The Thot Sign List (TSL)

An open digital repertoire of hieroglyphic signs[*]

Stéphane Polis (F.R.S.-FNRS / ULiège), **Luc Desert** (ULiège), **Peter Dils** (SAWL), **Jorke Grotenhuis** (ULiège), **Vincent Razanajao** (UBordeaux-Montaigne), **Tonio Sebastian Richter** (BBAW), **Serge Rosmorduc** (CNAM), **Simon D. Schweitzer** (BBAW), **Daniel A. Werning** (BBAW, HUBerlin), **Jean Winand** (ULiège)

IN THIS PAPER, we introduce the Thot Sign List (TSL), an online digital repertoire of hieroglyphic signs (http://thotsignlist.org) that records the graphemes attested in the Ancient Egyptian hieroglyphic texts, with special attention to their contextual functions and to their iconic variations across media and time. The paper is structured as follows. In Section 1, we present a brief history of the TSL and we specify its goals and targeted audience. In Section 2, we introduce the data model, which meets the requirements formulated in Meeks (2013) and Polis & Rosmorduc (2013), and discuss the use of shared thesauri for the metadata (Thesauri and Ontologies for documenting Ancient Egyptian resources; https://thot.philo.ulg.ac.be). Section 3 describes the user interface and discusses how data can be browsed, searched, and visualized by users depending on their level of access. We describe the way credits and citation work for this database in Section 4, and we give a list of the collaborators. In the conclusions, we argue that such an online resource will immensely benefit from monitored crowdsourcing: Egyptologists all around the world can enrich this digital repertoire with new sources and examples of hieroglyphic signs.

1. Introduction

1.1. *A (brief) history of the Thot Sign List*

The idea of this digital repertoire of hieroglyphic signs goes back to a workshop held on May 21-22, 2013 in Tours (France), entitled 'Gestion informatisée des écritures anciennes. État des lieux et perspectives.' Based on an analysis of thousands of spellings found in the Ramses corpus (http://ramses.ulg.ac.be), Polis & Rosmorduc (2013) analyzed the *Manuel de Codage* (Buurman et al. 1988) encoding scheme used for hieroglyphic texts, which developed organically, based on practical needs for text editions that were implemented in new software solutions (Gozzoli 2013, Rosmorduc 2015: 2–4). They argued that this *Manuel* is problematic for the development of text corpora (see also Nederhof 2013). Since this is also true for Unicode (Everson 1999), which takes over the hieroglyphic signs and their palaeographical variants found in existing sign lists without critical analysis (Schenkel 1999, Meeks 2013, Everson and Richmond 2007), it became clear that a new repertoire of hieroglyphic signs would be a worthwhile undertaking for the Egyptological community.

[*] We are very grateful to Eitan Grossman and Mark-Jan Nederhof for insightful comments and detailed feedback on first drafts of this paper.

In this early vision, the planned digital sign list would (minimally) have (1) to be structured, making a clear distinction between meaningful variants at the functional and iconic levels, and (2) to be referenced, including links to ancient sources displaying signs in context, with specific functions and forms. These general principles were outlined by Meeks (2013) and Polis & Rosmorduc (2013), and were discussed by Hafemann and van der Moezel (2018).

In 2015, the project was introduced at the 11th International Congress of Egyptologists (Florence) and it materialized thanks to the Anneliese Mayer research grant introduced by Joachim Fr. Quack (at that time director of the Berlin and Leipzig academies joint project 'Structure and Transformation in the Vocabulary of the Egyptian Language' with the *Thesaurus Linguae Aegyptiae*, the related electronic corpus) and awarded to Jean Winand (University of Liège). Henceforth, it has been conceived as a joint endeavor between the University of Liège and the Berlin-Brandenburg Academy of Sciences and Humanities. In 2015-2016, the data model was finalized by Jorke Grotenhuis, Stéphane Polis, Vincent Razanajao, Serge Rosmorduc, and Jean Winand, and a first version of the encoding tool was implemented by Luc Desert (CIPL) in Liège, while hieroglyphic sign functions were documented in Berlin within a specific database by Silke Grallert, Ingelore Hafemann, Simon D. Schweitzer, Dina Serova, and Lisa Seelau.

By January 2017, the online encoding tool was functional. Having imported hieroglyphic signs from JSesh (https://jsesh.qenherkhopeshef.org/) as well as digitized hieroglyphs from printed sign lists, such as the Berliner Zeichenliste (http://aaew.bbaw.de/archive/berliner-zeichenliste-1935-39) and the unpublished sign list by Hornung & Schenkel (kindly provided by Wolfgang Schenkel), and having integrated the preliminary data collected by the Berlin team, a systematic encoding of the signs, functions, and their sources could begin. From 2016 to present, Jorke Grotenhuis has been the main person in charge of the encoding in Liège, while several scholars from Germany enriched and emended the database, including Max Bader, Peter Dils, Silke Grallert, Tilmann Kunze, Lisa Seelau, Dina Serova, Jakob Schneider, Simon D. Schweitzer, Anja Weber, and Daniel Werning.

On January 18-19, 2018, a first workshop was held in Liège in order to evaluate the tool and data encoded in TSL thus far. The participants included Peter Dils, Jorke Grotenhuis, Ingelore Hafemann, Stéphane Polis, Vincent Razanajao, Daniel Werning, and Jean Winand. This resulted in an evolution of the database structure, and led to the implementation of several new features. It was followed by a second workshop on 15–16 May 2018 in Berlin, with Peter Dils, Silke Grallert, Jorke Grotenhuis, Ingelore Hafemann, Simon D. Schweitzer, Lisa Seelau, and Daniel Werning, which aimed at solving further structural issues and at organizing the encoding process. On November 6, 2019, a first version of the end-user interface was released online during the 12th International Congress of Egyptologists in Cairo.

1.2. *General goal and guiding principles*

As TSL is born out of the practical necessity of creating electronic corpora of hieroglyphic texts in a principled way (§1.1), its first aim is to document the functions attested for individual signs in order to indicate which standardized character may be used in which context. To meet this goal, two non-trivial steps must be taken.

First, ancient sources have to be identified for as many functions as possible. Referenced sign lists, such as Gardiner 1957: 438–549, Cauville 2001, Kurth 2009, I: 127–453, Schenkel 1983, I: 45–83, or Borghouts 2010, II: 10–195, greatly help in this respect, but finding an

actual image or accurate facsimile of individual sources turned out to be a challenge. Such images are a prerequisite for a sound approach of the written material, since handwritten transcriptions and glyphs in fonts are unfortunately not reliable enough (Meeks 2013). Luckily, the situation is quickly improving thanks to the digital turn in the field illustrated, for instance, by the Karnak project (http://sith.huma-num.fr) and the open access policy of several Egyptological museums (for instance, the Museo Egizio in Turin).

