Daniel Burt, Ahmad Al-Jallad and Michael C.A. Macdonald

8 The Online Corpus of the Inscriptions of Ancient North Arabia

Abstract: The *Online Corpus of the Inscriptions of Ancient North Arabia* (OCIANA) was created to make available in a fully searchable online database the texts and translations of all the inscriptions of ancient North Arabia, together with metadata and photographs. Developed in Filemaker Pro, it is consultable both online and as a series of fully searchable pdfs. All known inscriptions from ancient North Arabia have been entered, except the “Thamudic”, which pose particular problems, and will be entered in the next phase of the project.

Keywords: Ancient North Arabian, online database, lexicography, ancient literacy, glyph variation

This chapter is in three parts. In the first, Michael Macdonald describes the origins and purpose of the *Online Corpus of the Inscriptions of Ancient North Arabia* (OCIANA).¹ In the second, Daniel Burt describes its present structure and performance, and in the third Ahmad Al-Jallad looks forward to the aims of Phase 3 of the project, which he will direct.²

8.1 The Background to OCIANA

OCIANA aims to make available in one place an edition of all known inscriptions from ancient North Arabia. The term “ancient North Arabia” in this context refers geographically to the Arabian Peninsula north of Yemen,³ with a fluid northern

1 [<http://krc2.orient.ox.ac.uk/ociana/>].

2 Phase I was a preparatory stage lasting one year (2011–2012) and funded by the John Fell Fund of the University of Oxford. During this phase the *Ancient Arabia: Languages and Cultures* (AALC, <http://www.ancientarabia.co.uk/>) website was established and approximately 10,000 black-and-white negatives and colour slides of previously unpublished Safaitic inscriptions from the Basalt Desert Rescue Survey (BDRS) were scanned in preparation for their insertion in OCIANA during Phase 2.

3 Ancient South Arabian inscriptions have been collected and edited in the *Digital Archive for the Study of pre-Islamic Arabian Inscriptions* [<http://dasi.cnr.it/>], based at the University of Pisa and *The Sabaic Dictionary Online* which are described in Chapters 1 and 9 in this book.

Daniel Burt, Michael C.A. Macdonald, University of Oxford

Ahmad Al-Jallad, Universiteit Leiden

© 2018 Daniel Burt, Ahmad Al-Jallad and Michael C.A. Macdonald

This work is licensed under the Creative Commons Attribution 4.0 International license (CC BY 4.0)

border including modern Jordan, southern Syria and western Iraq. Chronologically, it refers to inscriptions in all languages and scripts from this area before the Islamic era.

This means that while the vast majority of the inscriptions in OCIANA are, and will continue to be, in the Ancient North Arabian [ANA] scripts (see below), it will also contain those texts in Akkadian, Old Aramaic, Imperial Aramaic, local forms of the Aramaic script, Nabataean, Palmyrene, Greek, and Latin that have been found in Arabia, north of Yemen.

The ANA scripts are varieties of the “South Semitic script-family”, which separated from the North-West Semitic (Phoenico-Aramaic) branch shortly after the invention of the alphabet, and developed in parallel to it. In antiquity, it was used solely in Arabia and its immediate surroundings, and its only modern survivor is the vocalized alphabet used in Ethiopia for Gə‘əz, Amharic and other languages (Macdonald, 2008). In antiquity, one form of the South Semitic script-family was used in southern Arabia – the *musnad*, or Ancient South Arabian [ASA] “monumental” script, from at least the tenth century BCE (Stein, 2013). From this then developed the *zabūr*, a form of the script used to carve everyday documents on the stems of palm-leaves or on sticks (Stein, 2005a, 2005b). In the east of Arabia, between the Saudi Arabian oasis of al-Ḥasā and the Oman Peninsula, the ASA script was used to express what may be a North Arabian language, “Hasaitic”, alongside Aramaic (Overlaet, Macdonald, & Stein, 2016, pp. 132–140).

However, in the western two-thirds of Arabia, north of Yemen, a number of different alphabets developed from the South Semitic script-family and these were used by the inhabitants of oases in north-west Arabia (Dadan – modern al-‘Ulā; Figures 8.1 and 8.2 – Taymā’, and probably Dūmah – modern Dūmat al-Jandal/al-Jawf).

