



HAL
open science

Editing New Testament Arabic Manuscripts on a TEI-base: fostering close reading in Digital Humanities

Claire Clivaz, Sara Schulthess, Martial Sankar

► To cite this version:

Claire Clivaz, Sara Schulthess, Martial Sankar. Editing New Testament Arabic Manuscripts on a TEI-base: fostering close reading in Digital Humanities. 2016. hal-01280627v3

HAL Id: hal-01280627

<https://hal.science/hal-01280627v3>

Preprint submitted on 1 Sep 2016 (v3), last revised 3 Jun 2017 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Editing New Testament Arabic Manuscripts on a TEI-base: fostering close reading in Digital Humanities

Claire Clivaz^{1*}, Sara Schulthess^{1,2}, Martial Sankar¹

1 Swiss Institute of Bioinformatics, VITAL-IT, Switzerland

2 University of Lausanne, Switzerland

*Corresponding author: claire.clivaz@sib.swiss

Abstract

Convinced by Franco Moretti that “quantitative research provides data not interpretation” [Moretti, 2005, 9], we consider close reading as the necessary step to cross big data and interpretation and core duty of the Humanities. To test its potential in a neglected field – the Arabic manuscripts of the Letters of Paul of Tarsus – we are developing an enhanced, digital edition as a further step in a Swiss National Fund project. This short paper presents the development of this edition and perspectives on the future of this project. Based on the *Vercelli Book Digitale* project and its tool EVT, the digital edition provides a transcription, a standardized and vocalized version, as well as French and English translations, with texts encoded in Arabic TEI. Further features are in preparation such as textlink and a philological commentary. The work will be extended to the unique New Testament trilingual manuscript Arabic-Greek-Latin, the Marciana Library Gr. Z. 11 (=379), 12th century.

Keywords

Digital edition, Arabic TEI, New Testament, Letters, Paul of Tarsus

INTRODUCTION

Digital Humanities are often understood as a switch towards big data studies. Convinced by Franco Moretti that “quantitative research provides data not interpretation” [Moretti, 2005, 9], we consider close reading as the necessary step to cross big data and interpretation and core duty of the Humanities. To test its potential on a neglected field – the Arabic manuscripts of the Letters of Paul of Tarsus – we would like to develop a digital enhanced edition in the next step of a Swiss National Fund 2013-2016 [http1]. This short paper presents the editorial tool we have begun to build [http2] and that we would like to develop in the continuation of this research fund.

I HUMANITIES PROJECT BECOMES DIGITAL

1.1 The case of the Arabic manuscripts of New Testament

Our interest for the Arabic manuscripts of the NT was raised by two observations. Firstly, the very low number of publications on the topic is intriguing; the reference book still used today goes back to 1944 [Graf, 1944]. In fact, while Western scholars started very early to study the Arabic versions [Raimundi, 1590], since the end of the 19th century biblical textual criticism lost increasingly interest for this tradition [Vollandt, 2013]. Only a few orientalists continued

to study the field [Kashouh, 2012, 9-33]. Secondly, investigating the digital evolution of the New Testament Textual Criticism (NTTC) on the Internet, we discovered Islamic websites that are studying Greek New Testament manuscripts [Clivaz, 2013; Schulthess, 2013]. These websites provide reference material for the discipline of NTTC and use images of manuscripts, which are today more and more accessible online, to underline the variants and scribal interventions [http3]. This phenomenon is a reactualisation of a long polemical tradition, the concept of the falsification of the Scriptures (*at-tahrīf*), using now digital media [Schulthess, 2016]. On the other side, several Christian Evangelical websites attempt to answer the accusations of *tahrīf*, and one can observe interactions between apologists from opposite communities. Interestingly, Arabic New Testament manuscripts are also integrated in these discussions. Since we started to prepare the project, we were able to observe a global renewed interest, with numerous new publications and projects [Kashouh, 2012; Arbache, 2012; Moawad, 2014; http4]. Furthermore, some identities issues are present in this contemporaneous research, for example the defence by HikmatKashouh, a Christian Lebanese scholar, of the existence of pre-Islamic translations, a minority point of view (see 2.1).