Second, among the available signs in modern lists and fonts, one has to identify glyphs that are actualizations of the same character (or grapheme) and as such share the same function(s). This was the original intention of the *Manuel de Codage* that states explicitly (Buurman et al. 1988: 51): "there is a clear distinction made between graphemes and graphic variants. Code written as letter + figure (+ figure, + figure) refer to graphemes – i.e., hieroglyphs differing from one another in their reading or meaning. Codes written as letter + figure (+ figure, + figure) + letter refer to graphic variants – i.e., hieroglyphs differing neither in their reading nor in their meaning from those they come from." With the Latin alphabet, this amounts to stating that 'a' and 'A' are both instances of an abstract grapheme <a>, which is rather uncontroversial. However, defining and identifying graphemes is a notoriously difficult task (Klinkenberg & Polis 2018: 69–81) and is even more problematic for pictorial scripts, like the hieroglyphic script, where minor variations can turn out to be significant.

Fig. 1. "Hommes assis versant de l'eau" (Buurman et al. 1988: 57).

In the example of Fig. 1, one can deduce based on the codes that A6A, A6B, A6C, and A6D are groups of graphic variants of A6, which is identified as the standard grapheme. Each group is illustrated by a number of tokens (between 1 and 7) that are characterized by iconic differences: A6 [vase on top of the head, water forwards], A6A [water backwards], A6B [vase in the hand(s)], A6C [water in both directions], A6D [libation vase with several trickles of water]. Note that such descriptions have to be inferred from the catalogue and that some

tokens (e.g., the 4th of A6) do not fit the proposed analysis. Despite the iconic differences, these groups of variants should have the same reading (or have the same meaning), here presumably *wꜥb* '(to be) pure', but such information about the signs' function is not part of the *Manuel*.

Note however that the *Manuel* is not systematic here. A first illustration of this lack of systematicity is D7 (eye with painted lower lid) and D7A (eye with painted upper lid). According to the abovementioned principles, the latter should be a variant of the former. However, the signs are not only different from an iconic point of view, they also have distinct functions: D7A is generally not used as phonogram *ꜥn(ty)* or as a classifier with the semographic value [adornment, beauty] like D7, but rather as a [sight] related grapheme. The opposite scenario also occurs (here inherited from Gardiner): D19 (𓂝) and D20 (𓂞) receive two different codes, while Gardiner states explicitly that the latter is a 'semi-cursive variant of last (…) Use as last, but seldom in careful sculptures or paintings' (Gardiner 1957: 452). Accordingly, both hieroglyphs should rather be envisioned as 'graphic variants' than as distinct graphemes (as would seem to be the case based on their codes).

Fig. 2a. D7 vs. D7A (Buurman et al. 1988: 89).

Fig. 2b. D19 vs. D20 (Buurman et al. 1988: 90).

Furthermore, problems arise due to the fact that various influential sign lists display different sign shapes or even different glyphs under one and the same code [Fig. 2c].

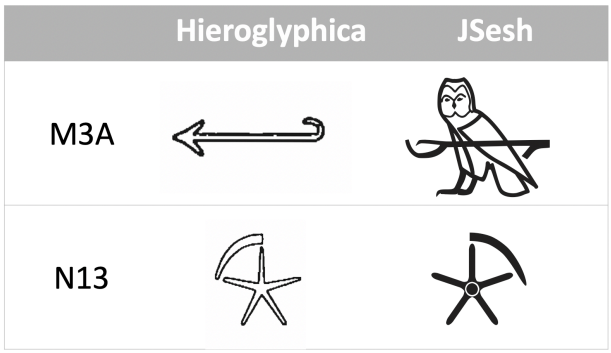

Fig. 2c. M3A and N13 in the *Hieroglyphica* vs. *JSesh* [default installation].

1.3. *Targeted audience*

While including a palaeographical dimension (since images and facsimiles of ancient sources are provided whenever possible), the primary goal of the TSL is not to provide Egyptologists with a palaeographical analysis of the hieroglyphic material (Servajean 2020). There are specific projects devoted to this question (such as the volumes of the *Paléographie hiéroglyphique* at the French Archaeological Institute in Cairo, initiated by Meeks 2004) and TSL does not intend to replace them. Rather, we principally have the following users in mind:

– Text editors, encoders, software developers, and others who look for stable IDs for hieroglyphic graphemes and sign shapes, with precise indications about their meaning and uses.

– Students, for whom TSL can be a convenient resource for learning about the possible readings and functions of hieroglyphic signs.

– Font specialists, who will find here a list of signs and variants that are duly documented and can safely be integrated in standards like Unicode.

2. Data model and metadata

2.1. *TSL data model*

The TSL data model has been designed in order to tackle the challenges outlined above (§1.2) in a principled way. The most straightforward way to explain this model is probably to describe the encoding process in a bottom-up fashion (in what follows, conceptual entities of the database are in bold face). One starts from a **Document** (temple, stela, etc.) in which ancient hieroglyphic **Sources** can be recognized (understood as a coherent chunk of hieroglyphic text). A **Source** contains hieroglyphic **Tokens**, a label that refers to actual hieroglyphs that are materialized on a medium. In this specific context, each Token has one **Function** (or 'reading', for instance, a phonographic function *x* or a logographic function *y*) and participates in one **Class** – defined as a group of Tokens with shared iconic features and functions; Classes are therefore conceptually identical to coherent groups of 'graphic variants' in the *Manuel de Codage* (see the discussion of Fig. 1 and 2 above). Finally, a **Grapheme** is a second-degree abstraction: it is envisioned as a cluster of Classes whose Tokens share the same Functions (and similar iconic features). The basic data model of Fig. 3 visualizes this

textual description.

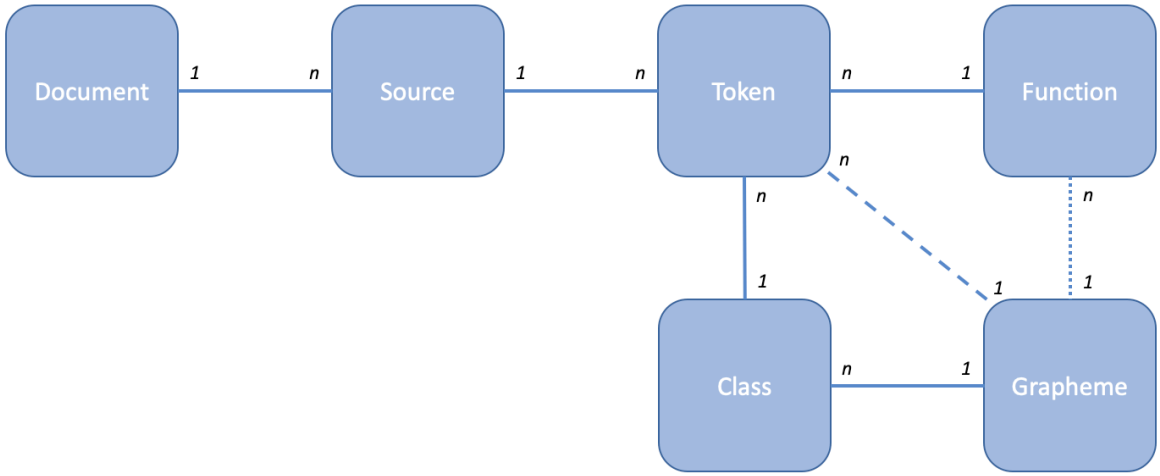

Fig. 3. Visualization of the TSL data model (main elements).

