Studies in Manuscript Cultures

Edited by
Michael Friedrich
Harunaga Isaacson
Jörg B. Quenzer

Volume 16

Creating Standards

Interactions with Arabic Script in 12 Manuscript Cultures

Edited by
Dmitry Bondarev
Alessandro Gori
Lameen Souag

DE GRUYTER

ISBN 978-3-11-063498-3
e-ISBN (PDF) 978-3-11-063906-3
e-ISBN (EPUB) 978-3-11-063508-9
ISSN 2365-9696

Library of Congress Control Number: 2019935659

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Contents

Preface

This volume grew out of the workshop 'Creating standards: orthography, script and layout in manuscript traditions based on Arabic alphabet' held at the Centre for the Study of Manuscript Cultures, University of Hamburg, on 10–11 October 2013. The convenors of the workshop (and the two first editors of the volume), followed the inspiring initiative of Michael Friedrich to compare standards in various manuscript cultures influenced by Arabic script. Our initial – and over-ambitious – plan was to (a) identify tendencies of standardisation in orthography, script and layout, (b) examine the extent to which these three domains of manuscript production are related and (c) delineate factors behind standardisation processes. During the workshop discussions and later in the process of editorial work, it became increasingly clear that the paths of standardisation in the domains of language, orthography and manuscript production are not necessarily connected, and the standards are perceived and measured differently in each of the domains. This is directly and indirectly confirmed by the chapters of this volume, most of which have more confident conclusions about standardisation processes in orthography rather than in other domains of manuscript production.

This book deals with various aspects of standardisation by stepping outside the disciplinary and regional boundaries and providing a typological cross-cultural comparison of standardisation processes in writing traditions influenced by Arabic where different cultures, languages and scripts interact. A wide range of case studies gives insights into the factors behind uniformity and variation in Judaeo-Arabic in Hebrew script (8th–12th centuries, *Esther-Miriam Wagner*), South Palestinian Christian Arabic (8th–9th centuries, *Paolo La Spisa*), New Persian (9th–11th century, *Paola Orsatti*), Aljamiado of the Spanish Moriscos (15th–17th centuries, *Nuria de Castilla*), Ottoman Turkish in the Arabo-Persian script (14th–19th centuries, *Jan Schmidt*), a single multilingual Ottoman manuscript (late 16th century, *Branka Ivušić*), Sino-Arabic writing *xiaojing* in Northwest China (18th–20th centuries, *Florian Sobieroj*), Malay Jawi script writing in the Moluccas (17th–19th centuries, *Jan van der Putten*), Kanuri and Hausa Ajami writing (17th–20th centuries, *Dmitry Bondarev* and *Nikolay Dobronravin*), the Berber language Kabyle in Algeria (19th–20th centuries, *Lameen Souag*), and Ethiopian *fidäl* script used in transliteration of Arabic (19th–20th centuries, *Alessandro Gori*).

A comparative analysis of pathways of standardisation in the twelve manuscript cultures addressed in this volume allows for some generalisations, as follows. Contact situations do not necessarily lead to the exchange of standardised orthographic principles. In many cultures, the co-existence of Standard Arabic and non-standardised languages spoken and written in Muslim communities poses a

https://doi.org/10.1515/9783110639063-202

paradox: such languages are profoundly influenced by Arabic, but their orthographies are not modelled on the principle of standardisation. This apparent paradox is resolved by the prediction that standards in orthography – one of the domains of manuscript culture – are conceptually different from standards in other domains, such as format, layout and script. Each domain of manuscript culture develops microsystems of standardisation and different domains have different 'areas' of uniformity and standardisation in a given manuscript culture. Thus, a general tendency observable at the level of physical features of manuscript production is that layout and script types tend to be unified, irrespective of orthographic norms and, vice versa, orthographic norms develop irrespective of norms applied to physical domains of manuscript production.

The editorial process took us longer than we planned, and we are immensely grateful to the contributors for their patience and trust in our collaborative work. Our gratitude goes to all the presenters and participants of the October 2013 workshop for the inspiring exchange of ideas many of which have materialised in this volume. It was a great pleasure to work with Carl Carter, Maya Kiesselbach and Joe McIntyre who meticulously copy-edited most of the contributions. We thank you sincerely for your most helpful corrections, remarks and suggestions. Our appreciation goes to the anonymous reviewers for their constructive comments on various parts of the volume. We owe an immeasurable debt to Cosima Schwarke who has been a guiding lantern during our long journey. It is thanks to your day-to-day support in all editorial matters that this book finally sees the light. Our appreciation also goes to Astrid Kajsa Nylander who greatly assisted with the final layout of the book. We are most grateful to the editors of the series *Studies in Manuscript Cultures* for taking an interest in this volume proposal. This publication project would not have been possible without the financial support of the German Research Foundation (DFG) which funds the Sonderforschungsbereich 950 Manuscript Cultures in Asia, Africa, and Europe.

Hamburg, Copenhagen, Paris. September 2018
Dmitry Bondarev, Alessandro Gori, Lameen Souag

Transliteration of Arabic and some Arabic-based Script Graphemes used in this Volume (including Persian and Malay)

Arabic script	Transliteration	Author
ا	’	de Castilla, Schmidt, Orsatti, Wagner,
	ʾ	Ivušić, Schmidt
ء	ʾ	Bondarev & Dobronravin, de Castilla, La Spisa, Schmidt, Sobieroj, Souag,
	’	Gori, Ivušić, van der Putten, Schmidt
	ʔ	Ivušić, Schmidt
ب	b	all
پ	b_1	Ivušić
	p	Orsatti
ت	t	all
ث	ṯ	all
ج	ǧ	Gori, La Spisa, Souag
	ǰ	Orsatti
	j	Bondarev & Dobronravin, Ivušić, van der Putten, Schmidt, Wagner
چ	$ǧ_1$	Ivušić
	č	Orsatti
	c	van der Putten
ح	ḥ	all
خ	ḫ	Gori, Ivušić, La Spisa, Wagner
	x	Bondarev & Dobronravin, Orsatti, Souag
	kh	Sobieroj
د	d	all
ذ	ḏ	Bondarev & Dobronravin, Gori, Ivušić, Orsatti, La Spisa, Souag, Wagner
	dh	Sobieroj
ر	r	all
ز	z	all
ژ	z_2	Ivušić
	ž	Orsatti
س	s	all
ش	š	Bondarev & Dobronravin, Ivušić, La Spisa, Orsatti, Souag
	sh	Sobieroj, Wagner
ص	ṣ	all
ض	ḍ	all
ط	ṭ	all

ظ	ẓ	Bondarev & Dobronravin, Orsatti, Wagner,
	ḏ	Ivušić
	đ	Souag
ع	ʿ	Bondarev & Dobronravin, Ivušić, La Spisa, Schmidt, Sobieroj, Wagner
	ʕ	Souag
	‘	de Castilla, Gori, Orsatti, Schmidt
غ	ġ	Ivušić, La Spisa, Orsatti, Wagner
	gh	Bondarev & Dobronravin, Schmidt, Sobieroj
	ɣ	Souag
ڠ	ng	van der Putten
ف	f	all
ڤ	b̲	Orsatti
	p	van der Putten
ق	q	all
ك	k	all
گ	g	Orsatti,
ڭ	g	Orsatti
	$k_{ʒ}$	Ivušić
ڬ	g	van der Putten
ل	l	all
م	m	all
ن	n	all
ڽ	ny	van der Putten
ه	h	all
و	w	all
ي	y	all

Dmitry Bondarev

Introduction: Orthographic Polyphony in Arabic Script

> [...] standardization emerges as a complex process whose many facets (linguistic, social, cultural, educational, political) we still do not fully understand, and which warrant further research from comparative, case-study and interdisciplinary perspectives. (Deumert and Vandenbussche 2003a, 11)

> Printing gave rise to a distinct literate culture, and the earlier scribal culture had many of the same limitations often attributed to oral culture: individual copyists produced texts with idiosyncratic formats, conventions and mistakes, whereas printing allowed a large number of identical texts. (Barton 1994, 124)

Manuscripts originate from literacy practices embedded in numerous social domains, such as education, administration, religion and trade. Expressed in spoken and written languages, various social activities prompt the development of organising principles and structures, which in turn serve as models for the agents and participants of literacy practices. The degree to which such organising models may develop varies from lax to strict. The strict models of literacy practices are usually regulated by sets of standards. When a certain literacy event takes place – writing a letter, copying a poem or commenting on a canonical text, for example – then the regulatory normative patterns (or their absence) may variously be reflected in the resultant manuscript. The size and form of the manuscript is one such indexical feature, the layout another, and the type and style of script and spelling conventions are yet another feature indicative of the degree of standardisation imposed on the scribe. If such features are examined in relation to each other rather than separately, and if the patterns of their relationship in one manuscript culture are compared to the patterns in other cultures, we may learn a great deal about the underlying forces of literacy and specifically about language in its relation to the manuscript medium.

A holistic comparison of the sort in line with the stance of the first quote in this chapter would shift manuscript studies to a previously unexplored vantage point. We would not only understand which components of a manuscript culture were historically more impervious to stabilisation and standardisation, and which disfavoured variation, but we would also come close to big WHY questions.

I am most grateful to Michael Friedrich and Lameen Souag for their constructive comments on earlier drafts.

https://doi.org/10.1515/9783110639063-001

Why do some cultural models lead to totalitarian types of standardisation of writing, such as Western societies? Why do others keep on with limited standardisation – strict in one social domain and lax in another? Such would be the cultures with fixed orthography and regulated reading of codified texts (e.g. the Qur'an) and non-codified writing practices in the languages or language varieties other than those of the codified texts. And why do the other cultural models exist without any standardisation? Or do such cultures exist at all?

However promising a holistic comparison might sound, we are still far from that illuminating vantage point. It might in principle be feasible to carry out a study on the multiplicity of factors behind standardisation (or failure thereof) in one manuscript culture, but a comparative study of several cultures seems an enormous task. This is because in order to make the comparison typologically valid, we need to identify social domains related to manuscript production for each culture (by no means static), then study norms, prescriptions and codes in each identified domain. Only then will we arrive at substantiated observations about the standardisation factors in the history of a manuscript culture. However, to the best of my knowledge, no studies have tried to treat all the possible factors of standardisation in one manuscript culture holistically yet, let alone comparative cross-cultural studies. This is not surprising, actually. In Western societies, the (positive) notion of a standard developed at the time of print, long after manuscripts ceased to be the prime medium of literacy practices. The manuscript age was seen as a pre-standard stage in the history of the development of written languages, this history culminating in standardised print culture.[1] So the study of standardisation was in the areas of human activity where it was expected, which excluded manuscripts.

The long history of successful attempts to eliminate variation in spoken and, especially, written European languages and to promote the primacy of a standard led to scholarly frameworks with dismissive attitudes to variation in written texts (manuscripts). Up until sometime in the middle of the 20th century, linguists considered textual variation in manuscripts as an uncomfortable situation resulting from 'idiosyncratic formats, conventions and mistakes', as expressed in the second quote in this chapter. Recognising variation as an important factor in understanding language in its spoken and written form and in literacy practices in general was a novelty in some disciplines in the 1970s and 80s, only recently gaining momentum in linguistics, sociolinguistics, literacy studies and

1 See the discussion *inter alia* in Eisenstein 1979, 1983, Williams 1981, Stubbs 1980, Bullough 1991, Barton 1994, Linn and McLelland 2002, Agha 2007, Sebba 2007, Stenroos 2018 and Van der Horst 2018.

manuscript studies.[2] It may seem a truism that understanding variation is essential for understanding what kills it, namely standardisation. But despite the development of variation-oriented studies (and thus concerned with standardisation in one or another way) in linguistics, sociolinguistics, philology and generally in manuscript studies, little has been done so far to approach variation/standardisation phenomena holistically, involving interdisciplinary dialogue.[3]

1 Standardisation: why sociolinguistics?

A manuscript is a meeting place of different cultural practices and domains. Some of these practices can be recognised visually in the manuscript's size, form and the material it is made of (all of these roughly corresponding to the crafts of bookmaking) or in the layout, script type and style, orthography and language (the scribal domain). Understandably, there has been a division of labour between specific disciplines dealing with these different sociocultural domains. The material, size, form and layout of manuscripts are common fields of investigation for codicologists, script type and style are in the scope of palaeography, and orthography and language are treated by philology and (socio)linguistics. It is instructive to learn that of all the disciplines, the only one that has developed a systematic approach to the study of standards (and the dichotomy between standard and variation) is sociolinguistics. That is not to say that codicology, palaeography and philology are not concerned with standardisation tendencies,

2 In linguistics, the quest for comparative cross-cultural research into standardisation started with Jespersen (1925, 46) and was coined 'comparative standardology' by Joseph (1987, 13) and resulted in a comprehensive comparative work on Germanic languages (Deumert and Vandenbussche 2003a) (see especially Deumert and Vandenbussche 2003b, 1), which is discussed in the following sections. For manuscript studies, see Sobieroj 2016 and his overview of a recent trend in Arabic studies to 'place variance itself in the focus of research' (Sobieroj 2016, 2).

3 One significant exception is a collection of articles edited by Jennifer Cromwell and Eitan Grossman (2018), *Scribal Repertoires in Egypt from the New Kingdom to the Early Islamic Period*. This work deserves special attention. With their focus on the exact opposite of the subject of this volume, there is a shared goal to study the connected phenomena – 'their' variation and 'our' standardisation – in the complex linguistic and extra-linguistic dimensions. Although the cultural and geographic scope of Cromwell and Grossman 2018 is Egypt, their book is in essence a cross-cultural and typologically oriented comparative study since it covers cultures in Egypt which co-existed or replaced one another in the course of four millennia, while the typological frame is given through the lens of historical sociolinguistics informed by European philology (especially studies of pre-modern English).

though. Standard(ised) practices are by all means mentioned or studied in literature from these disciplines (*inter alia* Beit-Arié 1992, 2017, George 2007, Déroche 2006, Gacek 2009, Pollock et al. 2015, etc.). In these fields, the word 'standard' is used in many different senses within its semantic domain, ranging from a source of authority to a level of achievement.[4] In sociolinguistics, however (and more generally, in linguistics), it has a narrower scope of 'language codification leading to elimination of variation' (more on this definition below). The causes and consequences of language codification have been discussed in various branches of (socio)linguistics across major topics such as social and linguistic identity (Milroy and Milroy 1992, Agha 2007), language varieties and dialects (Trudgill 1979, Biber and Finegan 1994, Ferguson 1994), language variation and change (Romain 1982, Milroy and Milroy 1985, Chambers and Schilling 2013), the distinction between speech and writing (Biber 1995, Biber and Conrad 2009, Lillis 2013, Lillis and McKinney 2013), the development of writing systems and language planning (Fishman 1974, Grenoble and Whaley 2006, Sebba 2007, 2009), and the sociolinguistics of reading and writing (Stubbs 1980, Street 1993, Blommaert 2005). It is a matter of course that without writing there would not have been any manuscript cultures, so the concept of writing seems to be the most natural node connecting manuscript studies and the discussion of standardisation in sociolinguistics. These research fields are not overtly connected, however. The following aims to reveal some interdisciplinary bridges.

Sociolinguistics emerged as a subfield of linguistics in the late 1960s/1970s, at a time when written language was only marginally considered worthy of linguistic investigation (Barton 1994, 2007, Lillis and McKinney 2013, Stenroos 2018). Naturally, there was not much discussion about writing, let alone about writing in the 'manuscript age'.[5] This changed a decade later with the ever-growing anthropological and sociolinguistic enquiry into literacy and the relationship between speech and writing (Scribner and Cole 1981, Stubbs 1980, Goody 1987) and with the formation of historical sociolinguistics, which focuses on extra-linguistic factors as a way of explaining language change (Weinreich et al. 1968, Romain 1982, Mattheier 1988).

4 An illuminating short overview of the historical and semantic scope of the term in English has been provided by Raymond Williams, one of the founding figures in Cultural Studies, in his vocabulary of culture and society (1983, 296–99). Also see Williams 1981 (esp. 87–118) on standards and standardisation in print cultures.

5 Mesthrie et al. (2013, 27) acknowledge that 'the study of writing as a social practice is a relatively new interest in sociolinguistics'.

Since then, the sociolinguistics of writing and literacy studies has evolved into New Literacy Studies (NLS), which pays great attention to the social and material context and modes of writing – and thus has the potential to extend its interest to manuscripts as well (Barton 1994, 2007, Blommaert 2005, 2008, Lillis 2013, Juffermans et al. 2014, Weth and Juffermans 2018). NLS's dynamic approach to the codification of written forms of language explores non-unidirectional dimensions in the development of literacy practices which counteract the 'tyranny of writing'.

In the meantime, historical sociolinguistics has grown into a diverse field, bringing together linguists, philologists and historians who work with manuscripts. The increased interest in manuscripts, not only as a mine of data, but as a subject of study in its own right, was prompted by the drive to make the 'best use of bad data' (Labov 1994, 11, referring to written artefacts with their scarcity of background information and their texts skewed to the registers of the educated). The result was a number of publications that were helpful across disciplines (Hernández-Campoy et al. 2012, Langer et al. 2012, Wagner et al. 2013b; Cromwell and Grossmann 2018).

Thanks to the interdisciplinary mergers, the field of sociolinguistics seen as a whole (with all its interrelated subdisciplines) seems to offer a set of terms and approaches relevant to the question of standardisation in manuscript cultures.

2 'Comparative standardology'

The written and spoken counterparts of language as topics of study have gone hand in hand in the history of European linguistics, with one hand pulling harder than the other at different points on this journey (see Barton 1994, 2007 for an overview). The major sociolinguistic concepts about standardisation grew from the study of spoken languages, which was the initial focus of the discipline. However, the written counterpart came onto the scene at a very early stage.

As said before, standardisation in sociolinguistic terms is generally understood as language codification leading to elimination of variation. This definition is a hybrid one, uniting both wider and narrower senses. In a wider sense, '[s]tandardisation refers to the process by which a language has been codified in some way' (Wardhough 2010, 31). In a narrow sense, 'the process of language standardisation involves *the suppression of optional variability in language*' (Milroy and Milroy 1999, 6, emphasis in the original).[6] The 'process' is key here. Many authors

6 A more functional and explicitly socially oriented definition has been provided by Garvin and

try to overcome 'the somewhat teleological orientation of traditional standardisation models' (Deumert and Vandenbussche 2003c, 457). The process may take unexpected turns and lead to de-standardisation, to cycles and to intricate relations between standards, sub-standards and non-standards against the backdrop of sociocultural domains – a complex which prompts the notion of 'standard language cultures' (Milroy 1999). The question of language standardisation received particularly comprehensive treatment in Germanic (socio)linguistics. A systematic comparative approach to the study of linguistic and extra-linguistic factors has been elaborated in Deumert and Vandenbussche (2003a). Several instructive points stemming from this study seem promising for an integrated analysis of related phenomena in manuscript cultures. Deumert and Vandenbussche (2003a,b,c), develop Haugen's (1966a,b) four-way model into a comprehensive framework for what they call 'comparative standardology' (following Joseph 1987, 13, cited in Deumert and Vandenbussche 2003b, 1).

Haugen's grid of standard language development consists of (1) norm selection, (2) norm codification, (3) norm implementation and (4) norm elaboration (Haugen 1966a,b summarised in Deumert and Vandenbussche 2003b, 4). What is especially interesting for the study of manuscripts is that a written variety of the language is typically considered a key agent of standardisation at all four stages: 'it is [a] significant and probably crucial requirement for a standard language to be written' (Haugen 1972, 246, cited in Deumert and Vandenbussche 2003b, 3). However, it is not necessarily the case that a written standard code initially selected as a model will be carried over to the codification phase associated with the creation of grammars and dictionaries which fix the norms in prescriptive mode. Initial written standards might be lost, as was the case with Old Frisian (Hoekstra 2003) and Low Middle German (Langer 2003). Linguistic competition between different available norms may lead to the suppression of one norm and elevation of another, resulting in the co-existence of standard (written) languages or language varieties, each covering different social domains and having a suppressive or enriching influence on each other.[7] Complex situations of contact between standard and non-standard

Mathiot (1960, 783, cited in Mesthrie et al. 2013, 20): 'codified form of a language, accepted by, and serving as a model to, a larger speech community'. Romain (2000, 14) defines a standard language as 'a variety that has been deliberately codified so that it varies minimally in linguistic form but is maximally elaborated in function'.

7 See Mattheier 2003 on the co-existence of Latin (in the clerical and literary domain), the written Alemannic dialect (between the 9th and the 13th centuries in a narrow domain of court poetry and epic) and other written vernaculars developing from the 11th to the 15th century and resulting in the formation and co-existence of four main uniform written linguistic norms, namely East Upper German (Bavarian-Austrian), West Upper German (Alemannic), East Middle

varieties may result in de-standardisation[8] and the emergence of new regional or local norms through the convergence of standard and non-standard norms or the convergence of non-standard varieties or through divergence, for example that of the Scottish regional norm from the Northern English dialect (Dossena 2003). In the process of divergence or convergence, various diglossic situations may arise, such as 'standard/dialect diglossia' (Deumert and Vandenbussche 2003c, 7), or 'medial diglossia' (writing in one language and speaking in the other [Mattheier 2003, 212]; Lüpke 2011 calls the same phenomenon 'exographia').

In summary, the 'comparative standardology' framework provides a useful interdisciplinary set of concepts for the study of standardisation, as follows: norm selection, codification, implementation and elaboration; co-existence of local norms; competition; loss (and thus vestiges); (dis)continuity; de-standardisation; centripetal and centrifugal cycles; interaction and contact; and divergence and convergence.

The important notion of 'standard language culture' as well as the distinction between a standard language *sensu strictu* and the process of standardisation developed within this framework invites connections with the study of manuscript cultures. In recent comparative studies on linguistic variation and change in manuscript traditions, the notion of a standard was seen critically as being too teleological and unidirectional and hardly applicable to the multilingual environment of earlier manuscript cultures with their variation of registers and linguistic codes. Thus, in her study of late Middle English scribal practices, Merja Stenroos observes that 'terms such as "standard" and "standardisation" may not be very useful when applied to fifteenth-century materials' and that there are cases which 'do not fit into a unidirectional view of the standardisation process' (2013, 160). The 'comparative standardology' approach helps in this respect as it offers epistemological scope to include all the cases in standardisation studies that are not covered by the models of standard languages *sensu strictu*.

3 Written language and orthography

Even though spoken standards typically develop hand in hand with their written counterparts, standardisation of writing differs considerably from standardisa-

German (Saxonian) and West Middle German (Franconian).

8 See Greenberg 1986 and Ferguson 1988 on 'standardisation cycles' understood as 'a succession of periods of focus with standardization and periods of diffusion with dialect differentiation' (Ferguson 1988, 121).

tion of speech. The 'writing system [...] is relatively easily standardised', whereas 'absolute standardisation of a spoken language is never achieved' (Milroy and Milroy 1999, 19). This is not surprising given the difference between the linguistics structures meant to be covered systematically by the writing system and the structures covered by the spoken language. The scope of writing systems is limited to a countable number of items: smaller numbers in phonographic systems (from phonemes to syllables) and much larger, but still finite ones in phono-logographic systems (from phonemes to words). The scope of language is a nearly infinite number of grammatical structures and variant forms. This difference is empirically observable in various alignment scenarios whereby standardisation of linguistic structures may develop without standardisation of orthography or at a different pace to it, be they interrelated (in some societies) or unrelated. Thus, English orthography has changed very little since the codification activity of the 18th-century prescriptivists, but codification of the spoken language has been less successful (Milroy and Milroy 1999, 28; Agha 2007 [chapter 4], Sebba 2007). In the case of Persian, the orthography was standardised together with the emergence of New Persian and its standardisation into Classical Persian (Perry 2012, *Orsatti*).[9] And in the case of Ottoman Turkish, 'there was no standard form for written Turkish and no standardised spelling until the 20th century' (Darling 2012, 174; *Schmidt*). Genre-specific standards in linguistic structures developed in epistolary writing in Judaeo-Arabic, featuring strong spelling variation (Wagner 2010, 2013, *Wagner*). Standard Spoken Tamil is reported not to have a standard orthography counterpart (Schieffman 1998).

A writing system in its visual graphic representation is the interface between linguistic structures and manuscripts.[10] Language is converted into manuscripts through a graphemic code, and it is through this code that linguistic structures are retrieved from manuscripts. This trivial remark is meant to remind us that orthography as a set of spelling conventions (be it strict or lax) is inseparable from written artefacts. So, the study of orthography should be intrinsic to research on

9 The authors of this volume are indicated in italics.

10 I avoid the simple dichotomy of speech vs manuscript (writing) because many linguistic structures are predominantly realised in writing and many speech discourses are not meant for writing. Potentially, any linguistic structure can be written down, but not all structures are feasible in spoken language. The study of the relationship between speech and writing has a long and rich history spanning more than half a century. For more recent treatment and an overview of the topic, see Barton 2007, Biber and Conrad 2009, Lillis 2013. In historical sociolinguistics, a productive approach is to treat written data in historical documents as 'text language' (Fleischman 2000) or 'manuscript language' (Stenroos 2018).

manuscript cultures, and insights from sociolinguistics are equally helpful in this respect.[11]

The notion of orthography has two terminological poles. The first defines orthography in the narrow sense as 'the standardized variety of a given, language-specific writing system' (Coulmas 2003, 35), the definition very closely linked to the word's etymology ('correct writing', German *Rechtschreibung* or Russian *pravopisanie* [правописание], etc.). The second definition has a wider scope: 'the set of conventions for writing words of the language', which leads to 'the notion of orthography as social practice' (Sebba 2007, 10–11, 13).[12] As Mark Sebba puts it, 'Orthography is *par excellence* a matter of language and culture' (2007, 7).[13]

Following the lead of literacy studies (Scribner and Cole 1981, Street 1984, Barton 1994, Gee 1990), Sebba's view of orthography is reminiscent of the 'comparative standardology approach' in that it sees orthography as a dynamic concept situated in social and cultural practices rather than as a fixed entity (Sebba 2007, 13). Such a sociocultural approach allows us to recognise orthography as part of changing literacy practices, in contrast to the 'autonomous models', which treat orthography as 'neutral technology that can be detached from specific social contexts', as defined by Street (1984, 1) regarding the notion of literacy and applied to orthography by Sebba (2007, 14). Discussing English orthography, Sebba makes a very important methodological statement:

11 In various philological fields, orthographic variation is not usually studied as a process or practice. Rather, it is seen as a means of reconstructing 'original' texts or pronunciation/sound systems (den Heijer et al. 2014) or as a means of studying language change (Wagner et al. 2013b).

12 The restrictive definition of orthography might be convenient to contrast institutionalised regulatory mechanisms with opposing tendencies of norm deviation and de-standardisation (in terms of struggling against the tyranny of writing; see Weth and Juffermans 2018 or Blommaert 2008, 7, who sets off orthography as normative, set against 'hetero-graphy'); or it might be helpful to appraise the effects of Western models of language standardisation when designing orthography for previously unwritten and/or minority languages (see Lane et al. 2017). From a historical perspective, the narrow sense of the term might not be useful, given that the rise of orthographies as 'absolute' standards is a recent phenomenon. The fact that the notion of orthography does not yield easily to the restrictive meaning can be seen ironically in Rutkowska and Rössler (2012, 214), who first define the term in the narrow sense ('a spelling norm which consists of all the standardized and codified graphic representations of a language'), but then use it along the wide continuum from an unstandardised orthography characterised by variance to a standardised orthography without any variance.

13 Lillis (2013, 24) widens the definition even further: 'Orthography: 1. A writing system specifically intended for a particular language 2. A particular way of performing/producing a writing system of a particular language (for example, types of handwriting, fonts, spelling conventions used to represent verbal language)'.

> This possibility of variation and deviation (licensed or unlicensed) from the conventional norms makes it reasonable to think of orthography as a social practice – a widespread and recurrent activity which involves members of a community in making meaningful choices, albeit from a constrained set of possibilities. (2017, 31)

'Meaningful *choices*, albeit from a *constrained set* of possibilities' (or 'repertoires') were indeed recurrently made by the scribes of the manuscript cultures discussed in this volume.

Orthography, as a set of conventions, may have patches of standard spelling within a system of internally organised sets conditioned by a multiplicity of linguistic and extra-linguistic factors. The *variable application of conventions* leading to combinations of *orthographic tendencies* which sometimes developed in a non-unidirectional way is demonstrated by many of the contributors to this volume.

4 Written language: terminology

Before expounding one crucial difference between orthographic standards and non-orthographic standards in manuscripts, it is worthwhile outlining the terms associated with orthography as they will frequently be evoked in this book. Coulmas (2003, 35–6) provides a useful set of terminology, in part summarised below and supplemented with other definitions – heuristically useful, if sometimes conflicting.

- *Writing system* refers both to 'the writing system of an individual language and to an abstract type of writing system' (Coulmas); or 'it is a means of representing graphically a language or group of languages' (Lillis 2013, 24).
- *Script* stands for 'the graphic form of the units of a writing system' (Coulmas). Sebba 2007, 11 and Lillis 2013, 24 consider *script* a synonym of 'writing system'.
- *Orthography* (as mentioned earlier) is 'the standardized variety of a given, language-specific writing system' (Coulmas) or it is a set of conventions for writing words of the language (Sebba 2007, 10).
- *Spelling* is 'the application of those [orthographic] conventions to write actual words'. Thus, 'I am spelling the words of this sentence according to the orthography of English using the Roman writing system (or script)' (Sebba 2007, 11), whereas in Coulmas' opinion the term is 'used interchangeably with orthography'.
- *Alphabet* has several meanings, but it should be restricted to systems 'where signs individually denote consonant and vowel phonemes' (Daniels 1997, 370). Sebba 2007 and Lillis 2013 use the attributive form 'alphabetic', referring to a system based on consonants and vowels as individual units. A fine-

tuned definition of related terms (alphabet, abjad, abugida) is proposed in Daniels (1990).

- *Letter*, in the most general sense, refers to 'the basic functional units of all writing systems'. In a narrow sense 'it refers to the basic symbols of Semitic-derived writing systems, including the Latin alphabet' (Coulmas).
- *Grapheme* refers to 'the abstract type of a letter and its position in a given writing system' (Coulmas). Some linguists do not draw a sharp line between letter and grapheme, thus, 'the smallest independent unit of the writing system — for example, a letter of the alphabet or a character in Chinese' (Sebba 2007, 169). In this volume, irrespective of individual approaches to what constitutes a letter and grapheme, angled brackets are used for transliteration of non-Roman-based letters/graphemes. For example, the Arabic letter ب is represented as <b>.
- *Graph* is 'a single visual sign or mark' or 'any written character or mark' (Boltz 1994, 19, 180).
- *Phoneme*, on which the term 'grapheme' is modelled, is an abstract notion denoting the smallest distinct unit of sound. Phonemes are represented in slanted brackets, as in /b/.
- *Phone*, or sound, is the acoustic realisation of a phoneme. Phonetic transcription is given in square brackets, so an aspirated pronunciation of the phoneme /b/ would be written as [b^h], for example.

Some confusion may arise because of palaeographic usage of certain related terms, such as 'graphic'/'graphical' and 'script'. Unlike 'graphemic', which refers to the abstract level of 'grapheme', '*graphic*' and '*graphical*' refer to the shape or visual depiction of a sign. Different realisations of graphic shapes result in different *script types* and *script styles*. From this viewpoint, palaeography is interested in visual patterns of a given script rather than in abstract structures of writing systems. The abstract components of writings systems are conceptualised as *types* in linguistics and philosophy, whereas concrete instances of types, or their spatio-temporal particulars (e.g. ink composition), are *tokens* (Wetzel 2006). However, the word 'script' in the palaeographic sense of the term also has abstract and concrete components, which are especially discernible in manuscript cultures with standardised script types. This leads me to the next point – the difference between orthographic and non orthographic standards in manuscripts.

5 Orthography and manuscripts: on the assessment of standards

Compare these two statements:

1. 'Paper is produced according to standard sizes [...] For instance, a quarto from Royal paper (4°R) is 30 × 22, and a folio from Common paper (2°C) is 31 × 22 (or *somewhat smaller*, owing to trimming in binding).' (Gumbert 2010, emphasis added);
2. The word 'standardisation' is spelled with an <s> in standard British English and with a <z> in standard American English.

In (1) there is a certain range within the standard sizes ('*or somewhat smaller*'). In case (2), it is an either/or principle.

Standards in manuscript form, layout and script are measured and perceived differently compared to standards in orthography. As regards layout, variation in ruling patterns and varying sizes between the edge of the paper and the edge of the text area would not contradict a general standard of using intended principles (Andrist et al. 2013, 94), even if the proportions and dimensions were regulated by geometry (Déroche et al. 2006, 169–71; George 2007). Rules governing standards of script style, that is, the shapes and proportion of graphic units, might also be strict and yet there is a certain amount of scope within which the inevitable variation in production (tokens) is permissible.

In orthography, what is regulated by a standard is the abstract graphemic representation (type). The abstract match between phonemes and graphemes has to be absolute, whereas script style does not have to be. For example, if the phoneme /b/ is prescribed to be written as <b>, then it should not be written as <d> even if the shape of the letter may seem similar; but if the style of script is prescribed to be executed in a certain shape, e.g. at an angle, as in *italics*, some deviations from that particular shape (the angle in the case of italics) will still be counted as the prescribed style. So, in writing, matches of abstract types (between phonemes and graphemes) are regulated by orthography, and matches of types and tokens (graphemes and their shapes, or a type of layout and its realisation, or a type of manuscript form and its realisation) by other domains.

This is not to say that script style does not have the abstract *type* as a conceptual counterpart of the shape of a letter – it certainly does. But what is important is that in prescribed realisation of graphemes as representors of phonemes, the *absolute* match is the requirement, whereas in realisation of shapes, what is required is *approximation* to the abstract.

Another way of illustrating the type/token distinction in relation to orthography and script is as follows. The script's coverage of phonemes is introduced through tokens: <b>, <*b*>, <B> and <b> all stand for /b/; the letter <b> can be conceptualised as a 'vertical ascender bar with a single bottom loop/two loops facing right'. Changing a single compositional component is enough to write the letter incorrectly. Thus, if the component 'facing right' is altered, this results in <d> or <d> instead of <b>, and the violation of 'vertical ascender' gives <p> or <*p*>.

In complex orthographic systems (like English), morphological, lexical or positional parameters define the correct representation of a sound. Thus, *fish* cannot be represented as *ghoti*, as suggested by Bernard Shaw, because <gh> is only /f/ when it is used at the end of a word (e.g. in 'enou**gh**'), <o> is just incidentally/i/ in the word 'w**o**men', and <ti> is only /ʃ/ in a single orthographic unit: <-tion> (e.g. 'nation').[14] The combination of rules may be complex, but a single misrepresentation at the level of types will be enough to invalidate the correct spelling. Thus, *enough* is not wrong because it is written in italics or as enougb or ***enough***, but because the graphic unit <gh> is faulted by a single abstract segment, <b>.[15]

The set of rules governing non-orthographic features in manuscripts is not just different, but much more complex. Many are familiar with the notorious problem of *describing* a script type/calligraphic style verbally without *showing* the examples, even if the given script type has a precise set of abstract features (as in the geometrical tradition of certain Arabic calligraphic hands). Even though the scribes who used such scripts employed geometry as an abstract model based on mathematics, a single deviation from one of the many parameters does not invalidate a script style, layout or paper form.[16] In contrast, a single deviation from one of the parameters for a given grapheme will invalidate a letter.

I took this digression about type/token relations to demonstrate that the dimension of orthography cannot be directly compared with the other dimensions of manuscript production. This difference predicts that standardisation in the domain of orthography should not necessarily lead to standardisation in the other domains, such as form, layout and script (and vice versa). A similar unrelatedness has been mentioned earlier regarding the disparity between standardisation of language and that of orthography.

14 See Stubbs 1980, 51 for his discussion of the spelling <ghoti> proposed for *fish* by Bernard Shaw.
15 The discussion of the ontological relationship between type and token is much more complex, questioning the existence of types, differentiating tokens from occurrences, etc. A more subtle model of orthographic validity could be described in terms of matches between types and occurrences (which are non-material instances of types) rather than between types. But since both types and occurrences are abstract entities, the general idea still seems relevant.
16 This argument might not stand a chance in (post-)print societies with totalitarian standardisation.

6 Arabic script for non-Arabic languages

Manuscript cultures based on Arabic script are especially interesting from the viewpoint of the theoretical considerations outlined above. Early codification and standardisation in many domains of manuscript production was *the* characteristic feature of Islam and Arabic as the language of the Qur'an. At the time when the Muslim Arabs started interacting widely with other cultures, introducing them to Islam or encompassing them within the realms of various Islamic polities, standard orthography and regulated ways of producing manuscripts were already part of the Islamic cultural package.

