numbers/counters for each entactic composition of units, since they all perform a meaningful relationship between two or more units.

Acknowledging, though, that such encoding system is not yet in place in Unicode standard, and assuming that it will never be included into the standard, then all relationships between entactic compositions of units described by Layer b – Synsemia will have to be handled by the writing software. The latter solution, however, would prevent the "portability" (in terms of compatible encoding) of any text across non-dedicated writing softwares: such a text could not be copied and pasted, because it would turn into linear, sequential sequence of entactic compositions of units without meaning and internal organization. This, in turn, would ultimately miss the purpose of encoding and describing Aztec writing, because relationships between units loose their semantic value which is an integral part of this script.

Our hypothesis is therefore that Aztec will be effectively supported in Unicode only if the standard is set to include also Layer b – Synsemia in the encoding process.

3.4. Layer b.1 – Orientation

In § 2.2 we already mentioned the peculiar case of character FOOTPRINT used in iteration to express *othli* ("road"), whose glyph is displayed as a series of footprints in sequence. Contrary to what we see in almost all glyphs as shown in other scripts, it is not located on a precise site of the written space, but rather it draws a linearly processed path oriented from the starting unit to the end unit (to both of which the *othli* glyph itself is supposed to be visually pinned). *Othli,* therefore, express a relationship: both the denotation of movement and actors and objects involved depend on a "start" and a "destination" position, so they are oriented according to these two pivotal points.

Another example of relevant synsemic role of orientation can be found in the same CM f. 2r: the gazes of noble warriors are oriented so as to point deictically towards the compositional and semantic center of the page (the founding myth of Tenochtitlan). However, graphic conventions make the reader feels that in the upper (east) and lower (west) quadrants of the city's division by canals the noblemen are looking at each other, and that the same happens with human figures in the north (the quadrant of the priest and chief Tenochtzin) and the south.

3.5. Layer b.2 – Size, hierarchical layers

The Layer b.2 – Size does not have a specific semantic value; rather it specifies, through synsemic relations, both textual syntax and information hierarchy. It also defines the positioning of elements in the writing space: in fact, setting differences in size while composing a text is generally equivalent to overlapping planes or visual "layers", so that the larger units are "closer" and thus carry information that must be read with priority. Therefore, in each text more than one reading plane/layer can be identified or articulated, suggesting the reader to proceed from the units largest in size (and usually also centrally located) to the layers below (smaller and progressively reduced in size).

3.6. Layer b.3 – Position, topological, inside/outside

A peculiar feature of the Aztec texts is that any unit acts as a sort of synsemic reference point for other units. In particular, in the often mentioned f. 2r, the landscape sectioned by water canals is an entactic unit that acts as a reference frame, in which Aztec noblemen are placed. The relative positioning of each nobleman within each spatial unit has semantic value, and the whole picture constitutes as such a synsemic composition of entactic units. The graphical positioning of each unit, of course, may also change slightly; but topological positions of elements are unique, and significant. Moreover, the positioning of the noblemen is meaningful also because it identifies which portion of the city is ruled by whom.

Vol 25 • No 30 • June 2024 • DOI: 10.57576/ocula2024-5

Figure 8. Folio 2r of the Codex Mendoza.

4. Current limitations and additional issues

Even though the proposed model for the encoding of Aztec writing is made to be flexible within a confined set of few rules, the complexity of the task brings some important limitations to be further explored.

Graphic composition of each basic unit should allow for "deformable" characters, to avoid drawing each and every combination with any of the possible modifiers and interactions with other composition units.

While we devised the matrix of layers in order to minimize the number of characters (and codes), notwithstanding we posited that the overall amount of basic units would have been high; at the same time, however, this figure in not comparable, in terms of "numerousness" and "digital space", to the space occupied by the over 70.000 codes of Han writing as implemented in the Standard since three decades.

It should be noted, as we said above, that the model for Aztec sketched so far, although coherently designed according to Unicode rules and option of encoding, is in fact going beyond the standard typographic logic to which Unicode Consortium still subscribes.

Commenting on the main features of Western movable type typography, twenty years ago Giovanni Lussu argued that it was "eminently alphabetical: its mechanics based on juxtaposition is perfectly molded on the arbitrary segmentation to which alphabet subsume the speech chain" (Lussu 2003: 46; our translation). While segmental and unilinear principles are completely unfit to provide a suitable analysis of Aztec writing, our proposal is thought to come to terms with the overall (and as such unmodifiable) architecture of Unicode by "softening" some rigid coding principles and, at the same time, increasing the intricacy or complexity of others features.

The introduction of emoji in the Standard has, in part, paved the way for such a work to account for Aztec writing, since emoji "open" system of "symbolic" notational units shows an internal articulation which is best described as a set of combinatorial features non-linear, entactically arranged in a visual frame motivated and iconical.

Finally, we suggest that the specific structure and features of our model devised for Aztec will be useful to face the coding of non-alphabetic devices in use in our texts to envision information such as graphs, statistical flow charts and other non-linear devices grounded on synsemic relations.

References

Bertin, Jacques
1967 *Sémiologie graphique. Les diagrammes Les réseaux Les cartes*, Paris, Gauthier-Villars. Eng. trans. *Semiology of Graphics: Diagrams, Networks, Maps*, Redlands CA, Esri Press, 2011.