Three remarks are in order.

1. The encoding workflow cannot systematically follow the bottom-up process outlined above. Indeed, it is only after having collected several Tokens for a given Grapheme that different Classes can be identified: the creation of Classes calls for an analysis of written variations. As such, Tokens are usually first connected to a (highly abstract) Grapheme (dashed-line in Fig. 3), represented by Class '00', and other Classes are only identified later on. Consequently, if Graphemes are essentially clusters of Classes, they may – in a first place – consist of a single broad Class to begin with.

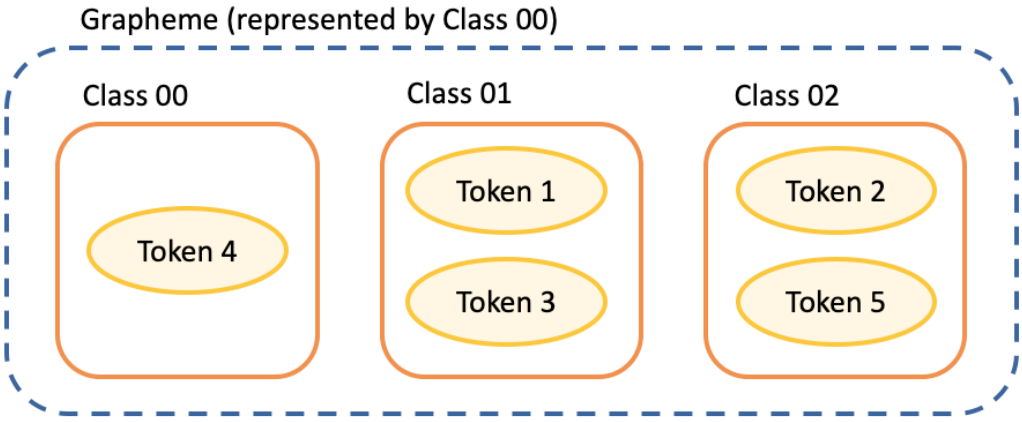

Fig. 4. Grapheme, classes, and tokens in TSL

2. Classes and Graphemes are both generalizations over empirical observations: there is no ontological distinction between the two categories, which share identical features (they can be represented by a prototypical glyph, have a code, be described with tags, etc.). Therefore, it was decided to represent the Graphemes with the Class 00 [Fig. 4]; the choice of the Class 00 that

represents the Grapheme is to a large extent arbitrary (it depends chiefly on the Egyptological tradition), and it can change based on empirical observations. Note that the IDs in TSL have either a prefix 1 for abstract Graphemes and Classes (e.g., TSL_1_82 for A1 or TSL_1_630_01 for A4C, etc.), or a prefix 3 for concrete Tokens.

3. A Function is first and foremost a property of an individual Token in context [Fig. 3], and only indirectly a feature of the Grapheme to which the Token belongs. However, in order to be able to document Functions that are not (or not yet) documented with ancient Sources (either because we did not yet manage to find an actual example or because the said function is found in the modern literature but has been proven wrong), Functions are also directly connected to Graphemes (see the dotted line in Fig. 3).

2.2. *Metadata*

The Documents and Sources are described with metadata provided by the Thot thesauri (https://thot.philo.ulg.ac.be), a set of resources for documenting ancient Egyptian sources in a shared and interoperable way developed by Vincent Razanajao. In the TSL, we record information about the object type, date, provenance, current location, material, script, technique of inscription, language and text content.

Fig. 5. The metadata of a Source (ID 425) in TSL and their link to the multilingual *Thot* thesauri.

As visualized by Fig. 5, metadata about a Source (ID 425) in the sign list are linked to the hierarchically organized concepts in *Thot*. Since individual Tokens in the TSL are systematically linked to a Source, they inherit from these metadata and can thereby be browsed and organised according to several criteria.

3. What is available, for whom, and how does the web-interface work?

For the first release of TSL, the goal was to have the digital repertoire 'Gardiner-ready', which means that the hieroglyphic Graphemes and Functions found in Gardiner's (1957: 438–548) sign list should be documented with at least one Token coming from an ancient Source. Version 1.0 is however not limited to Gardiner (1957): if an interesting Grapheme, Class or Function was documented in a Source processed for TSL, it was encoded. As shown by Fig. 6, the vast majority of Tokens and Functions encoded in the TSL so far are already available online, but many Graphemes still have to be processed.

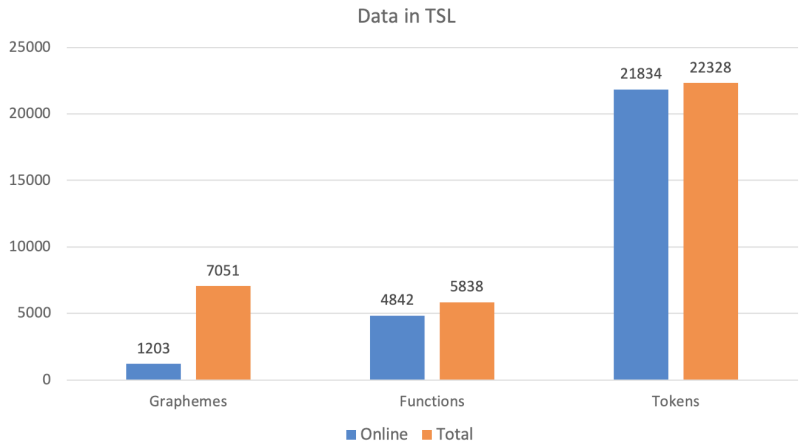

Fig. 6. Data in TSL (online vs total).

There are two basic ways to explore the TSL: Browsing (§3.1) and Searching (§3.2). Both approaches present the users with Graphemes – displayed bigger, with darker border – and Classes – displayed smaller, with lighter border. This distinction [Fig. 7] shows in a straightforward way which specific Class has been selected to represent a Grapheme (see Section 2.1 for more details about the data model).

Fig. 7. Grapheme vs. Class in TSL.

When clicking on the code associated with a Grapheme or Class, the user opens a new tab in the browser, which displays the relevant Grapheme, with its Functions, Classes and Tokens (§3.3). Note that clicking on a Class (e.g., A7A in Fig. 7) opens the Grapheme to which this Class belongs (represented by A7 in Fig. 7). Depending on the user's level of access (§3.5), the quantity of information that can be visualized varies (especially as far as the Sources are concerned, §3.4), but is essentially the same in qualitative terms.

3.1. *Browsing TSL*

When browsing the TSL, two types of filters are available: (1) the category to which a Grapheme or Class belongs – note that we use the cover-term Sign when referring to both Graphemes and Classes indistinctly – and (2) its basic shape. Both organizing principles are standard in Egyptology since Gardiner (1957), but for the thematic categories we follow the new structure suggested by Meeks (2004: XIX–XXII). These filters may be combined. In Fig. 8, for instance, the Signs belong to the 'Loafs and cakes' category and have a 'low narrow' shape.

Fig. 8. Filtering the signs when browsing in the TSL.