Figure 8.1: Dadanic inscriptions at al-‘Udhayb (al-‘Ulā, Saudi Arabia). (Photograph by C.J. Robin)

Figure 8.2: Detail of Fig. 8.1, Dadanitic inscriptions at al-‘Udhayb (al‘Ulā, Saudi Arabia). (Inscriptions U 011–019, 021–026, see OCIANA). (Photograph by C.J. Robin)

A number of different scripts of the same family were also used very widely among the nomads from southern Syria to Yemen, who were literate at different times in different areas during the second half of the first millennium BCE and the fourth century CE.⁴ By far the most numerous of these graffiti by nomads are the “Safaitic inscriptions” (Figures 8.3 and 8.4), which are found in their tens of thousands on the rocks of the deserts in southern Syria, north-eastern Jordan, and northern Saudi Arabia (Macdonald, 2010).

Almost certainly, more people in North Arabia were literate during this period than in any other part of the Middle East, and they have left us vast numbers of inscriptions. Yet, despite this, the history of Arabia is still largely known from external rather than from indigenous sources. Some of the reasons for this are set out below.

⁴ These dates are necessarily very approximate since the dating evidence for these inscriptions, almost entirely graffiti, is extremely slight.

Figure 8.3: Safaitic inscriptions at Jabal Says, southern Syria (C 25–32, see OCIANA). (Photograph by M.C.A. Macdonald)

Figure 8.4: Safaitic inscriptions on a stone at al-ʿĪsawī, southern Syria (C 3260–3264 see OCIANA). (Photograph by M.C.A. Macdonald)

8.1.1 Building a Digital Corpus: Challenges, Objectives and Perspectives

By the twenty-first century, the challenges faced by anyone trying to work with this material were considerable. Firstly, approximately 20,000 inscriptions had been published in scattered books and journal articles in many different languages. Additionally, an unknown number had been edited in unpublished dissertations, mainly in the Arab world, and many thousands were known to have been recorded but remained unedited and so unpublished.

Secondly, although Safaitic graffiti were first discovered in 1858, it was not until several decades later that some monumental ANA inscriptions were photographed or recorded by squeezes, and it was almost a century before epigraphic expeditions regularly photographed graffiti. Given that in the first 50 years the majority of texts were copied before the scripts had been deciphered, we are fortunate that, in general, the copyists were skilled, though they often made mistakes. Before the advent of digital photography, the number of films an expedition could take with it, and keep cool before and after use, was limited, so only a minority of inscriptions, particularly graffiti, was photographed. When editions were published, it was only possible to include a tiny number of photographs for reasons of cost. All this greatly hampered the progress of research into the languages, scripts, history and cultures of ancient North Arabia.

For this reason, there were virtually no research tools. The most recent grammatical sketch of the largest group (Safaitic) was published in 1943 (Littmann, 1943, pp. viii–xxiv), and of the second largest (Dadanitic) in 1954 (Caskel, 1954, pp. 60–77, repeated with minimal changes and corrections in Farès-Drappeau, 2005, pp. 61–77); there were no dictionaries, and the only list of names was published in 1971 (Harding, 1971).

Despite this situation, more and more inscriptions were recorded by Saudi and Jordanian academics and published in small handfuls, often with no photographs and little, often confused, information on their provenance.

If this situation was to be improved it was clear that all known inscriptions from ancient North Arabia needed to be sought out and brought together in a single, digital corpus, edited in a single international language to enhance access, with all available metadata, fully searchable texts and as many images as possible (preferably photographs). Between 1995 and 2003 Macdonald had begun a process of finding the sites in southern Syria where early travellers had copied Safaitic inscriptions, photographing the inscriptions, describing the sites and recording their location as accurately as possible (GPS was in its infancy and was forbidden in Syria at the time).

At the same time, with the help of Laïla Nehmé (CNRS Paris) and the late Geraldine King (independent scholar) he created a database of the Safaitic inscriptions (“The Safaitic Database”) using the platform, *4th Dimension*. This was maintained and expanded until 2012, when it was decided to use it as the basis for a corpus of *all* the inscriptions of ancient North Arabia, not just Safaitic. This was OCIANA, which was

based at the Khalili Research Centre, University of Oxford, and funded for three and a half years (September 2013 to March 2017) by the UK's Arts and Humanities Research Council (AHRC).