1.2 The project becomes digital

If the starting point of our interest was digital, we started with a classical Humanities project, and according to the rules still in use in the Faculties in Humanities, the PhD resulting from this project will be presented in autumn 2016 under the form of a paper book. We have chosen to study the Arabic manuscripts of Pauline letters, especially the *VaticanusArabicus* 13, a manuscript playing an important role in this controversial question of the existence of pre-Islamic biblical translations into Arabic. However, regarding the context we described, it seems to us very important to do a research accessible online, with defined methodologies and possibilities of discussions. The logical next step for this project is then to initiate a digital scholarly enhanced edition of the First Letter to the Corinthians in the Vat. Ar. 13.

The acquisition of encoding skills should today be one of the duties of a Humanist PhD student working on texts, but such trainings are not included at the moment in the Swiss PhD curricula. It was then a great opportunity to be welcomed at the Swiss Institute of Bioinformatics (VITAL-IT group), where we have begun to develop tools for the edition. We started from the open source visualization tool of the *Vercelli Book Digitale* project, which deals with XML-encoded text [http5]. It allows us notably to edit several layers of the text (image, transcription, standardized text, translation [http2]). Beyond all the questions of access and open science, it is all the more necessary to produce an open edition of this manuscript, considering that this neglected field is often misused as a weapon for identity battles. Intercultural meetings and confrontations of diverse perspectives can only occur online, since these worlds are still not joined in the academic sphere. This reason in particular motivates us to apply for the continuation of this fund which relies on the new SNF rules for innovations in project funding [http6].

1.3 Digital edition for Early Middle Arabic

Digital editions offer very challenging features for editing Arabic manuscripts. Among digital editions, different projects propose several versions of the same text, a process revealed by “layers” or “levels”. The *Queste del Saint Graal* project [http7] applies this method to Medieval French, *Vercelli Book Digitale* project [http5] to Old English, The *Coptic Scriptorium* [http8] to Coptic, and for Arabic there is the *Arabic Papyrology Database* (APD

[http9]). As demonstrated by the APD, the “edition by layers” method is particularly relevant for “non standardized” Arabic, as it is the case for the so-called “Early Middle Arabic”, used notably by Arabic-speaking Christians in the First Millennium. The text of Vat. Ar. 13 corresponds to the features described as Early Middle Arabic by Joshua Blau [Blau 1966], a helpful linguistic category despite some criticism of its historical development [Kouloughli, 2007]. A digital edition allows for the editing of the text as found in the document, and secondly the text with linguistic adaptations: missing diacritical points, *tashkīl* (with or without vocalization), and so forth. In our case, the aim was to edit the First Letter to the Corinthians in the Vat. Ar. 13, one of the oldest Arabic codices of the New Testament. We wanted the digital edition to provide the images of the manuscript, two layers of the text, an accurate diplomatic transcription (respecting graphic peculiarities for example) and a standardized and vocalized version, as well as a translation.

II DIGITAL ENHANCED EDITION OF VAT. AR. 13

2.1 The *VaticanusArabicus* 13

The manuscript, hold today in the Vatican Library, comes with great certainty from the Arabicized Greek Orthodox community in the region of Syria-Palestine. It is undated but some part of the manuscript probably goes back to the 9th century. Due to its early age, the manuscript was comparatively often object of study but was never edited. It contains parts of Matthew, Mark and Luke and the Pauline letters (with Hebrews at the end). Thanks to the colophon, we know that it originally contained the Psalms, the four Gospels, Acts, the Catholic letters and Pauline letters. This is very unusual for the early time and Graf described it as the possible oldest collection manuscript of the NT in Arabic [Graf, 1944, 138]. Kashouh argues in his thesis that the text of the oldest part of the Gospels goes back to the sixth or early seventh century [Kashouh, 2012, 142-170], a opinion that was questioned by several researchers [Griffith, 2013, 116-117; Monferrer-Sala, 2015a]. We focused on the First Letter to the Corinthians, because only the Gospels parts caught the attention of the research and we wanted to start the study of the second corpus. This situation is probably now evolving, since Juan-Pedro Monferrer-Sala published a study of the Letter to Philemon [Monferrer-Sala 2015b].