The codes displayed when browsing the TSL are meant to help the user identify quickly a sign based on the codes that he or she knows. They are based, in hierarchical order, on Gardiner (1957), JSesh, Hieroglyphica, Unicode, and the IFAO catalogue. This means, for instance, that a code from Unicode will only be displayed at this level if the hieroglyph did not receive a code in Gardiner (1957), JSesh or in the Hieroglyphica. Signs without a code ('no code') represent new Graphemes that are added to the TSL, but are not found in the above-mentioned sign lists.

If the mouse is positioned over the code of a Grapheme or Class, an overview of the functions associated with the Grapheme is shown. Fig. 9 illustrates this point. Two functions are available for the Grapheme linked to Gardiner code A35: based on the Tokens encoded so far, it can be used as a classifier with the meaning 'building' and as a logogram with the value *ḳd* 'to build'.

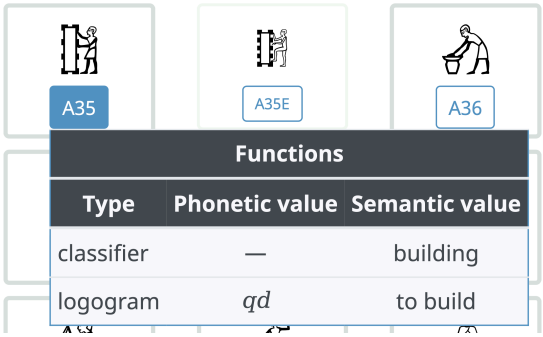

Fig. 9. Mouse-hover while browsing

3.2. *Searching TSL*

Users can use the search engine in order to look for a Grapheme or Class based on any combination of features relative to its Functions (type, phonological and semantic value), description (plain text, tag, basic form, and type), and codes. Furthermore, operators (equals, contains, does not contain) can be used for any feature so as to narrow down the search. Accordingly, one can build queries such as: [Function type = 'logogram' and Tag contains 'foreigner'], or [Phonetic value contains *mr* and Description does not contain 'canal'], etc.

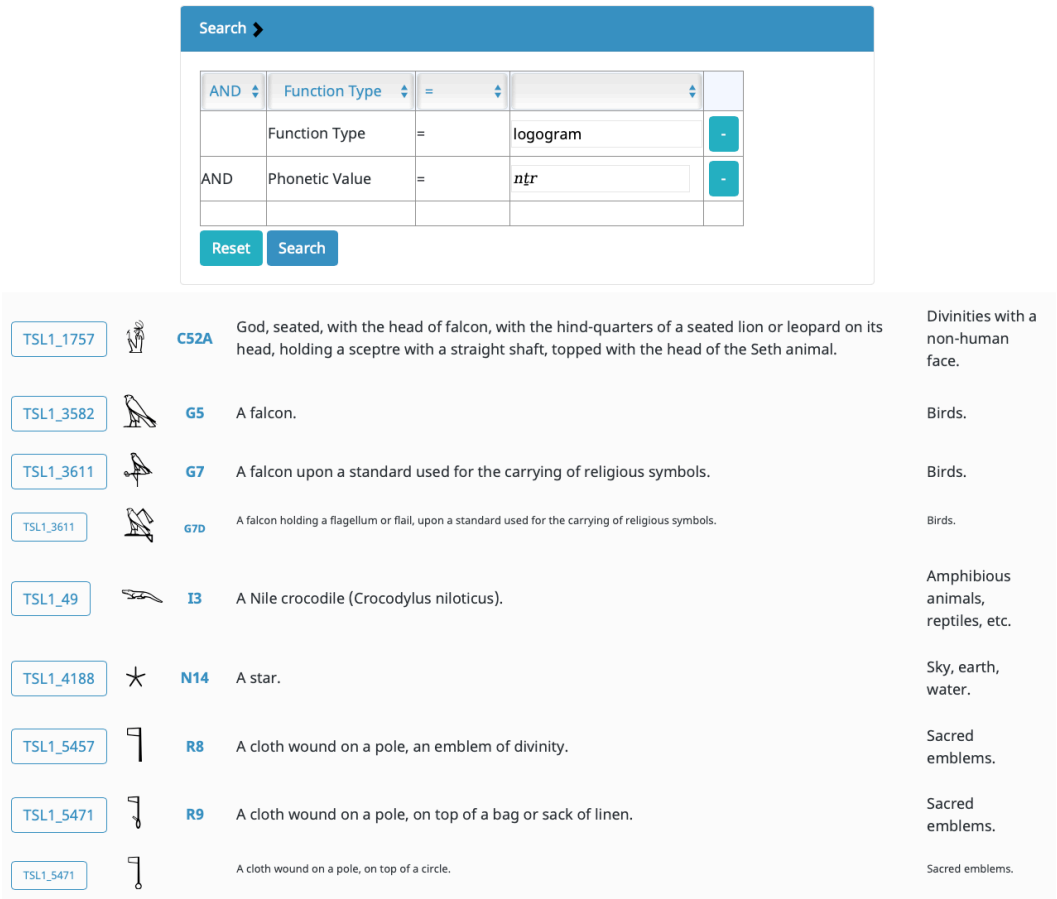

Fig. 10. Search [Function type = logogram and Phonetic value = *nṯr*].

The search of Fig. 10 lists all the Graphemes (bigger) and Classes (smaller) in TSL that are attested as logograms with the phonetic value *nṯr*. The results are sorted according to Gardiner codes, and TSL IDs give access to the relevant Grapheme (on click). Note that it is also possible to search for a TSL ID; the search engine will display all the Classes associated with a given ID.

3.3. *Visualizing a Grapheme, its Classes, and Tokens*

Fig. 11 illustrates how information is organized in TSL for individual Graphemes (here TSL_1_2177 = Gardiner D33). Four drop-down menus (on top) give access to the Description, Codes, Bibliography, and Credits for this Grapheme, while (up to) four tabs (below) gather information about the Functions, Classes, and Tokens (as well as how to cite data from the TSL).