The standard orthography of Arabic script was codified during the first one-and-a-half centuries of Islam as the outcome of the standardisation of the Qur'an text. By the late 2nd/8th century, 'the system of Arabic orthography was almost completed and [...] it has remained essentially the same ever since' (Versteegh 2001, 57). The process of standardisation of Classical Arabic was slower, but nevertheless it was completed by the 4th/10th century. The corpus of the language is believed to have been closed from that time onwards (Versteegh 2001, 64).[17]

Other domains of standardisation brought with Islam and visible in manuscripts are form, layout, script type and genre. However, unlike Arabic orthography, these changed with time and with cultures. Many remarkably unified types of format, layout and script tied to particular genres existed from the earliest centuries of Islam (George 2007) to the latest transitions from manuscript to print (Dobronravin 2017). Many of these types stemmed from the configuration of complex literacy practices born out of interaction between norms and standards of the contacting Islamic Arabic culture and non-Islamic non-Arabic cultures. The contributions in this volume – roughly organised in chronological order – deal with such contact phenomena, looking at various domains of the standardisation process. Eight chapters (2, 5–11) focus on writing traditions which adapted the Arabic script for non-Arabic languages, two chapters (3, 12) are respectively concerned with Hebrew and *fidäl* scripts used in the contexts of close contact with written and spoken Arabic, and one chapter (4) investigates possible influences of the Qur'an manuscript standards on Christian Arabic manuscripts (*La Spisa*). In some less well-studied cultures presented in this volume, the writing system based on Arabic script and orthography (in the general sense outlined earlier)

17 It should be noted that these dates are only helpful as general guidelines for periodisation of the norms of Classical Arabic because, as den Heijer (2012, 10) puts it, 'an overall history of Arabic orthography, which only partly overlaps with palaeography (a much better documented and studied issue!) is yet to be written'.

is the first entry point to the respective manuscript cultures (*van der Putten* 9, *Bondarev and Dobronravin* 10, *Souag* 11, *Gori* 12). In some better-studied cases, orthography and interaction between different writing systems and scripts are investigated with the new findings at hand (*Orsatti* 2, *Wagner* 3) and in the other traditions, some better studied, some little studied, standardisation of orthography is compared with that of script types, language, genre, layout and format (*de Castilla* 5, *Schmidt* 6, *Sobieroj* 8). The orthographic features of a single manuscript written in no less than seven languages, all in Arabic script, are discussed in chapter 7 (*Ivušić*).

The study of the Arabic script as the medium for writing non-Arabic languages and research into the interaction of Arabic script with non-Arabic languages and scripts are by no means a novelty, Mohammed Naim's (1971) survey being one of the earliest. However, previous research has touched upon these issues from the perspective of established disciplines and regional studies. Thus, Spooner and Hanaway (2012b) is a collection of papers on a wider topic of literacy in the Iranian cultural areas, with some articles addressing standardisation in various domains of language use and manuscript production, mostly in relation to Persian, but also dealing with Ottoman Turkish (Darling 2012) and giving comparative insights into the normativity of Arabic, Persian and Latin (Morton 2012). Some instances of Arabic script used for writing non-Arabic languages are discussed in *Script Beyond Borders: A Survey of Allographic Traditions in the Euro-Mediterranean World* (den Heijer et al. 2014). This collection of articles focuses on what the authors call 'allography': the phenomenon of writing a language in the script of another language. The comparative scope of *Script Beyond Borders* is vast, albeit restricted to the cultural areas prominently featured in the philological and historical disciplines. A collection of studies in Zack and Schippers (2012) looks into variation and development of standards in the context of interactions between religions, scripts and linguistic varieties of Arabic known as Middle Arabic and Mixed Arabic – the topic within the traditional scope of Middle East Studies.

Adaptations of Arabic script for writing non-Arabic languages have also been treated in specialised research, such as Daniels 1997, 2014, who provides a theoretical background from the perspective of linguistics, Kaye 1996, 2006, who maps the Arabic script in various world languages,[18] Dobronravin 1999, which is the

18 Kaye (2006) mentions Berber, the Dravidian language Moplah, a dialect of Malayalam (related to Tamil), the Indo-Arian languages Urdu, Sindhi and Kashmiri; the Iranian languages Balochi, Pashto, Persian and Kurdish; the Austronesian languages Malagasy, Malay and Sulu; and Turkic and Caucasian languages. He also mentions special aspects of eleven languages: Persian,

first comprehensive study of the application of Arabic script to African and some other languages, and Mumin and Versteegh 2013, a collection of papers covering an impressive range of African languages written in Arabic script.[19]

What makes this volume different from previous literature is its attempt to study Arabic-based writing systems from the perspective of 'comparative standardology', stepping outside the traditional disciplinary and regional boundaries and treating such systems in the context of manuscript production and (reconstructed) social practices. A typological cross-cultural perspective is provided by a wide range of case studies – albeit limited – presenting twelve distinct writing traditions set up in contact situations, whereby different languages, cultures and scripts interact. These are as follows: Judaeo-Arabic in Hebrew script (8th–12th century, *Wagner*), South Palestinian Christian Arabic (8th–9th century, *La Spisa*), New Persian (9th–11th century, *Orsatti*), Aljamiado used by the Spanish Moriscos (15th–17th century, *de Castilla*), Ottoman Turkish in the Arabo-Persian script (14th–19th century, *Schmidt*), a single multilingual Ottoman manuscript (late 16th century, *Ivušić*), Sino-Arabic writing in Northwest China (18th–20th century, *Sobieroj*), Malay Jawi script writing in the Moluccas (17th–19th century, *van der Putten*), Kanuri and Hausa Ajami writing (17th–20th century, *Bondarev and Dobronravin*), the Berber language Kabyle in Algeria (19th–20th century, *Souag*), and Ethiopian *fidäl* script (19th–20th century, *Gori*).[20]

7 Factors of standardisation

As mentioned, the paths of standardisation in the domains of language, orthography and manuscript production are not necessarily connected, and the standards are perceived and measured differently in each of the domains. With our limited knowledge of social practices and manuscript production in earlier cultures,

Kurdish, Pashto, Kashmiri, Urdu, Sindhi, Ottoman Turkish, Uyghur, Malay (Jawi), Hausa and Swahili.

19 Arabic script applied to non-Arabic languages is referred to as Ajami, from Arabic *ʿajamī* 'non-Arab' (and also 'Persian'), derived from a collective noun, *ʿajam*, meaning 'barbarians, non-Arabs; Persians'.

20 Due to terminological inconsistency across different disciplines, some of the terms denoting written cultures require preliminary clarification. Thus, Judaeo-Arabic means Arabic texts written in Hebrew script, Arabo-Persian is Persian in Arabic script, Christian Arabic stands for Christian texts written in the Arabic language and Arabic script, and Sino-Arabic indicates texts written in Arabic script but influenced by the Chinese writing practices. Further details are given in the respective chapters.

there is no simple way to demonstrate such relationships. This is directly and indirectly confirmed by the chapters in this volume, most of which come to more confident conclusions about standardisation processes in orthography rather than in other domains of manuscript production.

Nonetheless, one possible way of seeing a larger interconnected picture would be to identify the factors behind the standardisation of writing. Various factors of this kind are presented in this volume. Thus, six chapters are the first studies of orthographic conventions in the given cultures (*Ivušić*, *Sobieroj*, *van der Putten*, *Bondarev and Dobronravin*, *Souag*, *Gori*), whereas five others refine, critically analyse or summarise received understanding of the better-documented cultures (*Orsatti*, *Wagner*, *La Spisa*, *de Castilla*, *Schmidt*). For the latter, it is easier to identify connections between norms in social structure, manuscript production and orthography. In the less-known cultures, such links are not as obvious. The factors outlined below are thus more of a selection of representative phenomena than a representative typological survey. But even the uneven comparison yields some interesting results.

I have grouped the phenomena identified as relevant for the standardisation process into seven umbrella classes. These are factors related to (1) *contact* situations, (2) *authority*, in the sense of top-down regulations, (3) *networks* of scribes, (4) *identity* – both communal and scribe-centred individual aspects of writing, (5) *genre*, as a cover term for socially identified literacy events, linguistic codes and registers, (6) *language*, in the sense of linguistic structures and features conditioning orthographic choices, and (7) *medium* (manuscript and print). This is a simplified classification: the conceptual scope of each of the seven keywords is wider and many phenomena cannot neatly be subsumed under a single group of factors, while some factors might better be grouped under a separate umbrella concept which I may have omitted. I shall try to cross-reference related groups of factors and point to possible overlaps as well.[21]

7.1 Contact

As discussed in the section on comparative standardology, the phenomenon of standardisation can be understood as a special type of language contact (Haugen

21 Unsurprisingly, this grouping has a bias towards sociolinguistics where such factors are typically in the focus of study. I hope that invoking familiar sociolinguistic notions and relating them to the phenomena in manuscript cultures creates a potential for bridging the disciplines on more recognisable common ground and thus for achieving better interdisciplinary compatibility.

1972, 247) and this is equally relevant for the other domains of culture, too, such as religion, writing and manuscript production.

What appears to be the most obvious factor behind standardisation in the Islamic cultures based on Arabic script is the standardised text of the Qur'an, which can be seen as the interface between Arabic and non-Arabic cultures in contact situations. As Morton (2012, 150) puts it:

> The orthography of New Persian has been remarkably stable considering that the language has been used for well over a millennium [...] An important factor here is the influence of Arabic, the orthography of which has been even more stable than that of Persian and over a longer period. In the case of Arabic, stability was encouraged in particular by the attention paid to the interpretation of the text of the Qur'an and religious concerns in general. [22]

In many cases, religion barriers were irrelevant for the impact of Arabic orthography, as argued by *Wagner*: 'the newly emerging Judaeo-Arabic standard was heavily Arabicised and written in an orthography where Classical Arabic spelling conventions were imposed on the Hebrew letters'. *La Spisa* equally shows that the influence of the orthography of the language of the Qur'an is discernible even in Christian Arabic manuscripts.

However, standardised Arabic (including the Qur'an and Classical Arabic texts) was not the sole force exerting influence on standardisation in the contact culture. Thus, '[t]he models for correct Persian usage emerged in the 9th century from the pre-Islamic heritage of the epistolographic practices of the Sasanian Empire (AD 224–651)' (Spooner and Hanaway 2012a, 17). The normalised orthography of New Persian written in Arabic script (Arabo-Persian) might have resulted from a convergence of pre-Islamic and Islamic standards of literacy practices. In turn, as *Orsatti* demonstrates, once it was established as a stable norm by the end of the 9th century, the orthography of New Persian influenced spelling in other scripts in Judaeo-Persian, Syro-Persian and Manichaean texts. A similar tendency for orthographic interaction across different scripts is suggested by *Ivušić* for some spelling conventions in Hungarian, Latin and German written in Arabic script in a 16th-century Ottoman manuscript which may have been influenced by German Latin-script orthography.[23]

22 However, as Lameen Souag observes in his comment on this chapter, 'while the text of the Qur'an is highly standardised, its orthography is much less so. The same word may be written in one *aya* with *alif* and in another without it; in one *aya* with ة, and in another with ت... And, of course, to the extent that it is standardised, its orthography often differs from what would become the much more standardised orthography of Classical Arabic'.

23 The Ottoman chanceries were essentially embedded in a multilingual environment: 'Turkish was not the only language used in the Ottoman chancery. Over the years the Ottomans emplo-

Contact-induced spelling conventions are also common in languages written in the same script. Such was the influence, albeit minor, of the non-Ottoman Turkic language Chagatai and Azeri Turkish on Ottoman Turkish orthography (*Schmidt*) and the influence of Kanuri spelling conventions on Hausa Ajami writing (*Bondarev and Dobronravin*).

A special case of the effect that standardised Classical Arabic had on the orthography of the contact/target language is retention of the (historical) spelling of Arabic words. This is reported for most of the manuscript cultures discussed in this volume, even those lacking unified spelling conventions, and can be seen as a micro-area of standardisation, whereby the diffusion of a standard spelling is confined to a restricted set of lexical items or grammatical structures (*Bondarev and Dobronravin*).

Convergence induced by contact is also manifest in palaeographic features of manuscripts, irrespective of any religious divides, as *La Spisa* demonstrates with Christian Arabic sources. A striking case of interaction is the influence of Chinese calligraphy on the Arabic script used in Sino-Arabic manuscripts (*Sobieroj*). In manuscript cultures sharing the same religion, the unifying force of contact is even stronger (as in the Kanuri and Hausa manuscripts, both cultures being Islamic).

A complex contact situation involving very different writings systems – 'consonantal' and 'alfasyllabary' – is discussed by *Gori* using the example of transliteration of Arabic texts in the Ethiopian script *fidäl*, written in the Muslim communities of Harar.[24] *Gori* argues that the standard Arabic orthography of the Arabic text written in *fidäl* does not prompt the spelling characteristics of Ethio-Arabic texts. Rather, it is oral recitation in Arabic that provides a reference point for spelling choices.

An important aspect of contact is the tendency of contrast and divergence. With increased socio-cultural tensions, the orthography of one language initially modelled on the norms of another may de-standardise and take on new normative principles, as observed in the change of Judaeo-Arabic from the stage of orientation based on Classical Arabic in the 11th century to the period of the 13th century influenced by Hebrew norms (*Wagner*). Similar divergence is often reported for manuscript cultures, with normative patterns in one religion triggering opposite patterns in the other, such as retention of the rollbook by the Jews 'in order to differ from the Christians' (Beit-Arié 1992, 11).

yed scribes who wrote in Latin, Greek, Italian, Uighur, Persian, Arabic, Serbian, Hungarian, and other languages [...]' (Darling 2012, 177).

24 Strictly speaking, Arabic is not purely consonantal, nor is *fidäl* alfasyllabic, but the systems differ in their treatment of vowels: Arabic makes short vowels optional, whereas it is obligatory to mark all the vowels in *fidäl*. Daniels (1990, 1997, 2014) calls the former 'not a perfect *abjad*' (2014, 30) and the latter *abugida*.

7.2 Authority

Authority is used here in the sense of top-down regulations. These include all sorts of centralised controlling mechanisms imposed by political, administrative, religious and other institutions. The contrasting side (in other words, the opposite value) of such top-down regulations is 'bottom-up literacy regimes' (Blommaert 2008, Juffermans et al. 2014, Weth and Juffermans 2018). The phenomena associated with authority are discussed in sociolinguistics literature as the most common and typical causes of language standardisation. However, top-down authority does not constitute an important factor in any of the twelve cultures presented in this volume. Spooner and Hanaway's (2012a, 14) survey of the history of written Persian is instructive in this regard:

> What is most remarkable is the lack of any central authority to govern usage or establish models of correctness. [...] For Persian, [...] there was neither a primary text nor any other type of authority besides the heritage of Sasanian bureaucrats, which was gradually succeeded by the evolving canon of secular Persian literary texts. In this connection it may be worth noting that Islamic civilization in general was characterized by a lesser degree of centralization than other parts of the world, until perhaps the later emergence of what Hodgson (1974) calls the Gunpowder Empires: the Ottomans, Safavids, and Mughals.

The emergence of written New Persian in the 9th century discussed by *Orsatti* was followed by a centrifugal spread of its standardised variety. But it was the prestige of the literary language in the Persian courts of eastern Iran rather than any authoritative centralised force which was responsible for the spread of the written standard.

The authority of influential scribes played a significant role in generating normative spelling and unifying script styles, which were propagated through specific networks (*Wagner*). However, such cases of authority fit into the categories of network and identity better, which are discussed below.[25]

25 It is interesting to mention the case of the Masoretic Syriac compilations of the 8th–13th centuries because they manifest a strict regulatory standard tradition developed at the time roughly coinciding with the time of the Judaeo-Arabic, Christian-Arabic and Arabo-Persian traditions discussed in this volume. 'Based on sample texts they [the compilations] standardize the orthographic representation of the pronunciation' (Juckel 2011, 276). What makes the Syriac case special is that the scale of influence imposing an exclusive standardised orthography is greater than individual influence propagated via networks and therefore, in its prescriptive force, it is comparable to the better-known regulated standardisation processes, such as those in European languages.

A special orthographic twist indirectly related to top-down regulations is discernible in Aljamiado manuscripts, which *de Castilla* discusses. There, centralised authoritative efforts of standardisation applied to written Spanish (Castilian and Aragonese)[26] in Latin script are visible in Aljamiado writings in Arabic script where an Aragonese variety of Spanish is spelled according to 16th-century conventions.

Centralised administrative power does not seem to have played a direct role in standardisation even in the Ottoman Empire, although the stability of the written language was a distinctive feature throughout the history of Ottoman manuscripts starting from the late 15th century. Thus, 'the stable language tradition in Turkish was [...] that of administrative and government documents' (Darling 2012, 179) and 'the tendency towards standardisation is visible in Ottoman texts' (*Schmidt*). However, before 1908 there was no formal policy which regulated standards, either in manuscript production or in orthography. It is noteworthy that what was least formalised was orthography, as opposed to the more fixed uniformity of the layout, script style, formulaic expressions and lexicon (each entity conditioned by specific written genres).

In sum, the positive value of top-down control almost seems irrelevant for standardisation processes in the cultures that are discussed in this volume. Looking at the category of authority from the opposite value – bottom-up literacy – might lead to insights, but these will be phenomena more adequate for the categories of networks, identity, genre and language.

7.3 Network

In a narrow sociolinguistic sense, a network is 'a boundless web of ties that reaches out through a whole society, linking people to one another, however remotely' (Milroy and Milroy 1992, 5). First-order networks are ties directly anchored to individuals, and second-order networks are ties linked through others. I expand the notion of network to a wider connotation standing for a communal base of social practices in contrast to the individual space. For example, the space where reading of a text is shared by people of different ages and social status – as in the recitation of canonical texts – creates a specific literacy event network which may serve the participating individuals as a reference to normative codes. The concept of a network is tightly linked to what I call 'identity' in the next section, the notion related to individual-based phenomena.

26 Wright 2006, 608.

In many pre-modern cultures, it is the role of scribal networks rather than central authority which drove (de)standardisation processes, as shown by *Wagner* and, *inter alia*, in Wagner et al. 2013a, b: '[…] local networks, scribal schools and gatherings of writers developed their own norms, which they then followed' (*Wagner*). The network factor, although not specifically addressed by *La Spisa*, was probably at work in the production of Christian Arabic manuscripts in the monastic communities of South Palestinian monasteries.

The standards in written Persian 'were maintained through the interaction within and between chancelleries and the court communities of multiple sultanates' (Spooner and Hanaway 2012a, 17). By the same token, the chancery networks were responsible for the formation and maintenance of standard forms in Ottoman Turkish (Darling 2012).

The network factor also seems to be important in Aljamiado manuscripts of the 15th–17th centuries. Although they were written by unidentified scribes, *de Castilla* argues 'that most of the manuscripts were produced by skilled copyists' to 'maintain the cohesion of their communities and to control them, at a moment when they were losing their rites and cultural practices'.

Multi-nodal networks were behind the diversification of manuscript traditions related to Malay in insular Southeast Asia of the 15th to the 19th century. *Van der Putten* explores (in)consistency in orthography and other domains of Malay manuscripts from the Moluccas, which was one of those nodes.

Souag, describing orthographic characteristics in Kabyle, the largest Berber variety of Algeria, shows that the absence of stable spelling conventions is commensurate with the patchy and troubled history of Kabyle educational institutions: 'At no point has any one writer's or school's Arabic-script work been sufficiently widely read to be imitated, and all but the most prominent of one generation's orthographic innovations have been forgotten by the next'. This negative finding only confirms the role of networks as one of the primary factors of standardisation processes.

Shared space in literacy practices may have played a role similar to networks in stabilising spelling conventions in Old Kanembu manuscripts, as suggested in *Bondarev and Dobronravin*, and in Bondarev and Tijani 2014.

7.4 Identity

Identity is a scalar concept, ranging from the wider sense of group identity to a narrower sense of personal identity, which stands for individual-centred phenomena. By explicitly narrowing down one side of the definition, I try to distinguish between communal-based group identities and individual-based identities,

however artificial such a distinction might seem. The heuristic reason for narrowing down the notion of identity to a strictly individual level lies in the fact that the first-tier relation between the reader of a manuscript and the scribe is basically a one-to-one relation. When I see a written word penned in a certain hand and style, what typically 'stands' on the other side of the word is a single hand of a single scribe. It goes without saying that any individual identity will most likely represent a larger picture of group identity upheld through networks. But it will be the individual scribal expression with all its idiosyncrasy which will provide the first entry point to whatever might be behind a unique instance of writing.

From the standpoint of this categorial division, it is informative to explore the extent to which the impact of group identity on standardisation differs from that of individual identity. The first type of impact might be easier to identify. The group identity will not necessarily be directly present (or explicitly named) in the artefact. Rather, it will have to be postulated as a result of historical reconstruction. If the identity of one group opposing another is historically proven, however, as is the case for Muslim vs Christian identities in late medieval Iberia, for example (cf. Catlos 2014), then it is plausible to see differences in writing between these two groups as identity markers. And if there are consistent ways of writing associated with one particular group, it is reasonable to postulate identity as a factor for such consistency.

Many chapters in this volume allude, in one way or another, to such reconstructed group identities. The emergence of the New Persian orthographic standard during the 9th century is related to the milieu of the Persian courts which flourished in eastern and north-eastern Iran (*Orsatti*), with the subsequent spread of the written New Persian standard driven by the prestige of the literary language. The identity of the Jewish middle-class writers as a factor of (de)standardisation process is reconstructable for the Judaeo-Arabic manuscripts from the Cairo Genizah, as discussed by *Wagner*. She also argues on the basis of her own and others' sociolinguistic research that standardisation may be blocked by the factors related to regional identity or associated with particular scribal schools. Religious identity might be a factor for a micro-standard set of features within a larger system of conventions, as demonstrated by *La Spisa*. Thus, the orthography of the Christian-Arabic manuscripts of the Melkite monastic environment, which is largely congruent to Islamic manuscript conventions, has some distinctive features specific to Christian Melkite writing. The cultural identity of Mudejars and Moriscos in Castile and especially Aragon strongly influenced the production of Aljamiado manuscripts and associated orthographic conventions (*de Castilla*).

Turning to the narrower side of identity, the most obvious instance of individuals playing a role in the standardisation process is the prestige of those who set up the models subsequently copied by the others. The history of standardisation

testifies to the importance of individuals standing behind codification processes (Agha 2007, Sebba 2007). Individual prestige – one of the most important factors 'controlling the assertiveness of linguistic forms' (*Wagner*) – is the long-standing subject of sociolinguistic investigation. Related to prestige, but not necessarily fully matched with it, are the instances of (identifiable) individual choices of norm or variation and generally, all sorts of idiosyncrasy. Den Heijer (2012, 11), discussing standardisation in what is known as Middle and Mixed Arabic, suggests that the deviations from normative systems 'reflect a conscious desire to mix registers and styles'. This point is elaborated by *Wagner*, who shows that individuals may consolidate microsystems shared by a limited group of manuscripts and persons. Similar small-scale individual substandards are observed by *van der Putten*, *Bondarev and Dobronravin* and *Souag*.[27]

When doing my research on Old Kanembu manuscripts in Nigeria, I often asked the Islamic scholars proficient in the variety of Old Kanembu called Tarjumo why people wrote the way they did. The typical answer was: 'people write however they want to, the way they hear it'. Yet on closer inspection, various idiosyncratic choices betray something more systematic than mere idiosyncrasy (*Bondarev and Dobronravin*). Even though the full scale of personal social relationships, as it might be manifest on the pages of manuscripts, will most likely remain out of reach due to the inevitable lack of contextual information, the factor of individual identity – however difficult it is to tackle – is central to the study of standardisation. This is because personal identity intersects with all the factors discussed here (as well as others that I may have omitted). Although the factors are certainly interrelated, the act of implementing a literacy practice – informed by a multiplicity of factors and scribal repertoires – is ultimately carried out by a single scribe in a single hand.

7.5 Genre

I use genre as a cover term for linguistic codes and registers embedded in a *variety* of socially identified and conventionalised literacy practices, framing a 'particular perspective on the world' (Barber 2007, 41). This is wide enough to include issues of education, types of texts and manuscripts, scribal repertoires, transmission and translation of canonical texts, culturally conditioned preservation of earlier varieties of language, distribution of registers and genres across literacy practices and

27 On the issue of deviation from the norms conditioned by distinctive group and individual identity, see Weth and Juffermans 2018.

many other related phenomena. By encompassing all this under one category, I argue for the intrinsic interrelatedness of such phenomena, which constitute a distinct family of factors.[28] Written language (as well as spoken) is never neutral. Each speech and literacy event requires a selection from the available linguistic repertoires of speakers and scribes. Some events are flexible and open to a greater variation of choices, whereas others are more conservative with a limited range of choices, for example in recitation and written transmission of canonical texts. One literacy output will result in elimination of variation and the other in the proliferation of variation. It is predictable that the standardisation process and its speed will be different in different registers and types of texts, and it will often be a single register which will activate the process of standardisation. As *Orsatti* argues, New Persian orthography came into being as part of the development of literary lyric poetry in the Persian courts. With the spread of the fixed orthography to other types of texts as well, the difference in genre did not disturb its stability. The subsequent irrelevance of genre for spelling conventions which became fixed is taken as proof of the antiquity of Arabo-Persian standardised orthography.

Genre variation is also a factor behind the divergence of standardisation. The genre-specific distribution of (sub)standard microsystems has been reported for many manuscript cultures, including those discussed in this volume. *Wagner* shows that from the 13th century onwards, the norms in the orthography of Judaeo-Arabic depended on the type of texts involved, unlike the previous period of the 10th–12th centuries when standardised orthography cut across the genres. Equally, *van der Putten* points to a higher degree of consistency in doctrinal Islamic texts as compared to secular texts in the contexts of Malay manuscripts. Genre-specific orthography is elaborated by *Sobieroj* using a number of examples in *xiaojing* writing. The differences in Kabyle orthographic subsystems presented by *Souag* are to a certain extent motivated by different types of texts and literacy contexts, such as the needs of students attending a *zaouia* (religious school).

A specialised standardised register may come into being as a result of translating holy scriptures. Thus, Saadiah Gaon's Bible translation facilitated the consolidation of the consistent Arabicised orthography of standard Classical Judaeo-Arabic, which also spread to utilitarian writing in Fatimid Egypt (*Wagner*).

In the context of Islam in West Africa, the translation of scriptures and canonical texts was a prominent factor in creating special literary registers in West African societies, leading to stable spelling conventions (Bondarev 2014, *Bondarev and Dobronravin*, Ogorodnikova 2017, Tamari and Bondarev 2013). One side-

28 For similar or related approaches to genre, see, *inter alia*, Bakhtin 1981, Barton 1994, Barber 2007, Biber and Conrad 2009, and Lillis 2013.

effect of translational practices is the preservation of earlier linguistic features of language, for example, archaic Aragonese in Aljamiado manuscripts (*de Castilla*), archaic Kanuri in Old Kanembu manuscripts (Bondarev 2013, *Bondarev and Dobronravin*) or archaic Hausa in Hausa manuscripts (Dobronravin 2013).

Different written genres could influence the distribution of linguistic norms and manuscript features, such as layout and script style. *Schmidt* discusses different script styles applied to different types of texts in Ottoman manuscripts. What makes the case of financial documents written in *siyakat* script especially interesting is that the language of the texts included fixed sets of formulaic expressions and vocabulary and was meant to be clear and unambiguous, while the script 'was intended to make the documents hard to falsify and hard for outsiders to understand or imitate' (Darling 2012, 180).

7.6 Language

I restrict this group of factors to a selected number of identifiable cases of linguistic structures motivating orthographic choices. These factors are followed by a separate sub-entry of 'orthographic design and its uses' as the topic typically discussed in linguistics. Many factors mentioned in this section deal with linguistic features expressed at individual levels, therefore this group relates to the wider category of 'genre' in a similar way to 'identity' relating to 'network'.

7.6.1 Linguistic structures: phonology, morphology, syntax

Many manuscript cultures have phonetic orthography at one stage of their development or the other (*Wagner*, *Ivušić*, *Sobieroj*, *van der Putten*, *Souag*, *Gori*). Phonetic transcription opens avenues for variation due to variation in spoken language conditioned by social status, register, dialect and other factors. An important feature of phonetic transcription is that it is largely based on scribes' linguistic intuition. Spelling choices informed by speakers/writers' intuition may form stable patterns at micro-levels of orthography with the resultant retention of such microsystems. What may first appear to be spelling inconsistency on closer examination turns out to be a stable orthographic tendency applied at micro-levels, which reflect fair guesses by the scribes about the structure of the language they were using in writing. The patterns of emergent spelling in manuscripts are in many respects comparable to the patterns discovered in the studies of children's invented spelling (see an overview in Read and Treiman 2013). In spite of great variation from child to child, some patterns of spelling are systematic

and uniform in children's writing. For example, one of the typical features is omission of nasals before consonants. This same systematic omission has been reported for many Arabic-based systems (*Bondarev and Dobronravin*, Mumin and Versteegh 2013, Souag 2010). The systematic patterns of scribes' linguistic intuition may be detected in various domains of the language: phonemic (e.g. *Souag*), phonological (e.g. *Orsatti, Ivušić*), morphological (e.g. *Ivušić*) or syntactic (e.g. Dobronravin 2006). Taking such patterns into account may help us see distinct microsystems or 'collections of unstandardised consistencies' (*van der Putten*) in manuscripts with considerable orthographic variation. In *Bondarev and Dobronravin*, we describe such microsystems as '*stable sets* of grapheme-phoneme correspondences', highlighting the fact that orthographies may be composed of a variety of stable subsystems, each with its specific multiplex correspondences.[29] For example, set 1 has a certain number of graphemes to cover a certain number of phonemes, whereas set 2 will have different graphemes to cover different phonemes, and set 3 will differ from sets 1 and 2, etc. Thus, the orthographic system which may be highly variable within one set might be remarkably stable in terms of contrast between the sets.

The conditioning factors behind such microsystems of standardisation include salience and frequency. For example, phonetically salient items such as stressed vowels or syllables at the end of a phrase (*Bondarev and Dobronravin*) have high potential for being encoded uniformly. The frequency of items may be conditioned by purely language-specific peculiarities such as the frequency of function words like prepositions. But frequency can also be register-specific, some registers having more of one type of item and less of another (see Biber 1995, for instance).

As mentioned under 'contact', words borrowed from Arabic typically contribute to a standardised orthographic subset as they are usually written in the original standard Arabic spelling, irrespective of their actual pronunciation in the recipient language. Specialised vocabulary is another factor of stability, as was the case with botanic items in Aljamiado (*de Castilla*).

In this random overview of linguistic factors playing a role in standardisation processes, one special case worth mentioning is linguistic economy. In Ottoman Turkish, it was the 'omission of whatever can be omitted, such as repetitive verb endings' (Darling 2012, 180), that counteracted standardisation.

29 This approach is comparable to the treatment of graphemic and phonemic variation in Middle English manuscripts. Scholars use *literal substitution sets* for graphemes and *potestatic substitution sets* for phonemes (Rutkowska and Rössler 2012, 222; LAEME – Linguistic Atlas of Early Middle English).

7.6.2 Orthographic design and its uses

The phenomena discussed here are more related to the orthographic system as a whole rather than to individual choices accounted for by scribes' phonological (and, generally, linguistic) intuition.

Phonological differences between Arabic and non-Arabic languages result in orthographies with underspecification (one letter standing for more than one sound) and overspecification (more than one letter for one sound), which could either have (de)standardisation effects or have no effect whatsoever. A common way of overcoming divergences between Arabic and other languages' phonologies was the invention of new letters. The introduction of new letters explicitly leads to a break from the standard model of Arabic orthography, as was the case in Persian, Turkish, Malay, Kanuri, Hausa, Berber and other languages. In some manuscript cultures such as Persian (*Orsatti*), the new letters entered pre-existing stable orthographic systems, while in others the invented letters initially increased the amount of variation in spelling, with subsequent stabilisation of orthography, as the history of Hausa writing shows (*Bondarev and Dobronravin*).

Historical orthography, often originating from certain literary registers, has a tendency to remain in manuscripts over a long period as standardised microsystems (*Orsatti*, *de Castilla*, *van der Putten*, *Bondarev and Dobronravin*).

Some orthographic designs derive from the Arabic model, but take on a different function, as, for example, the otiose use of the Arabic letter <h> to indicate the final -a of lexical items (*Orsatti* on Persian and *Bondarev and Dobronravin* on Kanuri).

Idiosyncratic unsystematic spelling might be seen as chaotic and outside any considerations of norms. However, many studies, including this volume (*Wagner*) and elsewhere (Vandenbussche 2002; den Heijer 2012, 11), show that variability of spelling – even of the same word on the same page or line – was the result of a conscious desire to mix linguistic styles, and therefore such inconsistencies are better described as 'distinct spelling systems' (Vandenbussche 2002, 32). Thus, in many manuscript cultures, a substantial variation in spelling may paradoxically be indicative of a standard norm.

7.7 Medium

7.7.1 Manuscript

Ideally, this set of factors should address the manuscript-related causes of the process of standardisation as it may be manifested in manuscript culture.

However, the great breadth of 'manuscript-related' issues makes it open to taxonomically different factors. One group of factors conditions the standardisation of manuscript production – technology and husbandry will condition paper production and parchment production respectively, for example. The other group of factors is generated by manuscript production itself and has both self-inflicting influence (for example, the production of writing material such as parchment will dictate the size of manuscripts) and external influence (upon orthography, for example). Since standardisation processes in the manuscript cultures presented in this volume are largely explored in the domain of orthography, the complex issues of the relationship between codicological and linguistic features have only been touched on lightly. In the following paragraph, I shall only mention some manuscript-related factors as they were presented in some of the chapters.

La Spisa points to the relationship between the norms of script type and layout: the change from the early Abbasid script to *naskhī* script led to a change in the whole structure of page layouts. *Sobieroj* observes that in some liturgical texts 'standardisation in handwriting and choice of format has been realised in mutual dependency'. He also discusses the influence of Arabic manuscript culture on the processes of standardisation in *Khaṭṭ-i ṣīnī* manuscripts (this factor overlapping with those discussed under 'contact'). *Schmidt* demonstrates that standardisation of script types correlated with standards regarding formats and layouts, and was prominent in Ottoman manuscripts. He also points to the ascendancy of calligraphy over orthography: 'Calligraphy was considered to be one of the highest forms of art, if not an esoteric science, compared to which mere orthography was insignificant'.

An important factor underlying the uniformity of scripts and uniform or varied orthographic conventions was copying, which is not discussed in this volume, but was vividly described by Darling 2012, again using Ottoman Turkish manuscripts as examples. In the end, copying variation in manuscripts reduces the potential to invent one's own spelling.

7.7.2 Print

Most of the cultures discussed in this volume existed before the print era or were outside its immediate sphere of influence. Some of the chapters here mention the interaction between manuscript and print cultures when it is observable (*Schmidt, Sobieroj, Bondarev and Dobronravin*) and one chapter's conclusions are mostly based on printed material (*Souag*). Other than that, the topic of print was considered beyond the book's scope. Firstly, standardisation related to print cultures, especially in the context of standardisation of European languages, has

been studied much more than standardisation in manuscript cultures. Secondly, the transition from manuscript to print and its implications for standardisation would require separate research. 'Manuscript vs print' is not a dichotomy. Contrary to the second quote in this chapter (Barton 1994, 124), there is no clear divide between 'the earlier scribal culture' with 'idiosyncratic formats' produced by 'individual copyists' on the one hand and printing on the other hand, which 'allowed a large number of identical texts' (for more subtle views on the problem, see the literature cited in the section on 'comparative standardology'). Upon the advent of print, manuscripts were the models for book production, and with its advancement, printed material influenced manuscripts, the latter having persisted up to the present day in many corners of the world.

8 Conclusions

Continuing the dialogue between 'comparative standardology' and variation studies, this volume offers comparative data which allows us to delineate large groups of factors behind the standardisation process in manuscript cultures. The grouping into factors suggested here is by no means definitive and requires further elaboration. That said, I hope that the groups of factors discussed here will be helpful as a preliminary checklist for standardisation in interdisciplinary research.

The discussion in this chapter draws on the assumption that standardisation can be understood better if it is considered a non-unidirectional process. The groups of factors can thus be seen as channels or pathways of (de)standardisation. A comparative analysis of such pathways in the twelve manuscript cultures addressed in this volume permits us to postulate some generalisations, as follows.

Contact situations do not necessarily lead to the exchange of standardised orthographic principles. In many cultures, the co-existence of Standard Arabic and non-standardised languages spoken and written in Muslim communities poses a paradox: such languages are profoundly influenced by Arabic, but their orthographies are not modelled on the principle of standardisation.

This apparent paradox is resolved by the prediction that standards in orthography – one of the domains of manuscript culture – are conceptually different from standards in other domains, such as format, layout and script. As there is a difference between the standardisation of spoken language and written language, there is also a difference between the standardisation of orthography and other domains of manuscript production.

Each domain of manuscript culture develops microsystems of standardisation. This observation seems especially pertinent to the understanding of variation and standardisation in orthography. Orthographic variation is rarely arbitrary. Rather, it is organised in distinctive grapheme-phoneme subsets. Thus, the orthographic system which may be highly variable within one set might be remarkably stable in terms of contrast between the sets.

Different domains have different 'areas' of uniformity and standardisation in a given manuscript culture. Layout and script type are one such area, which is often subdivided into two: layout, having a wider regional and cultural scope of uniformity, and script type, having a narrower scope of uniformity. The other area of uniformity, which is not necessarily linked to the layout/script domain, is orthography with its own principles of (non-/de-)standardisation.