OCIANA's first objective was to identify all the known inscriptions of ancient North Arabia, whether published, edited in an unpublished dissertation, or recorded but unpublished. Ali Al-Manaser, a member of the project, among much else traced a large number of dissertations and sought and received permission from the authors and universities to include new editions of the inscriptions in OCIANA. Numerous scholars generously made available their photographs of published and unpublished inscriptions and gave the project permission to edit or re-edit them in OCIANA.

The second objective was to produce an up-to-date edition or re-edition of all the known inscriptions from ancient North Arabia in a single international language, English. This would be achieved by checking, and if necessary revising, a previous edition of the inscription in light of the most up-to-date knowledge. Previous readings and interpretations would be given in the *apparatus criticus* (translated into English if necessary), and there would be commentaries on the reading, translation or content of the inscription where necessary.

Thirdly, every word, name, genealogy, narrative, and prayer would be tagged to make it possible to search for all examples of words, grammatical features, expressions, personal, divine, place and group names, genealogies, etc.; something which by this time was becoming increasingly impossible to do in all the scattered publications on paper and online. This also meant that it would be easy to search across corpora to find, for instance, whether a particular word, grammatical feature or name is found in both Safaitic and Dadanitic, or Hismaic or Taymanitic. This, of course, provides the basis for the research tools that will be one of the major outcomes of the project (see below).

Fourthly, the bibliography of the inscriptions of ancient North Arabia was already large, with the added problem that scholars in the West had great difficulty hearing about the publication of books and articles produced in the Middle East, and to a lesser extent *vice versa*. Clearly, there was an urgent need for a regularly updated bibliography.

At the end of the second phase in March 2017, the database contained 42,672 (previously published and unpublished) inscriptions, the metadata of which had been entered and their data tagged over the course of three and a half years. Furthermore, over 100,000 negatives, prints and colour slides had been scanned and entered into the database.

The scripts bundled together under the heading "Thamudic" are only partially deciphered and it is therefore clearly necessary to find a way to enter the texts in a form which does not prejudice efforts to make a more satisfactory decipherment possible. It will therefore be necessary to develop a system of glyphs that, in a formalized fashion, imitate the shapes of the original glyphs in the inscriptions. This will then allow the exploration of repeated patterns of glyphs that is essential to the

decipherment of scripts. This is one of the prime aims of the third phase of the project. Once the decipherments have been made and thoroughly tested, the glyphs will be converted to roman script so that the “Thamudic” inscriptions can be searched along with all the other corpora.⁵

Having established the database and provided, for the first time, a fully-searchable corpus of most of the known inscriptions from ancient North Arabia, it is now time to use it to provide the basis for research into the languages, scripts, history and cultures of ancient North Arabia and to produce up-to-date research tools for the subject.

In the next phase, OCIANA will be used to produce concordances of words in context as the basis for the creation of online dictionaries and grammars. The *Dictionary of the Inscriptions of Ancient North Arabia* (DIANA) project based on OCIANA has already started this work (see section 8.3).

An up-to-date and easily updatable onomasticon will be produced both within each corpus and across all the corpora of the inscriptions. This will provide the material for a thorough study of the names within and across the various corpora.

Concordances of genealogies will also be produced. The Safaitic inscriptions, carved by nomads, form by far the largest corpus in the database (79%) and almost all provide genealogies varying from two to nineteen generations. The concordance will then be used not only to show the relationship of one author to another but, when combined with the provenance data of individual inscriptions, will be used to provide a picture of the movements of these nomads.

One of the most urgent and difficult problems to be faced is keeping the database up-to-date, given that several thousand previously unknown inscriptions are discovered each year, and that future research will inevitably require the revision of interpretations of individual texts. We are therefore working on the establishment of OCIANA within a university or other academic environment. It is hoped this will involve an endowment, making it possible to attract students and post-docs to continue work on the content, including both the editing of newly discovered inscriptions and the output of regularly updated research tools, as well as studies of the inscriptions and their contexts, and a constant updating of the bibliography.

8.2 The Development of OCIANA

In 2012, Michael Macdonald’s Safaitic Database was converted from the program *4th Dimension* to Filemaker Pro in preparation for Phase 2 of the OCIANA project. From September 2013, the OCIANA database was built in this application, although it also makes use of HTML, XML, SQL, and JavaScript. When planning for the development of the OCIANA database, we were aware that many projects in the field of digital

⁵ For the technical problems that we faced and resolved see section 8.2.

epigraphy had opted to base the development of their databases on XML, specifically the TEI (Text Encoding Initiative)⁶ and EpiDoc standards⁷ (developed for tagging collections of text-based data). We chose not to base our new system on this format in the input process, but to ensure that any data within the database we were developing could be outputted as marked-up XML, for the purposes of Open Access, and opted to develop OCIANA in Filemaker Pro.