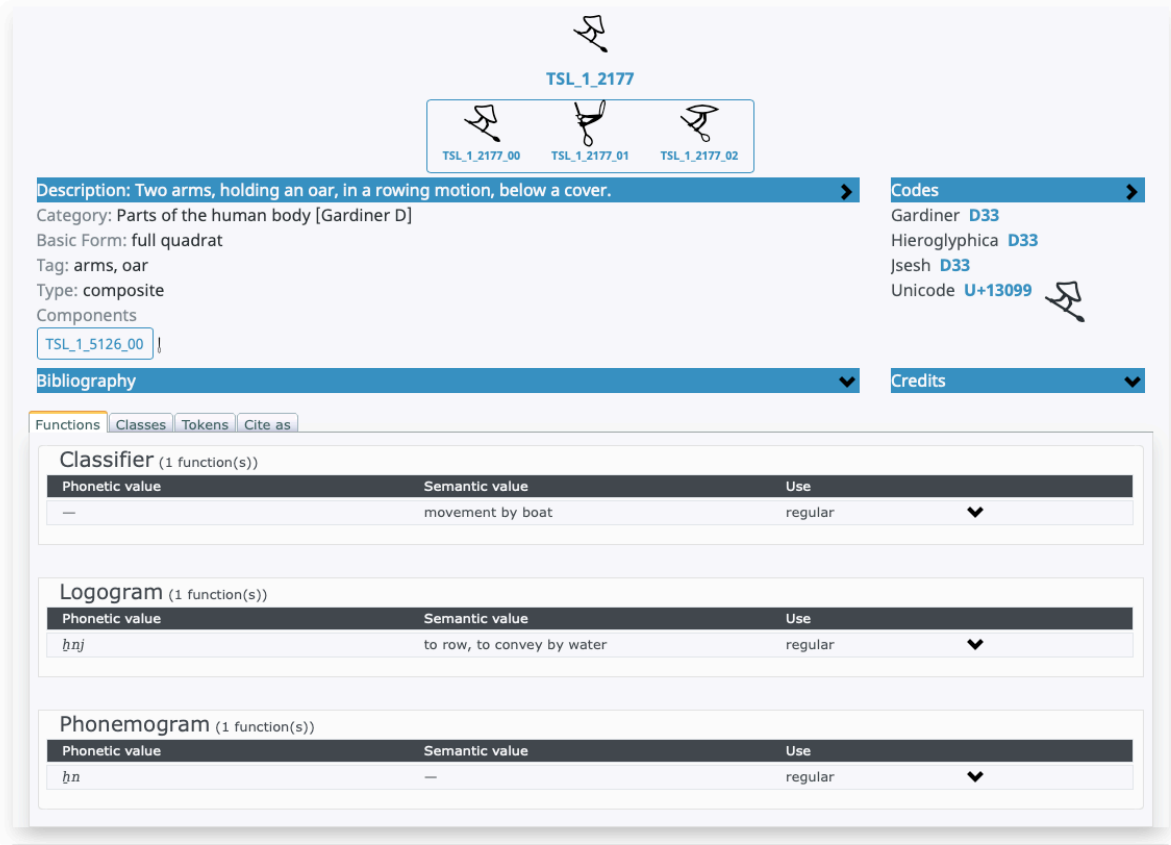

Fig. 11. Grapheme TSL_1_2177 (= Gardiner D33).

– Description provides a plain text description of the iconic features (and referent) of the Grapheme, which is represented by the Class 00. Below are its Category (based on Meeks 2004: XIX–XXII, see Section 3.1), Tags (coming from a non-hierarchical thesaurus) that describe the Class (and its components) with keywords that are intended to help users find hieroglyphs easily, and Type (we distinguish between 'simple', 'compound', and 'composite' signs; see Polis 2018: 328, Fig. 35). If a Class is analyzed as a compound or composite hieroglyph, its component(s) are given (in Fig. 11,

TSL_1_5126 (the oar) is a component of TSL_1_2177).

– Codes lists the codes attributed to the Grapheme in Gardiner (1957), Hieroglyphica, Jsesh, and Unicode (when available).

– Bibliography cites references that are relevant for the entire Grapheme. Note that if references deal with specific aspects (e.g., a Function, a Class, etc.), they are quoted under the relevant entry.

– For Credits, see §4.

Every Grapheme is accompanied by three tabs (Functions, Tokens, Cite as), and a fourth tab (Classes) is displayed when several Classes are available for the Grapheme.

– Functions. The functions are grouped by type (classifier, logogram, radicogram, phonemogram, interpretant, phono-repeater (a.k.a., phonetic determinative); see Polis & Rosmorduc 2015 and Hafemann 2018) and receive a phonetic and/or semantic value. Fig. 11 shows that three functions are documented for TSL_1_2177 as of September 2020: classifier ‘movement by boat’, logogram for *ḫnj* ‘to row, to convey by water’, and phonogram *ḫn*.

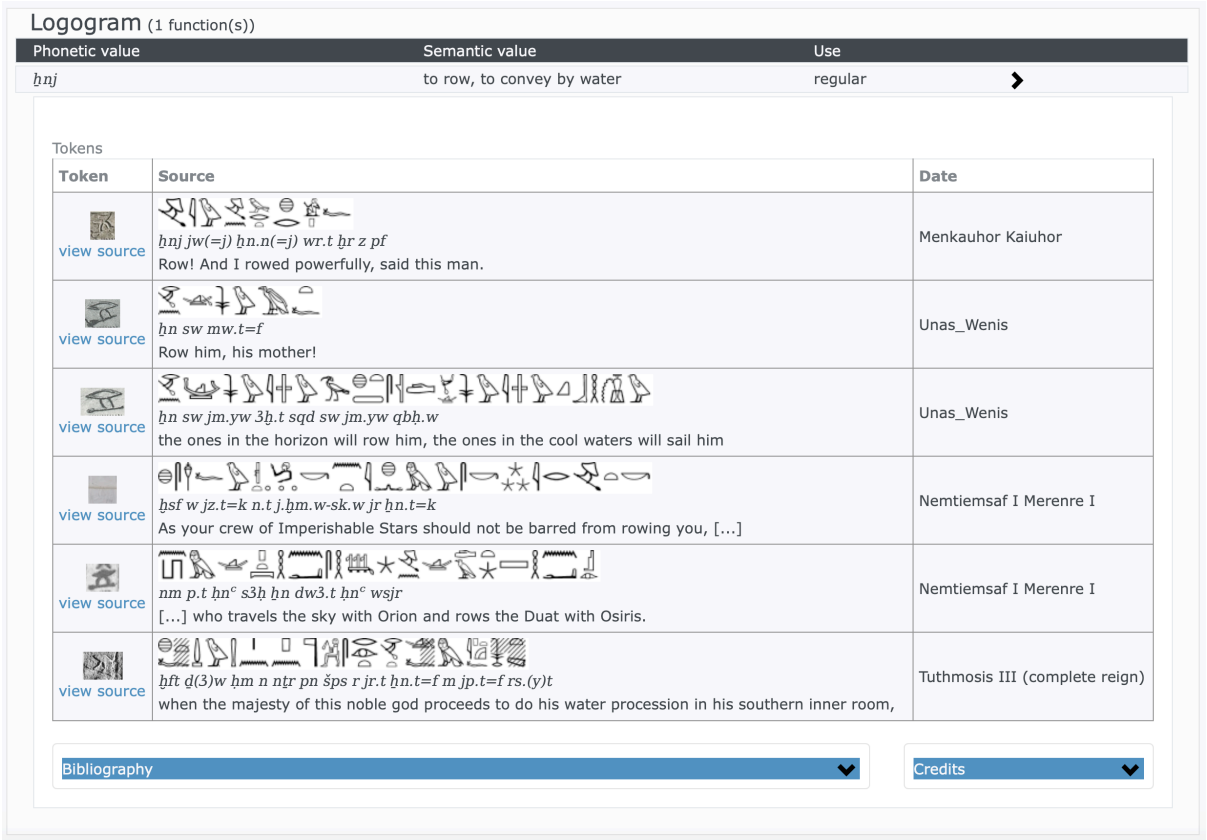

Fig. 12. Tokens for TSL_1_2177 used as logogram.