Thus, a general tendency observable at the level of physical features of manuscript production is that layout and script types tend to be unified, irrespective of orthographic norms and, vice versa, orthographic norms develop irrespective of norms applied to physical domains of manuscript production.

References

Agha, Asif (2007), *Language and Social Relations* (Studies in the Social and Cultural Foundations of Language). Cambridge: Cambridge University Press.

Andrist, Patrick, Paul Cana.t and Marilena Maniaci (2013), *La syntaxe du codex : essai de codicologie structurale*. Turnhout: Brepols.

Bakhtin, Mikhail M. (1986), *Speech Genres and Other Late Essays*, trans. by V. W. McGee, Austin, Texas: University of Texas Press.

Barber, Karin (2007), *The Anthropology of Texts, Persons, and Publics*. Cambridge: Cambridge University Press.

Barton, David (1994), *Literacy. An Introduction to the Ecology of Written Language*. Oxford: Blackwell.

Barton, David (2007), *Literacy. An Introduction to the Ecology of Written Language* (second edition). Oxford: Blackwell.

Beit-Arié, Malachi (1992), *Hebrew Manuscripts of East and West* (Panizzi Lectures). London: The British Library.

Beit-Arié, Malachi (2017), *Hebrew Codicology: Historical and Comparative Typology of Hebrew Medieval Codices based on the Documentation of the Extant Dated Manuscripts Using a Quantitative Approach* [pre-print internet version in English, v0.1]

Biber, Douglas (1995), *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press.

Biber, Douglas, and Edward Finegan (eds) (1994), *Sociolinguistic Perspectives on Register*. Oxford: Oxford University Press.

Biber, Douglas, and Susan Conrad (2009), *Register, Genre and Style*. Cambridge: Cambridge University Press.

Blommaert, Jan (2005), *Discourse*. Cambridge: Cambridge University Press.
Blommaert, Jan (2008), *Grassroots Literacy*. London: Routledge.
Boltz, William G. (1994), *The Origin and Early Development of the Chinese Writing System*, New Haven: American Oriental Society.
Bondarev, Dmitry (2013), 'Language change induced by written codes: a case of Old Kanembu and Kanuri dialects' in Esther-Miriam Wagner, Ben Outhwaite and Bettina Beinhoff (eds), *Scribes as Agents of Language Change* (Studies in language change). Boston, Berlin: De Gruyter Mouton, 291–323.
Bondarev, Dmitry (2014), 'Old Kanembu and Kanuri in Arabic script: phonology through the graphic system', in Meikal Mumin and Kees Versteegh (eds), *The Arabic Script in Africa: Studies in the Use of a Writing System*. Leiden, Boston: Brill, 107–42.
Bondarev, Dmitry (2017), 'Islamic Education and Ample Space Layout in West African Manuscripts', in Andrea Brigaglia and Mauro Nobili (eds), *The Arts and Crafts of Literacy: Manuscript Cultures in Muslim Sub-Saharan Africa*. (Studies in Manuscript Cultures, 12). Berlin, Boston: De Gruyter, 105–142.
Bondarev, Dmitry and Abba Tijani (2014), 'Performance of multilayered literacy: Tarjumo of the Kanuri Muslim scholars', in Kasper Juffermans, Yonas M. Asfaha and Ashraf K. Abdelhay (eds), *African Literacies: Ideologies, Texts, Education*, Newcastle upon Tyne: Cambridge Scholars Publishing, 118–146.
Bullough, Donald. A. (1991), *Carolingian Renewal: Sources and Heritage*. Manchester: Manchester University Press.
Chambers, J. K., and Natalie Schilling (eds) (2013), *The Handbook of Language Variation and Change* (second edition) (Blackwell Handbooks in Linguistics). Malden, MA/Oxford: Wiley-Blackwell.
Catlos, Brian (2014), *Muslims of Medieval Latin Christendom, c. 1050–1614,* Cambridge: Cambridge University Press.
Coulmas, Florian (2003), *Writing Systems: An Introduction to their Linguistic Analysis*, Cambridge: Cambridge University Press.
Cromwell, Jennifer, and Eitan Grossman (eds) (2018), *Scribal Repertoires in Egypt from the New Kingdom to the Early Islamic Period*. Oxford: Oxford University Press.
Daniels, Peter (1990), 'Fundamentals of grammatology', in *Journal of the American Oriental Society* 110, 4, 727–31.
Daniels, Peter T. (1997), 'The Protean Arabic Abjad', in Afsaruddin, Asma and A. H. Mathias Zahniser (eds), *Humanism, Culture, and Language in the Near East: Studies in Honor of Georg Krotkoff*. Winona Lake: Eisenbrauns, 369–384.
Daniels, Peter T. (2014), 'The Type and Spread of Arabic Script', in Meikal Mumin and Kees Versteegh (eds), *The Arabic Script in Africa: Studies in the Use of a Writing System*, Leiden, Boston: Brill, 25–39.
Darling, Linda T. (2012), 'Ottoman Turkish: Written Language and Scribal Practice, 13th to 20th Centuries', in Brian Spooner and William L. Hanaway (eds), *Literacy in the Persianate World: Writing and the Social Order*, Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology, 171–195.
Den Heijer, Johannes (2012), 'Introduction: Middle and Mixed Arabic, a New Trend in Arabic Studies', in Zack, Liesbeth and Arie Schippers (eds) (2012), *Middle Arabic and Mixed Arabic. Diachrony and Synchrony*, Leiden, Boston: Brill.

Den Heijer, Johannes, Andrea Schmidt and Tamara Pataridze (eds) (2014), *Scripts Beyond Borders: a Survey of Allographic Traditions in the Euro-Mediterranean World*. Leuven: Peeters.

Déroche, François et al. (2006), *Islamic Codicology: An Introduction to the Study of Manuscripts in Arabic Script*; translated by Deke Dusinberre and David Radzinowicz, edited by Muhammad Isa Waley, London: Al-Furqān Islamic Heritage Foundation.

Deumert, Ana, and Wim Vandenbussche (eds) (2003a), *Germanic Standardisation. Past to Present.* Amsterdam, Philadelphia: John Benjamins.

Deumert, Ana, and Wim Vandenbussche (2003b), 'Standard Languages: Taxonomies and Histories', in Ana Deumert and Wim Vandenbussche (eds), *Germanic Standardisation. Past to Present.* Amsterdam, Philadelphia: John Benjamins, 1–14.

Deumert, Ana and Wim Vandenbussche (2003c), 'Research Direction in the Study of Language Standardization', in Ana Deumert and Wim Vandenbussche (eds), *Germanic Standardisation. Past to Present.* Amsterdam, Philadelphia: John Benjamins, 455–469.

Dobronravin, Nikolay (1999), Arabograficheskaya pis'mennaya traditsiya Zapadnoy Afriki [The Written Tradition of West Africa in Arabic Script], St Petersburg: St Petersburg University.

Dobronravin, Nikolay (2006), 'TAM markers in "Littafen nahwowin hausance": interpreted as verbs?', in Bernard Caron and Petr Zima (eds), *Sprachbund in the West African Sahel*, Louvain: Peeters, 129–43.

Dobronravin, Nikolay (2013), 'Classical Hausa' glosses in a Nineteenth-century Qur'an Manuscript: A Case of "Translational Reading" in Sudanic Africa?', in *Journal of Qur'anic Studies*, 15/3: 84–122.

Dobronravin, Nikolay (2017), 'Design Elements and Illuminations in Nigerian "Market Literature" in Arabic and ʿAjamī', in *Islamic Africa*, 8, 1–2: 43–69.

Dossena, Marina (2003), 'Scots', in Anna Deumert and Wim Vandenbussche (eds) (2003), *Germanic Standardisation. Past to Present.* Amsterdam, Philadelphia: John Benjamins, 383–404.

Eisenstein, Elizabeth L. (1979), *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early-Modern Europe*. 2 vols. Cambridge: Cambridge University Press.

Eisenstein, Elizabeth L. (1983), *The Printing Revolution in Early Modern Europe*. Cambridge: Cambridge University Press.

Ferguson, Charles Albert (1988), 'Standardization as a form of language spread', in Peter Lowenberg (ed.), Language Spread and Language Policy: Issues, implications, and case studies. Georgetown University Round Table on Languages and Linguistics 1987. Washington, D.C.: Georgetown University Press, 119–132.

Ferguson, Charles Albert (1994), 'Dialect, Register and Genre: Working Assumptions about Conventionalization', in Douglas Biber and Edward Finegan (eds), *Sociolinguistic Perspectives on Register*. Oxford: Oxford University Press, 15–30.

Fishman, Joshua A. (ed.) (1974), *Advances in Language Planning*. The Hague, Paris: Mouton.

Fleischman, Suzanne (2000), 'Methodologies and Ideologies in Historical Linguistics: On Working with Older Languages', in Susan C. Herring, Pieter van Reenen and Lene Schøsler (eds), *Textual Parameters in Older Languages*. Amsterdam: Benjamins. 33–58.

Gacek, Adam (2009), *Arabic Manuscripts: A Vademecum for Readers*. Leiden: Brill.

Garvin, Paul L. and Madeleine Mathiot (1960), 'Urbanisation of the Guarani language: a problem in language and culture', in A. C. Wallace (ed.), *Men and Culture*. Philadelphia: University of Pennsylvania Press, 783–90.

Gee, James Paul (1990), *Social Linguistics and Literacies: Ideology in Discourses* (third edition). London, New York: Routledge.
George, Alain (2007), 'Geometry of early Qur'anic manuscripts', in *Journal of Qur'anic Studies*, 9,1: 78–110.
Goody, Jack (1987), *The Interface between the Written and the Oral*. Cambridge: Cambridge University Press.
Greenberg, Joseph. H. (1986), 'Were there Egyptian Koines?', in Joshua A. Fishman, A. Tabouret-Keller, M. Clyne, B. Krishnamurti and M. Abdulaziz (eds), *The Fergusonian Impact: In Honor of Charles A. Ferguson on the Occasion of His 65th Birthday*. Vol. 1: *From Phonology to Society*. Berlin: Mouton de Gruyter, 271–290.
Grenoble, Lenore A., and Lindsay J. Whaley (eds) (2006), *Saving Languages: An Introduction to Language Revitalization*. Cambridge: Cambridge University Press.
Cromwell, Jennifer, and Eitan Grossman (eds) (2018), *Scribal Repertoires in Egypt from the New Kingdom to the Early Islamic Period*. Oxford: Oxford University Press.
Grossman, Eitan, and Jennifer Cromwell (2018), 'Scribes, repertoires, and variation' in Jennifer Cromwell and Eitan Grossman (eds), *Scribal Repertoires in Egypt from the New Kingdom to the Early Islamic Period*. Oxford: Oxford University Press, 1–19.
Gumbert, Johan Peter (2010), 'On too many formats', MS (working paper of the Encyclopaedia of Manuscript Cultures, CSMC).
Hernández-Campoy, Juan Manuel and Juan Camilo Conde-Silvestre (eds) (2012), *The Handbook of Historical Sociolinguistics* (Blackwell Handbooks in Linguistics). Malden, MA/Oxford: Wiley-Blackwell.
Haugen, Einar (1966a), *Language Conflict and Language Planning: The Case of Modern Norwegian*. Cambridge, MA: Harvard University Press.
Haugen, Einar (1966b), 'Dialect, language, nation', in *American Anthropologist*, 68 (6): 922–935.
Haugen, Einar (1972) [1968], 'The Scandinavian languages as cultural artefacts', in *The Ecology of Language,* selected and introduced by A. S. Dil. Stanford: Stanford University Press, 265–286.
Hoekstra, Eric (2003), 'Frisian', in Anna Deumert and Wim Vandenbussche (eds) (2003), *Germanic Standardisation. Past to Present.* Amsterdam, Philadelphia: John Benjamins, 193–210.
Jespersen, Otto (1925), *Mankind, Nation and Individual from a Linguistic Point of View*. Oslo: H. Aschehoug & Co.
Joseph, John E. (1987), *Eloquence and Power. The Rise of Language Standards and Standard Languages.* Oxford, New York: Basil Blackwell.
Juckel, Andreas (2001), 'Masora', in Brock, Sebastian P. Aaron Michael Butts, George Anton Kiraz and Lucas van Rompay (eds), *Gorgias Encyclopedic Dictionary of Syriac Heritage*. Piscataway, NJ: Gorgias Press, 276–279.
Juffermans, Kasper, Yonas M. Asfaha and Ashraf K. Abdelhay (eds) (2014), *African Literacies: Ideologies, Scripts, Education*. Newcastle-upon-Tyne: Cambridge Scholars.
Kaye, Alan S. (1996), 'Adaptations of Arabic Script', in Daniels, Peter T. and William Bright (eds), *The World's Writing Systems*. New York: Oxford University Press, 743–762.
Kaye, Alan S. (2006), 'Arabic Alphabet for other Languages', in Versteegh et al. (eds), *Encyclopaedia of Arabic Language and Linguistics, Vol. 1*, Leiden, Boston: Brill, 133–136.
Labov, William (1994), *Principles of Linguistic Change*. Volume 1: Internal Factors. Oxford: Blackwell.

LAEME – *A Linguistic Atlas of Early Middle English*: http://www.lel.ed.ac.uk/ihd/laeme2/laeme2.html (accessed on 18 March 2018).

Lane, Pia, James Costa and Haley De Korne (eds) (2017), *Standardizing Minority Languages: Competing Ideologies of Authority and Authenticity in the Global Periphery* (Routledge Critical Studies in Multilingualism). London: Routledge.

Langer, Nils (2003), 'Low German', in Anna Deumert and Wim Vandenbussche (eds) (2003), *Germanic Standardisation. Past to Present*. Amsterdam, Philadelphia: John Benjamins, 281–302.

Langer, Nils, Steffan Davies and Wim Vandenbussche (eds) (2012), *Language and History, Linguistics and Historiography: Interdisciplinary Approaches*. Oxford: Peter Lang.

Lillis, Theresa (2013), *The Sociolinguistics of Writing* (Edinburgh Sociolinguistics). Edinburgh: Edinburgh University Press.

Lillis, Theresa and Carolyn McKinney (2013), 'The Sociolinguistics of Writing in a Global Context: Objects, Lenses, Consequences', in *Journal of Sociolinguistics*, 17, 4: 415–439.

Linn, Andrew and Nicola McLelland (eds) (2002), *Standardisation: Studies from the Germanic Languages*. Amsterdam: John Benjamins.

Lüpke, Friederike (2011), 'Orthography development', in Peter K. Austin and Julia Sallabank (eds), *Handbook of Endangered Languages*. Cambridge: Cambridge University Press, 312–336.

Mattheier, Klaus J. (1988), 'Nationalsprachenentwicklung, Sprachstandardisierung und historische Soziolinguistik', in Klaus J. Mattheier (ed.), *Standardisierungsentwicklungen in europäischen Nationalsprachen: Romania, Germania* (Sociolinguistica. Internationales Jahrbuch für Europaïsche Soziolinguistik, 2), Tübingen: Niemeyer, 1–9.

Mattheier, Klaus J. (2003), 'German', in Anna Deumert and Wim Vandenbussche (eds) (2003), *Germanic Standardisation. Past to Present*. Amsterdam, Philadelphia: John Benjamins, 211–244.

Mesthrie, Rajend, Joan Swann, Anna Deumert and William L. Leap (2013), *Introducing Sociolinguistics* (second edition). Amsterdam: Benjamins.

Milroy, Lesley (1999), 'Standard English and Language Ideology in Britain and the United States', in T. Bex and R. J. Watts (eds), *Standard English. The Widening Debate*, London: Routledge, 173–206.

Milroy, James and Lesley Milroy (1985), 'Linguistic change, social network and speaker innovation', in *Linguistics*, 21: 339–384.

Milroy, Lesley and James Milroy (1992), 'Social network and social class: Toward an integrated sociolinguistic model', in *Language in Society*, 21: 1–26.

Milroy, Lesley and James Milroy (1999), *Authority in Language: Investigating Standard English*. London: Routledge.

Morton Alexander H. (2012) 'The Transmission of Persian Texts Compared to the Case of Classical Latin', in Brian Spooner and William L. Hanaway (eds), *Literacy in the Persianate World: Writing and the Social Order*, Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology, 143–158.

Mumin, Meikal and Kees Versteegh (eds), T*he Arabic Script in Africa: Studies in the Use of a Writing System*. Leiden, Boston: Brill.

Naim, C. Mohammed (1971), 'Arabic Orthography and some Non-Semitic Languages', in Girdhari L. Tikku (ed.), *Islam and Its Cultural Divergence: Studies in Honor of Gustave E. von Grunebaum*, Urbana: University of Illinois Press, 113–144.

Ogorodnikova, Darya (2017), ''Ajamī Annotations in Multilingual Manuscripts from Mande-Speaking Areas: Visual and Linguistic Features', in *Islamic Africa*, 8, 1–2: 111–143.
Perry, John R. (2012), 'New Persian: Expansion, Standardisation, and Inclusivity', in Brian Spooner and William L. Hanaway (eds) (2012), *Literacy in the Persianate World: Writing and the Social Order*, Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology, 70–94.
Pollock, Sheldon, Benjamin A. Elman and Ku-ming Kevin Chang (eds) (2015), *World Philology*. Cambridge, MA: Harvard University Press.
Read, Charles, and Rebecca Treiman (2013), 'Children's invented spelling: What we have learned in forty years', in M. Piattelli-Palmarini and R. C. Berwick (eds), *Rich Languages from Poor Inputs*, New York: Oxford University Press, 197–211.
Romaine, Suzanne (ed.) (1982), *Sociolinguistic Variation in Speech Communities*. London: Edward Arnold.
Romain, Suzanne (2000), *Language in Society: An Introduction to Sociolinguistics* (second edition). Oxford: Oxford University Press.
Rutkowska, Hanna, and Paul Rössler (2012), 'Orthographic Variables', in Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre (eds), *The Handbook of Historical Sociolinguistics. Blackwell Handbooks in Linguistics*. Malden, MA, Oxford: Wiley-Blackwell, 213–36.
Schieffman, Harold (1998), 'Standardization or Restandardization: The Case for "Standard" Spoken Tamil', in *Language in Society,* 27: 359–85.
Scribner, Sylvia and Michael Cole (1981), *The Psychology of Literacy*, Cambridge, MA, London: Harvard University Press.
Sebba, Mark (2007), *Spelling and Society: The Culture and Politics of Orthography around the World*. Cambridge: Cambridge University Press.
Sebba, Mark (2009), 'Sociolinguistic Approaches to Writing Systems Research', in *Writing Systems Research* 1, 1: 35–49.
Sobieroj, Florian (2016), *Variance in Arabic Manuscripts: Arabic Didactic Poems from the Eleventh to the Seventh Centuries – Analysis of Textual Variance and Its Control in the Manuscripts* (Studies in Manuscript Cultures, 5). Berlin, Boston: De Gruyter.
Souag, Lameen (2010), 'Ajami in West Africa', in *Afrikanistik Aegyptologie Online* https://www.afrikanistik-aegyptologie-online.de/archiv/2010/2957 (last accessed on 29/03/18).
Spooner, Brian, and William L. Hanaway (2012a), 'Introduction: Persian as Koine: Written Persian in World-historical Perspective', in Brian Spooner and William L. Hanaway (eds), *Literacy in the Persianate World: Writing and the Social Order*, Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology, 1–68.
Spooner, Brian, and William L. Hanaway (eds) (2012b), *Literacy in the Persianate World: Writing and the Social Order*. Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology.
Stenroos, Merja (2013), 'Identity and intelligibility in Late Middle English scribal transmission: local dialects as an active choice in fifteenth-century texts', in Esther-Miriam Wagner, Ben Outhwaite and Bettina Beinhoff (eds) (2013), *Scribes as Agents of Language Change* (Studies in language change). Boston, Berlin: De Gruyter-Mouton. 159–181.
Stenroos, Merja (2018), 'From Scribal Repertoire to Text Community: the Challenge of Variable Writing System', in Jennifer Cromwell and Eitan Grossman (eds), *Scribal Repertoires in Egypt from the New Kingdom to the Early Islamic Period*. Oxford: Oxford University Press. 20–40.

Street, Brian V. (1984), *Literacy in Theory and Practice* (Cambridge Studies in Oral and Literate Culture). Cambridge: Cambridge University Press.
Street, Brian V. (1993), *Cross-cultural Approaches to Literacy*. Cambridge: Cambridge University Press.
Stubbs, Michael (1980), *Language and Literacy: The Sociolinguistics of Reading and Writing*. London: Routledge.
Trudgill, Peter (1979), 'Standard and Non-Standard Dialects of English in the United Kingdom: Problems and Policies', in *International Journal of the Sociology of Language*, 21: 9–24.
Tamari, Tal, and Dmitry Bondarev (2013), 'Introduction and Annotated Bibliography', in *Journal of Qur'anic Studies*, 15(3): 1–55.
Vandenbussche, Wim (2002), 'Dutch orthography in lower, middle and upper-class documents in 19th-century Flanders', in Andrew Linn and Nicola McLelland (eds), *Standardization. Studies from the Germanic languages*. Amsterdam: John Benjamins, 27–42.
Van der Horst, Joop (2018), 'The End of the Standard Language: The Rise and Fall of a European Language Culture', in Constanze Weth and Kasper Juffermans (eds), *The Tyranny of Writing: Ideologies of the Written Word* (Advances in Sociolinguistics). London: Bloomsbury, 47–61.
Vázquez, Nila, and Teresa Marqués-Aguado (2012), 'Editing the Medieval Manuscript in its Social Context', in Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre (eds), *The Handbook of Historical Sociolinguistics* (Blackwell Handbooks in Linguistics). Malden, MA, Oxford: Wiley-Blackwell, 123–139.
Versteegh, Kees (2001), *The Arabic Language*. Edinburgh: Edinburgh University Press.
Wagner, Esther-Miriam (2010), *Linguistic Variety of Judaeo-Arabic in Letters from the Cairo Genizah*. Leiden: Brill.
Wagner, Esther-Miriam (2013), 'Challenges of Multiglossia: the emergence of substandard Judaeo-Arabic registers', in Esther-Miriam Wagner, Ben Outhwaite and Bettina Beinhoff (eds), *Scribes as Agents of Language Change* (Studies in language change). Boston, Berlin: De Gruyter-Mouton, 259–273.
Wagner, Esther-Miriam, Ben Outhwaite and Bettina Beinhoff (2013a), 'Scribes and language change', in Esther-Miriam Wagner, Ben Outhwaite and Bettina Beinhoff (eds), *Scribes as Agents of Language Change* (Studies in language change). Boston, Berlin: De Gruyter-Mouton, 1–18.
Wagner, Esther-Miriam, Ben Outhwaite and Bettina Beinhoff (eds) (2013b), *Scribes as Agents of Language Change* (Studies in language change). Boston, Berlin: De Gruyter-Mouton.
Wardhaugh, Roland (2010), *An Introduction to Sociolinguistics* (sixth edition). Oxford: Blackwell.
Weinreich, Uriel, William Labov and Marvin I. Herzog (1968), 'Empirical foundations for a theory of language change', in Winfred P. Lehmann and Yakov Malkiel (eds), *Directions for Historical Linguistics. A symposium*. Austin, London: University of Texas Press, 95–195.
Weth, Constanze, and Kasper Juffermans (eds) (2018), The *Tyranny or Writing: Ideologies of the Written Word.* (Advances in Sociolinguistics). London: Bloomsbury.
Wetzel, Linda (2006), 'Type versus Token', in Keith Brown (ed.), *Encyclopedia of Languages & Linguistics, Vol. 14* (second edition). Amsterdam: Elsevier, 199–202.
Wright, Roger (2006), 'Spanish', in Keith Brown (ed.), *Encyclopedia of Languages & Linguistics, Vol. 12* (second edition). Amsterdam: Elsevier, 608–610.
Williams, Raymond (1981), *Culture*. Glasgow: William Collins & Sons.
Williams, Raymond (1983), *Keywords: A Vocabulary of Culture and Society*. New York: Oxford University Press.
Zack, Liesbeth, and Arie Schippers (eds) (2012), *Middle Arabic and Mixed Arabic. Diachrony and Synchrony*. Leiden, Boston: Brill.

Paola Orsatti

Persian Language in Arabic Script: The Formation of the Orthographic Standard and the Different Graphic Traditions of Iran in the First Centuries of the Islamic Era

Abstract: This paper offers a critical review of the orthography of the most ancient original New Persian texts written in the Arabic (11th century), as well as in other scripts (Hebrew, Syriac and Manichaean), to discover any indirect evidence about the beginnings and the formative period of Arabo-Persian orthography. On the basis of the New Persian original documents in non-Arabic scripts here examined, dating to an earlier period than documents in the Arabic script, we can tentatively date and localise the beginning of the Arabo-Persian orthographic influence on the other written traditions of Iran: northeastern Iran, end of the 9th– beginning of the 10th centuries. By then, Arabo-Persian orthography appears as already fixed. Though it cannot be excluded that some scattered and unsystematic attempts at adaptation of the Arabic script to Persian were accomplished here and there in different places of Iran, the hypothesis of a multi-centric origin of the Arabo-Persian orthographic canon seems less probable in the light of our documentation.

1 Introduction

The aim of this study is to carry out a critical review of the orthography of the most ancient original New Persian texts written in the Arabic, as well as in other scripts (Hebrew, Syriac and Manichaean), to discover any indirect evidence about the beginnings and the formative period of Arabo-Persian orthography.[1] By New Persian, or simply Persian, I am referring, in a broad sense, to the Persian language of the Islamic period.

Though agreeing with Giorgio Banti's criticism expressed during the Hamburg conference in 2013 about the use of the term 'standard' for pre-modern times, I retain it to underline the fact that New Persian orthography in Arabic script seems to follow a well-established norm from the time of its most ancient original attestations.

1 A general study of the most ancient New Persian documents in alphabets different from Arabic is given by Orsatti (2007, 107–172, with reference to previous studies), and – from a historico-cultural perspective – by Orsatti (2008). For the layout of the page in Islamic manuscripts

https://doi.org/10.1515/9783110639063-002

2 The adaption of the Arabic script to New Persian

As is well-known, towards the mid-7th century western Iran, the heart of the Sasanid empire, was conquered by the Arabs. This historical event had many important consequences: most Iranians, more or less gradually, gave up their ancient Zoroastrian religion and converted to Islam; and – given the strong link existing between writing and religion – they abandoned the ancient Pahlavi script, in which what we call Middle Persian was written, and adopted the Arabic alphabet to write their language.

The first original written records of Persian in Arabic script (I will speak of 'Arabo-Persian'),[2] both literary and not, go back to the 11th century: therefore they are not really so ancient. We do not know when, where and for what purposes (administration, literature, private documents etc.) the Arabic script was first adapted to write the Persian language. And we do not know in which way and through what stages Arabo-Persian orthography was fixed. From the extant documentation we get the impression that an orthographical norm was well-established from the very beginning. On the other hand, at least for literary texts which have undergone a long manuscript transmission, we do not know to what extent the normalising intervention of copyists may have contributed to suggest the existence of an early orthographical norm.

Historical sources can help find out when and for what purposes New Persian emerged as a written language, prevailing over Middle Persian, written in the Pahlavi script, and over Arabic, the language of religion and science. In the first centuries of Islam the prestige of Middle Persian was still strong, especially in south-western Iran. From a piece of information provided by the Arab historian Balāḏurī (9th century) we know that, in western Iran, Middle Persian in Pahlavi script was used for administration until the end of the 7th or the beginning of the 8th century, and in eastern Iran even longer, before being substituted by Arabic.[3] For administration, therefore, Middle Persian was replaced directly by Arabic, and only later was New Persian used as the language of administration. The same can be said of epigraphy (Bivar 1986): for inscriptions, Middle Persian in Pahlavi

---

and its relationship to the text copied, manuscript typology and destination, I again refer to some observations in Orsatti (1993, 281–282, 319–323). On the historical orthography of New Persian texts in Arabic scripts, apart from scattered remarks on the orthography of manuscripts in the introductions to the editions of texts, see Širvānī 1974 and Hashabeiky 2005, 79–85 (remarks based on an early study by Jalāl Matīnī).

2 The term 'Muslim Persian' is less appropriate because different minorities in Islamic Iran, especially Christians, also commonly used the Arabic alphabet (see below).

3 See Xānlarī 1986, I, 307-314.

script was gradually replaced by Arabic, written in the so-called Kufic script, one of the most ancient inscriptions in New Persian being the Kufic inscription in Persian verses in the palace of Masʿūd III at Ghazna (Afghanistan), dated 505/1111 (published by Bombaci 1966).

What is sure is that, when in the 9th century, during the period often referred to as the 'Iranian renaissance', New Persian lyric poetry of the literary type – i.e. patterned after Arabic poetry – came into being in the Persian courts of eastern Iran, it could not be written down other than in the Arabic script. Indeed, New Persian lyric poetry arose as an experiment in the Persian courts, by then independent or semi-independent from the caliphate of Baghdad; an experiment that consisted of substituting Persian for Arabic within the pattern of Arabic poetry.[4] We can suppose that the establishment of New Persian orthography was a part of this experiment.

The anonymous author of a local chronicle, the *Tārīkh-i Sīstān* 'History of Sistan' (11th century, with later additions), recounts an interesting story about what he presents as the first Persian poem of the Islamic era, which – he says – was told for Yaʿqūb b. Layṯ of the Saffarid dynasty of Sistan (south-eastern Iran). In 251/865 Yaʿqūb gained an important victory over one of his enemies; on this occasion poets recited poems in Arabic in front of him, extolling his victory. Yaʿqūb, who was a man of arms and not a learned person, said then the famous phrase: 'Why should a poem be composed which I cannot understand?' Muḥammad b. Waṣīf, his correspondence secretary, was in the audience; he had the idea of composing a poem in Persian, of which six lines, full however of Arabic words and expressions, are quoted (*Tārikh-e Sistān*: 166–167).

The story of Yaʿqūb the Saffarid and his secretary and court poet Muḥammad b. Waṣīf is an important piece of information about when, where and in which way the new Arabicised poetry in the 'vulgar tongue' of the time, that is New Persian, came into being: during the 9th century, in the Persian courts which flourished in eastern and north-eastern Iran.[5] Prose began later: the first prose texts in literary New Persian, in Arabic script, go back to the mid-10th century – the most ancient dated New Persian manuscripts of literary works in Arabic script, however, belong to the middle of the following century.

4 On these themes cf. in particular Bausani 1960, 307–311 and the studies by Lazard 1971 and Lazard 1975.

5 For an evaluation of this and similar accounts concerning the birth of New Persian literary poetry, cf. in particular Lazard 1975, 607–610.

2.1 Persian in Arabic script: general characters

The following table shows the most important characteristics of the adaptation of the Arabic alphabet to the Persian language (mid-11th – mid-13th century):

A. Conservation of the 'only Arabic' letters, i.e. letters found only in Arabic loanwords:

Arabic letters	Arabic value	Phonetic value in Persian
ث <ṯ>	θ	s
Arabic ذ <ḏ>	δ	z (perhaps pronounced δ in Early New Persian)
ح <ḥ>	voiceless pharyngeal fricative ħ	h
ص <ṣ>	pharyngealised s	s
ض <ḍ>	pharyngealised δ or d	z
ط <ṭ>	pharyngealised t	t
ظ <ẓ>	pharyngealised z or δ	z
ق <q>	q	introduction of the new phoneme /q/, in later times merged into /ɣ/ <ġ>, but retained till now in Afghani Persian
ع <ʿ>	voiced pharyngeal fricative ʕ	not pronounced or pronounced as a glottal stop (merging with the Arabic *hamza* [ʔ])

B. Introduction of new letters, mainly by addition of diacritical points

Letters	Phonetic value	Commentary
پ	p	already introduced, as the three other 'Persian' letters (see below), in the most ancient manuscripts, especially of fine workmanship, but normally seldom used
چ	č	see above
ژ	ž	see above
گ	g	see above
ڤ <b̲>	β	spirant allophone of /b/. This letter is used only in the most ancient manuscripts
'Persian' ذ <ḏ>	δ	spirant allophone of post-vocalic /d/, used in manuscripts until the middle of the 13th century

C. One digraph, which represents the phoneme /xw/ of Early New Persian (this being the only case of historic orthography in the New Persian writing system, conserved till now):

خو	xw	later merged with /x/ with dropping of the labial articulation of the ancient phoneme

One of the main differences between the Arabo-Persian writing system and the Arabic one is the regular spelling of final short *-a*, which in Arabic is not written, by means of <h>. This is well explained by a passage from the *Muʿjam* by Šams-i Qays (first half of the 13th century), where the great Persian philologist says that: 'In Persian orthography (*dar xaṭṭ-i pārsī*), whenever a word ends in *-a*, you have to add a *hā* to it [...] These *hā*s in the Arabic language are clearly pronounced [i.e. in their value as *h*] [...] Instead in Persian (*pārsī*) they are by no means pronounced except, by poetic license, if they are in rhyme position; (only) in this case they are counted as a quiescent letter and are feebly articulated' (Šams-i Qays 1981, 243–244).

2.2 Written and not-written morphemes

An important characteristic feature of Arabo-Persian orthography is the way two important and very frequent morphemes of the Persian language are represented: the *iḍāfa*, that is the particle *-i* of the *status constructus*, and the coordinative conjunction *u* (English 'and'). Both consisted in a (probably) short vowel, and therefore – according to the Arabic writing system – should not have been written; but, while the first one – in contrast with the pre-Islamic graphic traditions of Iran – is not represented in the Arabo-Persian orthography, the second one is written <w> as a stand-alone word.

The *iḍāfa* particle is an enclitic *-i* vowel, originally long but probably shortened from the beginning of the New Persian period, which follows the head of a nominal phrase: *kitāb-i muʿallim* 'the book of the teacher', *kitāb-i buzurg* 'the big book'. In Arabo-Persian orthography this short *-i* is not written, except after words ending in a vowel: only in this case is the *iḍāfa* represented as <y> (the letter *yā*) attached to the preceding word – what is represented in such cases is the glide *y* which develops between the final vowel of the word and the *iḍāfa* vowel: *banda-yi tu* 'your slave'. In the most ancient Persian manuscripts this orthographic rule is always observed, apart from very rare cases in which the *iḍāfa* is represented by <y> also after a word ending in a consonant (Lazard 1963, 200, § 162). A couple of such cases, pointed out by Minorski (1942, 188), seem to be present also in one of

the most ancient non-literary Arabo-Persian documents, the deed for the sale of land found in Khotan (Central Asia) (line 2), dated 501/1107. Though it is generally admitted that in New Persian the vowel representing the *iḍāfa* had already been shortened, these occasional spellings, as well as the fact that in poetry the *iḍāfa* can count as a short or as a long vowel, seem to point to the presence – in Early New Persian – of a long variant of the *iḍāfa* (Meier 1981, 131–132).[6]

An ancient spelling of the *iḍāfa* particle by means of an isolated or proclitic *alif*, attested in the quotation of Persian phrases in works of some Iranian authors (Abū Nuwās, Hamza al-Iṣbahānī, the author of the *Ta'rīx-i Qumm*) of the 9th–10th centuries writing in Arabic, may represent an ancient spelling, soon fallen into disuse, preceding the establishment of Arabo-Persian orthography (Henning 1958, 88–89); this is even more relevant in the light of the fact that the authors in which this spelling is to be found are from central and western Iran.

A group of fifteen Judaeo-Persian tombstones edited by Gnoli (1964) found in the Ghūr (Afghanistan) and dated – according to the Seleucid era – to the years between the second half of the 12th till the beginning of the 13th century CE, show the gradual falling into disuse of the writing of the *iḍāfa*, perhaps due to the influx of coeval Arabo-Persian orthography, and possibly also as a consequence of the disappearance of the long *iḍāfa*. In the initial formula *wafāt-i 'adan* 'Edenic death', the *iḍāfa* is written as <-y> joined to the preceding word in tombstones IV and VIII (dated 1484 Seleucid/1172 and 1502/1190 respectively), and is not written in tombstones III (also dated 1484 /1172), V, VI, VII, IX, X, XI (this being the most recent one, dated 1526/1214), and XIV.

Besides the *iḍāfa* particle – which, being a (short) vowel, very soon disappeared from orthography though remaining well alive in the language – other morphemes are not represented in the Arabo-Persian orthography. One of them, consisting in an originally long *ō* afterwards shortened, is the directional Middle Persian preposition *ō* 'towards, to', well-known for the New Persian linguistic period from the most ancient Judaeo-Persian documents originating from south-western Iran (Lazard 2009; see below, Group C). Its disappearance from Arabo-Persian orthography was perhaps due to the fact that this morpheme was considered too dialectal, colloquial or archaic to be allowed into the writing; or, rather, its disappearance may well correspond to its gradual falling into disuse from the north-eastern variety of Persian which was at the basis of literary New

6 See also an inverse spelling in a New Persian fragment in Manichaean script, pointed out by Provasi (2011, 162, 164): in fragment M 595a+, *dāng*, 'a small coin', followed by *iḍāfa* is written <d'ngyh>, with the abstract suffix *-ī* (spelled -yh) representing a probably long *iḍāfa*.