Whilst XML is a useful format for sharing data and data outputs, it suffers from being fairly unforgiving in terms of data entry, which will often mean that it takes considerably longer to enter and “mark up” content in platforms that use this standard. XML is also generally inefficient when dealing with very large sets of data, as it needs to load in the entire content of its file before users can start working on individual elements. Developers can create ways to work around these limitations, but the fact remains that XML is not really a database application, but rather a well-structured and delineated flat file of text. Its strength is its platform-independence, and the self-contained nature of its content, making it a great resource for data outputs; however, its strength does not extend to data management and manipulation, which is better managed by other applications and formats.

The benefits of using Filemaker Pro were considerable. It had proved to be a robust and stable platform, storing a great amount of data, serving a considerable number of concurrent users, with very low maintenance overheads.⁸ Filemaker Pro could manage all the needs of the project, whilst ensuring that the corpus could be exported in formats that would be platform-independent and allow for sharing in line with Open Access standards.

Filemaker Pro offers an intuitive development environment, which allows for the rapid development of database solutions, whilst also offering a depth and flexibility that make it an extremely good fit for humanities data. A major benefit of the platform is that it is possible to publish Filemaker databases online with very little customisation, via the platform’s built-in web publishing engine.⁹ The database had to be accessible and searchable via the internet, and the project’s AHRC funding stipulated that the data had to be freely available as an open-access resource, and its

⁶ Further information about TEI is available online at [<http://www.tei-c.org>].

⁷ Details about EpiDoc can be found at [<https://sourceforge.net/p/epidoc/wiki/Home/>].

⁸ I have been working with Filemaker Pro since the late 1990s, and have used it to develop many different databases and applications: a Patient Information and Chemotherapy Management System for Cancer Research UK, used by the Medical Oncology Unit at Oxford’s Churchill Hospital for a period of almost ten years; and a number of databases for research projects at the Khalili Research Centre, University of Oxford.

⁹ My work on the databases for the Pitt Rivers Museum, Oxford, had made use of this functionality, and the online catalogues have proved to be stable, and surprisingly speedy, even though the two catalogues, one for museum objects, and the other for photographic collections, each contained more than 250,000 records.

content reusable through Creative Commons licensing. Filemaker Pro allowed us to achieve this without the need for developing separate platforms from those used by the internal research team.

Central to the initial development of the database was the need to convert the flat-file structure of Michael Macdonald's Safaitic Database into a relational database, in order to handle efficiently the planned addition of an estimated 20,000 additional inscription records, as well as upwards of 100,000 images. In addition to these central requirements, the project team did not set out to create a simple searchable repository, but rather to build further functionality into the system, for the generation of research outputs, including the tagging of individual elements of inscriptions, such as nouns, verbs, adjectives, place names, personal names, divine names, and so on. The aim of this work was to lay the groundwork for grammars and dictionaries of these ancient languages, as well as searchable concordances of genealogies, and other important outputs that would help to shed light on the milieu from which the inscriptions came.

As the project progressed, it was possible to identify further tools and outputs we could develop on top of the core foundations we were building, and work to build them into the developing platform. I will touch on many of these over the course of this section, but the most significant development was the ability to output the entire corpus in pdf format, and its subsequent publication on the Bodleian Library's Online Research Archive (ORA). This dataset and publication is, at the time of writing, the single largest repository of data that is hosted on the ORA.

As outlined previously, the central core element, or record type, stored in OCIANA is the information about each individual inscription. This forms the basis for the whole database. It is possible to perform very detailed searches of the inscriptions contained in OCIANA, and each inscription record contains a large amount of information, split into a series of fields.

Clearly, the main element of any inscription record is the text of that inscription, and this textual content is recorded in both transliteration and translation within OCIANA. Transliterations are presented both in roman characters and in glyphs imitating the original letter forms,¹⁰ with the customary editorial apparatus. The translations, *apparatus criticus*, commentaries and all information about each inscription are presented in a single international language, English, regardless of the language in which the inscriptions were first edited.

