



**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-01280627v1

**HAL Id: hal-01280627**

**<https://hal.science/hal-01280627v1>**

Preprint submitted on 29 Feb 2016 (v1), 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<sup>1\*</sup>, Sara Schulthess<sup>1,2</sup>, Martial Sankar<sup>1</sup>

1 Swiss Institute of Bioinformatics, VITAL-IT, Switzerland

2 University of Lausanne, Switzerland

\*Corresponding author: [claire.clivaz@isb-sib.ch](mailto:claire.clivaz@isb-sib.ch)

### Abstract

Since “quantitative research provides data not interpretation” (F. Moretti), close reading should be considered as the necessary bridge between big data and interpretation, as well as being the 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 a digital, enhanced edition of the *1<sup>st</sup> Letter to the Corinthians* in the *Vaticanus Arabicus* 13. This short paper presents an editorial tool based on the *Vercelli Book Digitale* project; it 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 development such as TexLink and a philological commentary. We hope to extend the project to the unique New Testament Arabic-Greek-Latin manuscript from the Marciana Library Gr. Z. 11 (=379), 12th century.

### Keywords

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

### INTRODUCTION

The Digital Humanities are often understood as a switch towards big data studies. Having been convinced by [Moretti, 2005, 9] that “quantitative research provides data not interpretation”, we consider close reading as the necessary step to bridging the gap between big data and interpretation as well as being the 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 a digital, enhanced edition of the *1<sup>st</sup> Letter to the Corinthians* of Paul of Tarsus in a Swiss National Fund 2013-2016 project [http1]. This short paper presents the editorial tool we have begun to build [http2] and would like to develop for the 12<sup>th</sup> century Gr. Z. 11 (=379), the unique New Testament Arabic-Greek-Latin manuscript.

### I HUMANITIES PROJECT BECOMES DIGITAL

#### 1.1 The case of the New Testament Arabic manuscripts

Our interest in the Arabic manuscripts of the New Testament was piqued 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, despite Western scholars starting to study the Arabic versions very early on [Raimundi, 1590], since the end of the 19th century biblical textual criticism increasingly lost interest in this tradition [Vollandt, 2013]. Only a few orientalist continued to study the field [Kashouh, 2012, 9-33]. Secondly, investigating the digital evolution of New Testament Textual Criticism (NTTC) on the Internet, we discovered

Islamic websites that are studying Greek New Testament manuscripts [Clivaz, 2012, 7-9; Clivaz, 2013; Schulthess, 2013]. These websites provide reference material for the discipline of NTTC and use images of manuscripts, which are increasingly accessible online, to underline the variants and scribal interventions [http3]. This phenomenon is a reactualization of a long-standing polemical tradition, the concept of the falsification of the Scriptures (*at-tahrīf*), now using digital media [Schulthess, *forthcoming*]. Furthermore, 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 into these discussions. Since we started our research, we were able to observe a global renewed interest as evidenced by the appearance of numerous new publications and projects [Kashouh, 2012; Arbache, 2012; Moawad, 2014; http4]. Furthermore, some identity issues are present in this contemporaneous research, for example: Hikmat Kashouh, a Christian Lebanese scholar, and his defense of the existence of pre-Islamic translations, a minority point of view (see 2.1).

### **1.2 The project becomes digital**

Despite the starting point of our interest being 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 in the form of a paper book. We have chosen to study the Arabic manuscripts of Pauline letters, especially the *Vaticanus Arabicus* 13, a manuscript playing an important role in the controversial question of the existence of pre-Islamic biblical translations into Arabic. However, with respect to the context, it is very important to do research that is accessible online, with defined methodologies and the possibility of discussions. The logical next step for this project is then to initiate a digital, scholarly enhanced edition of the *1st Letter to the Corinthians* in the Vat. Ar. 13.

Today, the acquisition of encoding skills should be one of the duties of a Humanist PhD student working on texts, however such training is not currently included in the Swiss PhD curricula. It was therefore a great opportunity to be welcomed to the Swiss Institute of Bioinformatics (VITAL-IT group), where we have begun to develop our editing tool [http2]. We started with the open source visualization tool, the *Vercelli Book Digitale* project, which deals with XML-encoded text [http5]. It allows for the editing of several layers of text (image, transcription, standardized text, translation). Beyond all the questions of access and open science, it is all the more necessary to produce an open edition of this manuscript, especially 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 particular reason is a motivation for us to apply for the continuation of this grant. We want to further explore the impact of online research in this field.

### **1.3 Digital edition for Early Middle Arabic**

Digital editions offer challenging features with respect to 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 [http6] applies this method to Medieval French; the *Vercelli Book Digitale* project [http5] to Old English; the *Coptic Scriptorium* [http7] to Coptic; and for Arabic there is the *Arabic Papyrology Database* (APD [http8]). 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. We will not discuss the notion of Early Middle Arabic here. Developed by [Blau 1966], the text of Vat. Ar. 13 corresponds to the described features, even if we don’t agree with the historical development according to the criticism of [Kouloughli, 2007]. A digital edition allows for the editing of the text as found in the document and with its linguistic adaptations such as 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 *Vaticanus Arabicus* 13

The manuscript, held today in the Vatican Library, certainly comes from the Arabized Greek Orthodox community in the region of Syria-Palestine. It is undated but some parts of the manuscript probably go back to the 9th century. Due to its age, the manuscript was relatively often the object of study but was never edited. It contains parts of Matthew, Mark and Luke as well as the Pauline letters (with Hebrew at the end). Thanks to the colophon, we know that it originally contained the Psalms, the four Gospels, Acts, and the Catholic and Pauline letters. This is very unusual for this early time period and [Graf, 1944, 138] describes it as possibly the oldest collection of NT manuscripts in Arabic. In his thesis, [Kashouh, 2012, 142-170] argues that the text of the oldest part of the Gospels goes back to the sixth or early seventh century, an opinion that has been 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 has been evolving since [Monferrer-Sala 2015b] published a study of the Letter to Philemon.