Persian, the so-called *darī* or *pārsī-yi darī*.[7] Indeed, in the extant New Persian documentation in Manichaean script, mainly originating from Samarkand and eastern Iran between the 10th and 11th centuries, there is no trace of it; and it is rarely attested in the ancient Judaeo-Persian texts from north-eastern Iran (Paul 2003, 183; Shaked 2009, 452). It is possible therefore that the preposition *ō*, shortened in New Persian and not represented in Arabo-Persian orthography, was gradually replaced in common usage by other more fortunate prepositions, as the polysemous *ba* 'in, to, on, by, with, for'; and that it was then expunged from literary New Persian, though surviving in many Persian dialects especially in the western part of the Iranian plateau (Browne 1895; Filippone 2011, 198). However, this preposition seems to have been transmitted as a legacy to the common language based on literary New Persian. An indirect trace of it – no longer written nor pronounced – can be detected in a number of cases, typical of the contemporary spoken language (ex. *miram xune* 'I am going home', instead of *be xāne miravam*), where a preposition is lacking or has fallen away (Lazard 1986, 252; Lazard 1990, 189).

The same can be said of the use of the *iḍāfa* particle as a relative pronoun, well-known from New Persian texts in scripts different from Arabic. In this case too, more modern forms replaced the outcome of the old Middle Persian relative pronoun *ī*; but scattered traces of it can be detected in literary Early New Persian texts (Lazard 1963, 490–491, §§ 855–856), as well as in the modern spoken variety (Lazard 1990, 188–189).

Arabo-Persian orthography does not distinguish between some originally different morphemes: the Middle Persian conjunction *kū* 'that, than', the relative/interrogative pronoun *kē* 'who, which', and the temporal/conditional conjunction *ka* 'when, if', all three merged in what seems to be a single new form *ki*, spelled <ky> or <kh>, or simply <k> joined to a preceding pronoun (<'nk> *ān-ki* 'that which') or to a following word beginning with a vowel (<k'md> *k-āmad* 'who came') in early Arabo-Persian orthography.[8] These morphemes were still distinct

7 For the meaning of these glottonyms, and the distinction between *pārsī* and *darī*, see the studies by Gilbert Lazard reprinted in the book *La formation de la langue persane*, Paris 1995. Roughly speaking, *pārsī* means generically 'the Persian language', while *darī*, '(language) of the court', indicates – at the beginning of the Islamic era – the north-eastern variety from which literary New Persian sprang.

8 Šams-i Qays (1981, 249) says that both spellings, with final <y> or <h>, are only a graphic device intended to represent a preceding short *i* vowel (*kasra-yi mā-qabl*). When *ki* is used with an interrogative and abstract (*istifhām-i mujarrad*) value, as in the expression *ō kī-st?* 'Who is he?', however, Šams-i Qays recommends the spelling with final <y> which, he says, 'is well perceptible in pronunciation too'. Perhaps he meant that, when this form had the value of an interrogative

– at least graphically – in some of the most ancient Judaeo-Persian texts, but often appear to be used interchangeably as far as their function is concerned; and *ka* 'when, if', spelled <k'>, has been retained only in the text of religious controversy referred to as Argument, published by MacKenzie in 1968 (Paul 2013, 151–152, § 185c, and 168–169, § 207). Manichaean New Persian too, despite its adherence to the graphic tradition of Middle Iranian languages, especially Sogdian and Middle Persian, in Manichaean script, shows some traces of confusion (see de Blois 2006, 106, *s.v.* k'; Provasi 2011, 166-167, *s.vv.* k', kw, ky). This means that – during the formative period of Arabo-Persian orthography – such morphemes were losing, or had already lost, their distinct meaning due to a possible formal coalescence caused by the shortening of the vowels (*kū* > *ku*, *kē* > *ki/ke*). This probably helped their merging into one single form.

Some of the earliest Judaeo-Persian documents – from both south-western Iran (as for example *Argument*; but not the two dated documents referred to as Group C below) and Central Asia (for example the two letters from Dandan Uiliq in Central Asia: Group B below) – give evidence of the existence of two different prepositions, corresponding to the sole Arabo-Persian preposition *ba(d)* written <b-> (or <bd-> before a vowel):[9] *pa(d)* from Middle Persian *pad* 'to, at, in, on', spelled <pd>, <p'>, or <p-> joined to the following word; and <by>, to be probably read *bē*, with a directional value 'to, towards'. These two prepositions are also attested – respectively spelled <b-> and <by> and with the same distinct meaning – in an interlinear translation of the Qur'an, in Arabic script, written in a particular variety of New Persian rich in dialectal features attributable to Sistan (south-eastern Iran). The manuscript of this text, referred to as *Qur'ān-i Quds* and dated by Lazard to the end of the 11th century (1990, 188, 192), has been edited, in facsimile and in a diplomatic edition, by Rivāqī (1985).[10]

As to the genesis of the Arabo-Persian preposition *ba(d)*, apart from a possible coalescence between the two prepositions *pa(d)* and *bē* mainly evidenced by *pa(d)* also acquiring a directional value,[11] contamination with the almost synonymous Arabic preposition *bi-* 'with, in, by, at, near' can also be cited, at least as

pronoun, it was still articulated with a long *ē* or *ī*, as it is today. On the form of this/these morpheme(s) in the most ancient New Persian prose texts in Arabic script see Lazard 1963, 237, § 253 (pronoun); 473, § 809 (conjunction).

9 Lazard 1963, § 145, p. 191 shows that vocalised manuscripts also attest an occasional *bi* or *bu* (in labial context) pronunciation.

10 A clear study of the repartition of the two prepositions' usages has been given by Lazard (1986, 245–247 and 1990, 187–188).

11 Paul (2003, 179–185, and especially the table on p. 185) shows that it is only in New Persian that ancient *pa(d)* also acquired a directional meaning.

far as its spelling with initial <b-> is concerned (de Blois 2006, 109, n. 8).[12] Indeed, the Arabo-Persian preposition is spelled with initial <b-> – instead of <p-> attested by coeval texts in scripts different from Arabic – even in the most ancient manuscripts which make at least occasional use of the four 'Persian' letters. The Manichaean New Persian fragment M 595a+ edited by Provasi shows a curious inverse spelling for the verbal prefix *bi-*, written <pd> as the preposition (Provasi 2011, 161–162, 166). This might indicate that the rather inaccurate scribe of this fragment confused the two morphemes, perceiving them as homophonous (at least as regards their initial consonant); or that he was influenced by the Arabo-Persian orthography, in which the verbal prefix and the preposition were both spelled <b->.

2.3 The rule of *dāl* and *ḏāl*

As seen above, a clear normalizing aim seems to have presided over the establishment and development of Arabo-Persian orthography. This ended up with the exclusion of all forms considered too dialectal or colloquial from writing, and therefore from literary New Persian; and with the simplification of a number of forms which had lost their distinct original form or meaning.

Among the normalizing choices connected with the development of early Arabo-Persian orthography there is – to my mind – the so-called 'rule of *dāl* and *ḏāl*' well-known from the most ancient literary manuscripts. According to this rule, all *d*s after a vowel or a diphthong, within the same morpheme, were spelled <ḏ>, i.e. as an interdental voiced fricative (ex. <pḏr> *piδar* 'father'); instead, after a consonant or at the beginning of a word (ex. <drd> *dard* 'ache') or after morpheme boundary (ex. <bd'n> *bi-dān* 'know'), they continued to be spelled <d>. The complementary distribution of *d* and *δ* implied by this rule concerned only the words of Persian (or Iranian) origin and the loanwords, such as those from Greek, entered into Persian at an ancient date. By contrast, the development of *δ* from postvocalic *d* did not concern Arabic loanwords, probably because in Arabic /d/ and /ḏ/ were (and are) two distinct phonemes. As manuscripts follow this orthographic usage approximately until the middle of the 13th century, scholars have supposed that, around that time, the complementary distribution of the outcome of historical /d/ in Persian words, as *d* or as *δ*, came to a halt and all

12 Martin Schwartz (personal communication) suggested another possible explanation for *pa(d)* > *ba(d)* in New Persian: alignment with the initial *b-* of many other prepositions (*abar*, *abāg*, *abē*, *abāz*), that by then had lost their initial vowel.

'Persian' *δ*s (but not Arabic ones) became *d* again, as they are today, with the exception of a small number of words which retained *δ* – afterwards pronounced *z* – from historical *d*: *guδaštan* 'to pass', *guδāštan* 'to put', *paδīruftan* 'to receive', and a few others.

To explain this change in the manuscript tradition, Pisowicz (1985, 109) formulated another hypothesis, namely that from the second half of the 13th century the allophonic development of the ancient postvocalic *d* to *δ* – which he considered characteristic of central and western dialects – was ousted from the literary language due to the influence of north-eastern dialects, where this phenomenon had not occurred. Indeed, in a famous passage from the *Muʿǰam*, Šams-i Qays says that 'in the language of the people of Ghaznīn, Balkh, and Transoxiana there are no *ḏāl*s, all of them being pronounced as the letter without points [i.e. *dāl*]' (Šams-i Qays 1981, 221); this – according to Šams-i Qays – would explain a number of irregular rhymes, which do not comply with the distinction between *dāl* and *ḏāl*, in the works of poets coming from these regions. Meier (1981, 104), however, had already shown that the spirantisation of postvocalic *d* was not unknown to the north-eastern dialects, considered as a whole; according to him the boundary between the areas affected and not affected by this phenomenon passed between Marv in Khorasan (spirantisation), and Bukhara in Transoxiana (without spirantisation). Pisowicz's hypothesis appears problematic also because the influence of north-eastern dialects on the literary language was certainly relevant from the very beginning of the history of literary New Persian; therefore it is unlikely that such an influence had had consequences only in the 13th century.[13] When Šams-i Qays says that 'in the language of the people of Ghaznīn, Balkh, and Transoxiana [i.e. Afghanistan and Central Asia] there are no *ḏāl*s', he is simply noting that not all north-eastern dialects have a spirant allophone of /d/, this causing some difficulties for the poets coming from these regions; but he is not stating anything about the complementary distribution of *d* and *δ* implied by the orthography of the most ancient manuscripts.

Recently Filippone, in the context of a study of the language of the *Qurʾān-i Quds*, where no instances of spirantisation of postvocalic /d/ in words of Persian origin are to be found, has conducted research on Persian dialects to find out the dialect or the dialects from which the complementary distribution of *d* and *δ* represented in literary ancient manuscripts could have originated. She concludes: 'In analyzing these phenomena, I think that one should avoid reconstructing

13 Recent summaries of this complicated question have been given by de Blois 2006, 94–96; Orsatti 2007, 94–98 (with translation of the relevant passage from Šams-i Qays); Filippone 2011, 185–186.

highly standardised realities, with a homogeneity in time and place which has probably never existed. The tendency to a certain (contextual or free) variability of *d/δ/z* appears as a constant factor throughout the history of West Iranian. [...] But cases of *d/z* fluctuation are mostly bound to single words. In many cases they remain unexplained' (Filippone 2013, 186). This is exactly the situation of modern standard New Persian, where *ḏāl* (pronounced *z*) from historical postvocalic /d/ is only found in a limited number of words.

At this point, a possible solution would be to consider the complementary distribution of *dāl* and *ḏāl* attested in early literary manuscripts as the result of the application of a (mainly orthographic) rule intended to set order into the multiform realisations of /d/ in the spoken or dialectal varieties of Persian. Šams-i Qays, in the previously mentioned passage, attributes the complementary distribution of *dāl* and *ḏāl* to the 'correct *darī* language', that is to literary New Persian, and not to Persian *tout court*. Throughout Chapter Two of Section Two of the *Muʿǰam,* devoted to 'The letters of the rhyme' (where – concerning the rhymes of the letter *ḏāl* – the passage in question is to be found), he distinguishes carefully the Persian language, called *pārsī* or *luġat-i pārsī*, from the *luġat-i darī*, 'the *darī* language', often qualified as *ṣaḥīḥ* 'correct'. For example, he says that in the Persian language (*pārsī*) most words end with a quiescent letter, i.e. – roughly speaking – in a consonant (Šams-i Qays 1981, 209); and that among the peculiarities of the Persian language (*luġat-i pārsī*), there is the fact that the clause is not complete without the copula (Šams-i Qays 1981, 215). When he notes that in Persian the final long *-ā* is normally shortened, he refers this linguistic notation to the colloquial form of Persian (*muḥāwarāt-i pārsī*) (Šams-i Qays 1981, 211). When the author uses the term *darī*, instead, he refers to a rule or a canon, often even complaining of the lack of a clear criterion on which to rely (Šams-i Qays 1981, 205).

That the rigorously complementary distribution of *d* and *δ* was a phenomenon only affecting the literary language is proved by the fact that it seems to be unknown in New Persian texts of non-literary character. Neither the ancient New Persian texts in Hebrew script coming from south-western Iran (see Group C below), nor the non-literary *Qurʾān-i Quds* in Arabic script, from Sistan, bear any trace of a spirant pronunciation of postvocalic /d/ (Lazard 1995, 136), and even less of a complementary distribution of *d* and *δ*. Among the New Persian texts written in scripts different from Arabic only one shows a clear complementary distribution of *d* and *δ*: it is the fragment of the bilingual (Syriac, and New Persian in Syriac script) Psalter from Central Asia (edited by Müller 1915, Sundermann 1974, and Sims-Williams 2011). Sundermann, however, ascribes the orthographic usage in this manuscript to the influence from the coeval Arabo-Persian orthography (1974, 450). The Manichaean New Persian orthography,

though having at its disposal a distinct letter <δ>, attests only occasional renderings of Persian postvocalic /d/ as *δ* in <ʾryδ> *āraδ* 'he brings' and <nbwδm> *nabūδam* 'I was not' in the fragment of the *qaṣīda* published by Henning (1962, ll. 10 and 34); and in <gwδrʾndg> *guδaranda* 'passing' in the so-called 'Catechism' (*Lehrtext*) published by Sundermann (2003, c2, where this form occurs twice, this verb being one of the few Persian words that have retained the letter *ḏāl* till now): a total of only three words.

Therefore, the complementary distribution of *d* and *δ* should probably not be considered as a genuine linguistic phenomenon originating from one or more Persian dialects and hence entering into the literary language, but as a rule – in fact, ancient authors often speak of 'rule' – that was supposed to be applied to literary New Persian (*darī*) and had its main scope and field of application in orthography and in the scholastic and artificial pronunciation characteristic of poetry (see also Section 2.5). Probably around the mid-13th century such a rule ceased to be extensively applied in the copying of literary manuscripts.

2.4 The most ancient New Persian texts in Arabic script: the Codex Vindobonensis and the marriage contract from Bāmiyān

One of the most ancient dated Persian literary manuscripts is the Codex Vindobonensis (cod. A.F. 340 of the Österreichische Nationalbibliothek in Vienna), containing a pharmacological tractate by Abū Manṣūr Muwaffaq b. 'Alī al-Hirawī (therefore from Herat in modern Afghanistan), copied by the poet Asadī of Ṭūs in Šawwāl 447 / 24 December 1055 – 21 January 1056 (facsimile editions: Muwaffaq 1972, 2009).[14] A relatively short span of time separates the composition of the work, which can be dated to the second half or end of the 4th/10th century, from this copy. Both the author and the copyist were from Khorasan, in the east of the historical Iranian territory.

The orthography of the manuscript is not too different from today's. The Arabic loanwords are spelled as in Arabic, though some orthographic usages of Arabic are not retained: a principle which has presided over Arabo-Persian orthography over the years. The *tā marbūṭa* alternates with the *tā ṭawīla* 'long *tā*', both read *-at*, or with final <h>, read *-a*. In particular, taking into consideration the second double-page of the manuscript (fols 2v–3r), the *tā marbūṭa* is used in *jihat*

14 For a critical evaluation of the manuscript for linguistic studies and a bibliography until then cf. Lazard 1963, 45–48, N. 4.

Fig. 1: Codex Vindobonensis, fols 1v–2r (from Muwaffaq 1972).

Fig. 2: Codex Vindobonensis, fols 2v–3r (from Muwaffaq 1972).

(2v 9), *saʿādat* (2v 10), *ḥaḍrat* (3r 1 and – probably with loss of the two points – 2v 5), and *ziyādat* (3r 1); the *tā ṭawīla* is found in *quvvat* (2v 1, followed by the plural suffix *-hā*), in *quvvat u maḍarrat u manfaʿat-aš* (3r 5, before the suffix pronoun *-aš*), and in *qismat* (3r 7); instead, the writing with final <h>, certainly read *-a* (in fact, it is followed by the <y> indicating the *iḍāfa* after a vowel), is found in two Arabic words: *xizāna* (3r 2) and in *daraja* (3r, 5, 8–12). *Hamza* is not written in *muʾayyad* (2v 5), *taʾammul* (2v 7) and *taʾlīf* (3r 2). The latter spelling could actually represent a linguistic feature of the language of the text: the dropping of the glottal stop with compensatory lengthening of the preceding vowel (Meier 1981, 128, 133–134).

In this manuscript, final <y> – besides its normal form(s) – also has a small form, mostly used to represent the *iḍāfa* after a word ending in *-a* (see 1v 4,5; 2r 2; 3r 2, 5, 8–12). This small *yā* can be considered as the origin of the 'Persian' *hamza*, i.e. a *hamza* placed over or by a final <h> indicating *-a*. In fact, in fol. 1v, a second hand seems to have begun to replace these 'little *yā*s' with more modern *hamza*s (ll. 4 and 5).

The manuscript is fully vocalised and provided with orthographic signs, and complies with the 'rule of *dāl* and *ḏāl*'. As to the so-called *majhūl* 'unknown' vowels, i.e. the vowels *ē* and *ō* unknown to the Arabs, a vocalisation for *ē* different from *ī* is well attested (Meier 1981, 86–87), whereas I found no examples for *ō*. A massive use is also made of the distinctive signs (additional diacritical points and little letters written above or below the main letters) intended, in addition to the usual diacritic points, to differentiate letters of the same form:[15] three points under *sīn* and *kāf* to differentiate them from *šīn* and *gāf*; one point under *rā*, *dāl*, *ṣād* and *ṭā*, to differentiate them respectively from *zā*, *ḏāl*, *ḍād* and *ẓā*; a letter underwritten, to differentiate *'ayn* from *ġayn* and *ḥā* from *xā*. Final <y> has often two subscribed points, as in Arabic writing, a graphic usage afterwards abandoned in Arabo-Persian orthography.

Among the other more relevant orthographic usages of this manuscript there are the following:

- the 'Persian' letters <p>, <č> and <g> are occasionally used (I have found no instances of <ž>), all three written with three points below, but often replaced respectively by <b>, <j>, <k>;
- letter ڤ representing the spirant allophone of /b/ occurs quite regularly in verbal forms from *aʙzūdan* 'to add' and *aʙgandan* 'to throw', and in the suffix *- ʙām* 'color';

15 Grohmann 1971, 42–46, § 4.

- the preposition *ba-* 'in, to, by' (<*pa*/*paδ*) is always written with< b> – and not with the more ancient <p> – attached to the following word, a usage continued until recent times (now discouraged in favor of the 'separate' spelling <bh>);
- *īn* 'this' and *ān* 'that' are often written, without initial *alif*, joined to a preceding preposition (*az-īn* 2v 3; *andar-īn* 3r 3) according to a usage frequent until very recent times;
- the conjunction 'that'/relative pronoun 'who, that, which' are always written <ky>;
- the verbal durative prefix has the form *hamē* and is always written as a standalone word;
- *alif-madda* is only occasionally used; initial *ā-* is often written as a simple *alif*, or as two *alif*s next to one another.

Apart from minor fluctuations continuing until recent times, this already normalised orthography is attested not only in the most ancient literary manuscripts, but also in the two most ancient private documents in Arabo-Persian: a marriage contract dated 470/1078 found in Bāmiyān (Afghanistan),[16] and a deed for the sale of land found in Khotan, dated 501/1107.[17]

Among the orthographic features of the marriage contract, written only about 20 years after the Codex Vindobonensis and originating from Afghanistan, the following should be noticed:

- the orthography of the text is not too different from modern orthography, with the exception that here diacritical points are often omitted;
- all Arabic loanwords maintain their original spelling;
- the four 'Persian' letters are not used: cf. for ex. <sbyd'r> *sapēdār* 'white poplar' (ll. 2, 3, 12) and <krftn> *giriftan* (l. 25);
- the 'rule of *dāl* and *ḏāl*' is not respected; in particular, two Persian words which even today are spelled with *ḏāl* are written with <d>: <bbdrfth> *bipadrifta* 'he has accepted' (l. 28), corresponding to literary New Persian *bipaḏīrufta* and <kwdšth> *gudašta* 'passed, elapsed' (l. 31), corresponding to literary New Persian *guḏašta*;

16 Published in Latin transcription/transliteration by Scarcia (1966). The same author gave a study of this text together with a photographic reproduction of it in a previous article (Scarcia 1963).

17 This document was first studied by Margoliouth (1903), who also offers a photographic reproduction of it. Minorsky (1942) gave a transcription and translation of the text and a new study, also correcting Margoliouth's reading of its date: not 401 but 501 of the Hegira.

- in l. 1 the preposition *bar* 'on' is written <vr> or <fr> (it is not easily readable from the photograph published by Scarcia 1963), a spelling which probably represents a dialectal form;
- the verbal prefix *mē-* (< *hamē*), which always occurs in this more modern shortened form, is written attached to the following verbal form;
- the preposition *ba* 'in, to, at', is written as <b-> attached to the following word;
- the conjunction *ki* 'that'/relative pronoun *ki* 'who, which' are both written <kh>;
- the plural suffix, always occurring in the form *-hāy*, is written attached to the preceding word, except after words written with final <y>; e.g. <sr'y h'y> *sarāyhāy* 'houses', ll. 18, 25 and *passim* (there are no plural words ending in <h>);
- *alif-madda* is never used;
- final <y> has, though very rarely, two points below, thus indicating that the 'Arabic' writing of final <y> with two points below, later abandoned, was still in use;
- the coordinative conjunction *u/wa* 'and' is not written, apart from in eight cases (ll. 24, 25, 27, 28 twice, 29, 31 twice), mostly placed at the beginning of a new sentence.

The latter feature is very interesting and represents the sole real divergence from the orthography of literary texts (and from modern orthography). It seems to suggest that in the manuscript of the marriage contract the short *u* of the conjunction was not written, except when it was at the beginning of a sentence or after a pause. We can suppose that precisely at the beginning of a sentence or after a pause the Persian conjunction *u* (< Middle Persian *u*, *ud*) begun to be pronounced *wa*, as in Arabic. These, therefore, would be early attestations of the new form, probably influenced by the Arabic conjunction *wa*, of the Persian conjunction, even now pronounced *va* or *o* according to its syntactic position and elocution speed.

Ancient literary manuscripts too – less carefully copied than the Codex Vindobonensis – show some traces of a failure to write the coordinative conjunction. For example, in the ancient fragment of 'Unṣurī's poem *Wāmiq wa ʿAẕrā* (datable to the 11th–12th centuries) the conjunction is occasionally not written and has been integrated into the edition (Hägg and Utas 2003, 79).

Considering that the two ancient manuscripts analyzed here, that is the Codex Vindobonensis and the marriage contract, pertain to different textual typologies – a beautiful copy of a scientific-literary text and a legal private document – the orthographic differences between them are not so great. The lack of

vocalisation and of any orthographic sign in the marriage contract, as well as the defective diacritical pointing of the letters, are clearly to be connected with the practical scope of the document, devoid of any aesthetic pretension, and with the formulaic character of the text, written in a highly standardised language. The comparison between the orthography of the two texts shows a high degree of normalisation from early times, at least as far as Eastern Iran and (today) Afghanistan are concerned.

In this regard one fact is particularly meaningful: in the two most ancient non-literary documents, i.e. the marriage contract from Bāmiyān and the deed for the sale of land from Khotan, the Persian word *pānṣad* 'five hundred' (Scarcia 1963, ll. 3, 22; Margoliouth 1903, l. 12) is already written with Arabic <ṣ>, as it is now – *ṣad* 'hundred' being one of the few Persian words written with this 'only Arabic' letter. The reason for this spelling is perhaps to avoid confusion with other homophonous words such as *sad(d)* 'rampart, obstruction' or – as Perry (2002) suggests – with the very common word *šud* '(he) became' which, due to the frequent defective writing of the diacritical points, could be confused with the word for 'hundred'. The fact that both non-literary documents already offer this 'normalized' spelling is a clear proof of the high degree of standardisation of Arabo-Persian orthography from ancient times.

2.5 The spelling of the Arabic loanwords

As we have seen, one of the most striking features of Arabo-Persian orthography since its beginning is the preservation of the original spelling of the Arabic words which entered into Persian, though we can suppose that – once established in the Persian language – the Arabic loanwords were pronounced, as they are today, according to Persian phonology. The only exception are the introduction of the new phoneme /q/, in Early and Classical New Persian still clearly distinguished from /ɣ/ <ġ> (Pisowicz 1985, 111–117) and, perhaps, the pronunciation of Arabic <ḏ>: indeed, given the existence, in words of Iranian origin, of a fricative postvocalic allophone of historical /d/, the letter *ḏāl* could well have been pronounced, in Arabic loans, as an interdental voiced fricative (de Blois 2006, 94).

The preservation of the original orthography of the Arabic loanwords is certainly a consequence of their scholarly origin: the Arabic loanwords entered the Persian language mainly from books, from the written Arabic language; and only gradually, by 'osmosis from above' (Bausani 1978, 13–14), did they penetrate into the everyday language. It was mainly the Persian (or Iranian) bilingual scholars, who knew and used Arabic as a scholarly language (as was Latin for European scholars), who were responsible for the introduction of a great quantity of Arabic

learned vocabulary into their works. If the spelling of the Arabic loanwords in Persian has to be taken as referring to real pronunciation, and not only to the written form of the words, it probably represents mainly a scholarly or literary pronunciation of Classical Arabic, i.e. an artificial and altogether scholastic pronunciation. The same artificial pronunciation of Arabic loanwords is characteristic of the metrical reading of poetry. In Persian poetry, for example, the Arabic *ʿayn* is always 'pronounced', i.e. counts as a consonant, even in positions where, in normal speech, it is (and probably was) never pronounced. Likewise, the *hamza* (glottal stop) of Arabic words is counted as a consonant; and this can happen even at the beginning of a word, before a vowel, when the *hamza* is and was neither written nor pronounced. Moreover, <z>, <ḏ>, <ḍ> and <ẓ> never rhyme together, nor can <s> rhyme with <ṯ> or with <ṣ>, <h> with <ḥ>, or <ʾ> with <ʿ> (Meier 1981, 103). This seems to point to a sort of artificial and scholarly pronunciation of Arabic loanwords in Persian poetry.

The preservation of the original Arabic orthography of loanwords has an important implication (and was also probably dictated by this need): it makes the Arabic loanwords immediately recognisable, without destroying the kinship between words pertaining to one and the same Arabic root. A kind of consciousness of the original written form of the Arabic lexicon within Persian has always been maintained, and is proved not only by the fact that Arabic words can be uttered (in poetry or in scholarly contexts) with a literary pronunciation, approximating to that of Classical Arabic, but also by the fact that even now, in standard New Persian, the intervocalic glottal stop /ʾ/ can be replaced by the glide *y* only if it is represented in the original writing by a *hamza*, but never if it is represented by the homophonous (in Persian) letter *ʿayn*: for ex. /lāʾeq/ 'worthy' can be uttered as [lāʾeq], [lāyeq] or [lāeq]; but /šāʾer/ 'poet', from Arabic *šāʿir*, cannot be pronounced **šāyer*, given the origin of /ʾ/ from an Arabic *ʿayn* in this word (Pisowicz 1985, 20, 102).

The orthography of Arabic loanwords in Persian has remained virtually unchanged throughout the entire history of the New Persian written tradition, remaining impermeable to any influence from the different diachronic and dialectal varieties of Arabic (a relevant exception is represented by the Arabic loanwords ending in *tā marbūṭa*, on which cf. Perry 1991 and 1995).

An accurate preservation of the original Arabic spelling for Arabic loanwords is to be found not only in Arabo-Persian orthography,[18] but also – where possi-

18 One exception is the Persian orthography of words with *hamza* and their phonetic realisation, on which thorough historic-linguistic research is still lacking, apart from useful remarks in the excellent work on the history of the Middle and New Persian phonology by Pisowicz (1985, 20, 47–51, 102).

ble – in the orthography of the New Persian texts in scripts different from Arabic. Indeed, for the redaction of New Persian texts, with their rich Arabic lexicon, the Hebrew alphabet had the possibility of transliterating many of the 'only Arabic' letters: *ḥēṭ*, as opposed to *hē*, was used to transliterate Arabic *ḥā*; *ṭēṭ*, as opposed to *tāw*, was used for Ar. *ṭā; ʿayin* was available for Ar. *ʿayn*; *ṣādē*, which in the adaptation of Aramaic-based scripts to the Iranian languages had already been employed for Iranian *č* (Skjaervo 1996, 516), and partially for *ǰ*, was of course also suited to represent Ar. *ṣād*; and *qōph* was available for Ar. *qāf*. Moreover, for some of the Arabic and Persian sounds not represented in the Hebrew alphabet, the Judaeo-Persian writing system resorted to the possibility of representing the Hebrew spirant allophones of the plosives by means of diacritic signs: hence *tāw* was used to represent both /t/ and, with or without diacritics, the letter *ṯā* of the Arabic loans; *dālet* was used for /d/ and, with or without diacritics, for Ar. *ḏāl*; *kaph* for /k/ and for /x/ (in both Arabic and Persian words); *pē* was used for Persian /p/ and for Arabic and Persian /f/. The only Arabic letters which had no possible graphic equivalents in the Hebrew alphabet were *ḍād* and *ẓā*, for which – as well as for Persian and Arabic /ǰ/ – a series of different solutions were adopted (amply described by Paul 2013, 30–33, §§ 11–12).

For Persian texts in Syriac script, too, the 'only Arabic' letters were easily transliterated by means of the corresponding letters <ḥ ṭ ʿ ṣ q> of the Syriac alphabet. The Syriac letters <t> and <d>, which could also represent, with a point under the letter, the spirant allophones of /t/ and /d/, were also used to transliterate the Arabic letters *ṯā* and *ḏāl*; likewise <k>, <p> and <g> with a point below were used to represent Arabic and Persian /x/, /f/ and /ɣ/. On the model of Arabo-Persian orthography, however, in some texts /f/ was represented by <p> with a point above, instead of below. For Arabic <ẓ>, for which no letter was at hand, some Syro-Persian texts used <ṭ> with a dot above, again a clear calque of the Arabic letter *ẓā*.

The same can be said for the New Persian texts in Manichaean script. By the time Manichaean script was being adapted to New Persian (to be tentatively placed at the end of the 9th – first half of the 10th century), six new letters had already been added to the original 22 of the Aramaic alphabet to write other Iranian languages, and were already present in Manichaean texts in Sogdian: <β γ δ f j x>. For the Arabic words entering New Persian, therefore, the letter <δ> was used to transliterate *ḏāl*; and the same letter, single or more often doubled, <δδ>, was used to transliterate *ṯā*. Two new letters introduced by punctuation (<k> and <q> with two dots above) – already used in Turkish texts in Manichaean script – were used in Manichaean New Persian to transliterate *qāf*, given that simple <q> (without dots) had already been used (as an alternative to <k>) to represent *k*. Moreover <ʿ> with two dots above was introduced to represent *ʿayn* in

Arabic words, because in Manichaean orthography <ʾ> had already been used to represent an initial palatal vowel. For <ḍ> of Arabic words different solutions were adopted: <z>, as in <zʾwbt> for *ḍābiṭ* 'commander; chaste' in the 'Catechism' edited by Sundermann (2003, c19; a different reading for this word is suggested by de Blois 2006, 114, *s.v.*; see also Shokri-Foumeshi 2014, 202–203), and <d> in <hawwd> *ḥawḍ* 'basin, cistern' (Sundermann 2003: e18).[19] On the other hand, the 'Arabic' letters <ṭ>, <ṣ> and <ḥ> could not be transliterated, because the corresponding letters of the original Aramaic alphabet had already been used to note sounds of the Iranian languages.[20] Indeed, Manichaean orthography was already so loaded with graphic habits fixed for other Iranian languages that it became impossible to render all the 'only Arabic' letters into this script. The creation of new letters had in the meantime come to an end, soon to be followed by the disappearance of the Manichaean religion from the pages of history.[21]

3 New Persian texts in scripts different from Arabic

Let us now look at the most ancient New Persian texts in non-Arabic scripts – which cover exactly the period for which we have no original documents in Arabo-Persian – in the hope of discovering some indirect evidence about time and place of the formation of the Arabo-Persian orthographic canon.

3.1 Judaeo-Persian texts

Some Judaeo-Persian texts, that is Persian texts written by means of the Hebrew alphabet, are among the most ancient written New Persian documents. The Jewish minorities living in Iran spoke Persian, or one of the various Persian

19 The latter spelling is interpreted by de Blois (2006, 96) as being dictated by 'the 'Persian' convention of representing a postvocalic interdental as d rather as δ', exceptionally applied also to an Arabic word. Filippone (2011, 186), instead, thinks that this spelling reflects a dialectal pronunciation.

20 No loanword with Arabic <ẓ> is attested in Manichaean documents published up till now: cf. the glossaries by de Blois 2006 and Provasi 2011, 163–168. A glossary of all Arabic loanwords in Manichaean New Persian texts has been published by Shokri-Foumeshi 2014.

21 On the adaptation of the Manichaean script to the Iranian languages cf. Henning 1958, 73–75; and for writing New Persian, cf. Henning 1962, 89–91, Orsatti 2007, 150–164.

(or Iranian) dialects spread throughout the Iranian linguistic area, as their mother tongue (Yarshater 1974); and, for written purposes, they used Persian (more or less tinted with dialectal features) written in Hebrew characters. In the past, scholars thought that the texts emanating from Jewish minorities revealed a Persian dialect different from the language of their Muslim neighbors. Recent studies, however, especially since the discovery and publication of the manuscript of the Persian dialectal translation of the Qur'an known as *Qur'ān-i Quds* (see above), have shown that this is not so, apart – of course – from the presence in Judaeo-Persian texts of some Hebrew loanwords and expressions. A number of linguistic features known until then only from Judaeo-Persian texts were also found in this Muslim text (Lazard 1990). In general, the Judaeo-Persian texts, as well as the New Persian texts in Syriac and Manichaean scripts, reflect written varieties of New Persian differing from literary New Persian. Recent studies, moreover, have focused on a difficult task: analyzing the dialectal variations within the Judaeo-Persian and other written traditions of Iran.[22]

We do not know when the Hebrew alphabet was adapted to write Persian.[23] As only texts dated or datable to the Islamic period are extant, it is generally supposed that the adaptation of the Hebrew alphabet to the Persian language occurred only in Islamic times. But some scholars think that a Judaeo-Persian literature (in particular translations of the Bible into Persian) in Hebrew characters must have already existed in the Sasanid period (Bacher 1904). An answer to this question would be essential in order to ascertain the linguistic value of Judaeo-Persian orthography and the possible presence of historical spellings. In what follows I will present the main orthographic characters of some Judaeo-Persian texts, grouped according to chronological and geographic criteria.

Group A. The most ancient dated Judaeo-Persian documents are three short inscriptions carved on a rock in a mountainous passage, Tang-i Azao, in Western Afghanistan, left by three merchants bearing Jewish names, who were coming from Kōban, the ancient name of the Qabul valley. These inscriptions (edited by Henning 1957) are dated, according to the Seleucid era, to 1064, corresponding to 752 CE.[24] They are all very short; no Arabic word is attested.

22 On Judaeo-Persian dialectology, after the groundbreaking article by Lazard 1968, see Shaked 2009 and Lazard 2014.

23 For a thorough analysis of the adaptation of the Hebrew alphabet to write Persian, cf. Paul 2013, 23–48.

24 Rapp (1967, 55–56) has unconvincingly questioned the dating proposed by Henning (1957, 338), proposing a much later date: 1299–1300 CE.

The main orthographic features of these inscriptions are the following:

- the *iḍāfa* particle is written with a *yōd* attached to the following word, unlike the Arabo-Persian orthography: <(y)'r y'wy> *yār-i ōy* 'his Friend' (C3);
- the suffix pronoun of the 3rd singular person *-aš* is written separated from the preceding word, i.e. with initial *'āleph*: <y'r 'š 'w b'd> *yār-aš ō bād* 'May He be his helper' (A3, B2-3);
- the letter *qōph* of the Hebrew alphabet is used to represent the Persian sound /k/: <qnd> *kand* '(he) incised' (A2, B2), <qy> *ki* 'who' (A2).

The use of <q> to represent /k/ implies that the Hebrew *kaph* could be left to represent the voiceless uvular spirant /x/ of Persian (no occurrence of *x* is to be found in these short inscriptions, however), for which sound no letter of the Hebrew alphabet was at hand. This can be taken as evidence that, in all probability, Arabic words had not entered the Persian language yet, or at least not the language represented by these inscriptions (Lazard 1968, 82).