Moreover, alongside the text, it was necessary to record many related items of metadata. Whilst every inscription in the database has a unique OCIANA identifier, it will also have a siglum assigned to it. This will usually indicate the original publication or survey from which the inscription hails, as well as providing some

¹⁰ This was done for those used to non-roman scripts, such as Arabic, many of whom have said that they find it easier to read these glyphs than the transliteration into roman letters with diacritical marks.

indication of its geographical provenance. These are elements that would be difficult to replicate via a system of database-generated sequential numbering. Inscriptions that had already been published were known by established reference sigla, such as those published in Winnett & Harding, 1978 (WH); Littmann, 1943 (LP); and from the *Corpus Inscriptionum Semiticarum Pars V* (C).¹¹ Where we were entering previously unpublished inscriptions we opted to follow this model, and created new sigla to assign to these collections.

Inscription records also contain several other fields containing research notes specific to that inscription, such as an alternative siglum, commentary and *apparatus criticus*. However, a large number of other fields for each inscription record contain data common to large groups of inscriptions, and many of these fields are therefore connected via a relational model to other database tables within OCIANA.

Table 8.1: Some of the fields assigned to each inscription in the database, with links to other tables within OCIANA

Field	Role	Linked?
inscription_recordID	The auto-generated unique identifier for each record	Yes
inscription_siglum	The historical reference identifier for each record	No
inscription_script	The script that the inscription is written in	Yes
inscription_fullText	The transliterated text of the inscription	No
inscription_translation	The English translation of the inscription	No
inscription_appCrit	The <i>apparatus criticus</i> of the inscription record	No
inscription_commentary	Research notes and commentary on the inscription	No

In the database's table of inscriptions there are, of course, many more fields than those listed in Table 8.1 (146, in fact), with some fields containing legacy data from the Safaitic Database, and others applying calculations and alterations to other fields. For example, the "inscription_fullText" field allows the text of an inscription to be input complete with editorial marks indicating where some of the glyphs are illegible, scratched out, or otherwise uncertain, but a second field removes this mark-up information in order to allow users to search the inscriptions without these being included. In addition to these further fields, a number of fields focus on the internal administration of records, including the date a record was created, when it was last modified, which researcher entered the data, and so on. A final set of calculation fields allows the database to output its content in HTML, for the publication of unique web pages (every inscription in the database has its own static web page, allowing for indexing by Google and other search engines), as well as XML. This facilitates sharing of the data with researchers via the Bodleian Online Research Archive.

¹¹ See Ryckmans, 1950–1951.

As mentioned earlier, the table of inscriptions is one of several tables in the database, and is linked via several internal relationships to a series of other tables. These tables serve a number of important functions, including links to images of inscriptions, bibliographic references, and other data common to the corpus. Relational table links are particularly important for one of the key functions of OCIANA as a research tool, which is the functionality that allows editors to tag elements of each inscription. Individual elements of inscriptions within the corpus contain a number of differing functions or characteristics. Individual words may be tagged as grammatical elements, and names are tagged for genealogical and onomastic searches. Collections of words or phrases may be tagged as narrative elements or prayers, and genealogies are also tagged. When editors select a word or phrase, they then apply an appropriate tag to their selection, and this will then create a new related record in the appropriate supporting table within the database, allowing us to create tables of unique grammatical content, genealogies and genealogical concordances, and a detailed list of onomastic elements within inscriptions. These supporting tables have then allowed us to create comprehensive word lists, complex concordances of genealogies and words, and to work towards creating grammars and dictionaries for the scripts and languages contained within OCIANA.

Not all of this functionality is yet available to online users of OCIANA, but they still have access to the entire corpus of inscriptions, and the ability to search its content, and the contents of a number of supporting tables, in great depth.¹²

Online users of OCIANA can freely search all of the published inscriptions (42,672 in total), the table of tagged grammatical elements (123,062), the tagged onomastics (95,673), and both the tagged genealogies and their concordances (37,222). As mentioned earlier, each individual inscription has its own unique web page containing all the information about it, details of its tagged elements, a list of the bibliographic references, and all available photographs of that inscription. The citation URL is listed on the page for each record, and we would encourage anyone making use of OCIANA to include this in their publications. An example of an inscription record from the online database is shown below, with the URL for citation and linking to the record indicated at the end. The example does not show the glyphs or images here, and the list of tagged grammar and onomastics has been omitted.¹³

¹² The online version of the OCIANA database can be freely accessed at [<http://163.1.184.24/fmi/webd/OCIANA>], and an overview of the online functionality of the database is covered in a talk I gave at the Digital Humanities Summer School (DH@OxSS) in July 2015, which can be viewed at [http://krc.orient.ox.ac.uk/resources/ociana/ociana_dhoxss.mp4].

