The Arabic papyrology database

Kaplony, A
The dry climate of Egypt has preserved about 300,000 documents from the period up to the Ottoman conquest in 1517, mostly on papyrus or paper, divided more or less equally between Old Egyptian, Greek, and Arabic, plus some more in Hebrew, Aramaic, Pehlevi, etc. These documents allow us unparalleled insights into 4,500 years of administrative, social, economic, and legal history. Yet the Arabic documents, the earliest of which date from about 640, remain by far the least researched and least used.

As in the history of the ancient world and medieval Europe, using documents will radically challenge what literary sources tell us. Up to now, historical research mostly relies on literary sources, while, due to the bad state of conservation and availability, the use of the abundant documentary material is limited. Yet, where original documents have been used, they have contributed to a new understanding of medieval Islamic history. For example, the official letters sent by Qurra ibn Sharık, governor of
Egypt from 709 to 714, preserved on papyrus, put his rule in quite a different light than the later anti-Umayyad biased sources. Similarly, the ongoing debate on the date of the Koran is challenged by early Koranic manuscripts found in Yemen.

If we compare Arabic papyrology to the much more developed Greek papyrology, we notice that although there is no difference in the number of Arabic and Greek documents, many fewer Arabic documents have been published, there is much less research on Arabic documents, and the results of Arabic papyrology have been much less used in Arabic and Islamic studies than the results of Greek papyrology have been in Greek and ancient studies.

Most strikingly, Arabic papyrologists have almost no reference tools. Greek papyrologists have the *Sammelbuch* and the *Berichtigungsliste*, both published for almost a hundred years. These collect the texts of edited papyri, which are often published in difficult-to-access journal articles, as well as all later suggestions about how to read them. Greek papyrologists likewise have strong web-based tools, such as the *Checklist* of all editions. One can search for content, document type, date, etc. in a number of metadata databases such as the *Heidelberger Gesamtverzeichnis*, the Cologne-based *Trismegistos* database and the New York-based *Advanced Papyrological Information System*. There is the *Duke Data Bank*, a searchable full-text database giving access to the full text of all published papyri, and the Heidelberg *WörterListen* giving the occurrences of months and days, personal names and goddesses. Greek papyrologists also have at their disposal the *Thesaurus Linguae Graecae*, a giant database offering the text of a great part of literary texts written in Greek from Homer to the fall of Byzantium in 1453.

So what about reference works for Arabic papyrology? Besides the glossaries which accompany the editions, there is the *Dictionary by Werner Diem and Hans-Peter Radenberg*, basically an index to the late Shlomo D. Goitein’s *magnum opus*, and Simon Hopkins’s grammar.

2 *Sammelbuch griechischer Urkunden aus Ägypten* (1915-).
3 *Berichtigungsliste der griechischen Papyrusurkunden aus Ägypten* (1922-).
5 *Heidelberger Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens* einschliesslich der Ostraka usw., der lateinischen Texte, sowie der entsprechenden Urkunden aus benachbarten Regionen, by D. Hagedorn et al. (http://www.rzuser.uni-heidelberg.de/~gv0/gvz.html).
6 *Trismegistos*, by Mark Depauw et al. (http://www.trismegistos.org).
12 S. D. Goitein, *A Mediterranean Society: The Jewish Communities of the Arab World as
The only web-based tools relevant in our context are the Checklist of editions by the International Society for Arabic Papyrology (ISAP) and the wordlist of the Princeton Geniza Project. As for dictionaries of medieval Arabic, there are only two, the Wörterbuch der klassischen arabischen Sprache by Manfred Ullmann in Tübingen which, unfortunately, covers only two letters of the alphabet, and the Dictionary of Mediaeval Judaeo-Arabic Texts recently published in Hebrew by Joshua Blau. And there are the outdated dictionaries like Edward Lane’s Lexicon which continue the tradition of medieval Arabic dictionaries like Ibn Manẓūr’s (13th c.) Lisān al-‘Arab.

To cut a long story short: Arabic papyrology has next to no reference tools. But because papyrology depends on comparing parallels in other texts, Arabic papyrologists have so far been reduced to making their own collections of rare words, formulas, idiomatic phrases, grammatical features, etc. This present state of Arabic papyrology – of a small group of scholars working individually with limited interaction – can be described as, compared to Greek papyrology, a structural delay of about one hundred years.

To change this situation, we have started to build up a virtual research community of scholars. To attract and form young scholars, we developed in Zurich the Arabic Papyrology School, an internet-based interactive tutorial, which introduces students and professionals to the reading of Arabic documents. Developing this tutorial was much helped by the International Society for Arabic Papyrology (ISAP) and became an interactive process with the worldwide community of Arabic papyrologists. Since January 2004, the Arabic Papyrology School is freely accessible on the internet. So far, 684 users have enrolled themselves and made some use of the School, and twelve of them have been awarded an academic certificate.