Group B. Chronologically, after the inscriptions of Tang-i Azao there follow two letters discovered at Dandan Uiliq (Central Asia, northeast of the Khotan oasis), and referred to as DU1 and DU2, datable to the second half of the 8th century. Of the first one, a fragment of a commercial letter, a continuous reading cannot be given, as the left and right margins of the sheet have been badly damaged.[25] The other, also coming from the same area and certainly written in Khotan, has recently been published by Zhang and Shi (2008) with a study in Chinese (which I have not been able to read). Both represent the same language and the same orthographic usage.

The main orthographic features of these letters are:

1. <q> represents Persian /k/, and <k> represents /x/: <qrdwm> *kardum* 'I made' (DU1, 2), <kwdh> *xudah/xudāh* 'God' (DU1 and DU2 *passim*);
2. sometimes long *ā* is not written: <yptwm> *yāftum* 'I found' (DU1, 28), <sd hzr> *sad hazār* 'hundred thousand' (DU2, 1), <kwhrq> *xwāharak* 'sister' (DU2, 3);
3. the *iḍāfa* particle/relative pronoun is variously represented:
   - <'y>: <'z swy 'y mn> *az sōy-i man* 'from me' (DU1, 18), <'z swy 'y dwyd r'> *az sōy-i Dawīd-rā* 'from David' (DU2, 34) <q'r 'y prmwdy> *kār-i farmūdī* 'the work which you ordered' (DU1, 29, with *iḍāfa* as relative pronoun);

25 Published by Utas (1968), with a bibliography of the previous studies. Lazard (1988) has given a valuable contribution to the reading and interpretation of a number of passages from this text.

- not written before a palatal vowel, with which it probably blended: <pnnʾm yzyd> *pannām(-i) īzid/ēzid* 'in the name of God' (DU2, 1), <tn yšmʾ> *tan(-i) išmā* 'yourself' (DU1, 23),[26] <nʾmh yšmʾ> *nāma(-i) išmā* 'your letter' (DU1, 28, 33), <mrdwmʾn yšmʾ> *mardumān(-i) išmā* 'your people' (DU2, 3);
- <y->: <kwdh yqrbqr> *xudah-i kirbakkar* 'the beneficent God' (DU2, 1);
- <-y> attached to a preceding demonstrative pronoun *ān* 'that': <ʾny *ān-i*> 'belonging to' (DU1, 4, 13), <wbr kwndwm ʾny nbyšt bwdy> *u bar xwāndum ān-i nibišt būdī* 'and I read what you had written' (DU2, 7, with *-i* relative pronoun and an old past participle, without final *-a*)

4. the conjunction is written <w> attached to the following word: <swd wzyʾn ʾy man> *sūd u-ziyān-i man* 'my profit and loss' (DU1, 14), <bzwrg wqwdq> *buzurg u kōdak* 'young and old' (DU2, 3);
5. <ṣ> is used for *č* and *ǰ*: <ṣwn> *čūn* 'as' (DU1, 13), <ṣmh> *ǰāma* 'clothing' (DU1, 10), <pnṣ> *panǰ* (DU2, 14, 27);
6. <by> representing the directional preposition *bē/bi* is well attested: <by šmʾ> *bē šimā* (or *bē-šmā* with contraction?) 'to you' (DU2, 27);
7. final *-a* is regularly written <h>: <prwkth> *furōxta* 'sold' (DU1, 10), <nʾmh> *nāma* 'letter' (DU1, 28, 32), except than in the monosyllables <pʾ> *pa* 'in, to' (DU1 and DU2 *passim*), <mʾ> *ma-* (verbal prohibitive prefix) <mʾ kwr> *maxwār* 'do not suffer' (DU1, 30), <nʾ> (verbal negative prefix) <nʾ dʾnwm> *nadānum* 'I do not know' (DU2, 27), and in the pronoun *išmā*/*šimā*;
8. <ʾ> is sometimes omitted before an initial vowel other than long or short *a*: <dwr by wptʾd> *dūr biyuftād* '(it) was delayed' (DU1, 7), <ydwn> *ēdūn* 'so' (DU1, 22, 24, 31), <yzyd> *īzid/ēzid* 'God' (DU1 and DU2 *passim*);[27]
9. in the first letter (DU1) there are no Arabic loanwords, but in the second one (DU2) a few are attested. Their spelling does not reproduce the original Arabic writing: *ḥakīm* 'doctor' is spelled <hqym> (Du2, 4, 13), *ḥarb* 'war' is spelled <hrb> (DU2, 33), without the Hebrew letter <ḥ> being used to represent the Arabic emphatic *ḥ*;

26 In these texts the personal pronoun for the second person singular is probably to be read *išmā* (also perhaps alternating with *šimā*), a form occasionally attested in Early New Persian (Lazard 1988, 208).

27 See the word <myd> *umēd* 'hope' in the *Qurʾān-i Quds* in Arabic script (Filippone 2011,190) and <mwd>, with dialectal *u*/*ū* for *ē*, in the Judaeo-Persian inscription A (l. 3) from Tang-i Azao (Henning 1957, 342 and n. 2).

10. perhaps the ancient directional preposition *ō* (reduced to *o*/*u* or *a* and represented by <’> attached to the following word) is represented in <’pyš> *a-pēš* ‘near, before’ (DU2, 8).

Therefore – except for the coincidence with the Arabo-Persian orthography in the spelling of final *-a* as <h> – the occasional defective spelling of *ā* and of initial *ʾāleph*, the different ways the *iḍāfa* particle is written, the use of letters <q> for /k/ and <k> for /x/ are all features representing an altogether different orthographic tradition, compared to the Arabo-Persian one.

Group C. Two dated documents pertaining to the Ahvaz or Southwestern group of Judaeo-Persian documents (10th – first half 11th centuries) are considered here: a legal document dated 1262 of the Seleucid era/ 950 CE, edited by Shaked (1971); and a legal document dated Ahvaz (Khuzistan, south-western Iran) 1332 Seleucid/1020 CE. In the latter (edited by Asmussen 1965, and generally referred to as *Law report of Ahvaz*):

1. <k> represents *k*, and <q> represents *q* of Arabic or Hebrew words, in contrast to the usage in the first two Judaeo-Persian texts already discussed: <knd> *kand* ‘he snatched’ (l. 5), <qw’my> *qiwāmī* ‘right, lawful’ (l. 12);
2. the letter *ṣādē* <ṣ> is used for Persian *č* and *ǰ*, but is also used to represent the Arabic letters *ḍād* and *ṣād*:
   - <ṣ> = *č*: <bṣ’> *bač(č)a* ‘baby’ (l. 4)
   - <ṣ> = *ǰ* in both Persian and Arabic words: <ṣwml’> *ǰumla* ‘whole’ (l. 2), <pnṣ> *panǰ* ‘five’ (l. 5), <ṣwpt> *ǰuft* ‘couple’ (l. 5), <ṣw’b> *ǰawāb* ‘answer’ (l. 7); only once *ǰ* is written <g>: <w’gyb> *wāǰib* ‘necessary’ (l. 8)
   - <ṣ> = *ṣ*: <myṣr> *Miṣr* ‘Egypt’ (l. 5) (for Persian *s*, the letter *samek* <s> of the Hebrew alphabet is used: <s’l> *sāl* ‘year’)
   - <ṣ> = *ḍ*: <ṣrwr’> *ḍarūra(t)*, ‘need’ (l. 8), <r’ṣy> *rāḍī* ‘content’ (ll. 10–11)
3. <h> e <ḥ> of the Arabic alphabet are carefully distinguished: <šhwtwm> *šahwat-um* ‘my desire’ (l. 5), <ḥwṣt> *ḥuǰǰat* ‘proof’ (l. 15);
4. <t> and <ṭ> are carefully distinguished: <byst> *bīst* ‘twenty’ (l. 5), <slṭ’ny> *sulṭānī* ‘sultanial’ (l. 12);
5. the Arabic *ʿayn* <ʿ> is always written: <ʿwṣ> *ʿiwaḍ*, *ʿawaḍ* ‘compensation’ (l. 11);
6. the *iḍāfa*: is spelled as <y-> attached to the following word or is not written:
   - <y->: <kwd’wnd’n ymylk> *xudāwandān-i milk* ‘the owners of the property’ (l. 9), <ʿwṣ y’yn drh’> *ʿawaḍ-i īn durhā* ‘as recompense for these pearls’ (l. 11), <p’yn ykyrdy> *p-īn-i kirdī* ‘in this that you have done’ (l. 7, with written *iḍāfa*/relative pronoun)

- not written: <ʾwṣ ʾyn dynʾr> *ʾawaḍ-i īn dīnār* 'in place of these dinars' (l. 12); <bd kyrdy ʾyn kyrdy> *bad kirdī īn(-i) kirdī* 'you did wrong (what) you did' (ll. 6–7, with not written *iḍāfa*/relative pronoun);

7. the directional preposition outcome of Middle Persian *ō*, already reduced to *u*/*o* or probably *a*, is spelled <'> attached to the following word: <ʾpyš> *a-pēš* 'near, before' (ll. 1, 3, 9, 10); it can also introduce a direct object: <ʾḥṣr' kyrd ʾdny'l> *iḥḍār kird a-Daniel* 'she cited Daniel' (l. 3);[28]
8. the directional preposition <by> is not attested.

Group D. Here only one text is considered, the *Tafsīr* of Ezekiel, that is a translation and commentary of the Book of Ezekiel. Its manuscript (ms. Firkowicz I 1682 of the St Petersburg Public Library, edited by Gindin 2007) is datable to a period between the late 10th and the early 11th century. The so-called 'Part 1' of the manuscript (pp. 1–169 and 221–226) seems to represent a northeastern dialectal variety of New Persian probably originating from northeastern Iran or Afghanistan (Gindin 2007, 23–26).[29] In Part 1:

1. the Arabic words are transliterated, and retain their original Arabic spelling:
   - <ṭ> (and not <ṣ>, as in the Ahvaz document) is used for Arabic *ḍ* (cf. Paul 2013, 33)
   - <g>, or <g> with a stroke (and not <ṣ>, as in the Ahvaz document) is used for *ǰ*, in both Arabic and Persian words: <gʻwd ʾz ʾw> *ǰud az ō* 'different from that' (5.8, Gindin 35), <gʻʼmʾyhʾ> *ǰāmayihā* 'clothes' (38.4, Gindin 73)
2. the orthography of the *iḍāfa* particle is the same as in the Arabo-Persian writing system, apart from rare cases in which it is written <-y> also after a word ending with a consonant: <šrḥy ʾn ʾydr by krd> *šarḥ-i ān ēdar bikard* 'he explained it here' (3.14, Gindin 33)
3. final *-a* is written <h> mainly in past participles, but retains the spelling with final <'> in monosyllables and in many words:
   - final *-a* is written <h>: <bwdh> *būda* 'been' (35.21, Gindin 69),
   - final *-a* is written <'>: <hm'> *hama* 'all' (*passim*), <ṣʻyhr'> *čihra* 'face' (38.7, Gindin 73), <p' zmʾnʾy mšh> *pa zamāna-yi Mōšeh* 'at the time of Moses' (5.10–11, Gindin 35)

28 The same form <'ḥṣr'>, probably to be read *iḥḍār*, is also found in the other legal document dated 950 (Shaked 1971, l. 1).

29 Recently Lazard (2014, 91–92) has instead argued that the language of this *Tafsīr* has a northwestern or central-northern origin. Quotations are given according to the page and line of the manuscript, followed by the page in Gindin's edition.

- monosyllables: <m' kwn> *makun* 'do not do' (5.6, Gindin 35), <n' kwnd> *nakunad* 'does not do' (5.9, Gindin 35)

4. the defective writing of long ā is very rare, not to say absent, and words are written as in Arabo-Persian.

From this brief sketch it is possible to conclude that the orthography of the Judaeo-Persian texts taken into consideration seems to show a clear trend: as the language becomes richer in Arabic loanwords, orthography shows an effort to find the best way to transliterate them. The few Arabic loanwords attested in the second letter from Dandan Uiliq are still written as if they were Persian words, without distinguishing the 'Arabic letters' (Group B). Instead, the numerous Arabic loanwords in the Ahvaz legal document (Group C) are carefully transliterated, despite the multiple values given to the letter *ṣādē* <ṣ>, which seems to represent the weak point in the writing system represented in this text. In this text, however, the word <w'gyb> *wājib*, written with <g> for /ǰ/, is interesting: it shows that for /ǰ/ in the increasing number of Arabic words, a new spelling was gaining ground, perhaps taken from Judaeo-Arabic orthography. In the *Tafsīr* of Ezekiel (Group D) the multiple values letter *ṣādē* has in the Ahvaz legal document have been made less ambiguous by further differentiating *ḍād* of the Arabic words from *ṣād*, and Persian *č* from *ǰ* of both Arabic and Persian words.

The first part of this *Tafsīr* shows an influence from Arabo-Persian orthography. Indeed, about the way the Hebrew alphabet is used in this manuscript, David Neil MacKenzie wrote: 'It is clearly based on a familiarity with the normal Arabo-Persian script, and indeed an excellent knowledge of Arabic' (MacKenzie 2003, 103–104). If we consider that the *Tafsīr* manuscript is more or less coeval or even older than the Law report of Ahvaz, it is possible to measure the importance that the north-eastern origin of this text has in regard to its orthography.

3.2 Syro-Persian texts

At this point I will speak very briefly about the Syro-Persian documentation, that is, New Persian texts in Syriac script, emanating from Christian milieus in Iran. Not many texts belong to this group because – in contrast to Persian Jews – Christians in Iran have very often used the Arabic script; it is interesting to note that, for example, the translations of the Bible into Persian in Hebrew milieus are written in Hebrew script, whereas the Persian Gospels are written in Arabic script. Christians even dated their manuscripts according to the Muslim era (see for example the manuscript of the Persian lectionary studied by Richard 1981).

Many Syro-Persian manuscripts are not dated and, even when they are datable with reasonable certainly, they have undergone a long transmission which may have exerted a normalizing effect on orthography; therefore they are not so useful for the purpose of giving indirect evidence on Arabo-Persian 'parallel' orthography.

Only the previously quoted fragment from the bilingual (Syriac and New Persian in Syriac script) Psalter from Central Asia is probably attested by an old manuscript. It shows a clear influence from Arabo-Persian orthography, in particular in the complementary distribution of *d* and *δ* (see above). Therefore, again, a text from northeastern Iran gives – through its orthographic characteristics – good evidence of how the coeval Arabo-Persian orthography was by then well-established, and exerted a strong influence on the other graphic traditions of Iran.

3.3 Manichaean New Persian texts

For the purpose of writing their texts, the followers of Manichaeism used the Manichaean alphabet, named after the founder of this religion, Mani (3rd century). Middle Persian and Parthian, two Iranian languages, were the languages of liturgy for Manichaeans. Therefore, New Persian Manichaean orthography shows an influence from both these languages' orthography, as well as from Sogdian Manichaean orthography, to the extent that New Persian in Manichaean script can be considered as the last heir of the graphic traditions of pre-Islamic Iran.

From historical sources we know that during the 10th century Manichaeans were obliged to leave Iraq and western Iran taking refuge in northeastern Iran, and especially in Samarkand, the ancient capital of Sogdiana. This was one of the capital cities of the Persian Samanid dynasty, which ruled over eastern Iran in the last quarter of the 9th century and during the entire 10th century. Persian poetry and prose had its first important blossoming under Samanid patronage. Samarkand also had strong links with the Sogdian colonies of Central Asia, in the Turfan oasis and other places in Chinese Turkestan where, at the beginning of the last century, all the extant texts in Manichaean script were discovered.

Two fragmentary poetical New Persian texts in Manichaean script were published by Walter Bruno Henning in 1962. Henning dates the first of these manuscripts, a fragment of the poem Bilawhar and Būdīsaf, to the first half of the 10th century, during the lifetime of the first great poet of Persian literature, Rūdakī. The second, a fragment of a monorhymic lyrical composition (*qaṣīda*), could be – judging from its script – even older than the first one (Henning 1962, 99; see also below).

These texts are of great interest. From the orthographic point of view they are a proof of the historical and conservative character of New Persian Manichaean orthography. Indeed, the two poetic texts become metrically readable only supposing – behind their conservative spelling – the new forms fixed through the Arabo-Persian orthography. Regarding the second text, the *qaṣīda*, Henning made the assumption 'that the poem had originally been written in Arabic script and was then transliterated into Manichaean script by a man who did not understand it properly' (1962, 99). In fact, an occasional mistake in the spelling of some words attached to one another is evidence of poor knowledge of the language by the copyist: <cwzg'hyy> for *juz gah-i*, in the expression *juz gah-i šumār* 'except at the time of the (last) reckoning' (l. 20).

Henning's brilliant hypothesis of an original Arabo-Persian version from which the text we possess has been copied is useful to explain the metrical form of the text (or of both texts), composed according to the new quantitative prosody deeply influenced by Arabic prosody. This hypothesis, beyond permitting Henning's masterly reading of a text so full of gaps, can certainly account 'for the omission of the Iḍāfe-particle (at least four times) and the word "and"' (Henning 1962, 99); but cannot account for all the other places where the *iḍāfa* particle has been correctly inserted (ll. 9 after *bād*, 10 after *pēš* and *šarāb*, 20 after *gah*, 21 after *zēr* and 30 after *sōy*). How could a copyist unfamiliar with Persian have inserted it correctly? Moreover, how could a copyist unfamiliar with Persian have used – for *juz* 'except' in the phrase above – the historical Manichaean orthography <cwz>, with <c> being the transliteration of the ancient Aramaic <ṣ>? In the Arabo-Persian orthography he would have found, rather, <j> instead of <č> (see Table in Section 2.1 above).

Another mistake made by the copyist, the writing <rrd'> instead of <drd'> *dard-ā* 'o grief!' (l. 1), can likewise be evidence that the antigraph from which the copyist was copying was written in Manichaean script. A confusion between <r> and <d> is possible in the Manichaean writing, where the two letters only differ for a point above <r> or below <d>. Instead, though in later styles of the Arabic script <r> and <d> could be confused, it was probably not so in early styles, where these two letters had a markedly different form, and <r> could rather be confused with final <n>.[30]

30 For <r/z>, <d/ḏ> and <-n> in the Codex Vindobonensis, see the first line of the text (fol. 1v): *sipās bāδ yazdān-i dānā wa tawānā-rā*. The confusion between <r/z> and final <n> in the earliest Arabic writing styles is proved by some mistakes in the quotation of Persian words by Arabic authors, as <gm'z> (with <z> = <r> with a point above) for Persian *gumān* 'doubt', and <'yḏn> for *ēδar* 'here' (Tafazzoli 1974, 339, 343).

Though an original redaction of the text of the *qaṣīda* in Arabic script cannot be excluded, we can suppose that the text we possess was copied – by an inaccurate copyist who was probably not completely familiar with Persian – not from a manuscript in Arabic script, but from one in Manichaean script already showing a strong influence from the new Arabo-Persian orthography. Evidence of such an influence must be seen – as pointed out by Henning – in the cases in which the *iḍāfa* particle, and in the one case (l. 12) in which the coordinate conjunction are not written.[31] An influence of the coeval Arabo-Persian orthography on the orthography of this text is also proved by the unusual spelling of the *iḍāfa* particle in the already quoted phrase *juz gah-i šumār* (l. 20): in <cwzg'hyy> the *iḍāfa* is spelled <yy> instead of <ʿyg> or <ʿy>, which are the normal spellings for it in Manichaean orthography. Lastly, 'the scribe of the Qaṣīde allows an occasional *-δ-* for postvocalic *-d-* (*'ryδ* 'he brings', *nbwδm* 'I was not')', instead of *-d-*, given throughout by the copyist of the other poetical text in Manichaean script (Henning 1962, 90): perhaps this too is an interference from the coeval Arabo-Persian orthography. For the preposition *pa(d)* and the negative verbal prefix *nē* (or already *na*), the copyist of the *qaṣīda* prefers the joint spellings <p-> and <n-> respectively, as in the new Arabo-Persian orthography, instead of the spellings <pd> and <ny>.[32]

As already stated, Henning considers the fragment of the *qaṣīda* as even older than the other manuscript, the fragment of the poem Bilawhar and Būdīsaf in Manichaean script, which he had dated to the first half of the 10th century. Therefore the fragment of the *qaṣīda*, and even more so its antigraph, should be dated at an early date, possibly between the end of the 9th and the beginning of the 10th centuries CE. The text of the *qaṣīda* turns out to be not only one of the earliest surviving poems of the kind, but an early original document attesting to the parallel development and fixing of the Arabo-Persian orthography.

31 Cases of unwritten *iḍāfa*s are also to be found in other New Persian Manichaean texts (whereas they are quite rare in Middle Persian Manichaean texts): an unwritten *iḍāfa* should probably be recognised in an otherwise incomprehensible passage in the Catechism (*Lehrtext*) published by Sundermann (2003, c13–14): <q'lbd k' x'n'g (h)wm'n'g 'st ''b rwšn ny> *kālbad ka xāna humānā ast āb(-i) rōšan nē* 'the body, which is like a house, is not bright water'. Also in the New Persian Manichaean texts published by Provasi there is at least one instance of an unwritten *iḍāfa* (2011, 149). As to the leaving out of the conjunction, we already saw that this is not a rare occurrence in the most ancient Arabo-Persian manuscripts.

32 A full discussion of Henning's hypothesis about a possible Arabo-Persian antigraph for the *qaṣīda* is given by Orsatti (2007, 161–164).

4 Conclusions

On the basis of the New Persian original documents in non-Arabic scripts here examined, we can tentatively date and localise the beginning of the Arabo-Persian orthographic influence on the other written traditions of Iran: northeastern Iran, end of the 9th – beginning of the 10th centuries. By then, this orthographic tradition appears as already fixed. Though it cannot be excluded that some scattered and unsystematic attempts were accomplished here and there in different places of Iran, the hypothesis of a multi-centric origin of the adaptation of the Arabic alphabet to Persian, favored for example by Akimuškin (1987, 332), seems less probable in the light of our documentation.

References

Akimuškin, Oleg F. (1987), 'Persidskaja rukopisnaja kniga' [The Persian manuscript book], in Jurij A. Petrosjan et al. (eds), *Rukopisnaja kniga v kul'ture Narodov Vostoka. Ocherki. Kniga pervaja.* Moskva: Nauka, 330–406.

Asmussen, Jes P. (1965), 'Judaeo-Persica II: The Jewish-Persian Law Report from Ahwaz, A. D. 1020', in *Acta Orientalia*, 29: 49–60.

Bacher, W. (1904), 'Judaeo-Persian', in *The Jewish Encyclopedia.* Vol. VII, 313–317.

Bausani, Alessandro (1960), 'La letteratura neopersiana', in *Storia della letteratura persiana*. Milano: Nuova Accademia (Reprint Roma: Nuova Cultura 2011, La Sapienza Orientale – Ristampe).

Bausani, Alessandro (1978), 'Incontri linguistici in Asia. Esperienze di un orientalista', in *Lingue a contatto nel mondo antico. Atti del Convegno della Società Italiana di Glottologia, Napoli 12 e 13 maggio 1978*. Pisa: Giardini, 11–33.

Bivar, Adrian David H. (1986), 'Kitābāt. 9. Iran and Transoxiana', in *Encyclopedia of Islam*. Second edition. V, 228–231.

de Blois, François (2006), 'Glossary to the New Persian texts in Manichaean script', in François de Blois, Erica C.D. Hunter and Dieter Taillieu (eds), *Dictionary of Manichaean Texts. Vol. II: Texts from Iraq and Iran (Texts in Syriac, Arabic, Persian and Zoroastrian Middle Persian)*. Turnhout: Brepols, 89–120.

Bombaci, Alessio (1966), *The Kūfic Inscription in Persian Verses in the Court of the Royal Palace of Mas'ud III at Ghazni*. Rome: Istituto per il Medio ed Estremo Oriente.

Browne, Edward G. (1895), 'Some Notes on the Poetry of the Persian Dialects', in *Journal of the Royal Asiatic Society*, 773–825.

Filippone, Ela (2011), 'The Language of the Qor'ān-e Qods and Its Sistanic Dialectal Background', in Mauro Maggi and Paola Orsatti (eds), *The Persian Language in History*. Wiesbaden: Reichert, 179–235.

Gindin, Thamar E. (2007), *The Early Judaeo-Persian Tafsīrs of Ezekiel: Text, Translation, Commentary.* Volume I: *Text*. Vienna: Österreichische Akademie der Wissenschaften.

Gnoli, Gherardo (1964), *Le iscrizioni giudeo-persiane del Ġūr (Afghanistan)*. Rome: Istituto Italiano per il Medio ed Estremo Oriente.
Grohmann, Adolf (1971), *Arabische Paläographie*. II. Teil. *Das Schriftwesen. Die Lapidarschrift*. (Forschungen zur islamischen Philologie und Kulturgeschichte. Band II), Vienna: Hermann Böhlaus.
Hägg, Thomas and Bo Utas (2003), *The Virgin and her Lover: Fragments of an Ancient Greek Novel and a Persian Epic Poem*. Leiden, Boston: Brill.
Hashabeiky, Forogh (2005), *Persian Orthography. Modification or Changeover? (1850-2000)*. Stockholm: Uppsala University Library.
Henning, Walter B. (1957), 'The inscriptions of Tang-i Azao', in *Bulletin of the School of Oriental and African Studies*, 20: 335–342.
Henning, Walter Bruno (1958), 'Mitteliranisch', in *Handbuch der Orientalistik*, Erste Abt., Band IV: Iranistik, 1. Abschnitt: *Linguistik*. Leiden, Cologne: Brill, 20–130.
Henning, Walter Bruno (1962), 'Persian Poetical Manuscripts from the Time of Rūdakī', in *A Locust's Leg. Studies in Honour of S.H. Taqizadeh*. London: Percy Lund, Humphries & Co., 89–104.
Lazard, Gilbert (1963), *La langue des plus anciens documents de la prose persane.* Paris: Klincksiek.
Lazard, Gilbert (1968), 'La dialectologie du Judéo-Persan', in *Studies in Bibliography and Booklore*, 8: 77–98 (= Lazard, La formation de la langue persane, pp. 27–48).
Lazard, Gilbert (1971), 'Les origines de la poésie persane', in *Cahiers de civilisation médiévale* (Xe-XIIe siècles), 14/4: 305–317.
Lazard, Gilbert (1975), 'The Rise of the New Persian Language', in R. N. Frye (ed.), *The Cambridge History of Iran, IV: The Period from the Arab Invasion to the Saljuqs*. Cambridge: Cambridge University Press, 595–632.
Lazard, Gilbert (1986), 'Les prépositions pad et bē (ō) en persan et en pehlevi', in Rüdiger Schmitt and Prods Oktor Skjaervo (eds), *Studia grammatica iranica. Festschrift für Helmut Humbach*. Munich: R. Kitzinger, 245–255 (=Lazard, *La formation de la langue persane*, pp. 149–156).
Lazard, Gilbert (1988), 'Remarques sur le fragment judéo-persan de Dandān-Uiliq', in W. Sundermann, J. Duchesne-Guillemin and F. Vahman (eds), *A Green Leaf. Papers in Honour of Professor Jes P. Asmussen*. Leiden: Brill, 205–209 (=Lazard, La formation de la langue persane, pp. 157–161).
Lazard, Gilbert (1990), 'Lumières nouvelles sur la formation de la langue persane: une traduction du Coran en persan dialectal et ses affinités avec le judéo-persan', in Shaul Shaked and Amnon Netzer (eds), *Irano-Judaica II*. Jerusalem, 184–198 (=Lazard, La formation de la langue persane, pp. 107–121).
Lazard, Gilbert (1995), *La formation de la langue persane.* Paris: Peeters.
Lazard, Gilbert (2009), 'Qu'est devenue la préposition *ō*?', in Eric Pirart and Xavier Tremblay (eds), *Zarathushtra entre l'Inde et l'Iran: Études indo-iraniennes et indo-européennes offertes à Jean Kellens à l'occasion de son 65e anniversaire*. Wiesbaden: Reichert, 169–176.
Lazard, Gilbert (2014), 'La dialectologie du persan préclassique à la lumière des nouvelles données judéo-persanes', in *Studia Iranica*, 43 (1): 83–97.
MacKenzie, David Neil (1968), 'An Early Jewish-Persian Argument', in *Bulletin of the School of Oriental and African Studies*, 31 (2): 249–269 [Reprinted in Carlo G. Cereti and Ludwig Paul (eds), Iranica diversa. Roma 1999, vol. II, 315–337 + Addenda et corrigenda, 671–673].
MacKenzie, David Neil (2003), 'The Missing Link', in Ludwig Paul (ed.), *Persian Origins: Early Judaeo-Persian and the Emergence of New Persian*. Wiesbaden: Harrassowitz, 103–110.

Margoliouth, D.S. (1903), 'Early Documents in the Persian language', in *Journal of the Royal Asiatic Society,* 761–765.
Meier, Fritz (1981), 'Aussprachefragen des älteren Neupersisch', in *Oriens*, 27–28: 70–176 [Reprinted in E. Glassen and G. Schubert (eds), *Bausteine. Ausgewählte Aufsätze zur Islamwissenschaft*. Istanbul 1992, vol. II. 1057–1164].
Minorsky, Vladimir (1942), 'Some Early Documents in Persian', in *Journal of the Royal Asiatic Society*, 184–194.
Müller, F. W. K. (1915), 'Ein syrisch-neupersisches Psalmenbruchstück aus Chinesisch-Turkistan', in Gotthold Weil (ed.), *Festschrift Eduard Sachau zum siebzigsten Geburtstage*, Berlin: Reimer, 215–222.
Muwaffaq b. 'Alī al-Hirawī (1972), *Das Buch der Grundlagen über die wahre Beschaffenheit der Heilmittel: vollständige Faksimile-Ausgabe des Codex Vindobonensis A. F. 340 der Österreichischen Nationalbibliothek.* Einführung Charles H. Talbot, Franciscus R. Seligmann, Graz: Akademische Druck- u. Verlagsanstalt.
Muwaffaq bin ʿAlī al-Hirawī, Abū Manṣūr (2009), *Kitāb al-Abniya ʿan Ḥaqāyiq al-Adwiya* [...]. Facsimile Copy of the Original Manuscript A. F. 340 Austrian National Library. Transcribed by ʿAlī bin Aḥmad Asadī Ṭūsī. Copied 447 Hijri. Tehrān: Mirāṯ-e maktub & Vienna: Austrian Academy of Sciences Press, 1388.
Orsatti, Paola (1993), 'Le manuscrit islamique : caractéristiques matérielles et typologie', in M. Maniaci and P. F. Munafò (eds), *Ancient and Medieval Book Materials and Techniques* (Erice, 18–25 September 1992) (Studi e Testi 358, Vol. II). Città del Vaticano: Biblioteca Apostolica Vaticana, 269–331.
Orsatti, Paola (2007), *Appunti per una storia della lingua neopersiana. Parte I: parte generale – fonologia – la più antica documentazione*. Rome: Nuova Cultura (La Sapienza Orientale – Manuali).
Orsatti, Paola (2008), 'Minoranze religiose e scrittura nel primo Iran islamico', in D. Bredi, L. Capezzone, W. Dahmash and L. Rostagno (eds), *Scritti in onore di Biancamaria Scarcia Amoretti*. Rome: Edizioni Q. Vol. III, 867–885.
Paul, Ludwig (2003), 'Early Judaeo-Persian in a historical perspective: The case of the prepositions *be*, *u*, *pa(d)*, and the suffix *rā*', in Ludwig Paul (ed.), *Persian Origins: Early Judaeo-Persian and the Emergence of New Persian*. Wiesbaden: Harrassowitz, 177–194.
Paul, Ludwig (2013), *A Grammar of Early Judaeo-Persian*. Wiesbaden: Reichert.
Perry, John R. (1991), *Form and Meaning in Persian Vocabulary: The Arabic Feminine Ending*. Costa Mesa: Mazda.
Perry, John R. (1995), 'Lexical doublets as a derivational device in Persian: The Arabic feminine ending', in *Acta Orientalia Academiae Scientiarum Hungaricae*, 48: 127–153.
Perry, John R. (2002), 'Arabic Language. V. Arabic Elements in Persian', in *Encyclopaedia Iranica*. Online edition (accessed on 10/10/2014).
Pisowicz, Andrzej (1985), *Origins of the New and Middle Persian Phonological Systems*, Kraków: Nakładem Uniwersytetu Jagiellońskiego.
Provasi, Elio (2011), 'New Persian Texts in Manichaean Script from Turfan', in Mauro Maggi and Paola Orsatti (eds), *The Persian Language in History*. Wiesbaden: Reichert, 119–177.
Rapp, Eugen Ludwig (1967), 'The Date of the Judaeo-Persian Inscriptions of Tang-i Azao in Central Afghanistan', in *East and West*, 17 (1–2): 51–58.
Rivāqī, ʿAlī (1985), Qurʾān-i quds. Kuhantarīn bargardān-i Qurʾān ba fārsī? [The Holy Koran. The most ancient translation of the Koran into Persian?]. Tehrān: Muʾassasa-yi farhangī-yi Šahīd M. Rivāqī, 1364. 2 vols.

Richard, Francis (1981), 'Un lectionnaire persan des Évangiles copié en Crimée en 776 H./1374', in *Studia Iranica*, 10: 225–245.

Šams al-Dīn Muḥammad b. Qays-i Rāzī (1981), *Al-Mu'jam fī ma'āyīr aš'ār al-'ajam* [The clarified work about the criteria of the poetry of the Persians]. Ed. Muḥammad Qazwīnī & Muḥammad-Taqī Mudarris Raḍawī. Tehrān: Zavvār, 1360, 3rd edn.

Scarcia, Gianroberto (1963), 'A Preliminary Report on a Persian Legal Document of 470-1078 found at Bāmiyān', in *East and West* N.S. 14 (1–2): 73–81.

Scarcia, Gianroberto (1966), 'An Edition of the Persian Legal Document from Bāmiyān', in *East and West* N.S. 16 (3–4): 290–295.

Shaked, Shaul (1971), 'Te'uda qara'it qeduma be-parsit yehudit' [An ancient Karaite document in Judaeo-Persian], in *Tarbiz*, 41 (1): 47–58.

Shaked, Shaul (2009), 'Classification of Linguistic Features in Early Judeo-Persian Texts', in Werner Sundermann, Almut Hintze and François de Blois (eds), *Exegisti monumenta. Festschrift in Honour of Nicholas Sims-Williams*. Wiesbaden: Harrassowitz, 449–461.

Shokri-Foumeshi, Mohammad (2014), 'Wāžahā-yi daxīl-i 'arabī dar kuhantarīn matnhā-yi fārsī-yi naw bar asās-i dast-nivīshā-yi mānawī-yi Tūrfān' [The Arabic loanwords in the earliest New Persian texts in Manichaean manuscripts from Turfan], in *Tārīx-i Īrān/Iran History* 72 (5). Autumn 1392/2013, published Spring 1393, 161–210.

Sims-Williams, Nicholas (2011), 'Early New Persian in Syriac script: Two texts from Turfan', in *Bulletin of the School of Oriental and African Studies*, 74 (3): 353–374.

Šīrvānī, Kobrā (1974), *Barrasī-yi imlāyī-yi dastnivīs-ī az* Tafsīr-i Sūrābādī [Orthographical analysis of a manuscript of the *Tafsīr-i Sūrābādī*]. Tehrān: Entešārāt-e Farhangestān-e zabān-e Irān, 1353.

Skjaervo, Prods Oktor (1996), 'Aramaic Scripts for Iranian Language', in P.T. Daniels and W. Bright (eds), *The World Writing Systems*. New York, Oxford: Oxford University Press, 515–535.

Sundermann, Werner (1974), 'Einige Bemerkungen zum syrisch-neupersischen Psalmenbruchstück aus Chinesisch-Turkistan', in Ph. Gignoux and A. Tafazzoli (eds), *Mémorial Jean de Menasce*. Louvain: Imprimerie orientaliste, 441–452.

Sundermann, Werner (2003), 'Ein manichäischer Lehrtext in neupersischer Sprache', in Ludwig Paul (ed.), *Persian Origins: Early Judaeo-Persian and the Emergence of New Persian*. Wiesbaden: Harrassowitz, 243–274.

Tafazzoli, Ahmad (1974), 'Some Middle-Persian Quotations in Classical Arabic and Persian Texts', in Ph. Gignoux and A. Tafazzoli (eds), *Mémorial Jean de Menasce*, Louvain: Imprimérie orientaliste, 337–349.

The Tārikh-e Sistān, transl. Milton Gold. Rome: Istituto Italiano per il Medio ed Estremo Oriente, 1976.

Utas, Bo (1968), 'The Jewish-Persian Fragment from Dandān-Uiliq', in *Orientalia Suecana*, 17: 123–136.

Xānlarī, Parvīz Nātil (1986), *Tārīx-i zabān-i fārsī* [History of the Persian language], 1st revised edn. Tehrān: Firdaws, 1365. 3 vols.

Yarshater, Ehsan (1974), 'The Jewish Communities of Persia and their Dialects', in Ph. Gignoux and A. Tafazzoli (eds), *Mémorial Jean de Menasce*. Louvain: Imprimerie orientaliste, 453–466.