Every function (Fig. 12) is illustrated at least by one token, i.e., an actual example, for which a context of use is provided (with a hieroglyphic transcription, a transliteration, and a translation) as well as an image for the registered users (see §3.5). For more information about a token, registered users can access the Source by clicking on 'view source' (see §3.4).

– Classes. If a Grapheme has multiple Classes, they can be visualized in the dedicated tab 'Classes' (Fig. 13). A Class is illustrated by a prototypical hieroglyph (vector graphics) and is described exactly like the main Grapheme (which is represented by Class 00, see above §2.1), with codes and literature, if relevant.

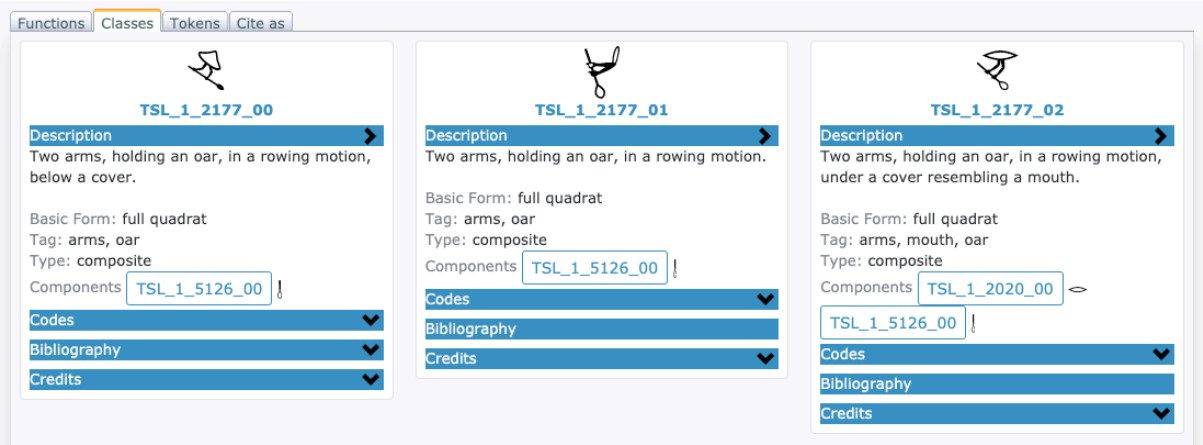

Fig. 13. Classes for TSL_1_2177.

– Tokens. The Tokens of a Grapheme can be visualized in the dedicated tab. The number in the lower-right corner corresponds to the Class to which this Token belongs (see Fig. 14).

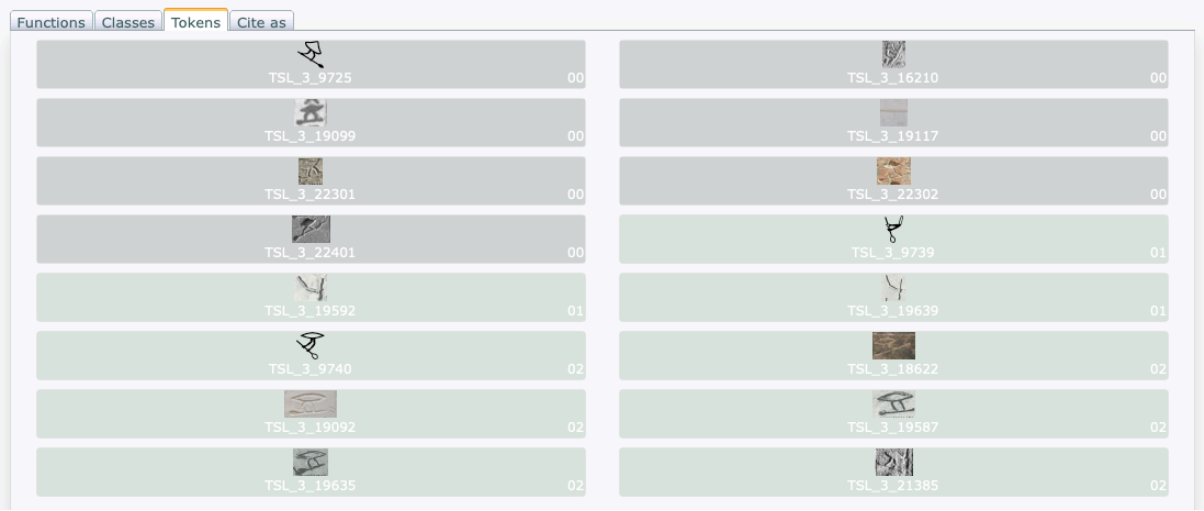

Fig. 14. Tokens of TSL_1_2177.

Details about an individual Token can be displayed by clicking, providing its Function and value, as well as its context of use (see Fig. 15). For more information about a Token, registered users can access the Source by clicking on 'view source' (see Sections 3.4 and 3.5).

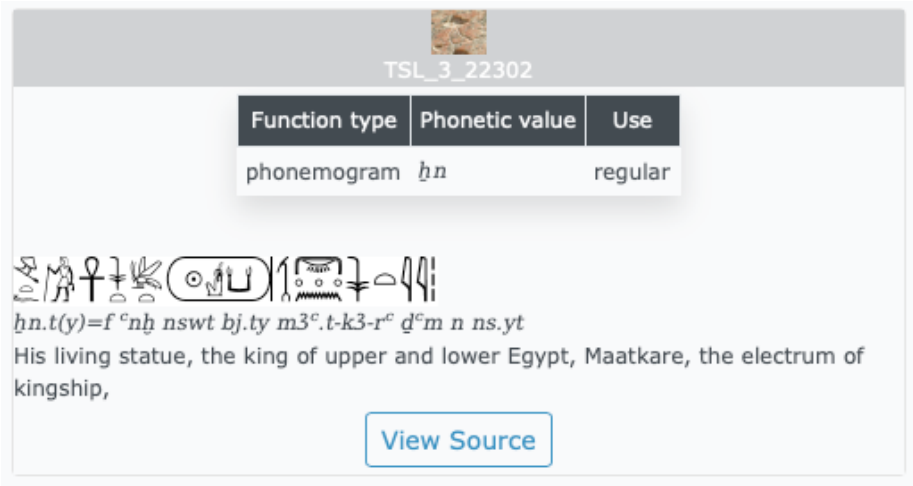

Fig. 15. Details about Token TSL_3_22302 of TSL_1_2177.

3.4. *Visualizing a Source*

Registered users may access the Sources that have been validated by an internal reviewing process. A Source is a section of a hieroglyphic (cursive hieroglyphic or hieratic) inscription; its length is not determined a priori, its purpose being to visualize a Token in context so as to be able to assess its Function. As such, it corresponds at least to a word, but more often to a phrase or a sentence.

Fig. 16. Source ID 1978.

As illustrated by Fig. 16, a Source consists of at least one image (picture, facsimile, etc.), accompanied by its standardized hieroglyphic transcription, transliteration and English translation. It is documented with metadata coming from the *Thesauri and Ontology for Documenting Ancient Egyptian Resources* (https://thot.philo.ulg.ac.be/), as discussed in §2.2.