¹³ The complete record, including these elements can be viewed at [http://krc.orient.ox.ac.uk/ociana/corpus/pages/OCIANA_0033109.html].

Sigla: AH 001; Sima 1999: 35–36; D 134 **Script:** Dadanitic **Language:** Dadanitic

Transliteration

- 1: *bn[w]d/w whb'm/w* '–
- 2: *wd/w lb'n/bnw*
- 3: *s¹'d'l/d yf'n/'z–*
- 4: *llw/zll/h- nq/l-*
- 5: *dġbt/frd -hm*

Translation

- 1: {Bnwd} and Whb'm and '–
- 2: wd and Lb'n sons of
- 3: S¹'d'l of the lineage of Yf'n per–
- 4: formed the zll-ceremony of the top of the mountain for
- 5: Dġbt and so favour them

Apparatus Criticus

TEXT

Line 1: Abū l-Ḥasan followed by Farès-Drappeau: *bnd w* rather than *bn[w]d*; Sima: *w- 'tb'm* rather than *w- whb'm*. The latter is clear on the photograph, although the *h* is slightly damaged and was copied and read as *t* by Abū l-Ḥasan.

Line 2: Abū l-Ḥasan: *wm* for *wd*; Sima: *wg* for *wd*.

Lines 3–4: Abū l-Ḥasan: *'tllw tll* for *'zllw zll*.

TRANSLATION

Lines 3–4: *'zllw h- zll*, Sima: 'they covered ???'; Farès-Drappeau: '(they) offered the sacrifice'.

Line 4: *h- nq*, Abū l-Ḥasan: 'the female camels'; Sima does not translate it; Farès-Drappeau: 'the female camel'.

DISCUSSION

Hidalgo-Chacón Díez 2016: 128, for the divine name *Dġbt*.

Commentary

The restored *[w]* in the first personal name is based on the existence of the personal name *Bnwd* in the inscription AH 011/1. A root *bnd* has not been found in Semitic (Cohen et al. 1970–: 71). The translation of the construct phrase *zll h-nq* is based here and in other texts on interpreting *h-nq* either as a place name or as a common noun from the Arabic word *nīq* 'mountain-top'. Given that the inscriptions mentioning *h-nq* are located on the way up Ġabal 'Ikmah or at the top of Ġabal Umm Daraġ, it seems unlikely that they would be recording the sacrifice of female camels (*nāq* or *nūq*), as suggested by Abū l-Ḥasan and Farès-Drappeau (see the *apparatus criticus*).

Subjects: Genealogy Lineage Religion Deity Prayer Topographic features

Country: Saudi Arabia

Region: Al-Madīnah

Site: Oasis of al-‘Ulā

Latitude: 26.616667

Longitude: 37.916667

Present Location: In situ

Notes: Al-‘Uḏayb (Ġabal ‘Ikmaḥ)

References:

- [AH] Abū ‘l-Ḥasan, Ḥ. ‘A.D. *Qirā’ah li-kitābāt liḥyāniyyah min ḡabal ‘akmaḥ bi-miṭṭaqat al-‘ulā*. Al-Riyāḍ: Maktabat al-malik faḥd al-waṭaniyyah, 1997. Pages: 53–61 Plates: 1
- Cohen, D., Bron, F., Lonnet, A. *Dictionnaire des racines sémitiques: ou attestées dans les langues sémitiques: comprenant un fichier comparatif de Jean Cantineau*. Paris: Mouton (fasc. 1-2)/Leuven: Peeters (fasc. 3–), 1970–. Pages: 71
- [D] Farès-Drappeau, S. *Dédan et Liḥyān. Histoire des Arabes aux confins des pouvoirs perse et hellénistique (IVe–IIe s. avant l’ère chrétienne)* (Travaux de la maison de l’Orient 42) de la maison de l’Orient, 42). Lyon: Maison de l’Orient et de la Méditerranée — Jean Pouilloux, 2005. Pages: 212
- Hidalgo-Chacón Diez. M. del C. The divine names at Dadan: a philological approach. *Proceedings of the Seminar for Arabian Studies* 46, 2016: 125–136 Pages: 128
- Sima, A. *Die lihyanischen Inschriften von al-‘Uḏayb (Saudi-Arabien)*. (Epigraphische Forschungen auf der Arabischen Halbinsel, 1). Rahden/Westf.: Leidorf, 1999. Pages: 35–36