DeFrancis, John
1989 *Visible Speech: The Diverse Oneness of Writing Systems*, Honolulu, University of Hawaii Press.

Fedorova, Liudmila
2009 "The Emblematic Script of the Aztec Codices as a Particular Semiotic Type of Writing System", *Written Language & Literacy*, 12.2, 258-274.

Fedorova, Liudmila; Antonio Perri
In press "Emblematic Techniques as Textual Strategies in Non-Linear and Linear

Scripts», in *Proceedings of the "Grapholinguistics in the 21st Century 2022 Conference"*, Yannis Haralambous (ed.), Brest, France, Fluxus Editions.

Goodman, Nelson
1968 *Languages of Art: An Approach to a Theory of Symbols*, Indianapolis, IN, Bobbs-Merril.

Haralambous, Yannis
2007 *Fonts & Encodings*, Beijing-Sebastopol CA, O'Reilly Media.

Harris, Roy
1998 *L'origine della scrittura*, Viterbo, Stampa alternativa & Graffiti.
2003 *La tirannia dell'alfabeto. Ripensare la scrittura*, Viterbo, Stampa Alternativa & Graffiti.

Klinkenberg, Jean-Marie; Stéphane Polis
2018 "On Scripturology", *Signata. Annales des sémiotiques/Annals of Semiotics*, 9, 57-102.

Korpela, Jukka K.
2006 *Unicode Explained*, Sebastopol, CA, O'Reilly.

Lussu, Giovanni
2003 "La forma del linguaggio", *Progetto grafico*, 1, 1, luglio, 40-49.

Perri, Antonio
2006 "Dire o mostrare? L'immagine-testo e i paradossi della pittografia come notazione", in A.-M. Christin, D. Barbieri, A. Catellani, A. Perri, *Immagini del testo. Per una semiotica dell'ideogramma*, Centro internazionale di Semiotica e Linguistica, Università di Urbino, Documenti e Pre-pubblicazioni, 350-351-352, 51-75.
2010 "Il problema delle unità minime nella scrittura azteca. Contributo a una teoria integrata della scrittura", in *Testo e Senso*, 11, <https://testoesenso.it/index.php/testoesenso/article/view/35>. Online June 2, 2023.

Perondi, Luciano
2012 *Sinsemie. Scrittura nello spazio*, Viterbo, Nuovi Equilibri.

Searle, John R.
1980 "Minds, brains, and programs", *Behavioral and Brain Sciences*, 3, 417-424.

Unicode Inc.
2022 Unicode Standard Version 15.0 – Core Specification, Chapter 4.8, page 180 <https://www.unicode.org/versions/Unicode15.0.0/ch04.pdf>). Online June 2, 2023.
2022 "The Unicode® Standard Version 15.0 – Core Specification. Chapter 9 Middle East-I", <https://www.unicode.org/versions/Unicode15.0.0/>. Online June 2, 2023.
n.d. a *Unicode® Emoji*, <https://unicode.org/emoji/techindex.html>. Online June 2, 2023.
n.d. b *Glossary of Unicode Terms, Glyph*, <https://unicode.org/glossary/#glyph>. Online June 2, 2023.

Vaillant, Pascal
1999 *Sémiotique des langages d'icônes*, Paris, Honoré Champion.

Zamora Corona, Alonso Rodrigo
2022 "Towards a Complex Theory of Writing: The Case of Aztec and Mixtec Codices", *Signata. Annales des sémiotiques/Annals of Semiotics*, 13, <https://doi.org/10.4000/signata.3866>. Online June 2, 2023.

Antonio Perri is associate professor at the University of Naples Suor Orsola Benincasa. His main research interests deal with the anthropological and linguistic features of writing systems and notations (in particular Aztec writing) and translation theory (more specifically, intersemiotic translation). His last book on writing is *Per una semiotica della scrittura* (Graphofeel, 2013).

Luciano Perondi is associate professor at Università IUAV di Venezia. His main fields of interest are reading process and the non linear use of writing (synsemia). His publications include: *Sinsemie: Scritture nello spazio* (Stampa alternativa/Nuovi equilibri, 2012); "Inter-letter spacing, inter-word spacing, and font with dyslexia-friendly features: testing text readability in people with and without dyslexia" (With J. Galliussi, G. Chia, W. Gerbino, P. Bernardis, *Annals of Dyslexia*, 2020).

Daniele Capo is graphic designer and type engineer (Cast Type Foundry). He is lecturer of typography at University La Sapienza, Rome.

Roberto Arista is a typographic designer and a font engineer. He builds tools and fonts at Typotheque, and he takes good care of RoboFont editor documentation. He graduated from ISIA Urbino and KABK type]media master program. In 2017 he published pythonfordesigners.com, an open-source award winning manual to teach designers how to program in Python.

Giampiero Dalai is Usability Expert and Information Designer (Alpaca Società Cooperativa). He does research in the development of interactive communication artifacts, with a focus on design for education and design for all. He published, with G. Bonora, D. De Rosa, A. Imperato, B. Martini, & L. Perondi: "Sinsemia as a tool for designing interactive artifacts for teaching. The case study of The topography of Dante's Inferno and the detailed definition of *reference frame*", *2CO Communicating Complexity*, Contributions from the 2017 Tenerife Conference, 2020, 94-104. <https://doi.org/10.25145/b.2COcommunicating.2020.012>