Zhang, Zhan and Shi, Guang (2008) [2009], 'Yijian xin faxian Youtai Bosiyu xinzha de duandai yu shidu' [The dating and interpretation of a newly discovered Judaeo-Persian letter], in *Dunhuang Tulufan yanjiu*, 11: 71–99, i color pl.

Esther-Miriam Wagner

Writing Judaeo-Arabic

Abstract: The transcoding of Arabic in Hebrew script found in Judaeo-Arabic texts presents a unique case study in the investigation of standardisation processes. Three major orthographical phases can be discerned. A *phonetic orthography* was used to write early Judaeo-Arabic. A new *Arabicised orthography* emerged in the 9th century known as Classical Judaeo-Arabic, setting the norms used by a large middle class of Jewish Egyptian writers. The economic deterioration and political changes which began in the 12th and 13th centuries resulted in the breakdown of the Jewish middle class. Segregation between Muslims and the minorities in the Islamicate increased, leading to a return to religious values and a more inward-facing attitude within the Jewish communities. This is also reflected in the written language, which shows an increase in markedly Jewish forms, and eventually led to the *Hebraisation* of Judaeo-Arabic orthography.

In addition, vast differences in writing standards can be observed in the various genres of Judaeo-Arabic from the 13th century onwards, with a widening gap between literary and utilitarian writing. This indicates the absence of an acknowledged supra-regional standard across the Jewish communities. In turn, this perhaps points to the lack of a central authority, which the Babylonian and Palestinian Academies had been in the Geonic period, issuing norms to which writers felt it necessary to adhere. The disappearance of universal standards may have been aided by the employment of Jews in government offices and the accompanying education of the Jewish elite with respect to the Arabic script. With segregation and the lack of Jewish clerks in official government functions in the late Ayyubid and Mamluk periods, this particular alley of secondary education became unavailable, which produced a knock-on effect on scribal practice within the Jewish community.

https://doi.org/10.1515/9783110639063-003

1 Judaeo-Arabic

In sociolinguistics, language is seen as a means of identity for speakers and writers. Writers of particular social groups employ particular forms of speech and written text, sometimes consciously, sometimes unconsciously, that mark them as members of a certain social entity. Religious communities are a particularly interesting object for sociolinguistic studies because they are among the groups that draw the strongest sense of communal belonging.

The sociolinguistics of medieval Arabic and the diverse religious sociolects of Arabic present intriguing topics as there is a clear dichotomy in the Muslim part of the population between the spoken vernacular, on the one hand, and the literary language, so-called Classical Arabic, based on the prescriptive norms collected by the grammarians of the 9th century. Classical Arabic probably emerged from a register associated with poetry and divination, and became the language variety in which the Qur'an and the religious literature were codified in the first centuries of Islam. It forms the prescriptive standard for Muslim literary writing, is relatively uniform and exhibits only minimum variability. Due to its link with the Qur'an, Classical Arabic was and is considered 'correct' Arabic, while the vernacular has, or at least used to have, little prestige in comparison.[1]

The situation in the Christian and Jewish parts of the medieval population was slightly different as they were not bound to the same degree as their Muslim counterparts to the literary ideal of *al-ʿarabiyya*. As a result, the norms of Arabic in the writings of Christians and Jews varied, at certain times to a considerable degree, from those standards employed by their Muslim compatriots.

Jews in particular created their own language standards as their primary education – focussed on enabling children to read the Bible – was conducted in Hebrew.[2] The Hebrew alphabet was thus part of Jewish linguistic ability as well as of Jewish identity, although there may have also been pressures from outside not to use Arabic script for Jewish community affairs. The Pact of 'Umar[3], for example, contains the promise 'not to engrave Arabic inscriptions on our seals'.[4]

1 Admittedly, we may be at a watershed moment in time as this attitude appears to be changing among younger Arabic speakers, due to the use of mixed registers that include many vernacular forms in mobile phone and internet communication, such as on Twitter or Facebook, which is written in a modified Latin alphabet.

2 See Olszowy-Schlanger 2003, and also Wagner 2018.

3 I will not discuss the disputed historic authenticity of the Pact of Umar here, but refer to the article by Cohen (1999) on the topic, and to the discussion pertaining to the general dubiousness of historic sources from the Early Islamic period as put down in Astren 2009.

4 See Cohen 1999, 107.

Although this may only reflect social attitudes of the 8th and 9th centuries, it could be suggested that the principles of this text, created as a precedent of relations between Muslims, Christians and Jews, still shaped linguistic policies at a much later time. Whether it was a matter of identity, literacy or political pressure, or most likely a combination of all three factors, Jews mostly wrote their Arabic texts in the Hebrew alphabet, creating what we call Judaeo-Arabic.[5] Our earliest sources of written Judaeo-Arabic are a small number of papyri from the 8th–9th century, whereas the earliest extant texts on parchment and paper date from the 10th century.

It was not only the written language forms that differed; speech too was specific to the respective faith communities. Specific spoken Jewish sociolects may possibly have existed as early as the 7th century, as reported in Islamic sources, but scholars such as Fred Astren have pointed out that these early accounts have to be read with caution and may possibly reflect much later attitudes (Astren 2009). Khan (2007, 526) has advised similar caution, since the surviving works of pre-Islamic poets 'do not exhibit anything that distinguishes them from the works of their non-Jewish contemporaries'. While it is therefore very likely that Arabic-speaking Jews spoke a sociolect among themselves throughout the centuries, the degree to which it differed from their non-Jewish neighbours must have varied considerably. The difference between Jewish and non-Jewish varieties of Arabic might have been both marginal and limited to the lexicon in pre-Islamic times, but this probably changed with the Islamisation of Arab society and the increasing association of Arabic with Islam, which led to the linguistic segregation of non-Muslim communities. The less Jews and Christians were part of the state, the

5 Several definitions for the term 'Judaeo-Arabic' have been proposed, which appears to be part of a larger terminological problem when discussing any Judaeo-X language. Criteria suggested by a variety of scholars include: 'written by Jews', 'written for Jews', 'using the Hebrew alphabet', 'using Hebrew loanwords' etc. Controversy is often caused by texts that are written in Hebrew characters but which present a simple one-to-one transcription of a text in a different alphabet, and as to whether these can be defined as being written in a Jewish sociolect based solely on the fact that the Hebrew alphabet was used. Such one-to-one transcription of Arabic works into Hebrew script, for example in scientific manuscripts, are not viewed as Judaeo-Arabic by some scholars because they are seen as mere renderings of an Arabic text in Hebrew characters. These texts are therefore sometimes referred to as 'Arabic written in Hebrew characters'. Yet the rendering of a text in Hebrew script shows in itself that the writer and intended audience of a piece of texts must have been of Jewish background. If definition requires that authorship and audience be discussed for every piece of writing, this raises highly controversial issues. Geoffrey Khan (2007) has, therefore, suggested that the nomenclature of Judaeo-Arabic should be based on a purely descriptive criterion: the use of Hebrew script. We follow his suggestion, and call Judaeo-Arabic all texts that are written in Hebrew characters.

more their language was removed from that of Muslims. In Fatimid Egypt on the one hand, where the Shiite rulers created a relatively tolerant atmosphere with a large Jewish middle class, in a state in which members of the minorities could rise to influential positions in the bureaucracy, the language of the correspondence of Jewish merchants would betray little difference to that written by contemporary Muslim traders. The linguistic non-conformity of Jewish writing in comparison to contemporary Muslim language norms during the later Ayyubid and Mamluk rule, on the other hand, is conspicuous.[6]

It is thus not surprising that the best known examples of religiously marked speech forms are the spoken early 20th-century Arabic dialects. Blanc (1964) describes how in Baghdad Muslims spoke a variety based on rural bedouin Arabic (the so-called *gilit* dialect) whereas Christian and Jews used to converse in an older, urban variety (of the so-called *qeltu* type). Initially, following the Islamic conquests, all three communities probably spoke the same emerging Baghdadi form of spoken Arabic. The variety that Muslims spoke, however, changed after subsequent waves of immigration from Arabia brought Arab Bedouin who settled in and around Baghdad. Due to the prestige that the language of Bedouin had among the Muslim population – Bedouin were seen as the arbiters of Arabic, preserving its true original character – their speech patterns spread in the Muslim population of Baghdad, who subsequently adopted this more rural dialect as their own. Christians and Jews, on the other hand, continued to speak the older urban dialect that had emerged following the original Muslim conquests.

An additional factor for the development of different forms of speech and written language particular to their religious groups may have been the desire of Christian and Jewish speakers to segregate themselves linguistically from the Muslim population and to create their own way of speaking and writing. Perhaps this happened not only on account of minority-internal linguistic politics, but also due to pressure from the religious majority. In the Pact of 'Umar, mentioned above, one of the tenets put down by the minority communities contains the phrase ولا نتكلم بكلامهم 'we shall not speak as they [Muslims] do'.[7] Although there is a debate as to what is really meant by this statement, and editions such as the 1990 version of Muḥammad Ibn al-Walīd al-Ṭarṭūšī 's *Sirāj al-Mulūk* (the most commonly used source for the Pact of 'Umar) by Jaafar al-Bayati omit the phrase,

6 See Wagner 2010, 229–233; 2017.

7 Taken from page 136 of the *Sirāj al-Mulūk* shelved as Arab.d.58 in the Bodleian Library Oxford, which is catalogued as the 1872 Cairo edition, but differs in page numbers from the numbering (pp. 229–230) given in Cohen (or any of the popular Pact of 'Umar webpages), and also does not agree with any of the page numbers provided by Cohen (1999, 104) for the other known editions.

we could interpret it as a deliberate attempt to linguistically separate Christians and Jews from Muslim speakers. The specific cultural environment and segregation of non-Muslim communities thus led to the emergence of sociolects particular to religious communities such as Jews and Christians.

2 The Cairo Genizah

The transcoding of Arabic in Hebrew script that we find in Judaeo-Arabic texts presents a unique case study for the investigation of standardisation processes. The most suitable sources for this sort of research are documents as they can usually be reliably dated and have not been subject to copying and re-editing in the same way that literary sources are. Documents also allow us to investigate the role of scribes for the standardisation of languages.

The majority of Judaeo-Arabic documents have emerged from the Cairo Genizah. A *genizah* is a storeroom where old manuscripts are discarded and stored. Every synagogue has one, because Jewish customs dictate that anything with the name of God written on it cannot be destroyed but must be stored away, or buried. In medieval penmanship, almost every piece of writing would mention God, and so there was reason to handle them all with care. Through various historical circumstances, the manuscripts of the Ben Ezra synagogue in Cairo were never interred but for nine hundred years, Egyptian Jews deposited anything they wrote into a large, walled-off chamber, where they were unearthed at the end of the 19th century.

The total number of leaves and fragments from the Genizah comes to *c*.350,000, of which about two thirds are currently stored in the manuscript rooms of Cambridge University Library. The manuscripts date from the 7th to the 19th century, and are mostly composed in Hebrew, Judaeo-Arabic, and Arabic, alongside other languages such as Greek, Judaeo-Persian and Yiddish. Many of the Genizah manuscripts are Bible fragments, pieces of religious literature such as Talmud or Mishnah, religious poetry and ethical treatises, but there is also an astonishing wealth of letters and legal documents. It has been estimated that about one seventh of all the Genizah fragments are documentary.

Because of the wealth of data gained from the sources of the Genizah, we also have sufficient background information about the writers of the documents. The biographies and connections between the more prolific scribes of legal documents are known as well as those of the merchants belonging to the network of traders exchanging business letters. Observations on the linguistic behaviour of these protagonists have informed the following sections.

3 Standardisation vs. de-standardisation

In our present world, an innate desire of individuals to standardise language forms appears to be taken for granted, whereas linguistic variation is often looked upon with disdain and resentment. This was not always the case in the past. In the documents and letters of the Cairo Genizah we find that in many cases scribes displayed 'an astonishing degree of inconsistency' or 'a predilection for variety' which the scribes 'must have regarded as a virtue' (Goitein 1971, 236). This can also be seen in spelling. Even the most prolific scribes, such as the court clerk Hillel b. Eli who, by Goitein's estimation, was the second most prolific Genizah scribe and who composed hundreds of documents and letters preserved in the Cairo Genizah 'would spell the same word in two adjacent lines in two different ways' (Goitein 1971, 237).[8] Modern sociolinguists, such as Kretzschmar (2009), have shown that, in speech, variation is a natural part of the linguistic system. In the variations in the writing of the Genizah scribes we may find the equivalent in the realm of the written language.

It is not only the lack of desire to homogenise one's own language, in itself and in comparison to others, that contributes to variation in language. Another factor that counteracts standardisation is the desire of scribes to give texts regional flavours, or to mark them as having emerged from a particular school of scribes. Stenroos (2013) has shown this phenomenon for Middle English scribes who wilfully introduce linguistic forms into legal documents which do not conform to the supralocal standardised variety but which are part of a local dialects. They do this in order to produce 'copies with a local identity', perhaps as a means to assert their authenticity by giving documents regional rooting, or out of a sense of regional pride.

Standardisation therefore cannot be seen as an unavoidable, linear process. Rather, there are factors that may aid de-standardisation, such as regionalisation, and deliberate efforts of scribes to vary the language in which they are writing, perhaps similar to how we in modern times paraphrase particular words in adjacent sentences. Yet, at the same time, within a network of writers certain people may attempt to alter their own individual writing style. Nissim b. Ḥalfon, an 11th-century Genizah trader who left behind dozens of mercantile letters, chooses to write Arabic *ẓā'* with the Hebrew letter ט for the first decades of his career, but in his later letters switches to the more popular צ employed by the majority of contemporary 11th-century Jewish Egyptian writers of mercantile correspondence

8 In my years of cataloguing Genizah documents I have seen many such examples, sometimes three or four different variations of the same word appear in one document by the same scribe.

(see Wagner 2010, 30). So within networks of writers we may indeed find efforts to produce a linguistic standard universal to that particular group.

4 Early Judaeo-Arabic orthography

The earliest works composed in Judaeo-Arabic are papyri from the 8th/9th century,[9] which have been analysed and published by Blau and Hopkins (1987, 1988). These extant Judaeo-Arabic papyri are composed in the Early Judaeo-Arabic phonetic orthography, which is based on phonetic principles and shows no influence from Classical Arabic orthography (Blau and Hopkins 1984, 124 and Hary 1996, 731). This lack of influence is not surprising yet the orthographical choices of those early Jewish writers cannot reasonably be compared to medieval material as we still do not know enough about the spelling conventions in the Muslim sources in the early Islamic period.[10] The Arabic language reforms of the 9th century anachronistically superimpose an impression of Classical Arabic normative rules onto the earlier centuries, yet the linguistic reality was very different. Ideas on this topic have been expressed by various scholars in the last years, in particular in the collected volume on Middle Arabic edited by Jérôme Lentin and Jacques Grand'Henry (2008), which clearly demonstrate that many phenomena normally attributed to Middle Arabic are in fact early Islamic writing conventions. Further research is therefore still urgently needed to inform our understanding of early Islamic written Arabic.

In the early, phonetically based writing efforts of Judaeo-Arabic the consonantal correspondences between Arabic and the Hebrew alphabet are as follows:

9 The papyri were initially thought to have been written in the 9th century, but were subsequently estimated as being of earlier provenance.

10 Early Muslim documents have been analysed and edited by Khan (1992, 1993, 2007), Hopkins (1984) and Grob (2010), but most early texts come from the middle of the second century of the Islamic era, i.e. the late 8th century.

ا	ʾ	א	ض	ḍ	ד
ب	b	ב	ط	ṭ	ט
ت	t	ת	ظ	ẓ	ד
ث	ṯ	ת	ع	ʿ	ע
ج	j	ג	غ	ġ	ג
ح	ḥ	ח	ف	f	פ
خ	ḫ	כ	ق	q	ק
د	d	ד	ك	k	כ
ذ	ḏ	ד	ل	l	ל
ر	r	ר	م	m	מ
ز	z	ז	ن	n	נ
س	s	ס	ه	h	ה
ش	š	ש	و	w	ו
ص	ṣ	צ	ى	y	י

The letter *dalet* ד thus represents Classical Arabic ظ <ẓ>, ض <ḍ>, ذ <ḏ> and د <d>. The difference between the various consonant qualities in Arabic is not marked with supralinear signs on top of the Hebrew letter, nor is it marked in the letter *tav* ת representing both ت <t> and ث <ṯ>. Yet it is difficult to assess to what extent this reflects the phonetic nature of contemporary spoken Arabic as there is also a lack in graphic distinction between phonetically distinct ج <j> and غ <ġ> both written as ג. In addition, it is hardly imaginable that ظ <ẓ>, ض <ḍ>, ذ <ḏ> and د <d> would all have had the same place of pronunciation. As Blau and Hopkins (1987, 133) point out 'dalet must have polyphonic function' in the papyri. Yet, the writers chose to use the same letter for all the phonetic realisations of various letters they perceived were best represented by *dalet* ד.

The Arabic article *al* is spelled phonetically such as ארחמן for <*al-raḥmān*> *ar-raḥmān*. This is a main point of difference to the later emerging Classical Judaeo-Arabic, which only very rarely spells the article in non-morphophonematic spelling, i.e. for example אלרחמן, often also separated אל רחמן. Short vowels, in particular [i] and [u], are frequently spelled *plene*, whereas long ā is spelled defectively (also a common feature of Early Islamic Arabic).[11]

11 For a more extensive in-depth analysis, see Blau and Hopkins 1987, in particular 124–151.

5 Transcoding medieval Judaeo-Arabic

The early phonetic writing appears to have largely been replaced by the 9th and 10th centuries, depending on the geographic region. The change came from the East, associated with the transition in writing materials. The 9th century in the area that covers modern Iraq was characterised by a surge in literary production. One of the factors that aided this increase in intellectual activity was probably the introduction of paper to the country. Paper as a writing material is transformative as it facilitates the education of a large mass of people. It is probably no coincidence that the 9th-century sees the rise of the Arabic grammarians, and advent of the standardisation of the Arabic language.

Thus, with the arrival of paper and the surge of literary production, efforts were made to standardise writing. Not only do we see this development in Arabic, but also in Judaeo-Arabic. The new Judaeo-Arabic norms that were being introduced in the 9th century were influenced by contemporary Arabic writing conventions. In contrast to the earlier Judaeo-Arabic writing efforts, the newly emerging Judaeo-Arabic standard was heavily Arabicised and written in an orthography where Classical Arabic spelling conventions were imposed on the Hebrew letters. Relatively little is known from which scribal schools this new standard emerged, but it is clear that the one book that made this style of writing popular was the Bible translation by Saadiah Gaon, a resident of Sura and Baghdad in what is modern Iraq. His Judaeo-Arabic version of the Holy Scriptures spread far beyond his own country, and was 'quickly found everywhere throughout the communities of the Near East, North Africa and Muslim Spain, which attest to the fact that it acquired an authoritative, almost canonical, status among all Arabic speaking Rabbanite communities' (Vollandt 2014, 69). Saadiah's Bible translation was not entirely original and probably based on earlier Bible translations, but the 'pinnacle in an evolutionary process in which first oral and then written Judaeo-Arabic translation had emerged since the early days of the Arabic conquests' (Vollandt 2011, 11).

Like the translations of scriptures in so many other languages, Saadiah's Judaeo-Arabic version of the Holy Scriptures became linguistically extremely influential. Because of its popularity it started to suppress the previous, phonetic writing system people had used, set the new standards of writing Arabic across the Islamic empire and became the norm for literary Judaeo-Arabic for the next centuries. This normative, Arabicised spelling is called Classical Judaeo-Arabic by most scholars.

The use of the Hebrew alphabet meant that particular Arabic orthographical traditions were easily abandoned, and it also facilitated the influence of Hebrew norms on written Judaeo-Arabic. On the other hand, Muslim standards were occa-

sionally applied as a means to alter the register of written Judaeo-Arabic texts, and thus, the proximity to Muslim traditions varies considerably in the different genres, and throughout time.

As in the earlier phonetic alphabet, the writers of Judaeo-Arabic had to accommodate the fact that the Hebrew alphabet consists only of 22 graphemes, while Arabic possesses 28, in Judaeo-Arabic a number of Hebrew graphemes were each called upon to represent two (or more) Arabic graphemes in the newly emerged Judaeo-Arabic orthography. This concerns the graphemes ג, ד, כ, צ, ט, and ת. In most cases, one of the set is provided with a dot above the grapheme whereas the other is not. Those graphemes supplied with a dot are in many cases those equivalent to Classical Arabic ث <ṯ>, ج <j>, خ <ḫ>, ذ <ḏ>, ض <ḍ> and ظ <ẓ> (ת̇ ג̇ כ̇ etc.), whereas Classical Arabic ت <t>, د <d>, ص <ṣ>, ط <ṭ>, غ, <ġ>, and ك <k> are without dot (ת ג כ etc.). Some writers, however, indicate *ġ* with the dot and not *j*. Others point both *ġ* and *j*.

ا	ʾ	א	ض	ḍ	צ̇
ب	b	ב	ط	ṭ	ט
ت	t	ת	ظ	ẓ	צ̇ ט̇
ث	ṯ	ת̇	ع	ʿ	ע
ج	j	ג̇	غ	ġ	ג̇
ح	ḥ	ח	ف	f	פ̇
خ	ḫ	כ̇	ق	q	ק
د	d	ד	ك	k	כ
ذ	ḏ	ד̇	ل	l	ל
ر	r	ר	م	m	מ
ز	z	ז	ن	n	נ
س	s[12]	ס	ه	h	ה
ش	š	ש	و	w	ו
ص	ṣ	צ	ى	y	י

In contrast to the earlier Judaeo-Arabic orthography, it is very obvious that the Arabic alphabet undoubtedly served as a template. In Early Judaeo-Arabic, ض <ḍ>[12]and ظ <ẓ> for example, would have been written with *dalet* ד, which was perhaps closer to the actual pronunciation in the spoken dialects. However, the literary standards of Classical Arabic exerted their influence in Classical Judaeo-Arabic, and thus the letter that corresponded in form to the one used in the Arabic alphabet was chosen, i.e. the graphemes צ or צ̇ were employed for ض. A slightly

12 In particular in letters from Byzantium, we also find ש for s/ س, see Outhwaite 2009, 214.

more complex situation is found in the spelling of Classical Arabic ظ. Because ض <ḍ> and ظ <ẓ> had merged in the spoken language probably fairly early, or perhaps, more precisely, because the difference had phonetically never been realised at all to start with in spoken dialects outside the Arabian peninsula, ذ was used to represent both consonants. However, ظ is also found spelled as ث, in analogy with the shape it takes in the Arabic alphabet.

Most importantly, in contrast to the phonetic orthography of the earlier period, Classical Judaeo-Arabic to a large degree also observes Classical Arabic rules concerning the spelling of short and long vowels. Short vowels are only rarely spelled, at least in the early medieval period, whereas long vowels are usually represented in spelling, as in Classical Arabic.

Not all medieval writers, however, follow the rules of Classical Judaeo-Arabic. In particular in letters written in crude hands we very often encounter spellings that are more reminiscent of the earlier phonetic writing conventions, see Blau and Hopkins (1984). Since crude letter writing is associated with a lack of formal scribal education, it is only natural that their writers would not be affected by Arabicised orthographical conventions that are clearly a result of educational standards of the upper classes (Goitein 1971, 346).

6 A middle class of medieval writers

We have an abundance of information concerning the education of medieval Jewish Egyptians from the Genizah sources. Children's education in Hebrew was mentioned above, where all boys but also girls went to schools, with literacy 'exceptionally high' (Olszowy-Schlanger 2003, 47), in particular in urban centres. Once more grown up, the Genizah writers had their skills honed by various channels of secondary education. Goitein (1971, 183–185) has distinguished three different types of professional scribes: first the government clerks working for the chanceries, who were proficient in writing Arabic script, and who received their training within these government institutions (*k*ātib). Secondly, there are scribes who wrote legal documents and letters for the Jewish community and its legal institutions, mostly in Hebrew script (*sofer*). These scribes were trained within the Jewish scriptoria; Wagner (2018) has shown the similarity in handwriting between teachers and students, and the way students appear to have learned by copying out older documents. Arad and Wagner (2013) have also demonstrated how professional scribes working for community dignitaries were the final arbiters of linguistic style used in documents, rather than their socially and politically more prestigious masters. Thirdly, we have scribes who copied books (*nāsiḫ*),

which seems to have been a specialised branch of scribal work. Added to this are traders and physicians, who – out of need for their professions often proficient in Hebrew and Arabic scripts – developed their own writing standards slightly outside of the norms and habits of professional scribes, which includes different styles of handwriting.[13]

The effect of high educational standards for all professional classes is convincingly demonstrated in mercantile letter writing from the 11th- century, which represents by far the largest percentage of medieval Genizah writing.[14] Religious tolerance and economic prosperity in 11th-century Fatimid Egypt produced a large middle class of Jewish traders who conducted business around the Mediterranean. These merchants created their own writing conventions in their extensive business correspondence, which is astonishingly homogeneous on a linguistic level and also displays a strong influence of contemporary Muslim letter writing norms. This is all the more astonishing as traders are usually known for their 'pragmatic literacy' (Parkes 1973, 555) and lack of adherence to superimposed literary norms.

From the way Arabic is used in merchants' letters and from information stated by the traders in their correspondence we can infer that many were very familiar with the Arabic alphabet. Muslim-Jewish business partnerships were common under the Fatimids, making it necessary for the Jewish merchants to deal with Arabic mercantile documents on a daily basis. Letters often have addresses in both Arabic and Hebrew script, displaying how fluid many merchants were in both alphabets.

Not only merchants do display their aptitude in the different alphabets. Some authors use different alphabets almost playfully. For example, Daniel b. Azaryah, an 11th-century leader of the Jewish community, employs Arabic script in letters to particular recipients as a way to save space at the end of line (writing the same expression in Arabic script is a lot more space effective than it would be in Hebrew script, see Fig. 1). Others, for example a writer called Judah b. Abraham, switch between Arabic and Hebrew script when they alternate between the Hebrew and Arabic languages (Fig. 2). We can thus assume schooling in both Hebrew and Arabic scripts for the protagonists of particular professions.

Education in Arabic must be responsible for the linguistic behaviour of particular writers, who follow particular Muslim conventions that are not ordinarily part of the Judaeo-Arabic repertoire. These are typically linguistically conservative

13 For traders, see Wagner 2017.

14 Circa 1000 letters have been edited in Gil's monumental works *In the Kingdom of Ishmael* alone (Gil 1997).

Fig. 1: TS-013-J-026-002-F. © Reproduced by kind permission of the Syndics of Cambridge University Library.

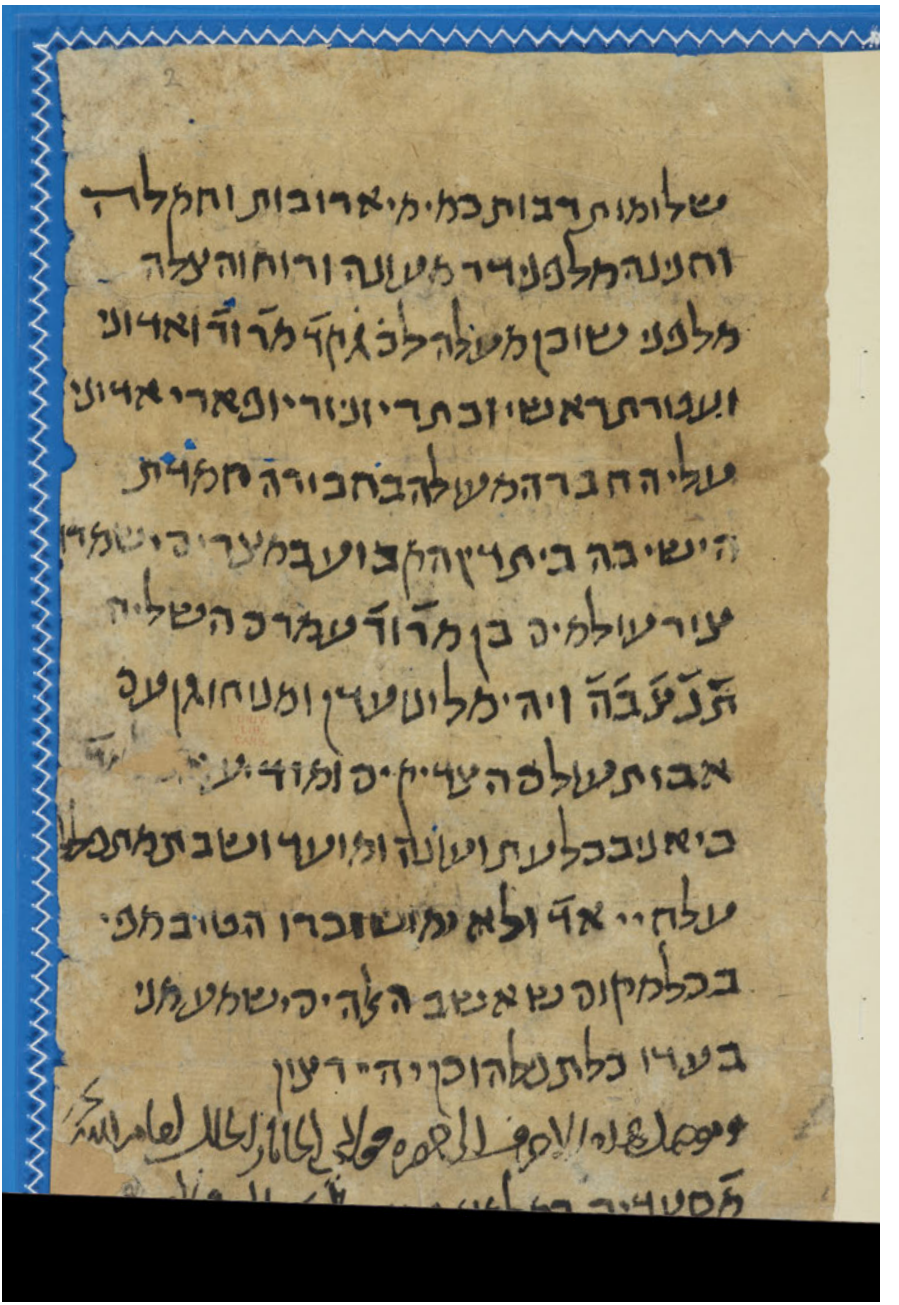


Fig. 2: T-S 13J13.2. © Reproduced by kind permission of the Syndics of Cambridge University Library.

forms, which can be found in otherwise linguistically quite progressive material. This is not unusual: Nevalainen (2013) has shown that individual Middle English writers may adopt some progressive forms, but shun others. A good example is presented by otiose *'alif* – the silent *'alif* at the end of particular verbal forms such as 3rd person plural perfect. This otiose *'alif* features in Saadiah's Bible translation and following that becomes part of Judaeo-Arabic literary texts emulating Saadiah's linguistic standards, yet it is never actually a full part of the repertoire in epistolary writing. In a corpus of more than a hundred medieval Judaeo-Arabic letters, otiose *'alif* only features in the correspondence of three writers: in a circle of educated writers surrounding the 13th-century judge Elijah b. Zechariah. Thus, within Judaeo-Arabic letter writing, otiose *'alif* was used exclusively by small networks of people, perhaps as a means to demonstrate a certain standard of education. The prestige of a Classical Arabic form that is very markedly a part of a high register thus also appears to have held appeal in Jewish writing networks.

Prestige and precedent generally play an important role in the letters. Once a form has been brought to paper, it is then often used by other authors. Writers with high prestige will find the forms they employ being used by those of lower social standing. A good example demonstrating the importance of prestige may be found in the handwriting of scribes and individuals of the late Fatimid and early Ayyubid period. From the 12th-century onwards, the handwriting of Egyptian Jews became considerably more cursive compared to the earlier, squarer way of writing Hebrew script. It has been suggested that this is due to the influx of Spanish Jews into Egypt. In Spain, through the closer contact with Arabic culture, the way of writing Hebrew had been permeated by the cursivity of the Arabic script. Due to the high prestige of Spanish Jews (among them the most famous personality of the Genizah, Moses Maimonides), this trend of cursivity also caught on among the Egyptian Jewry.

7 Changes from the 13th century onward: the 'Hebraisation' of Judaeo-Arabic writing

The homogeny and orientation along the lines of Classical Arabic found in 11th-century mercantile writing contrasts sharply with letter writing in the Ayyubid empire in the 13th century. During that time, we see the economic worsening already experienced in the 12th century coming into full effect, which results in the breakdown of the Jewish middle class. Segregation between Muslims and the minorities in their states increased, which led to a return to religious values and

a more inward-facing attitude within the Jewish communities.[15] This can also be observed in writing: correspondence of the 13th century begins to exhibit a much stronger influence of Hebrew norms, counteracting the Arabicisation imposed in the earlier centuries. These differences can be correlated to the advancing lack of integration of Jews within Egyptian society. In the Jewish communities, increasingly more isolated from other faith communities, the Hebrew influence on language became much stronger and where we used to find Arabic formulae in the introduction and blessings of letters, they have now been largely replaced by their Hebrew counterparts.

On an orthographic level, Hebrew influence is most noticeable in the increase of *plene* spelling of vowels, in the spelling of reduplicated [w] and [y] and in the way particular morphemes are written. For example, the early medieval sources usually follow Classical Arabic conventions and do not mark the reduplication graphically within the *rasm* (graphic line) of the word (although we also find *šadda* used above the Hebrew letters by certain writers). In the spelling of the 3rd person singular masculine suffix pronoun, the spelling ו- <-w> that denotes the suffix in Hebrew occurs as frequently as (and in particular texts even more frequently than) the Classical Judaeo-Arabic spelling ה- <-h>. It is worth noting that the Hebrew spelling conventions that are applied in the latter example are closer to the actual pronunciation of the suffix [-u] in the spoken Egyptian Arabic than the Standard Arabic orthography of [-h]. These phenomena can already be observed in material from the 10th–12th centuries but the increase in frequency over the century from the 13th century onwards is very noticeable (see the table in Wagner 2010, 40).

In addition, vast differences in writing standards can be observed in the various genres of texts from the 13th-century onwards and in particular in the Late Judaeo-Arabic of the Ottoman period.[16] The norms start to be mostly dependent on the type of text, in contrast to the period of the 10th–12th centuries during which Judaeo-Arabic across the genres was largely standardised. For example, the gap between utilitarian prose (in particular letters) and literary works is widening; letters are often written in a much more colloquial language than literary texts, which still follow Classical Judaeo-Arabic standards to a degree. At the same time, there is also enormous variation of forms when different writers are compared to

15 Goitein 1978, 161–162. For the changing attitudes towards minorities, also see Leiser 1976, in particular 68–88.

16 This article follows Khan's (2007, 526) periodisation, who proposes three major phases: Early (9th century), Classical (10th–15th centuries) and Late Judaeo-Arabic (from 15th century onwards).

one another, even though the Genizah contains much fewer sources from the 13th-century onwards as compared to material from the 10th–12th centuries.

8 Conclusions

To sum up: Language provides identity for its speakers and writers, and members of the different confessional groups employ particular forms of speech and written text that mark them as members of a certain social entity. Jews in particular created their own language standards as they, because of education, religious identity and perhaps political pressure, wrote their Arabic texts in the Hebrew alphabet, creating what we call Judaeo-Arabic in the process.

The transcoding of Arabic in Hebrew script found in Judaeo-Arabic texts presents a unique case study for the investigation of standardisation processes, and the most suitable corpus can be found in the documents of the Cairo Genizah. Standardised and varying forms used by the scribes of the Genizah demonstrate that standardisation cannot be seen as an unavoidable, linear process. Various factors may lead to de-standardisation, whereas network-internal linguistic standardisation may prove to be the driving force behind general standardisation processes. The main factor controlling the assertiveness of linguistic forms is perhaps the social prestige of those protagonists setting precedent.

In Judaeo-Arabic, three major orthographical phases can be discerned. In the early Islamic centuries, a *phonetic orthography* is used to write early Judaeo-Arabic. Affected by the standardisation efforts of Classical Arabic in the 9th century, a new *Arabicised orthography* emerged for Judaeo-Arabic too, aided by the spread of Saadiah's Bible translation; this initiated the period of so-called Classical Judaeo-Arabic. Classical Judaeo-Arabic was not only a literary standard but can be found in utilitarian prose, too: the efficiency of high educational standards in the relatively tolerant and prosperous 11th-century Fatimid Egypt produced a large middle class of Jewish traders who produced their own writing conventions in their extensive mercantile correspondence, displaying a strong influence of contemporary supra-communal letter writing norms.

The economic deterioration which began in the 12th and 13th centuries resulted in the breakdown of the Jewish middle class, and eventually led to the *Hebraisation* of Judaeo-Arabic orthography. Segregation between Muslims and the minorities in the Islamicate increased, leading to a return to religious values and a more inward-facing attitude within the Jewish communities. This can also be observed in writing: correspondence of the 13th century begins to exhibit a much stronger influence of Hebrew norms.