URL of this record (for citation):

http://krc.orient.ox.ac.uk/ociana/corpus/pages/OCIANA_0033109.html

In Phase 3 of the project, we intend to enhance the database in a number of ways, with perhaps the most interesting development relating to the group of inscriptions known as “Thamudic”. We will need to work on them by allowing the entry of inscriptions in glyph format, and then provide researchers with the ability to assign transliterations in roman characters to these glyphs at a later date, as they work to complete the decipherment of these scripts. Additional work would include moving some of the provenance data into separate related tables, as the work completed in Phase 2 has allowed us to develop a list of important sites of inscriptions in the region. At present, we are still at the early stages of planning for Phase 3, but we hope to begin work on the next stage of developments in 2018.

8.3 The Future of OCIANA

Phase 3 of OCIANA has three goals. The first is to keep the database up to date. The rapid pace of discovery requires the constant entry of new inscriptions and bibliography to ensure that it can be used to the maximum degree as a research tool. Hundreds of new Ancient North Arabian inscriptions are published each year, often in difficult-to-access publications of Middle Eastern universities or unpublished Master's and Ph.D. theses from Jordan and Saudi Arabia. Not only must these new texts be sought out, but also the photographs must be scanned, readings verified, along with the standard insertion of metadata.

In Phase 3, OCIANA will also fill an important gap in the database's current documentation: "Thamudic". This is a pending category covering the various corpora of Ancient North Arabian inscriptions that have not yet been subjected to thorough study. Thamudic now includes four categories: B, C, D, and F,¹⁴ spanning from Syria in the north to Yemen in the south. What is more, each of these classifications includes a remarkable amount of variation in letter shapes and in some cases the identification of a glyph with a phoneme is unclear – in other words, some of the scripts in the "Thamudic pending file" have not yet been fully deciphered. This fact presents a challenge to inserting data into OCIANA, in particular when it comes to transliteration. Rather than transcribing the glyphs by their assumed phonemic equivalent in the roman alphabet, of which we are often unsure, the transliteration of the Thamudic material will encode the actual letter shapes themselves. To illustrate, the glyph # represents *ḏ*, *ḏ*, and *ṭ* in Ancient South Arabian, Safaitic, and Himaic, respectively. Rather than assuming one of these values in a poorly understood Thamudic text, we will simply encode the glyph with a standardized version of the glyph itself. This neutral representation will then allow the researcher to revisit the patterns of distribution of problematic letter shapes across the entire corpus, allowing for a more precise classification of scripts and ultimately a clearer understanding of these enigmatic corpora. It is anticipated this process will eliminate, or greatly reduce, the "Thamudic" pending category and permit the recognition of new, properly understood, scripts.

The third goal of Phase 3 is the use of OCIANA as a research tool. The first sub-project of this goal is currently being realised as the *Dictionary of the Inscriptions of Ancient North Arabia* (DIANA), an online open-access supplement to the OCIANA database. The dictionary will include every lexical item contained in OCIANA. Each word will have a dedicated entry, with a full etymological discussion, ample

¹⁴ Some of the "Thamudic F" or "Southern Thamudic" graffiti have recently been deciphered by C.J. Robin and have been removed from the "Thamudic pending file" and relabeled "Himaic" from the area in southern Saudi Arabia where they are found. See Robin & Gorea, 2016.

illustrative examples, as well as synonyms. Users can easily follow a link to OCIANA to see all the attestations of a given lexical item in the corpus.