2.2 The Edition Visualisation Tool

As mentioned before, we started our Humanities project as a pre-digitized one. We made the choice to provide a transcription of the text (1) as well as a standardized text with vocalization (2). The reader can in this way fully follow the interpretations we made for the translation and has also all the information needed to understand the commentary. As a result, we can consider that we have two versions of the same Arabic text, because the transcription and the standardized text, notably through the vocalization, present constantly differences. Here an exemple:

- (1) بولوس المدعى رسول يسوع المسيح بمسره الله وسناتين الاخ
(2) بُوْلُوسُ الْمُدْعَى رَسُولَ يَسُوعَ الْمَسِيحِ بِمَسْرَةَ اللَّهِ وَسُنَّتَيْنِ الْآخِ

Thus, we started the conception of the digital edition having already two Arabic texts: a transcription (1) and a standardized and vocalized version (2), as well as a French translation; the English one is forthcoming. To begin with the tool of the *Vercelli Book Digitale* project (VBD) was a very interesting option for us. The VBD, much like the *Queste del Saint Graal* project and others, is open source and uses TEI XML for encoding the data, two aspects

we stand for. In fact, there is till now very few editions of Arabic manuscripts following the TEI standards [http10; http11]. The VBD team develops their project with the aim to provide an open source software tool to deal with TEI XML-encoded text [Rosselli del Turco et al., 2014-2015]. The software, called *Edition Visualisation Tool* (EVT), was built on open and modern web technologies such as HTML/CSS3 and JQuery framework. The beta version of the tool was released at the end of 2014 and uploaded under public license [http12]. The main benefits of EVT is the ability to automatically render TEI-encoded text into a modern user-friendly front-end. We took advantage of this key feature to wrap together the Arabic text (transcription (1) and normalized text (2)) and the translation of the images of our manuscripts. To make it possible, few modifications were performed into the XSL style sheet template files, mainly for html-rendering of TEI-tag or attributes. Considering the manuscript as our starting point, we used the embedded transcription [http13], made possible by the development of XSLT-based TEI XML formatting in EVT [Rosselli del Turco et al., 2014-2015]. As we had two different versions of the text and not only some character or word changes, we chosen to use the element <choice> and the sub-element <orig> for (1) and <reg> for (2), for each line. It allows the reader to compare easily (1) and (2) (see Figure 1 below). In addition, we marked the biblical verses with the <milestone> element, a very “neutral” way of marking something not present in the manuscript. Other elements as <unclear> and <add> were used in case of a scribal intervention.

III FURTHER PERSPECTIVES AND CONCLUSION

3.1 Enhancement of the Vat. Ar. 13 edition

The first goal regarding this project is to provide the edition of the complete First Letter to the Corinthians, that is 33 folios. Luckily, we are able to provide high-quality images. To offer the best reading experiment, we want to use the “Textlink” feature provided by the EVT, which offers the possibility to link zones in the manuscripts to the corresponding text and to highlight them. Furthermore, we want to add to the more interesting passages in the text the philological commentary that we have made in comparing with the Greek and Syriac versions, which will allow the reader to access to the pre-digitized Humanities research that we mentioned earlier.

3.2 Continuation of the project with a trilingual manuscript

As a potential extension of this project, we would like to test a continuous data and research publication model of a digital, enhanced edition of the unique trilingual New Testament manuscript Greek/Latin/Arabic (12th century, Marciana Library Gr. Z. 11 (=379)). The features of the edition of the Vat. Ar. 13 will be adapted to the manuscript with three columns in these three languages. Peculiar efforts should be allocated to deliver an optimal user experience to scholars interested to compare online the three versions. Software development process should involve several round of scholar/user feedback iterations through interdisciplinary collaborations. Finally, we will compare, on some parts of this manuscript, the manual transcription, and a handwritten text recognition one, with the tool *Transkribus*[http14], provided by the Digitisation and Digital Preservation group University of Innsbruck.

IV TABLES AND FIGURES

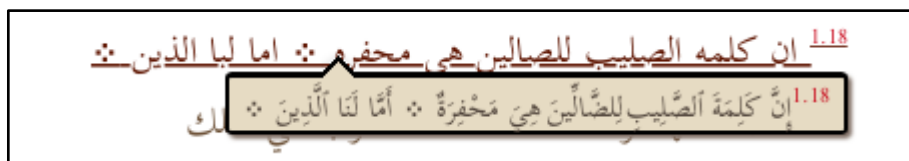


Figure 1. Screenshot of http2. Transcription and standardized text.