3.5. *Registration and levels of access*

There are two levels of access to the TSL (for users who are not collaborators of the project): unregistered and registered (https://thotsignlist.uliege.be/Account/Register). Registration is free and email addresses are collected exclusively in order to update users about evolutions of the database and website. Unregistered users have access to all the Graphemes, Classes, and Functions in TSL, but cannot visualize the Sources from which the Tokens are extracted.

3.6. *Technicalities*

The hieroglyphs in the TSL are displayed with the JSesh (https://jsesh.qenherkhopeshef.org/) wrapper developed by Dmitry Nikolaev (https://github.com/macleginn/jsesh-web) and the Sources are described with the Thot metadata (https://thot.philo.ulg.ac.be/concept/) using APIs developed by Vincent Razanajao (https://thot.philo.ulg.ac.be/api/index.html). The TSL database and front-end have been implemented by Luc Desert (CIPL / ULiège).

4. Credits, citation, and collaborators

In the TSL, credits are mentioned at three levels: Graphemes, Functions, and Sources. For each level, we identify the 'creator', namely the person who created a Grapheme, Function or Source in the sign list, and the 'editor(s)', i.e., the scholars who modified the content of the said entry at some point. To quote information about an entire Grapheme, please follow the pattern below:

> **Model**. Grapheme TSL_1_ID <http://thotsignlist.org/mysign?id=ID>, in Thot Sign List, edited by Université de Liège and Berlin-Brandenburgische Akademie der Wissenschaften.
>
> **Example**. Grapheme TSL_1_82 <http://thotsignlist.org/mysign?id=82>, in: Thot Sign List <http://thotsignlist.org>, edited by Université de Liège and Berlin-Brandenburgische Akademie der Wissenschaften.

Additional information about the creator and editor(s) of Graphemes may be quoted as follows:

> **Example**. Grapheme TSL_1_82 <http://thotsignlist.org/mysign?id=82> (created by L. Seelau and edited by J. Grotenhuis, I. Hafemann, L. Seelau, A. Weber, M. Bader, T. Kunze, D.A. Werning), in: Thot Sign List <http://thotsignlist.org>, edited by Université de Liège and Berlin-Brandenburgische Akademie der Wissenschaften.

When referring to a specific Function, the following citation system should be adopted:

> **Model**. Function FunctionName with value ‘values’ of Sign TSL_1_82 <http://thotsignlist.org/mysign?id=82>, in: Thot Sign List <http://thotsignlist.org>, edited by Université de Liège and Berlin-Brandenburgische Akademie der Wissenschaften.
>
> **Example**. Function Logogram with value ‘z(j) – man’ of Sign TSL_1_82 <http://thotsignlist.org/mysign?id=82>, in: Thot Sign List <http://thotsignlist.org>, edited by Université de Liège and Berlin-Brandenburgische Akademie der Wissenschaften.

Additional information about the creator and editor(s) of individual Functions may also be quoted:

> **Example**. Function Logogram with value ‘z(j) – man’ (created by TSL and edited by I. Hafemann, J. Grotenhuis, and D. Werning) of Sign TSL_1_82 <http://thotsignlist.org/mysign?id=82>, in: Thot Sign List <http://thotsignlist.org>, edited by Université de Liège and Berlin-Brandenburgische Akademie der Wissenschaften.

When referring to Source, the following citation system should be followed:

> **Model**. Source SourceID <http://thotsignlist.org/mysource?id=SourceID>, in: Thot Sign List <http://thotsignlist.org>, edited by Université de Liège and Berlin-Brandenburgische Akademie der Wissenschaften.
>
> **Example**. Source 375 <http://thotsignlist.org/mysource?id=375>, in: Thot Sign List <http://thotsignlist.org>, edited by Université de Liège and Berlin-Brandenburgische Akademie der Wissenschaften.

Additional information about the creator and editor(s) of individual Sources may be quoted:

> **Example**. Source 375 <http://thotsignlist.org/mysource?id=375> (created by I. Hafemann and edited by I. Hafemann and J. Grotenhuis), in: Thot Sign List <http://thotsignlist.org>, edited by Université de Liège and Berlin-Brandenburgische Akademie der Wissenschaften.

Below is a list of the individuals who have contributed to the TSL so far:

Name	Institution	Function	Date
Bader, Max	BBAW	Encoder	2018
Desert, Luc	ULiège	Designer, IT	2015–present
Dils, Peter	SAWL	Designer, Editor, Encoder	2016–present
Grotenhuis, Jorke	ULiège	Designer, Editor, Encoder	2016–present
Hafemann, Ingelore	BBAW	Designer, Editor, Encoder	2015–2019
Kunze, Tillmann	Berlin	Encoder	2019
Polis, Stéphane	FNRS / ULiège	Coordinator, Designer, Editor	2013–present
Razanajao, Vincent	UBM	Designer	2015–present
Richter, Tonio Sebastian	BBAW	Coordinator	2015–present
Rosmorduc, Serge	CNAM	Designer	2013–present
Schneider, Jakob	Berlin	Encoder	2018
Schweitzer, Simon D.	BBAW	Editor, Encoder	2018
Seelau, Lisa	BBAW	Encoder	2016–2019
Serova, Dina	BBAW	Encoder	2013–2014
Weber, Anja	BBAW	Encoder	2018–present
Werning, Daniel A.	HUBerlin; BBAW	Coordinator, Designer, Editor	2016–present
Winand, Jean	ULiège	Coordinator	2015–present

BBAW = Berlin-Brandenburgische Akademie der Wissenschaften.

CNAM = Conservatoire National des Arts et Métiers (Paris).

FNRS = Fonds National de la Recherche Scientifique (Belgium).

HUBerlin = Humboldt-Universität zu Berlin.

SAWL = Sächsische Akademie der Wissenschaften zu Leipzig.

UBM = Université Bordeaux-Montaigne.

ULiège = Université de Liège.

5. Conclusions

The TSL is a long-term project that aims at gathering more and more data over the years about the hieroglyphic signs and their uses in Ancient Egyptian texts. In this paper, we described how the foundations were laid for this project (§1) and provided information about the tool and the current state of the database (§2-3). In order to meet the scholarly needs of the

Egyptological community, however, we are aware that more signs need to be documented, more functions need to be recorded, and more sources need to be encoded in order to provide a more detailed picture of the hieroglyphic writing system throughout its diachrony.

Given the limitations in terms of human resources within our respective institutions, this goal can only be reached by opening up the encoding of new materials to Egyptologists who are willing to integrate Ancient Egyptian documents or corpora on which they work in the TSL. To this end, we have an encoding interface that allows for a monitored crowdsourcing: Egyptologists can access this tool and receive an 'encoding guide' by contacting us at thotsignlist@gmail.com. The data will be duly credited (§4) and will be available online after a reviewing process that ensures that the data are correct and consistent. We firmly believe that this new mode of collaboration and publication is the way to go for the Egyptological community in a digital age.

References

Borghouts, Joris F. 2010. *Egyptian: An introduction to the writing and language of the Middle Kingdom*, 2 vols. (*Egyptologische Uitgaven* 24). Leiden/Leuven: Nederlands Instituut voor het Nabije Oosten; Peeters.