In addition, vast differences in writing standards can be observed in the various genres of Judaeo-Arabic from the 13th century onwards, with a widening gap between literary and utilitarian writing. This could be interpreted as indicating that, in the absence of an acknowledged supraregional standard across the Jewish communities, local networks, scribal schools and gathering of writers developed their own norms, which they then followed. In turn, this perhaps points to the lack of a central authority, such as the Babylonian and Palestinian Academies had held in the Geonic period, issuing norms to which writers felt the urge to adhere. In comparison, during the Classical Judaeo-Arabic period, the writing standards used in the various literary and utilitarian genres were relatively homogeneous, and fairly close to those norms used by Muslim contemporaries. This may have also been aided by the employment of Jews in government offices, and accompanying Arabic script education of the Jewish elite. With segregation and the lack of Jewish clerks in official government functions in the late Ayyubid and Mamluk periods, this particular alley of secondary education became unavailable, which must have had a knock-on effect on scribal practice within the Jewish community.

The phenomenon of linguistic de-centralisation from the 13th century onwards is by no means restricted to the Jewish community; comparable processes can be observed in Muslim Arabic writing, too, as can be best seen in Muslim Arabic from the 16th century onwards.[17] Social, political and economic circumstances thus exerted a similar influence on all religious communities, regardless of their linguistic differences.

References

Arad, Dotan, and Esther-Miriam Wagner (2013), Online resource. Fragment of the month January 2013: *The preacher who speaks like a parrot*. http://www.lib.cam.ac.uk/Taylor-Schechter/fotm/january-2013/index.html

Astren, Fred (2009), 'Re-reading the Arabic sources: Jewish history and the Muslim conquests', in *Jerusalem Studies in Arabic and Islam*, 36: 83–130.

al-Bayati, Jaafar (ed.) (1990), *Sirāj al-Mulūk by Muḥammad Ibn al Walīd al-Ṭarṭūšī*. London: Riad el-Rayyes Books.

Blanc, Haim (1964), *Communal Dialects in Bagdad*. Cambridge, Massachusetts: Harvard University Press.

Blau, Joshua, and Simon Hopkins (1984), 'On Early Judaeo-Arabic Orthography', in *Zeitschrift für arabische Linguistik*, 12: 9–27.

17 See for example the analyses of Ottoman Arabic in Lentin 2008.

Blau, Joshua, and Simon Hopkins (1987), 'Judaeo-Arabic Papyri', in *Jerusalem Studies in Arabic and Islam*, 9: 87–160.

Cohen, Mark (1999), 'What was the Pact of Umar? A Literary-Historical Study', in *Jerusalem Studies in Arabic and Islam*, 23: 100–131.

Gil, Moshe (1997), *Be-malkhut Yishma`el bi-tekufat ha-ge´onim* [*In the Kingdom of Ishmael*], Jerusalem: Mosad Byalik.

Goitein, Shelomo Dov (1971), *Mediterranean Society. The Jewish Communities of the Arab World as Portrayed in the Documents of the Cairo Genizah. Vol II: The Community*, Berkeley: University of California Press.

Goitein, Shelomo Dov (1978), *Mediterranean Society. The Jewish Communities of the Arab World as Portrayed in the Documents of the Cairo Genizah. Vol III: The Family*, Berkeley: University of California Press.

Grob, Eva-Mira (2010), *Documentary Arabic and Private Business Letters on Papyrus: Form and Function, Context and Content*, New York, Berlin: de Gruyter.

Hary, Benjamin H. (1996), 'The Adaptations of Hebrew Script', in Peter T. Daniels and William Bright (eds), *The World's Writing Systems*, New York: Oxford University Press, 727–734.

Hopkins, Simon (1984), *Studies in the Grammar of Early Arabic*, Oxford: Oxford University Press.

Khan, Geoffrey (1992), *Arabic Papyri: Selected Material from the Khalili Collection*, London: The Nour Foundation in association with Azimuth Editions and Oxford University Press.

Khan, Geoffrey (1993), *Bills, Letters and Deeds: Arabic Papyri of the 7th to the 11th Centuries*, London: The Nour Foundation in association with Azimuth Editions and Oxford University Press.

Khan, Geoffrey (2007), 'Judaeo-Arabic', in *Encyclopedia of Arabic Language and Linguistics*, Leiden: Brill, 526–536.

Khan, Geoffrey (2013), 'The historical development of early Arabic documentary formulae', in Esther-Miriam Wagner, Ben Outhwaite and Bettina Beinhoff (eds), *Scribes as Agents of Language Change*, Berlin: de Gruyter-Mouton, 199–215.

Kretzschmar, William A. (2009), *The Linguistics of Speech*, Cambridge: Cambridge University Press.

Leiser, Gary La Viere (1976), *The Restoration of Sunnism in Egypt: Madrasas and Mudarrisūn 495–647/1101-1249*. PhD thesis, University of Pennsylvania.

Lentin, Jérôme (2008), 'Unité et diversité du moyen arabe au Machreq et au Maghreb: quelques données d'après des textes d'époque tardive (16ème–19ème siècles)', in Jérôme Lentin and Jacques Grand'Henry (eds), *Moyen arabe et variétés mixtes de l'arabe à travers l'histoire*. Louvain-la-Neuve: Peeters Press, 305–319.

Lentin, Jérôme and Jacques Grand'Henry (eds) (2008), *Moyen arabe et variétés mixtes de l'arabe à travers l'histoire*, Louvain-la-Neuve: Peeters Press.

Nevalainen, Terttu (2013), 'Words of kings and counsellors: register variation and language change in early English courtly correspondence', in Esther-Miriam Wagner, Ben Outhwaite and Bettina Beinhoff (eds), *Scribes as Agents of Language Change*, Berlin: de Gruyter-Mouton, 99–119.

Olszowy-Schlanger, Judith (2003), 'Learning to read and write in medieval Egypt: Children's exercise books from the Cairo Genizah', in *Journal of Semitic Studies*, 48: 47–69.

Outhwaite, Benjamin (2009), 'Byzantium and Byzantines in the Cairo Genizah: New and Old Sources', in Nicholas de Lange, Julia G. Krivoruchko and Cameron Boyd-Taylor (eds), *Jewish Reception of Greek Bible Versions. Studies in Their Use in Late Antiquity and the Middle Ages*, Tübingen: Mohr Siebeck, 182–220.

Parkes, Malcolm (1973), 'The literacy of the laity', in David Daiches and Anthony Thorlby (eds), *The Mediaeval World*, London: Aldus Books, 555–577.

Stenroos, Merja (2013), 'Identity and Intelligibility in Late Middle English scribal transmission: local dialect as an active choice in fifteenth-century texts', in Esther-Miriam Wagner, Ben Outhwaite and Bettina Beinhoff (eds), *Scribes as Agents of Language Change*. Berlin: de Gruyter-Mouton, 159–181.

Vollandt, Ronny (2007), *The Transmission of the Judaeo-Arabic Pentateuch Translation of Rav Saadiah Gaon in Arabic letters: A Case of Textual Diffusion*. MA Thesis Hebrew University of Jerusalem.

Vollandt, Ronny (2014), 'Whether to capture form or meaning: A typology of early Judaeo-Arabic Pentateuch translations', in Nadia Vidro, Irene Zweip and Judith Olszowy-Schlanger (eds), *A Universal Art. Hebrew Grammar across Disciplines and Faiths*. Leiden: Brill, 58–83.

al-Ṭarṭūšī, Muḥammad Ibn al Walīd (1872), *Sirāj al-Mulūk*. Cairo (Bodleian library Arab.d.158).

Wagner, Esther-Miriam (2010), *Linguistic Variety of Judaeo-Arabic in Letters from the Cairo Genizah*, Leiden: Brill.

Wagner, Esther-Miriam (2013), 'Challenges of Multiglossia: the emergence of substandard Judaeo-Arabic registers', in Esther-Miriam Wagner, Ben Outhwaite and Bettina Beinhoff (eds), *Scribes as Agents of Language Change*, Berlin: de Gruyter-Mouton, 259–273.

Wagner, Esther-Miriam (2014), 'Subordination in 15th- and 16th- century Judeo-Arabic', in *Journal for Jewish Languages*, 2: 143–164.

Wagner, Esther-Miriam (2017), 'The socio-linguistics of Judaeo-Arabic mercantile writing', in Esther-Miriam Wagner, Bettina Beinhoff and Ben Outhwaite (eds), *Merchants of Innovation. The Languages of Traders*, Berlin: de Gruyter-Mouton, 68–86.

Wagner, Esther-Miriam (2018), 'Scribal practice in the Jewish community of medieval Egypt', in Myriam Wissa (ed.), *Scribal Practices and the Social Construction of Knowledge in Antiquity, Late Antiquity and Medieval Islam*, Leuven: Peeters, 91–109.

Wagner, Esther-Miriam, Ben Outhwaite and Bettina Beinhoff (2013), 'Scribes as agents of language change', in Esther-Miriam Wagner, Ben Outhwaite and Bettina Beinhoff (eds), *Scribes as Agents of Language Change*, Berlin: de Gruyter-Mouton, 1–18.

Paolo La Spisa

Cross Palaeographic Traditions. Some Examples from Old Christian Arabic Sources

Abstract: This chapter deals with early Palestinian Arabic manuscripts from South Palestinian monastic collections now kept in the library of St Catherine's Monastery, Egypt. The aim of this study is to demonstrate that in the early Arabisation process of the Melkite Palestinian Church (8th–9th century), it is possible to find palaeographic, linguistic and layout features testifying to inter-faith interaction. Accordingly, the text of the holy book of Islam might have played an important role for the Arabised Melkite communities of Palestine. The early activity of translating the Bible and the Patristic and ascetic heritage into Arabic proved to be an important stage in the acquisition of the Arabic writing technique by Melkite monks living in the Caliphate. By comparing Islamic and Christian sources, I try to cross the all too narrow confessional boundaries in which 'Christian Arabic studies' have been confined for the last two centuries.

1 Historical introduction

The Melkite Church was the first eastern Christian church living in the Arab world that adopted Arabic as its liturgical language.[1] After the Arabic conquests, the Patriarchate of Jerusalem became part of the caliphate's territory; subsequently within the increasing 'Arabic-speaking Melkite community, Jerusalem and its monasteries effectively became the centre of a doctrinal development' (Griffith 2006, 185). Actually, even though Greek was the language that symbolically preserved the links with Byzantine orthodoxy beyond the borders of the caliphate, Palestinian monks had carried on the enterprise of translating the Bible into local tongues since pre-Islamic times (Griffith 1997; Briquel-Chatonnet and Le Moigne 2008). This means that Palestinian monasteries such as St Saba and St Kariton in the Judean desert had always preserved their local identity against the Greek culture of Constantinople. Accordingly, the Arabisation of the church of Jerusalem after the rise of Islam had a double function: to build an Arab Orthodox identity,

1 For a short history of the Arabic-speaking Melkite Orthodox Church see Griffith 2006, who clarifies why from the 8th century onward this church decided to translate its religious heritage from Greek into Arabic.

https://doi.org/10.1515/9783110639063-004

sociologically and culturally, albeit not doctrinally, distinguishable from their Greek Orthodox co-religionists on the one hand, and to be able to produce an apologetic literature in Arabic to cope with the new religious challenge of Islam, on the other. These are the reasons why, within the Melkite Jerusalem Patriarchate, South Palestinian monasteries were the cradle where Christian Arabic literature had its origins. The manuscripts that were once in their libraries, are now collected and preserved in St Catherine's Monastery in the Sinai Peninsula (Samir 1990, 1990–1991; Lafontaine-Dosogne 1996; Géhin 1998; Mouton 2000, 105–124; La Spisa 2008; 2012, 210–213).

In what follows I will examine some palaeographic features in manuscripts belonging to the St Catherine Library and directly originating from the South Palestinian monastic milieu. After having very briefly outlined the cultural framework in which they were produced, I will try to show to what extent the standardisation of the orthography, script and layout has been influenced by the orthography and language of the Qur'an even in the Christian Arabic manuscripts of the first millennium coming from the Arabic Melkite Church. Subsequently I will try to assess to what extent it is possible to speak of Christian Arabic features within a Muslim religious and cultural environment.

2 Qur'anic orthography and early Arabic manuscript tradition

The early Arabisation of the Greek Orthodox Church in Palestine led the monks to carry out the very first translations of the Bible into Arabic, so we have evidence of Palestinian Arabic translations dating back to the second half of the 8th century.[2] In this context it is legitimate to pose the following question: what kind

2 Opinions about the existence of a pre-Islamic Arabic translation of the Bible diverge: Baumstark (1929–1931) and Shahid (1995–2009) assumed that even though no material evidence is available, it is reasonable to think that such work had been accomplished at least during Muḥammad's lifetime. The issue of the existence of such translations received attention once again thanks to a recent study by Sidney Griffith, who concludes (2013, 41–42): 'no conclusive documentary or clear textual evidence of a pre-Islamic, written Bible in Arabic translation has yet come to light'. Nevertheless, the several Qur'anic references to biblical, hagiographical and homiletic literary traditions are undeniable. Griffith explains this phenomenon by stating that in pre-Islamic time and during Muḥammad's lifetime there was an oral transmission of the Jewish and Christian scriptural and homiletic traditions which were directly and spontaneously translated into Arabic for an Arabic-speaking audience. However, this does not exclude the existence

of Arabic did the scribes adopt for their translations, since the Arabic language before the 9th–10th centuries had not yet been normalised by the Iraqi philologists of Baṣra and Kūfa (Fleisch 1990, 1–15, Ferrando 2001, 117–133)? The first evidence of Arabic manuscripts surviving up to the present comes from a few copies of the Qur'an reportedly dating back to the 7th century (Déroche 2004, 16). So is it legitimate to suppose that the holy book of Islam had influenced even the Arabised Christian copyists?

If we have a look at some Qur'anic *Sūra*s, it is possible to single out some linguistic and palaeographic phenomena that western scholars have described as Middle/Mixed Arabic features (Lentin 1997, 2008, 2012). In the *Sūrat al-naḥl* (Q.16: 72) we read: *wa-bi-niʿmati-llāhi hum yakfurūn* 'do they repudiate the divine grace?' where the word *niʿmati* is written with a *tāʾ mabsūṭa* instead of *tāʾ marbūṭa*, the same phenomenon can be found in medieval Christian Arabic texts (Blau 1966, 115–116). However, in the *Sūrat al-shuʿarāʾ* (Q.26: 22) one can find the same word written with *tāʾ marbūṭa*: *wa-tilka niʿmat-un* 'is it a favour...?'. Both orthographic variants are well attested. One can suppose that the *tāʾ mabsūṭa* is used only in annexations, but in the *Sūrat al-ḍūḥā* (93: 11) we can read: *wa-ʾammā bi-niʿmati Rabbika fa-ḥaddiṯ!* 'but as for the favour of your Lord, report [it]!', where the same word in annexation is written with *tāʾ marbūṭa*.

The second example is taken from the *sūrat al-ʾisrāʾ*(17: 1) where we read: *subḥāna llaḏī ʾasrā bi-ʿabdihi layl-an mina-l-masǧidi-l-ḥarāmi ʾilā-l-masǧidi-l-ʾaqṣā* 'Exalted is He who took His servant by night from the Sacred Mosque to the Farthest Mosque'. The last word *al-ʾaqṣā* is an elative form of the adjective *qaṣiyy* 'faraway', which literally means 'the farthest'; however instead of *alif maqṣūra* at the end of the word as found in current Arabic orthography, there is an *alif ṭawīla* (cfr. Blau 1966, 81–82).

Finally, as far as syntax is concerned, in the *Sūrat al-māʾida* (Q. 5: 69) we read: *ʾinna llaḏīna ʾāmanū wa-llaḏīna hādū wa-l-ṣābiʾūna wa-l-naṣārā* 'Indeed, those who have believed and those who are Jews or Sabaeans or Christians'. Following the Classical and Modern Standard Arabic grammatical rule, one should expect to find the name 'Sabaean' to be in the oblique case since it is governed by *ʾinna*

of Arabic written notes by Christian literate monks and priests as *aides de mémoire*, as Schoeler 2002 has suggested. About the early Arabic translations of the Gospel see Guidi 1888, Arbache 2007, Griffith 1985, 2008, 2013, Schulthess 2018. The earliest Arabic Gospel has been recently identified by Kachouh (2012) in the *Vatican Arabic 13* which was copied in the Judean desert monastery of St Saba around the year 800 CE. The earliest New Testament Arabic version known so far is *Sinai Arabic 154*, whose second section contains the earliest Christian Arabic apologetic treatise, of 788 CE. Samir 1994; Swanson 1993; La Spisa 2014.

(cfr. Blau 1967, 326).[3] Another example could be taken from the *sūrat al-nisāʾ* (Q. 4: 162) where one can read: *wa-l-muʾminūna yuʾminūna bi-mā ʾunzila ʾilayka wa mā ʾunzila min qablika wa-l-muqīmīna al-ṣalāta wa-l-muʾtūna al-zakāta*... 'But the believers believe in what has been sent down to you and what was sent down before you, and those who perform the prayer and give alms [...]'. According to the Standard Arabic rules as well as to the context and the meaning of this verse, one should expect to find *wa-l-muqīmūna*, in the nominative case of the regular masculine plural (*al-marfūʿ bi-l-wāw wa-l-nūn*) as it is the case of the other nouns of the verse which have the same syntactical function (*wa-l-muʾminūna*, *wa-l-muʾtūna*). As we shall see from the following examples, all these variant forms are also frequent in written Middle Arabic of the pre-modern era.

Concerning the orthographic issue of the *tāʾ marbūṭa*, also in Christian Arabic texts, *tāʾ mabsūṭa* instead of *tāʾ marbūṭa* and also vice versa, is found: بقوت روح القدس 'by the strength of the Holy Spirit'; حيات يسوع 'the life of Jesus' (cfr. Blau 1966, 115).

Alif maqṣūra instead of *alif ṭawīla*: والثكلا اذا حزنت لبست السواد 'when the woman who lost her son is sad, she dresses black clothes'; as is well known, according to the standard orthographic rules, the feminine form of ثكلان is ثكلى with *alif maqṣūra*.

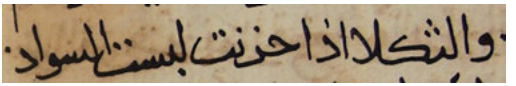


Fig. 1: *Dayr al-Muḫalliṣ 1807* (1643–44 CE) - Ǧūn (Lebanon), fol. 387v l. 7.

In the 10th–11th centuries Melkite bishop of Gaza Sulaymān al-Ġazzī's treatise on the holy Cross, one can read the following *incipit*: *iʿlamū ayyuhā l-mutaqallidīna nāmūs ṣalīb al-Masīḥ*... 'Know, you who abide by the Law of the Cross of Christ...' (La Spisa 2013, 1) where one should have expected to read *ayyuhā l-mutaqallidūna*.[4]

3 This verse should be compared with two others, which are very similar to each other: that of the *Surat al-baqara* (Q. 2, 62): *ʾinna-llaḏīna ʾāmanū wa-llaḏīna hādū wa-naṣārā wa-ṣābiʾīna, and that of the Surat al-ḥaǧǧ* (Q. 22, 17): *ʾinna-llaḏīna ʾāmanū wa-llaḏīna hādū wa-ṣābiʾīna wa-naṣārā*, where the word *ṣābiʾīna* is written according to the rule of *ʾinna*. For further details on these verses and their interpretation, see Burton 1988, 188–196 and Abdel Haleem 1992, 425–427.

4 All grammars of Classical and Standard Arabic agree by stating that أَيُّها and يا أَيُّهَا require after them a noun, singular, dual or plural, defined by the article, and in the *nominative* case.' (Wright 1962, 92–93, emphasis mine), see also Veccia Vaglieri (1937 [1996⁷], I, 135; II, 173). For having an idea on the discussions about the *nidāʾ* among Arab grammarians, see al-ʾAnbārī (1997, 128). For further examples in addition to the essential work in the field of Middle Arabic by Joshua Blau 1966–1967, 2002, see also Hopkins 1984; Lentin 1997, 2008, 2012; Grand'Henry 2006, which are

In 8th–9th century Arabic manuscripts nowadays kept in St Catherine's Monastery, it is possible to find all these aforementioned orthographic and linguistic features that lead us to think that, at the very beginning of Arabisation, the written language was the same for all religious communities (den Heijer 2012). This statement can be demonstrated by comparing sources dating back to the same period but emerging from different confessional and cultural backgrounds. The same conclusions could be formulated also for Arabic palaeography. In what follows I demonstrate that sources belonging to different religious milieus actually share the same palaeographic features.

In the first Abbasid era, the most widespread kind of Arabic script was the so-called *kūfī*, or, as Déroche (1987–1989, 353–354) has labelled it, *écritures abbassides anciennes* 'early Abbasid scripts' (Gacek 2009, 97–98), whose most relevant peculiarities are:

1. the isolated or final *alif* with a more or less developed extension below the line,
2. *dāl* with two parallel and horizontal rods,
3. the median *ʿayn* whose head is constituted by two antennas,
4. final *mīm* with a horizontal tail.[5]

In Christian Arabic manuscripts of St Catherine dating back to the same period, it is possible to find many examples of codices written in what scholars have called Sinaitic *kūfī* or Sinaitic-Palestinian *kūfī*. Some scholars have supposed that this kind of script was originated or directly influenced by the Syriac *estrangelo* script that was also used in Palestinian monastic *scriptoria*.[6] Nevertheless, by comparing different sources as Déroche has done, a great similarity between the 'Islamic' and 'Christian' variants of the so-called early Abbasid script comes to the fore. I will examine in detail some orthographic features in order to show this similarity.

only a few examples selected from the extended literature which developed in these last decades.

5 See also Déroche 2000, 234, Déroche and Sagaria Rossi 2012, 164–167.

6 Gacek (2009, 1) supposes that 'the origin of these scripts [Abbasid bookhand] are most likely traceable to the first century of Islam and some of them appear to have been influenced by the Syriac *sertā* script'. Many scholars tackled the issue of the origin of the Arabic script in the last century. Two hypotheses have been formulated: the first one identifies in the Nabatean inscriptions the origin of the Arabic letters (Cantineau 1930–32; Abbott 1939; Gruendler 2006); the second hypothesis says that the early Kūfī scripts are derived from the *estrangelo* Syriac script (de Sacy 1810; Starcky 1966; Troupeau 1991; Briquel-Chatonnet 1997; Noja Noseda 2006), without making any distinction between Christian or Islamic Arabic sources.

3 The case of the *qāf* and *fāʾ* diacritical points

Father Khalil Samir (1991) described for the first time some palaeographic features of early Christian Arabic apology as attested in the *Sinai Arabic 154*.[7] The two most important phenomena that struck him were the way of writing *qāf* and *fāʾ* and the presence of split words, elsewhere unusual in Arabic.

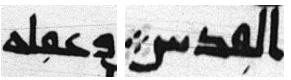


Fig. 2: Sin. Ar. 154 (fol. 101r ll. 1 and 7).

Samir remarked that, in this 8th century parchment codex, the *qāf* is always written with a dot below the line, while the *fāʾ* is written in the regular way (a dot above the letter). Samir (1994, 60) concluded that: 'The way the *qaf* is written seems to be absolutely unique in the Arabic script'. However, Monferrer Sala (2010, 197) carried out a little inquiry about this palaeographic feature within the same manuscript and pointed out the same phenomenon in at least two other Sinai codices: the parchment *Sinai Arabic NF perg. 17* belonging to the new finds of the St Catherine Monastery (Meimaris 1985, 27 [Greek] and 25 [Arabic]), and *Sinai Arabic 1* which is a translation into Arabic of some books of the Old Testament. Both codices date back to the 9th century. Within the same St Catherine manuscript collection, we can also add as an example the *Sinai Arabic 36*, which is a bilingual (Greek-Arabic) Psalter copied in the 8th–9th centuries having the same palaeographic features. Unfortunately reproductions of this precious codex are not available. Thanks to the *specimen* of fol. 10r published in Lafontaine-Dosogne (1996, 110), I could identify the same way of writing *qāf* and *fāʾ*. Monferrer Sala (2010, 197) concluded that this way of marking the *qāf* is a 'feature characteristic of early South Palestinian texts'.[8] Nonetheless it is noteworthy to remark that the same feature has also been found in Islamic sources and documents. Nabia Abbott (1967) has published some Islamic papyri dealing with Islamic traditional

7 For the edition of the Apology see Gibson (1899), with an English translation entitled 'Treatise of the Triune Nature of God'. On this Apology and its historical and religious context, see in particular Griffith 1985.

8 The conclusions which Monferrer Sala reached could induce one to think that this feature belongs exclusively to the Christian Arabic writing tradition (cfr. D'Ottone 2015, 271). Actually, this phenomenon seems to be cross-confessional.

literature nowadays kept at the Oriental Institute of the University of Chicago. The papyrus nr. 17.630, dating back to the 9th century, is a fragment of Islamic traditions (*ḥadīṯ*). Abbott (1967, 208) underlines the following palaeographic feature: 'Once each, *fa* and *qaf* have a dot above and below respectively'. Other non-Christian examples can be found in several other manuscripts kept in the National Library of Paris, in Istanbul and in Saint-Petersburg (Déroche and Sagaria Rossi 2012, 181).

It is interesting to remind that the *qāf* with the subscribed point can be also compared with the so-called *maghribī* way of writing *fāʾ* (with a point below) and *qāf* (with a point above). If we consider the history of the *maghribī* script and its origins, it is not astonishing to find several examples of the same typology even in eastern manuscripts, belonging both to Christian and Muslim traditions. It is noteworthy to mention an example of this script in what is considered a very old translation of the Gospel into Arabic, now kept at the Vatican Library: the *Vaticanus Borgianus 95*. Despite the relevance of this witness, an in-depth codicological description of the manuscript is still lacking.

This codex[9] is a parchment dating back to the 8th or the beginning of the 9th century. At the end of the 19th century Guidi (1888, 10) argued that it had probably been copied in the St Saba Monastery, in the Judaean Desert of Palestine. Belonging originally to the 'Collegio de Propaganda Fide', it was at a later stage part of the Borgian Museum taking the catalogue number *K. II. 31* before having the present number *95*. It is folded *in-quarto*, its dimensions are about 215 × 160 mm., the written area varies between 170 × 125 and 190 × 135 mm. with about 16–17 lines per page. Nowadays the codex contains 171 folia. I found 23 quires, all of which are quaternions.

Regarding the palaeographic features, it is possible to say that the codex presents a *maghribī* 'look'. Its script is an early Palestinian *naskhī*:[10] the *fāʾ* is written with a point below and *qāf* with a point above. As an example, see fol. 16v. line 4: *wa-yašfū*, line 8: *Yaʿqūb*, line 13: *qāyilīn* / *qāʾilīn*;[11] but the *qāf* is also written with two points above. The final *alif* is marked with a rod below the writing line. If one compared this witness with the western copies of the Qur'an, it should not be difficult to recognise many strong similarities with the script of the Arabic

9 Because of the antiquity of this manuscript, many scholars have discussed it: see as an example Guidi 1876–1877, 1888; Tisserant 1914, 55; 1924; Graf 1944, 142, 148; Metzger 1977, 262–263; Griffith 1985, 154–155; Orsatti 1996, 153.

10 Another example of the same kind of *naskhī* is in the *British Museum Or. 5019* (10th cent.).

11 For a sample of this very folio see Tisserant 1914, 55.

Gospel of the *Borgianus 95*.[12] Accordingly, as Déroche has clearly illustrated, it is possible to suppose that the script called *maghribī* today actually had an oriental origin. Afterwards the Maghreb preserved it with some minor regional changes.[13]

By way of some final considerations about this question, it should be mentioned that the diacritical points are randomly used in most of the quoted manuscripts. However, as regards the case of the *qāf*, it is possible to find it without points, with two points above and with a point below in the same document, if not in the same folio, as it is the case for the *Sinai Arabic 1*. As Monferrer Sala has rightly pointed out, in this manuscript the verb *qāla* is regularly written with a subscribed point, however in fol. 1r one can find the following words where the *qāf* is written with two points above: fol. 1r line 4: *ḫalaqa* 'he created'; fol. 1v line -4: *al-sarrāq* 'the thief'; fol. 2r line -1: *fawqa* 'above'. The same alternation can be found in the *Borgian 95*.

From what precedes it is possible to argue for the following hypothesis. Between the 7th and 9th centuries, the standardisation of diacritical points was not yet established. This explains why in manuscripts dating back to this period one can find at least four different ways of writing the letter *qāf* which alternate quite frequently: 1) without points, 2) with one point above (the so-called *maghribī* variant), 3) with one point below, 4) with two points above (which became the standard form). This alternation and fluctuation can exist even within the same document.[14] As Déroche (2004, 73) pointed out, in this very period there were constant movements of scribes between East and West. This may explain the eastern origin of the graphic variant to write the *qāf* which afterwards became characteristic of the *maghribī* script. On the other hand, in the East the standardisation of the language by Iraqi philologists stabilised the spelling of the *qāf* with two dots above, causing the disappearance of the other ways of writing this letter.[15] If this hypothesis is right, it is noteworthy that, within the Arabic written tradition up to the first millennium, there is no confessional difference and distinction.

12 See an example in Déroche 2004, 49. As for a Christian Arabic manuscript coming from the West and having the very same palaeographic peculiarities, see the bilingual (Greek-Arabic) parchment, Bibliothèque nationale de France, Suppl. Grec 911 (Géhin 1998, 166, 171).

13 Sijpesteijn (2008, 515a) came to the same conclusions based on papyrus documents dating back to 7th–8th centuries.

14 Although one cannot exclude *a priori* the possibility of the intervention of a second or later hand in order to explain this alternation — as Monferrer Sala supposed (2010, 197) — one could wonder why the later hand would not systematically intervene in every *qāf*.

15 It is also noteworthy to remark that, like in linguistics and textual criticism, the study of the early Arabic written tradition shows that peripheries are more conservative, as for the so-called *maghribī* script.

4 Early developments of the layout

As mentioned above, the presence of split words at the end of lines is another peculiar feature of 'early Abbasid script'. This feature seems to be characteristic only of codices dating back to the second half of the 8th century up to the beginning of the 9th. One can suppose that this special layout is due to the typical tendency of the early Abbasid era to fill the entire written area. In *Sinai Arabic 154*, whenever the text does not fill all the available space, it is possible to distinguish a stroke at the end of some lines: for instance in the following fols: 99v line 10; 109v -1; 110r line 5; 110v lines 3 and 8.

Fig. 3: *Sinai Arabic 154* (end of 8th c.), fols 110r line 5; 110v lines 3 and 8.

I think the copyist might have used this technique only when he was unable to stretch the last letter of the line (the so-called *mašq* technique), which however is widely used in the whole manuscript (Déroche 2000, 187; Gacek 2001, 135). The lack of space between words is another consequence of the tendency to fill all the available written area. In the case of the *Sinai Arabic 154* fol. 17v line 19 one finds a critical point which caused some problems of interpretation to philologists and editors. Samir rightly supposed that the words *wa-fakka riqābanā* 'and he untied our napes' were connected to each other due to an error of the copyist (Samir 1990–1991, 88–89; La Spisa 2014, 37).[16]

Vaticanus Borgianus 95 also shows several cases of words split at the end of lines. See for instance the following examples in fol. 16v. lines 2–3: *talā—mīḏahu* 'his disciples'; lines 4–5: *istir—ḫā* 'weakness'; lines 10–11: *wa-'a—marahum* 'and he ordered them'.

This very feature is widespread also in Islamic documents dating back to the same period; this confirms what we have already shown in the examples quoted above.[17]

16 It may be not by chance that this error occurred with a word whose first letter does not attach on its left as *rā'*; see Déroche and Sagaria Rossi 2012, 193.

17 See for instance the following documents: Oriental Institute of Chicago n. 14046, 17629, 17631, 17636, 17637, Vienna, Nationalbibliothek. Papyrus Erzherzog Rainer n. 734 (Abbott 1957, 32, 57, 80; 1967, 199, 207, 235).

Another point that should be highlighted is that when the script changes, as for instance in the passage from the 'early Abbasid script' to the longer-lasting *naskhī* script, the whole structure of the page changes consequently too. This change is probably strictly related to that of the material support of manuscripts. For instance, with the spread of paper in the Arab world and the progressive disappearance of parchment, layout techniques too were refined and improved. While we do not have any clear indication of the technique used for the justification of text before the introduction of paper, 'the progressive introduction of the *misṭara* led to a relative standardization of ruling types' (Déroche and Sagaria Rossi 2012, 123, Sagaria Rossi 2015, 102). The upside-down trapezoid or triangle form of the colophons at the end of the epistles, treatises and prose works in general (Déroche and Sagaria Rossi 2012, 207) denotes not only a greater availability of paper from the economic point of view, but also the development of the art of *mise en page*.

The technique of the trapezoid/triangle form at the end of the text (used not only for colophons) can be found in Christian Arabic texts too, as for example in the case of the Karšūnī-Arabic *Vatican Syrian 202* (17th cent.) and the *Šwayr 323 (123)* (18th cent.). So it is not difficult to conclude that only after the introduction of the paper in the Arab world did Muslim and Christian scribes alike feel that they could organise the layout of the page more freely.

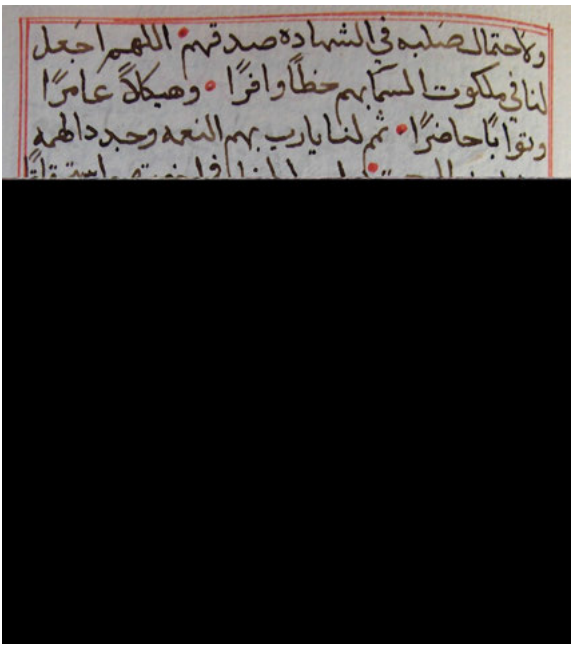


Fig. 4: Mār Yuḥannā al-Ṣābiġ Monastery (Lebanon) - *Šwayr* 323 (18th c.), fol. 87r.

5 A confessional Middle Arabic feature?

Islamic and Christian manuscripts share all the linguistic and palaeographic features mentioned so far. I now would like to deal with an orthographic peculiarity which I have found so far only in manuscripts belonging to the Christian Melkite milieu: the way of writing the prepositional group من أجل (*min ʾağli*) 'for the sake of' in Arabic, with disappearance of the consonant *hamza* (glottal stop) and the resulting coalescence between preposition and name (منجل *minağli*). Joshua Blau was the first who pointed out this feature in his *Grammar of Christian Arabic* (1966).[18] Blau described this phenomenon as an elision of the *hamza* (the glottal stop consonant) when it 'occurs at the beginning of a word governed by a preposition' (Blau 1966, 101–102). Another example is the case of *min ʾayna* which often becomes *minēn*. However, while this last change is also frequent in Modern Arabic dialects, the use of *minağli* such as attested in eastern Middle Arabic texts, might implies a hybrid register between *min ʾağli* and *mənšān* ('for, to, in order that') which is extensively used in modern Syrian dialect (Barthélemy 1936, 374; Cowell 1964, 491). Since the omission of the *hamza* is a typical Middle Arabic feature, we are facing here a classical example of mixed Arabic between *fuṣḥā* and *ʿāmmiyya*.[19] The same orthographic phenomenon occurs also in two other manuscripts belonging to the same Melkite monastic milieu: the *dayr al-Muḫalliṣ 1807* (Ğūn - Lebanon) and the *Balamand 135* (Tripoli - Lebanon) both coming from ancient Arabic-speaking Greek Catholic and Greek Orthodox monasteries respectively, of the 17th century (La Spisa 2012, 213). In conclusion it is noteworthy to remark that an orthographic phenomenon such as *minağli* for *min ʾağli*, which is typical for manuscripts coming from southern Palestinian monasteries, is shared also by texts written within the wide area that goes under the name of *Bilād al-Šām*.[20]

18 See also Blau 2002, 35 §26 and La Spisa 2012, 213.

19 In order to clarify this point it would be useful to remember that '[i]t would be wrong to suppose that every deviation in a written text is colloquial. Since people know that there is a difference between written and spoken language, they make a conscious attempt to write correctly but in doing so sometimes overreact using forms that are neither colloquial nor standard' Versteegh 2001, 115.

20 The sharing of this kind of linguistic phenomena within the Melkite Arabic tradition, led Joshua Blau to spoke about the Melkite Arabic literary *lingua franca* (Blau 1994; 2002, 72), on the same topic see also La Spisa 2012.