At the first conference of the International Society for Arabic Papyrology, held in Cairo in March 2002, Prof. Klaas A. Worp of Leiden University suggested that Arabic papyrology should recover lost ground by adopting as many as possible of the well-tried structures of Greek papyrology. He suggested that we should start to create our own Sammelbuch gathering all existing publications of Arabic documents, modelled after the Sammelbuch of Greek papyrology. While working on the Arabic Papyrology School, we revised his idea and started creating an Arabic Papyrology Database of published

16 M. Ullmann, Wörterbuch der klassischen arabischen Sprache, 2 vols. published so far (Wiesbaden, 1970–).
18 E. W. Lane, Arabic-English Lexicon, 8 pts. (Edinburgh, 1863–1893).
19 Lisān al-‘Arab li-... Ibn Manẓūr ..., ed. Aḥmad Fāris ([Cairo], AH 1300/[1882]).
documents with both full text and metadata.

We started from two weak points of Arabic papyrology which, at the same time, could be considered advantages. The first one was that up to now, only a small number of Arabic documents has been published. The Checklist mentions about 50 monographs and 100 articles, for a total of about 2,000 documents, so building up a database of all published documents, with all later readings added, would not be beyond our reach.

But what do Arabic papyrologists really need for their work? To answer this question, we presented the project at the second ISAP Conference in Granada, in March 2004, at a special workshop in Zurich in January 2005, and at the third ISAP conference in Alexandria in March 2006. On all three occasions, the emerging community of Arabic papyrologists suggested that we provide, first of all, the full Arabic text of each document, including all readings subsequently proposed. These should be accompanied by a translation, where available, as well as a link to a scan of the original document. In the end, we decided to distinguish five layers: (1) a Plain Line written in Arabic characters, with only those diacritical dots and vowels found in the original, and with all suggested emendations marked; (2) Plain Words, with words separated by spaces, whether these are in the original or not; (3) Full Dots, adding diacritical dots which are not in the original; (4) Full Dots and Vowels, with all vowel signs added, and (5) Latinized, transliterating the text into Latin characters. The five layers give variant readings which can all be seen at the same time and which are fully searchable. For linguistic research, we lemmatize each word by its radicals, grammatical form (case, tense, etc.) and word category (noun, verb, etc.). and mark certain word categories (colors, personal names, professions, etc.). Non-standard forms are also referred to.

In addition to that, we decided to retain, for each document and where possible, some metadata: name in publication; inventory number; place of origin; date provided in the document, i.e. Muslim date, Coptic date, indiction, and/or date of the week; date converted into the Julian calendar; material (papyrus, paper, etc.); width and height; scribe(s); document type (protocol, marriage contract, business letter, etc.); document title as given by the editors; language(s) used; edition(s); translation(s); image(s); and further remark(s). By using these categories, we follow the informal agreement on European papyrological metadata databases reached at the Cologne workshop held in July 2005.

The delay of Arabic papyrology compared to Greek papyrology as described above had the advantage that there was no traditional structure we needed to adhere to, so we were able to experiment with a new and innovative approach, creating an electronic database accessible on the web at any time to anyone from anywhere. The problem of transmitting Arabic script is definitely solved by Unicode. We have developed technical tools to make it possible to search for words with missing or limited diacritics exactly as they appear on the document or full diacritics, with vowels as on the document or full vocalization. Looking up a name of which some characters are known and others unknown, whatever the number and the
length of the known and unknown parts, is no longer a dream.

Since spring 2006, the Arabic Papyrology Database has been freely accessible on the internet. It already includes 638 documents, i.e., more than one fourth of all published. For strategic reasons, we focus on one hand on archives published, on the other hand on documents held by the Papyrussammlung of the Österreichische Nationalbibliothek in Vienna. Of the archives, we have already entered the administrative letters of Qurra ibn Sharīk (709–714) and the Marchands d’étoffe letters (9th c.), we are currently working on the Quseir merchant letters (14th c.), and we hope to add soon the Khurasan documents (8th c.). For the documents held in Vienna, the database includes those published in a good number of older articles by Adolf Grohmann, as well as

---

21 Y. Rāġib, Marchands d’étoffe du Fayyoum au IIIe IXe siècle d’après leurs archives (actes et lettres), 4 vols. published so far (Cairo, 1982–).
23 G. Khan, Arabic Documents from Early Islamic Khurasan (in press).
26 A. Grohmann, From the World of Arabic Papyri (Cairo, 1952).
27 A. Grohmann, Arabic Papyri in the Egyptian Library, 6 vols. (Cairo, 1934–62).