Sample Safaitic entry:

ts²wq *v.t2-stem. to long for; to feel longing.* Root: **s²wq**. [tas²awwaqa] [tas²weqa] HCH 44: ts²wq 'l-'b-h w 'l-'ḥt-h 'he longed for his father and for his sister'; C 95 wgd s¹fr dd-h f ts²wq 'he found the inscription of his paternal uncle and so was filled with longing' Variant: **ts²wqw**. [tas²awwaqaw(?)] RSIS 204: ts²wqw 'l-ṣḥ 'he longed for Ṣḥ' Variant: **ts²yq**. [tas²ayyaqa] KRS 124: ts²yq l-ḥbb 'he longed for a friend' Third feminine singular: **ts²wqt**. [tas²awwaqat] Damascus Museum 2786 = RyDamas 5537: l PN w ts²wqt 'l-nmn 'by PN and she longed for Nmn' Note: His: ts²wq (CH.R716) || The equivalent of Car ištāqa 'ilay-hi 'he was, or became, desirous of it ... [or he longed for it in his soul]' (Lane, 1620b).¹⁵

Other sub-projects of this Phase, pending funding, include the creation of an up-to-date onomasticon, and in-depth studies of the individual Ancient North Arabian corpora, such as Hismaic, the various properly identified scripts emerging from the study of "Thamudic", and the minor corpora.

Currently, DIANA and the maintenance of the database are progressing, but in order for phase 3 to be fully realized, the team is preparing applications for funding and institutional support, both at Oxford and Leiden University.

Bibliography

- Farès-Drappeau, S. (2005). *Dédan et Liḥyān. Histoire des Arabes aux confins des pouvoirs perse et hellénistique (IVe–IIe s. avant l'ère chrétienne)* (Travaux de la maison de l'Orient, 42). Lyon: Maison de l'Orient et de la Méditerranée - Jean Pouilloux.
- Caskel, W. (1954). *Liḥyan und Liḥyanisch* (Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen, Geisteswissenschaften 4). Köln: Westdeutscher Verlag.
- Harding, G.L. (1971). *An Index and Concordance of Pre-Islamic Arabian Names and Inscriptions* (Near and Middle East Series 8). Toronto: University of Toronto Press.
- Lane, E.W. (1863–1893). *An Arabic-English Lexicon, Derived from the Best and Most Copious Eastern Sources*. London: Williams & Norgate.
- Littmann, E. (1943). *Safaitic Inscriptions. Syria. Publications of the Princeton University Archaeological Expeditions to Syria in 1904–1905 and 1909. Division IV. Section C*. Leiden: Brill.
- Macdonald, M.C.A. (2008). The Phoenix of Phoinikēia: Alphabetic reincarnation in Arabia. In J. Baines, J. Bennet, & S. Houston (Eds.), *The Disappearance of Writing Systems: Perspectives on Literacy and Communication* (pp. 207–229). London: Equinox.
- Macdonald, M.C.A. (2010) Ancient Arabia and the written word. In M.C.A. Macdonald (Ed.), *The development of Arabic as a written language* (Supplement to the Proceedings of the Seminar for Arabian Studies volume 40) (pp. 5–27). Oxford: Archaeopress.
- Overlaet, B., Macdonald, M.C.A., & Stein, P. (2016). An Aramaic–Hasaitic bilingual inscription from a monumental tomb at Mleiha, Sharjah, UAE. *Arabian Archaeology and Epigraphy*, 27, 127–142.

¹⁵ The sigla of the inscriptions can be found in OCIANA.

- Robin C.J. & Gorea M. (2016). L'alphabet de Ḥimā (Arabie séoudite). In I. Finkelstein, C.J. Robin, & T. Römer (Eds.), *Alphabets, texts and artifacts in the ancient Near East. Studies presented to Benjamin Sass* (pp. 310–375). Paris: Van Dieren.
- Ryckmans, G. (Ed.). (1950–1951). *Corpus Inscriptionum Semiticarum. Pars V. Inscriptiones Saracenicas continens, Tomus 1. Inscriptiones Safaiticae*. (2 volumes). Paris: Imprimerie nationale.
- Stein, P. (2005a). The Ancient South Arabian minuscule inscriptions on wood: A new genre of pre-Islamic epigraphy. *Jaarbericht van het Vooraziatisch-Egyptisch Genootschap 'Ex Oriente Lux'*, 39, 181–199.
- Stein, P. (2005b). Stein vs. Holz, *musnad* vs. *zabūr* — Schrift und Schriftlichkeit im vorislamischen Arabien. *Die Welt des Orients*, 35, 118–157.
- Stein, P. (2013). Paleography of the Ancient South Arabian script. New evidence for an absolute chronology. *Arabian Archaeology and Epigraphy*, 24, 186–195.
- Winnett, F.V. & Harding, G.L. (1978). *Inscriptions from Fifty Safaitic Cairns* (Near and Middle East Series 9). Toronto: University of Toronto Press.