### 2.2 The Edition Visualization Tool

As mentioned previously, 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). In this way, the readers can follow the interpretations we made of the translation and they also have 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, constantly present differences. An example:

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

Thus, we started the conception of the digital edition with 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 developed tool of the *Vercelli Book Digitale* project (VBD) was an interesting option. 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 support. In fact, until now there have been very few editions of Arabic manuscripts following the TEI standards

[http9; http10]. The VBD team is developing their project with the aim of providing 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 [http11]. The main benefit of EVT is its 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 this possible, a few modifications were made to the XSL style sheet template files, mainly for the html-rendering of TEI-tags or attributes. Considering the manuscript as our starting point, we used the embedded transcription, made possible by the development of XSLT-based TEI XML formatting in EVT [Rosselli del Turco et al., 2014-2015; http12]. As we had two different versions of the text and not only 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 easily compare (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 used were <unclear> and <add> in case of a scribal intervention.

### III FURTHER PERSPECTIVES AND CONCLUSION

#### 3.1 Enhancement of the Vat. Ar. 13 edition

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

#### 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 in Greek/Latin/Arabic (13th century, Marciana Library Gr. Z. 11 (=379)). The features of the Vat. Ar. 13 edition will be adapted to the manuscript with three columns in these three languages. Peculiar efforts will be allocated to deliver an optimal user experience to scholars interested in comparing the three versions online. Software development processes should involve several rounds of scholar/user feedback iterations through interdisciplinary collaborations. Finally, certain parts of the manuscript will be compared with regards to the manual transcription and a handwritten text recognition using the tool *Transkribus* [http13], provided by the Digitisation and Digital Preservation group University of Innsbruck.

### IV TABLES AND FIGURES

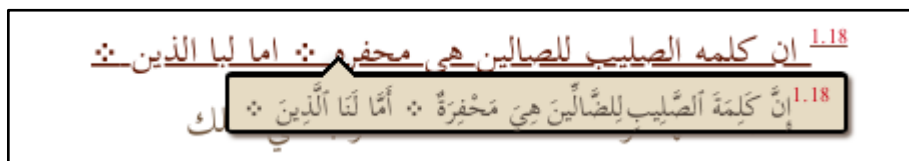


Figure 1. Screenshot of http2. Transcription and standardized text.

## 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., Homer and the New Testament as 'Multitexts' in the Digital Age, *Scholarly and Research Communication* 2012:3(3): 1-15, <http://src-online.ca/index.php/src/article/view/97>
- 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*. Bibliotheca 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://unil.ch/nt-arabe/edition-electronique/>
- http3 [http://alta3b.wordpress.com/2009/04/06/translation\\_jn-1-1/](http://alta3b.wordpress.com/2009/04/06/translation_jn-1-1/)
- http4 <http://biblia-arabica.com>
- http5 <http://vbd.humnet.unipi.it>
- http6 <http://portal.textometrie.org/demo/>
- http7 <http://data.copticscriptorium.org/>
- http8 <http://www.apd.gwi.uni-muenchen.de:8080/apd/project.jsp>
- http9 <http://www.ancientwisdoms.ac.uk/library/>
- http10 <http://dare.uni-koeln.de>
- http11 <http://sourceforge.net/projects/evt-project/>
- http12 <http://www.tei-c.org/release/doc/tei-p5-doc/fr/html/PH.html#PHZLAB>
- http13 <https://transkribus.eu>
- Kashouh H., *The Arabic Versions of the Gospels, The Manuscripts and their Families* (Arbeiten zur neutestamentlichen 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](http://cle.ens-lyon.fr/arabe/moyen-arabe-et-questions-connexes-30683.kjsp?RH=CDL_ARA120000).
- Moawad S. *Al-As'ad Abū Al-Faraḡ 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-Ingīl 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., Tahṛīf in the Digital Age. *Ancient Worlds in a Digital Culture* (Digital Biblical Studies 1). Ed. C. Clivaz, P. Dilley, D. Hamidović, in coll. with A. Thromas. Brill (Leiden), 2016, *forthcoming*.
- Vollandt R. Some Historiographical Remarks on Medieval and Early-Modern Scholarship of Biblical Version in Arabic: A Status Quo. *Intellectual History of the Islamicate World*. 2013;1: 25-42.

### ANNEX 1

Many thanks to Prof. Ioannis Xenarios, director of VITAL-IT, for his support to our work, to the Vatican Library for the digitized images of Vat. Ar. 13, and to Harley Edwards for his English proof reading.

### ANNEX 2

All hyperlinks have been checked on 02/29/16.