Buurman, Jan, Nicolas Grimal, Michael Hainsworth, Jochen Hallof, and Dirk van der Plas. 1988. *Inventaire des signes hiéroglyphiques en vue de leur saisie informatique: Manuel de codage des textes hiéroglyphiques en vue de leur saisie sur ordinateur / Manual for the encoding of hieroglyphic texts for computer-input / Leitfaden zur Verschlüsselung hieroglyphischer Texte für die Computer-Angabe*, 3rd consolidated ed. (*Informatique et Égyptologie* 2 ; *Mémoires de l'Académie des inscriptions et belles-lettres, nouvelle série* 8), Paris: Institut de France.

Cauville, Sylvie. 2001. *Dendara: le fonds hiéroglyphique au temps de Cléopâtre*, Paris: Cybèle.

Everson, Michael. 1999. "Encoding Egyptian hieroglyphs in Plane 1 of the UCS" (Internet resource: http://www.unicode.org/L2/L1999/N1944.pdf).

Everson, Michael & Bob Richmond. 2007. "Proposal to encode Egyptian hieroglyphs in the SMP of the UCS". Working Group Document ISO/IEC JTC1/SC2/WG2 N3237, International Organization for Standardization.

(Internet resource: http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3237.pdf).

Gardiner, Alan. 1957. *Egyptian grammar being an introduction to the study of hieroglyphs*, 3rd ed., Oxford: Oxford University Press.

Gozzoli, Roberto B. 2013. "Hieroglyphic text processors, Manuel de Codage, Unicode, and Lexicography". In Polis, Stéphane and Jean Winand (eds), *Texts, languages & information technology in Egyptology: selected papers from the meeting of the Computer Working Group of the International Association of Egyptologists (Informatique & Égyptologie), Liège, 6-8 July 2010* (*Aegyptiaca Leodiensia* 9), Liège: Presses Universitaires de Liège, pp. 89–101.

Hafemann, Ingelore 2018. "Die beschreibende und kommentierte hieroglyphische Zeichenliste als offenes System". In Gülden, Svenja A., Kyra van der Moezel & Ursula Verhoeven (eds), *Ägyptologische "Binsen"-Weisheiten III: Formen und Funktionen von Zeichenliste und Paläographie. Akten der internationalen und interdisziplinären Tagung in der Akademie der Wissenschaften und der Literatur, Mainz, im April 2016*, Stuttgart: Franz Steiner, pp. 19–49 (http://doi.org/10.25358/openscience-5178).

Hieroglyphica = Grimal, Nicolas, Jochen Hallof & Dirk van der Plas. 2000. *Hieroglyphica: Sign list, liste des signes, Zeichenliste*, 2nd ed., rev. and enlarged (*Publications interuniversitaires de recherches égyptologiques informatisées* 1). Utrecht: Centre for Computer-aided Egyptological Research.

JSesh = Rosmorduc, Serge. 2014. *JSesh Documentation*, available at: http://jseshdoc.qenherkhopeshef.org (accessed 2020, version 7.5.5).

Klinkenberg, Jean-Marie & Stéphane Polis. 2018. "On scripturology". Translated by Todd J. Gillen. *Signata: Annales des sémiotiques / Annals of Semiotics* 9, pp. 57-102 (https://journals.openedition.org/signata/1885).

Kurth, Dieter. 2007–2015. *Einführung ins Ptolemäische: eine Grammatik mit Zeichenliste und Übungsstücken*, 3 vols. Hützel: Backe (2009 for the 3rd ed. of vol. 1).

Manuel de Codage = Buurman et al. 1988.

Meeks, Dimitri. 2004. *Les architraves du temple d'Esna: paléographie* (*Paléographie hiéroglyphique* 1). Cairo: Institut français d'archéologie orientale.

Meeks, Dimitri. 2013. "Dictionnaire hiéroglyphique, inventaire des hiéroglyphes et Unicode", *Document Numérique* 16 (3), pp. 31–44.

van der Moezel, Kyra. 2018. "On signs, lists and standardization". In Gülden, Svenja A., Kyra van der Moezel, and Ursula Verhoeven (eds), *Ägyptologische "Binsen"-Weisheiten III: Formen und Funktionen von Zeichenliste und Paläographie. Akten der internationalen und interdisziplinären Tagung in der Akademie der Wissenschaften und der Literatur, Mainz, im April 2016*, Stuttgart: Franz Steiner, pp. 51-81.

Nederhof, Mark-Jan. 2013. "The Manuel de Codage encoding of hieroglyphs impedes development of corpora". In Polis, Stéphane & Jean Winand (eds), *Texts, languages & information technology in Egyptology: selected papers from the meeting of the Computer Working Group of the International Association of Egyptologists (Informatique & Égyptologie), Liège, 6-8 July 2010* (*Aegyptiaca Leodiensia* 9), Liège: Presses Universitaires de Liège, pp. 103-110.

Polis, Stéphane. 2018. "The Functions and Toposyntax of Ancient Egyptian Hieroglyphs: Exploring the iconicity and spatiality of pictorial graphemes", *Signata: Annales des sémiotiques / Annals of Semiotics* 9, pp. 291–363 (https://journals.openedition.org/signata/1920).

Polis, Stéphane & Serge Rosmorduc. 2013. "Réviser le codage de l'égyptien ancien: vers un répertoire partagé des signes hiéroglyphiques", *Document numérique* 16 (3), pp. 45–67 (https//orbi.uliege.be/handle/2268/163100).

Polis, Stéphane & Serge Rosmorduc. 2015. "The hieroglyphic sign functions: Suggestions for a revised taxonomy". In Amstutz, H., A. Dorn, M. Müller, M. Ronsdorf, and S. Uljas (eds), *Fuzzy boundaries: Festschrift für Antonio Loprieno*, vol. 1, Hamburg: Widmaier, pp. 149–174 (https//orbi.uliege.be/handle/2268/186322).

Rosmorduc, Serge 2015. "Computational linguistics in Egyptology". In Stauder-Porchet, Julie, Andréas Stauder & Willeke Wendrich. *UCLA Encyclopedia of Egyptology* 2015 (April).

Schenkel, Wolfgang. 1983. *Aus der Arbeit an einer Konkordanz zu den altägyptischen Sargtexten, Vol. I: Zur Transkription des Hieroglyphisch-Ägyptischen* (*Göttinger Orientforschungen, Reihe IV, Ägypten* 12). Wiesbaden: Harrassowitz.

Schenkel, Wolfgang. 1999. "Comments on the question of encoding Egyptian hieroglyphs in the UCS", in: http://std.dkuug.dk/jtc1/sc2/wg2/docs/n2096.pdf (accessed 2019-10-22).

Servajean, Frédéric. 2020. "Hieroglyphic palaeography". In Davies, Vanessa and Dimitri Laboury

(eds), *The Oxford handbook of Egyptian epigraphy and palaeography*. New York: Oxford University Press, pp. 536-549. DOI: 10.1093/oxfordhb/9780190604653.013.38.