V REFERENCES

- Arbache S. *L'Évangile arabe selon Saint Luc: Texte du VIIIe Siècle, copié en 897*. Safran (Bruxelles), 2012.
- Blau, J. *A Grammar of Christian Arabic, Based Mainly on South-Palestinian Texts from the First Millennium* (CSCO 267). Secrétariat du Corpus SCO (Louvain), 1966.
- Clivaz C. Internet Networks and Academic Research: the Example of the New Testament Textual Criticism. *Digital Humanities in Biblical, Early Jewish and Early Christian Studies* (Scholarly Communication 2). Ed. by C. Clivaz, A. Gregory and D. Hamidović. Brill (Leiden), 2013:151-173.
- Graf G. *Geschichte der christlichen arabischen Literatur, Vol. 1*. Biblioteca Apostolica Vaticana (Rome), 1944.
- Griffith S.H. *The Bible in Arabic. The Scriptures of the "People of the Book" in the Language of Islam*, Princeton University Press (Princeton), 2013.
- http1 <http://unil.ch/nt-arabe>
- http2 <http://tarsian.vital-it.ch>
- http3 http://alta3b.wordpress.com/2009/04/06/translation_jn-1-1/
- http4 <http://biblia-arabica.com>
- http5 <http://vbd.humnet.unipi.it>
- http6 <http://snf.ch/en/funding/projects/innovations-project-funding/Pages/default.aspx>
- http7 <http://portal.textometrie.org/demo/>
- http8 <http://data.copticscriptorium.org/>
- http9 <http://www.apd.gwi.uni-muenchen.de:8080/apd/project.jsp>
- http10 <http://www.ancientwisdoms.ac.uk/library/>
- http11 <http://dare.uni-koeln.de>
- http12 <http://sourceforge.net/projects/evt-project/>
- http13 <http://www.tei-c.org/release/doc/tei-p5-doc/fr/html/PH.html#PHZLAB>
- http14 <https://transkribus.eu/Transkribus/>
- Kashouh H., *The Arabic Versions of the Gospels, The Manuscripts and their Families* (Arbeitszurneute Testamentlichen Textforschung 42). De Gruyter (Berlin/Boston), 2012.
- Kouloughli D. E. Moyen Arabe et questions connexes. 2007. http://cle.ens-lyon.fr/arabe/moyen-arabe-et-questions-connexes-30683.kjsp?RH=CDL_ARA120000.
- Moawad S. *Al-As'ad Abū Al-Faraġ Hibat Allāh Ibn Al-'Assāl: Die arabische Übersetzung der vier Evangelien*. Alexandria School (Le Caire), 2014.
- Monferrer-Sala J.P. Translating the Gospels into Arabic from Syriac: Vatican Arabic 13 Restored Section, Strategies and Goals. *Arabica*. 2015a;62:435-458.
- Monferrer-Sala J.P., The Pauline Epistle to Philemon from Codex Vatican Arabic 13 (Ninth Century CE) Transcription and Study. *Journal of Semitic Studies*. 2015b;60(2):341-371.
- Moretti F. *Graphs, Maps and Trees*. Verso (London), 2005.
- Raimundi G. B. *Al-Ingil al Muqaddas. Evangelium sanctum*. Rome, 1590-1591.
- Rosselli Del Turco R., Buomprisco G., Di Pietro C. et al. Edition Visualization Technology: A Simple Tool to Visualize TEI-based Digital Editions. *Journal of the Text Encoding Initiative*. 2014-2015;8. <http://jtei.revues.org/1077>.
- Schulthess S. The Role of the Internet in New Testament Textual Criticism: The Example of the Arabic Manuscripts of the New Testament. *Digital Humanities in Biblical, Early Jewish and Early Christian Studies* (Scholarly Communication 2). Ed. by C. Clivaz, A. Gregory and D. Hamidović. Brill (Leiden), 2013:71-82.
- Schulthess S., Tahrīf in the Digital Age. *Ancient Worlds in Digital Culture* (Digital Biblical Studies 1). Ed. C. Clivaz, P. Dilley, D. Hamidović, in coll. with A. Thromas. Brill (Leiden), 2016 ;214-230..
- Vollandt R. Some Historiographical Remarks on Medieval and Early-Modern Scholarship of Biblical Version in Arabic: A Status Quo. *Intellectual History of the Islamic World*. 2013;1: 25-42.

ANNEX 1

Many thanks to Prof. Ioannis Xenarios, director of VITAL-IT, for his support to our work.

ANNEX 2

All hyperlinks have been checked on 08/31/16.