6 Concluding remarks

From all the data listed above it is possible to infer some considerations about the interactions between the early Arabised Christian communities in the monasteries of South Palestine and their Islamic religious environment. The first question I would like to pose is: is it possible to speak of new standards in the early Christian Arabic texts as compared with the Islamic standards (starting from the Qur'an and onward)? Just one century after the Arab conquests the Arabic spelling and grammar were almost the same for Christians and Muslims. Scholars having analysed the language of Christian Arabic manuscripts often concluded that there existed a so-called Christian Middle Arabic (Blau 1966–67, 1994, 2002, Grand'Henry 2006). Nevertheless Samir Arbache (2008), who studied the morphological verbal system in the *Sinai Arabic 72* codex, which dates back to the 9th century, drew the following conclusions:

> Les textes en moyen arabe ont existé depuis les origines, c'est-à-dire depuis le début de la littérature arabe écrite. [...] Si tel est le cas, le moyen arabe ne peut plus être envisagé comme une transformation ou une régression de l'arabe classique. Il sera plutôt objet d'analyse comme un état de la langue écrite au même titre que le dialecte ou la langue classique.[21]

Only after the normalisation process carried out by the Iraqi philologists of the 9th–10th centuries is it possible to speak of specific palaeographic and linguistic choices due to the need to build and consolidate a confessional identity. At the very beginning of the Arabisation, the 'Melkite Arabic' church distinguished itself by choosing the Arabic language, as it was spoken and written by all the Arabic speakers in that time, as the official language of the church. It stands to reason to believe that in Christian milieus *al-ʿarabiyya al-fuṣḥā* did not have any liturgical function or any specific religious meaning as it has in Islam. For this reason, Middle Arabic is much more pervasive even in liturgical, theological and patristic works of the 'Melkite Arabic' church.

On the other hand, some centuries after the Melkite Church, the other oriental churches living in Egypt and in the *Bilād al-Šām*, started to produce their literary, theological and patristic heritage directly in Arabic because their liturgi-

21 Arbache (2008, 19): 'Texts written in middle Arabic have existed since the origins, i.e. since the very beginning of written Arabic literature. [...] If this is the case, the Middle Arabic should no longer be considered as a sort of transformation or corruption of Classical Arabic. It is rather a special variety of the written language which has the same importance as the dialect and the Classical language' (my translation); for further reflections on the same question, see also Bettini and La Spisa (2012, viii–xii).

cal languages had become incomprehensible to Christian believers. The kind of Arabic they used is also called Middle Arabic, but in-depth studies trying to point out differences and similarities with the Melkite texts are still lacking.

References

Abbott, Nabia (1939), *The Rise of the North Arabic Script and its Ḳurʾānic Development* (University of Chicago Oriental Institute Publications, 50). Chicago: University of Chicago Press.

Abbott, Nabia (1957), *Studies in Arabic Literary Papyri. Vol. I: Historical Texts* (University of Chicago Oriental Institute Publication, 75). Chicago: University of Chicago Press.

Abbott, Nabia (1967), *Studies in Arabic Literary Papyri. Vol. II: Qurʾānic Commentary and Tradition* (University of Chicago Oriental Institute Publication, 76). Chicago: University of Chicago Press.

Abdel Haleem, Muhammad (1992), 'Grammatical Shift for Rethorical Purposes : *iltifāt* and Related Features in the Qurʾān', in *Bulletin of the School of Oriental and African Studies*, 55: 407–432.

Al-ʾAnbārī, ʿAbd al-Raḥmān (1997), *ʾAsrār al-ʿarabiyya*. Muḥammad Ḥusayn Šams al-Dīn (ed.), Bayrūt: Dār al-kutub al-ʿilmiyya.

Arbache, Samir (2007), 'Bible et liturgie chez les Arabes chrétiens (VIe-IXe siècle)', in David Thomas (ed.), *The Bible in Arab Christianity* (The History of Christian-Muslim Relations, 6), 37–48. Leiden, Boston: Brill.

Arbache, Samir (2008), 'Le texte du Sinaï Arabe 72 : éléments de morphologie verbale', in Jérôme Lentin and Jacques Grand'Henry (eds), *Moyen arabe et variétés mixtes de l'arabe à travers l'histoire. Actes du premier colloque international (Louvain-la-Neuve, 10-14 mai 2004)* (Publications de l'Institut Orientaliste de Louvain, 58), 1–20. Louvain-la-Neuve: Université catholique de Louvain – Institut Orientaliste.

Barthélemy, Adrien (1936), *Dictionnaire Arabe-Français. Dialects de Syrie : Alep, Damas, Liban, Jérusalem*. Paris: Librairie orientaliste P. Geuthner.

Baumstark, Anton (1929–1931), 'Das Problem eines vorislamischen christlich-kirchlichen Schrifttums in arabischer Sprache', in *Islamica*, 4: 562–575.

Bettini, Lidia, and Paolo La Spisa (eds), *Au-delà de l'arabe standard. Moyen arabe et arabe mixte dans les sources médiévales, modernes et contemporaines* (Quaderni di Semitistica, 28). Firenze: Dipartimento di Scienze dell'Antichità, Medioevo e Rinascimento e Linguistica.

Blau, Joshua (1966), *A Grammar of Christian Arabic Based Mainly on South-Palestinian Texts from the First Millennium* (Corpus Scriptorum Christianorum Orientalium, 27). Louvain: Secrétariat du CorpusSCO.

Blau, Joshua (1967), *A Grammar of Christian Arabic Based Mainly on South-Palestinian Texts from the First Millennium* (Corpus Scriptorum Christianorum Orientalium, 28-29). Louvain: Secrétariat du CorpusSCO.

Blau, Joshua (1994), 'A Melkite Arabic literary *lingua franca* from the first millennium', in *Bulletin of the School of Oriental and African Studies*, 57: 14–16.

Blau, Joshua (2002), *A Handbook of Early Middle Arabic*. Jerusalem: The Max Schloessinger Memorial Foundation/The Hebrew University of Jerusalem.

Briquel-Chatonnet, Françoise (1997), 'De l'araméen à l'arabe: quelque réflexions sur la genèse de l'écriture arabe', in François Déroche and Francis Richard (eds), *Scribes et manuscrits du Moyen Orient*. Paris: Bibliothèque nationale de France. 135–149.

Briquel-Chatonnet, Françoise, and Philippe Le Moigne (eds) (2008), *L'ancien Testament en syriaque* (Études Syriaques, 5). Paris: Geuthner.

Burton, John (1988), 'Linguistic Errors in the Qurʾān', in *Journal of Semitic Studies*, 33: 181–196.

Cantineau, Jean (1930–1932), *Le nabatéen : notions générales, écriture, grammaire*, vol. I. Paris: Presses universitaires de France.

Cowell, Mark W. (1964), *A Reference Grammar of Syrian Arabic based on the Dialect of Damascus*. Wasington D.C.: Georgetown University Press.

D'Ottone, Arianna (2015), 'Arabic palaeography', in Alessandro Bausi, Pier Giorgio Borbone, Françoise Briquel-Chatonnet, Paola Buzi, Jost Gippert, Caroline Macé, Marilena Maniaci, Zisis Melissakis, Laura E. Parodi and Witold Witakowski (eds), *Comparative Oriental Manuscript Studies. An Introduction*. Hamburg: COMSt, 271–276

de Sacy, Silvestre A. (1810), *Grammaire arabe à l'usage des élèves de l'école spéciale des langues orientales vivantes*. Paris: Imprimerie impériale.

Den Heijer, Johannes (2012), 'On language and religious identity: the case of Middle Arabic, with special reference to the Christian Arabic communities in the Medieval Middle East', in Gunvor Mejdell and Lutz Edzard (eds), *High vs. Low and Mixed Varieties. Status, Norms and Functions Across Time and Languages* (Abhandlungen für die Kunde des Morgenlandes, 77). Wiesbaden: Harrassowitz Verlag, 53–87.

Déroche, François (1987–1989), 'Les manuscrits arabes datés du III^{e}/IX^{e} siècles', in *Revue des études islamiques*, 55–57: 343–379.

Déroche, François (2000), *Manuel de codicologie des manuscrits en écriture arabe*. Paris: Bibliothèque nationale de France.

Déroche, François (2004), *Le livre manuscrit arabe. Préludes à une histoire*. Paris: Bibliothèque nationale de France.

Déroche, François, and Valentina Sagaria Rossi (2012), *I manoscritti in caratteri arabi. Al-maḫṭūṭāt bi-l-ḥarf al-ʿarabī* (Scritture e Libri del medioevo 9). Rome: Viella.

Fleisch, Henri (1990) (2nd edn.), *Traité de philologie arabe. Préliminaires, phonétique, morphologie nominale*. Beyrouth: Dar el-Machreq.

Ferrando, Ignacio (2001), *Introducción a la historia de la lengua árabe. Nuevas perpectivas*. Zaragoza: Navarro & Navarro.

Gacek, Adam (2001), *The Arabic Manuscript Tradition. A Glossary of Technical Terms and Bibliography* (Handbook of Oriental Studies, 58). Leiden, Boston, Cologne: Brill.

Gacek, Adam (2009), *Arabic Manuscripts. A Vademecum for Readers* (Handbook of Oriental Studies, 98). Leiden, Boston: Brill.

Géhin, Paul (1998), 'La bibliothèque de Sainte-Catherine du Sinaï. Fonds ancien et nouvelles découvertes', in Dominique Valbelle and Charles Bonnet (eds), *Le Sinaï durant l'antiquité et le Moyen Âge. 4000 ans d'histoire pour un désert*, 157–164. Paris: éditions errance.

Gibson, Margaret Dunlop (1899), *An Arabic Version of the Acts of the Apostles and the Seven Catholic Epistles From an Eighth or Ninth Century ms. in the Convent of St. Catharine on Mount Sinai, with a treatise On the triune nature of God, with translation from the same codex* (Studia Sinaitica, 7). London: C. J. Clay and Sons.

Graf, Georg (1944), *Geschichte der christlichen arabischen Literatur* (Studi e Testi, 118). Città del Vaticano: Biblioteca Apostolica Vaticana.

Grand'Henry, Jacques (2006), 'Christian Middle Arabic', in Kees Versteegh (ed.), *Encyclopaedia of Arabic Language and Linguistics* vol. 1, 383–387. Leiden, Boston: Brill.

Griffith, Sidney H. (1985), 'The gospel in Arabic. An inquiry into its appearance in the first Abbasid century', in *Oriens Christianus*, 69: 126–167.

Griffith, Sidney H. (1997), 'From Aramaic to Arabic: the languages of the monasteries of Palestine in the Byzantine and early Islamic period', in *Dumbarton Oaks Papers*, 51: 11–31.

Griffith, Sidney H. (2006), 'The church of Jerusalem and the "Melkites": the making of an "Arab Orthodox" Christian identity in the world of Islam (750-1050 CE)', in Ora Limor and Guy G. Stroumsa (eds), *Christians and Christianity in the Holy Land: From the Origins to the Latin Kingdoms* (Cultural Encounters in Late Antiquity and the Middle Ages, 5). Turnhout: Brepols, 175–204

Griffith, Sidney H. (2008), 'Les premières versions arabes de la Bible et leurs liens avec le syriaque', in Françoise Briquel-Chatonnet and Philippe Le Moigne (eds), *L'ancien Testament en syriaque* (Études Syriaques, 5). Paris: Geuthner, 221–245

Griffith, Sidney H. (2013), *The Bible in Arabic. The Scriptures of the 'People of the Book' in the Language of Islam*. Princeton, Oxford: Princeton University Press.

Gruendler, Beatrice (2006), 'Arabic Alphabet: Origin', in Kees Veerstegh (ed.), *Encyclopaedia of Arabic Language and Linguistics* vol. 1, 148–155. Leiden, Boston: Brill.

Guidi, Ignazio (1876–1877), 'Ragguaglio sui Codici arabi del Collegio Urbano de Propaganda Fide', in *Bollettino degli Studi Orientali*, 1: 198–202, 252–259.

Guidi, Ignazio (1888), 'Le traduzioni degli Evangelii in arabo e in etiopico', in *Atti della Reale Accademia dei Lincei. Memorie della classe di scienze morali, storiche e filologiche*, 4: 5–37.

Hopkins, Simon (1984), *Studies in the Grammar of Early Arabic Based Upon Papyri Datable to Before 300 A.H./A.D. 912*. Oxford: Oxford University Press.

Kachouh, Hikmat (2012), *The Arabic versions of the Gospel. The Manuscripts and their Families*. Berlin, New York: De Gruyter.

La Spisa, Paolo (2008), 'Fonti indirette e nuove fonti manoscritte nell'opera teologica di Sulaymān al-Ġazzī', in Davide Righi (ed.), *La letteratura arabo-cristiana e le scienze nel periodo abbaside (750-1250 d.C.). Atti del 2º convengo di studi arabo-cristiani, Roma 9-10 marzo 2007* (Patrionio Culturale Arabo Cristiano, 11), 299–315. Turin: Zamorani.

La Spisa, Paolo (2012), 'L'œuvre théologique de Sulaymān al-Ġazzī : un autre exemple de moyen arabe standard ?', in Lidia Bettini and Paolo La Spisa (eds), *Au-delà de l'arabe standard. Moyen arabe et arabe mixte dans les sources médiévales, modernes et contemporaines* (Quaderni di Semitistica, 28). Firenze: Dipartimento di Scienze dell'Antichità, Medioevo e Rinascimento e Linguistica. 209–225.

La Spisa, Paolo (2013), *I trattati teologici di Sulaymān Ibn Ḥasan al-Ġazzī* (Vol. 1). Leuven: Editions Peeters.

La Spisa, Paolo (2014), '*Excerptum* dalla più antica apologia araba cristiana', in Lorenzo Casini, Paolo La Spisa and Alba Rosa Suriano (eds), 'The Language(s) of Arabic Literature. Un omaggio a Lidia Bettini', in *Quaderni di Studi Arabi*, 9: 33–56.

Lafontaine-Dosogne, Jacqueline (1996), 'Le monastère du Sinaï. Creuset de culture chrétienne (X^{e} – XIII^{e} siècle)', in Krijnie Ciggaar, Adelbert Davids and Hermann Teule (eds), *East and West in the Crusader States. Context-Contacts-Confrontations. Acta of the Congress held at Hernen Castle in May 1993* (Orientalia Lovaniensa Analecta, 75). Leuven: Peeters, 103–129.

Lentin, Jérôme (1997), *Recherches sur l'histoire de la langue arabe au Proche-Orient à l'époque moderne*. Paris: Université de la Sorbonne Nouvelle III Ph.D. Thesis.

Lentin, Jérôme (2008), 'Middle Arabic', in Kees Versteegh (ed.), *Encyclopaedia of Arabic Language and Linguistics* vol. 3. Leiden, Boston: Brill, 215–224.

Lentin, Jérôme (2012), 'Reflections on Middle Arabic', in Gunvor Mejdell and Lutz Edzard (eds), *High vs. Low and Mixed Varieties. Status, Norms and Functions Across Time and Languages* (Abhandlungen für die Kunde des Morgenlandes, 77). Wiesbaden: Harrassowitz Verlag, 32–52.

Meimari, Ioannou (1985) Κατάλογος τῶν νέων ἀραβικῶν χειρογράφων τῆς ἱερᾶς μονῆς ἁγίας Αἰκατερίνης τοῦ Ὄρους Σινᾶ. Ἀθῆναι: Ἐθνικὸν Ἵδρυμα Ἐρευνῶν.

Metzger, Bruce Manning (1977), *The Early Versions of the New Testament. Their Origin, Transmission, and Limitations*. Oxford: Clarendon.

Monferrer Sala, Juan Pedro (2010), 'Once again on the early Christian Arabic Apology: remarks on a palaeographic singularity', in *Journal of Near Eastern Studies*, 69(2): 195–197.

Mouton, Jean-Michel (2000), *Le Sinaï médiéval. Un espace stratégique de l'islam*. Paris: Presses Universitaires de France.

Noja Noseda, Sergio (2006), 'L'origine della scrittura araba da quella dei nabatei: un'ipotesi da smontare', in Pier Giorgio Borbone, Alessandro Mengozzi and Mauro Tosco (eds), *Loquentes linguis. Studi linguistici in onore di Fabrizio A. Pennacchietti*. Wiesbaden: Harrassowitz Verlag, 535–553.

Orsatti, Paola (1996), *Il fondo Borgia della biblioteca vaticana e gli studi orientali a Roma tra Sette e Ottocento* (Studi e Testi, 376). Città del Vaticano: Biblioteca Apostolica Vaticana.

Sagaria Rossi, Valentina (2015), 'The layout of the page', in Alessandro Bausi, Pier Giorgio Borbone, Françoise Briquel-Chatonnet, Paola Buzi, Jost Gippert, Caroline Macé, Marilena Maniaci, Zisis Melissakis, Laura E. Parodi and Witold Witakowski (eds), *Comparative Oriental Manuscript Studies. An Introduction*. Hamburg: COMSt, 101–104.

Samir, Khalil S. (1990), 'La littérature melkite sous les premiers abbassides', in *Orientalia Christiana Periodica*, 56: 469–486.

Samir, Khalil S. (1990–1991), 'Une apologie arabe du christianisme d'époque Umayyade?', in *Parole de l'Orient*, 16: 85–106.

Samir, Khalil S. (1994), 'The earliest Arab Apology for Christianity (c. 750)', in Samir Khalil Samir and Jørgen S. Nielsen (eds), *Christian Arabic Apologetics During the Abbasid Period (750-1258)* (Studies in the History of Religions, 63), 57–114. Leiden, Boston, New York: Brill.

Schoeler, Gregor (2002), *Écrire et transmettre dans les débuts de l'Islam*. Paris: Presses universitaires de France.

Schulthess, Sara (2018), *Les manuscrits arabes des lettres de Paul. État de la question et étude de cas (1 Corinthiens dans le Vat. Ar. 12)* (Biblia Arabica, 6). Leiden, Boston: Brill.

Shahid, Irfan (1995–2009), *Byzantium and the Arabs in the Sixth Century*. 2 vols. Washington, DC: Dumbarton Oaks.

Sijpesteijn, Petra M. (2008), 'Palaeography', in Kees Versteegh (ed.), *Encyclopaedia of Arabic Language and Linguistics* vol. 3. Leiden, Boston: Brill. 513–521.

Starcky, Jean (1966), 'Petra et la Nabatène', in *Dictionnaire de la Bible. Supplement*. Vol. 7. Paris: Letouzey & Anén ed. 886–1017.

Swanson, Mark (1993), 'Some consideration for the dating of Fī taʦlīṯ Allāh al-wāḥid (Sin. Ar. 154) and al-Ǧāmiʿ wuǧūh al-īmān (London, British Library or. 4950)', in *Parole de l'Orient*, 18: 115–141.

Tisserant, Eugène (1914), *Specimina Codicum Orientalium* (Tabulae in usum scholarum, 8). Bonnae: A. Marcus & E. Weber.
Tisserant, Eugène (1924), *Inventaire sommaire des manuscrits arabes du fond Borgia à la Bibliothèque Vaticane* (Studi e Testi, 41). Città del Vaticano: Biblioteca Apostolica Vaticana.
Troupeau, Gérard (1991), 'Réflexions sur l'origine syriaque de l'écriture arabe', in Alan S. Kaye (ed.), *Semitic Studies in Honour of Wolf Leslau On the Occasion of his Eighty-Fifth Birthday, November 14th, 1991*. vol. 2, Wiesbaden: Otto Harrassowitz, 1562–1570.
Veccia Vaglieri, Laura (1937 [1996^{7}]), *Grammatica teorico-pratica della lingua araba*. 2 vols. Roma: Istituto per l'Oriente.
Versteegh, Kees (2001), *The Arabic Language*. Edinburgh: Edinburgh University Press.
Wright, William (1962), *A Grammar of the Arabic Language*. 2 vols. Cambridge: Cambridge University Press.

Nuria de Castilla

Uses and Written Practices in Aljamiado Manuscripts

Abstract: In the Aljamiado texts, we see a consistent system used for the transliteration of the Romance language in Arabic script. From the oldest copies preserved (15th century) until the last ones produced by the Moriscos at the beginning of the 17th century, a standardised graphic and orthographic system has been applied to these Spanish texts written in Arabic script.

1 Introduction

After seven centuries of Muslim presence in Spain, Boabdil had to surrender the keys of the city of Granada to Ferdinand, the Catholic King, in 1492. Despite the 'Capitulaciones de Granada', the agreement by which the Catholic kings showed an apparent kind understanding of the Muslim population (they could maintain their clothes, habits and rituals), in 1499, Cardinal Cisneros started to stress the necessity of religious unity.

In 1502 in Castile, and 1526 in Aragon, all the Muslims (the Mudejars) were obliged to be baptised (which is to say, they had to convert to Christianity) or to leave Spanish lands. The conditions set for exiting the country were so harsh that most of the people stayed on in the Iberian Peninsula. If they decided to be baptised and to stay in their lands, they were no longer called Mudejars, but Moriscos or 'convertidos de moro'. But what at the beginning was a religious matter became something wider: in 1504, Morisco communities had to pay special taxes (because they were converted), and from 1511, cultural differences were no longer accepted. Then, from 1516 the typical Morisco clothes were forbidden, as well as their music and *zambras* (dances), and of course, their food: not eating pork meant that you were a Muslim; at the same time, the Muslim way to slaughter the animals was banned too.

This chapter is the developed version of the paper presented in the conference 'Creating standards: Orthography, Script and Layout in Manuscript Traditions Based on Arabic Alphabet', held in Hamburg in October 2013. The research was carried out within the frame of the *Corana, Production and transmission of the Qur'ān in the Western Islamic World*, a project led by Nuria de Castilla, and financed by the Spanish Ministry of Science and Innovation (FFI2012-32294), and ERC Grant Agreement number 670628 SICLE (2016–2021).

https://doi.org/10.1515/9783110639063-005

In the second half of the 16th century, it was the turn of the books: in 1564, it was ordered to burn all the books in Arabic in Valencia. In 1564, Granadan Moriscos were enjoined by a Royal decree to learn Spanish within three years (Vincent 2006, 106). In 1567 in Castile, the use of Arabic language, written and spoken, as well as the right to own books in this language, was forbidden. Despite the interdictions, paradoxically, the greater part of the manuscripts which have remained from these Muslim communities were copied during this period, i.e., the second half of the 16th century, when the bans were more intense, and the use of the Arabic script acquired specific values of resistance (Bernabé Pons 2010, 30).

Spoken Arabic first disappears in a general way in Castile and Aragon, but these two kingdoms conversely witness the use of the Aljamía, a written variant of Castilian with specific linguistic features. Arabic letters are usually used, although it is sometimes in Latin letters. The manuscripts in this area, hidden in the ceilings or in the walls, were discovered in the 16th and the 17th century, but mainly from the 19th century to the present day. The most important trove took place in Almonacid de la Sierra. 'In 1302/1884 hundreds of manuscripts came to light during building works under the floor of a house in the village of Almonacid de la Sierra, close to Saragossa, the capital city of Aragon. The volumes were probably concealed there shortly before the expulsion of the Moriscos in 1018/1609 and, when discovered, were well-arranged, standing side by side, together with some bookbinder's tools' (Martínez de Castilla 2014, 89). Other similar finds took place in Sevilla, Pastrana, Cútar, Ocaña, Calanda, Torrellas, Novallas, Sabiñán, Urrea de Jalón, etc.; however, the number of manuscripts discovered in these places was smaller. The whereabouts of some of them, like that of Pastrana, are unknown (Martínez de Castilla 2016); others are in private hands, as those of Ocaña and Urrea de Jalón. Most are however kept in public collections: the extant manuscripts from the Almonacid de la Sierra trove are nowadays in the Tomás Navarro Tomás Library, CCHS-CSIC in Madrid (mainly) and Escuelas Pías of Saragossa; the manuscripts from Calanda are in the Library of the Cortes de Aragón (Saragossa), and so on (Villaverde 2010). Manuscripts hidden behind a wall and discovered during repair works in a house have been the major source of these witnesses of Morisco culture. However, other codices produced in the same area of Castile and Aragon were bought by bibliophiles in the 17th century. This was the case of two Aljamiado manuscripts in the Bibliothèque nationale de France (Castilla, forthcoming), both having been acquired by Antoine Galland and integrated into the King's library in Paris in 1686 according to the information collected by Morel-Fatio (1982, vii). These manuscripts produced by Mudejars and Moriscos from Aragon and Castile exhibit two graphic systems (Arabic and Latin script) and two languages (Arabic and Aljamía). In this chapter, I will focus on the Aljamiado texts (a variant of Spanish, usually in Arabic script).

The word 'Aljamía' is derived from the Arabic *al-'ajamiyya* (العجمية), meaning 'non Arabic language'. In our context, it was used among the Peninsular Muslims to designate the language used by their Christian countrymen. In the *Diccionario de Autoridades* (1726), it is said that Aljamía is 'the language or parlance the Arabs living in Spain used to speak with the Spanish Christians in order to make themselves understood in their contracts and dealings',[1] i.e. it refers to a *koiné* or Romance vehicular language.[2] In the Western texts of the 15th to 17th centuries, the word Aljamía is systematically identified with the meaning 'Romance', probably Castilian as can be deduced from the following example.[3]

> Thirdly, that they should not speak *Algarabía*, but all should speak *Aljamía*, and that all the documents and contacts which used to be written in *Arabic* should be carried out in *Castilian*.[4]

Or, as shown by the following account:

> He was carrying secretly with himself a certain silver coin on which were letters and characters in Arabic which read 'ley lehe ele Ala', i.e. they meant in *Aljamía* that there is no other Lord except God alone.[5]

The same applies to the Catalan translation of *Tirant lo Blanc*:

> Upon my honour, Sir, accompanied by these Moorish ladies came a very handsome girl who spoke very well *Aljamía*, and with much gracefulness, and if your Lordship would like to make me a favour, although I do not deserve it, when you take the city, you would make her a Christian and give her to me as wife.[6]

---

1 'La lengua o idioma que para entenderse en sus tratos y comercios hablaban los árabes que estaban en España con los cristianos españoles'.

2 For more information about the contexts in which this word was used during the Spanish Golden Age, see Martínez de Castilla 2006b, 235–246.

3 The edition of the texts and the use of italics in the given examples are mine, as well as their translation.

4 'Lo tercero, que no hablasen *algarabía*, sino que todos hablasen en *aljamía*, y que todas las escrituras y contratos que se solían hacer en *arábigo*, se hiciesen en *castellano*'. Fray Prudencio de Sandoval, *Historia de la vida y hechos del Emperador Carlos V,* 1604–1618, ed. and study by Carlos Seco Serrano, in www.cervantesvirtual.com (accessed 2 October 2014).

5 'Ocultamente traía consigo una cierta moneda de plata en que había en ella letras y caracteres en lengua arábiga, que decían "ley lehe ele Alá", que [en] *aljamía* querían decir que no hay otro señor sino Dios solo'. Diocesan archives of Cuenca (ADC), leg. 262, núm. 3573. See Martínez de Castilla 2006b, 237.

6 'Por mi fe, señor, en compañía de aquellas moras viene una donzella muy graciosa que habla muy bien el *aljamía* y con mucha gracia, y si vuestra señoría me quisiese hazer tanta merced que,

In some cases, the adjective *Aljamiado* is used as a designation of those speakers who were fluent in both Arabic and Castilian, or for those who were living in *aljamas*, i.e., in Muslim neighbourhoods:

> In order to identify the *aljamiados* who did not learn from a child our language and its pronunciation, they had them say 'cebolla' ['onion'] and the Morisco would say 'xebolla'.[7]

Today, when speaking of Aljamía, we are referring to a variant of Castilian exhibiting a series of linguistic features which can serve as a basis for its identification:
- influence of Arabic, not only in lexicon, but also in morphology and syntax;
- influence of Aragonese, the Romance linguistic variety used in Aragon;
- strong archaism. In some cases, it seems that the language belongs to the 13th–14th century rather than to the 16th century.

These are perhaps the characteristics Pedro de Herrera was alluding to when he asserts, in 1618, that the *moro* is ridiculous for his language:

> Two by two, as for a fight, they wanted to jostle when a ridiculous Moor (according to his garments, his physical appearance and the *aljamiado* language of his voice), singing, convinced them to relinquish the civil litigations since their revolt was known and that they had been ordered to be thrown out of Spain.[8]

However, in spite of the possible influences of the spoken language on this linguistic variety used by Mudejars and Moriscos, the Aljamía was foremost a written language.[9] One of the reasons that leads us to maintain this statement

---

aunque yo no lo tenga servido, que como toméys la ciudad la hagáys hazer cristiana y me la deys por muger'. Joanot Martorell, *Tirant lo Blanc*, Castilian translation, 1511, ed. by Martín de Riquer, Madrid, Espasa-Calpe, 1974, p. 172, see Corde, http://corpus.rae.es (accessed 2 October 2014).

7 'A los *aljamiados* que no habían desde niños aprendido nuestra lengua y su pronunciación, para conocerlos los hacían decir "cebolla", y el que era morisco dezía "xebolla"'. Bernardo de Aldrete, Antig. de Esp. [*Varias antigüedades de España, África y otras provincias,* Anvers, Juan Hasrey, 1614], I, 37. On the basis of this example, Cirot suggests that 'aljamiado' may have meant 'parlant arabe' (Georges Cirot, '"Ladino" et "aljamiado"', *Bulletin Hispanique* 38.4 (1936), p. 539; this meaning is already found in the *Diccionario de la lengua castellana de la Real Academia Española de 1770:* 'lengua árabe corrompida que hablaban los moros de España' (*s.v.* 'Aljamía').

8 'Dos a dos, como para reñir, querían acometerse, cuando un moro ridículo (por el vestido, figura personal y lenguaje *aljamiado* de su voz), cantando, les persuadió dexassen las pendencias civiles porque ya se sabía su rebelión y los mandavan echar de España'. Pedro de Herrera, *Translación del Santísimo Sacramento a la iglesia colegial de San Pedro de la villa de Lerma*, 1618, in Ferrer Valls 1993, 273.

9 The Aljamía 'was probably not a faithful mirror of the variant actually spoken by these people'

is the big linguistic difference between the texts translated from the Arabic – the most important part of the manuscripts produced by the Mudejars and Moriscos in Aljamía – and those which were a recent creation. When we speak of Aljamía, we speak thus about a written linguistic variant of Spanish, with a strong influence of Aragonese, and in a large part Arabised and Islamised as much from a linguistic point of view as for its contents.

2 The copyist and his *milieu*: time, place and society

Who produced this kind of codices? Aljamiado manuscripts were copied within the Muslim communities in Spain: Mudejars (before the forced baptism in 1502) and Moriscos (after this date), from the end of the 14th century until the beginning of the 17th century. The earliest manuscripts that have come down to us are probably from the 15th century, but the majority of the manuscripts date to the 16th and 17th centuries.

This production was localised in Aragon and Castile, not in Granada or in Valence. The various Spanish kingdoms have very different history, and the same applies to the Muslim communities living within each of them. Toledo was conquered in 1085, whereas Granada in 1492: four long centuries separate the two events. One can understand that the Castilian Muslim communities were much more assimilated in 1499 (the year when Cardinal Cisneros initiated the policy of instituting religious unity in the kingdom of Castile) than the Granadian ones. On the other hand, in the 14th century, which is probably when the Aljamiado phenomenon started, Granada was still a Muslim kingdom, and Valencia had been conquered one century before. This is the reason why Aljamía did not develop in every kingdom in Spain but only in Castile and above all in Aragon. In these kingdoms, Muslims had to struggle to preserve the knowledge of Arabic as well as their cultural identity.

These manuscripts were mainly written (copied or produced) by some unidentified scribes for the *faqihs* in order to maintain the cohesion of their communities and to control them, at a moment when they were losing their rites and cultural practices (de Castilla 2006a). Although we find familiar — or low quality copies — probably written within family circles, most of the manuscripts were produced by skilled copyists.

('no debía de ser un espejo fiel de la variedad hablada por aquellas poblaciones'). I am indebted to Olivier Brisville-Fertin for generously giving me access to his unpublished work-in-progress (p. 5).

3 Linguistic features

3.1 Graphic representation

Orthographic conventions are quite stable throughout the period concerned by this production, as we shall see later. However, the language itself varies according to the text typology, but it will remain in any case basically Romance: if the text is a translation of an earlier Arabic work, the text will exhibit a higher degree of Arabisation and archaism than a copy or adaptation of a European text or than a new composition produced in a Mudejar or Morisco context (Sánchez 1995, 339–348 and Montaner 2004, 99–100).

Since Spanish has a few phonemes that are not found in Arabic, Mudejars and Moriscos developed new graphemes in order to represent them:

1. Vowels: In Spanish there are five vowels (/a/, /e/, /i/, /o/, /u/), whereas in Arabic there are only 3 (/a/, /i/, /u/). For the vowel /e/, the Mudejars adopted a new orthographic combination, *fatḥa* plus *alif*, which was inherited by the Moriscos. Ex. مَاش = <mesa>. For the vowel /o/, there is no innovation, the *ḍamma* being used for both /o/ and /u/.
2. Consonants: The šadda, a diacritic used in Arabic system to geminate a consonant, is systematically used in order to provide solutions for rendering Spanish phonemes, as follows:
   - *bāʾ* with *šadda* = /p/ <p>.
   - *jīm* with *šadda* = /č/ <ch>.
   - *rāʾ* with *šadda* = /rr/.
   - *šīn* with *šadda* = /š/ <x>, adapted for the voiceless palato-alveolar fricative.[10]
   - *nūn* with *šadda* = /ñ/.

The various emphatic letters (ظ, ط , ض , ص and ق) are in most of the cases used for writing down Arabic words, not Spanish ones; examples like the following are common: <aˡṣṣala>, written with ص, but <açotea>, with س. However, in the Aljamiado texts in Latin script, a <ç> is used for both cases: 'açalá' and 'açotea', as the regular way to write these words in the 16th century Spanish.

10 Although this way of writing is the most commonly found, there are other manuscripts — *Poema de Yúçuf*, the manuscript of Urrea, BNE 5267, BNE 5305, BNE 5313, BRAH T12, BRAH T13, BRAH T18, RESC/13, RESC/30, RESC/33, RESC/37, RESC/52, RESC/64, RBME 1880, BPal 3226, BnF Arabe 774 or BnF Arabe 1163— in which the letter *šīn* without *tašdīd* is used exclusively to indicate both the apico-alveolar and the pre palatal fricative voiceless.

3.2 Consonant clusters

There is a tendency in Aljamía to maintain some of the features found in Arabic, such as the syllabic structure. Although some Arabic dialects have initial or final consonant clusters, Andalusi Arabic did not allow a sequence of two consonants in the same syllable without an intermediary vowel. In other words, the only possible syllabic structure is CV(C) (Corriente 2002, 35). The Aljamía usually introduced an epenthetic vowel, the same as in the following syllable, between the consonants in order to eliminate the tautosyllabic consonant clusters in the original Romance words, for example, *ap^{e}remiar*, *ap^{a}laçado*, *at^{e}revimiento*.

According to Quilis (1981, 298–300), in Aljamía, the epenthetic vowel was a merely graphic component unlike in Classical Arabic where the epenthetic vowel is usually phonetically realised. Nevertheless, this *svarabhakti* element (automatic vowel) seems to have had a phonetic representation, a conclusion reached by Corriente (2000–2001, 117) in relation to the botanical glossary of Abulxayr:

> The disjunctive vowel was not a mere graphic tool meant to avoid an aberrant orthography within the Arabic script, but a phonic reality due to the interference of the syllable taxonomical rules.[11]

This view, if applied to Aljamía, would entail that it was 'a tongue phonetically interfered with, in this case by Arabic'[12] (Corriente 2000–2001, 117). In the same way, Labarta (1982, 231) states in her analysis of an Aljamiado page from Tarazona that

> [t]he Castilian sequences formed by occlusive + liquid + vowel produce between their first two components a vocalic element with the same quality as that of the following vowel. Such a sound, which does not have for us a phonological value and does not have any graphic representation, was perceived and noted in Aljamía either because the Arabic does not allow for a syllable to begin with two consonants, or due to the nature of Aljamía itself, that reproduces through the Arabic script the spoken Castilian in an 'acoustic' (not 'phonologic') way with assimilations, contractions, and so on.[13]

11 'La vocal disyuntiva no era un mero recurso gráfico para evitar una ortografía aberrante en la escritura árabe, sino una realidad fónica debida a la interferencia de las reglas taxonómicas de la sílaba'.

12 'una lengua interferida, fonéticamente en este caso, por el árabe'.

13 'Las secuencias castellanas formadas por oclusiva + líquida + vocal crean entre sus dos primeros componentes un elemento vocálico de timbre similar al de la vocal que lo sigue. Tal sonido, carente para nosotros de valor fonológico y de representación gráfica, era percibido y marcado en la escritura aljamiada, ya como consecuencia de que el árabe no permite que una sílaba empiece por dos consonantes, ya debido al propio carácter de la escritura aljamiada, la

3.3 Sibilant consonants

Although the graphical difference between /ś/ = ش and /š/ = شّ is indicated in some manuscripts,[14] 60% of the corpus analysed does not express this distinction.[15] This implies that this phonetic difference is not so pertinent in Aljamía as has been argued[16] and it is not rare to find *šarabe* instead of *xarabe* ('syrup') – which could be explained by the Arabic etymology of the word *šarāb* ('beverage') –,[17] or *dišo* instead of *dixo* ('he said'). The cases which exhibit this graphical distinction could be interpreted as follows: either this is the result of the realisation of a phoneme non-existent today, or the copyist was very familiar with the texts in Latin script and tried to emulate them in Arabic script, introducing the *shadda* above the *šīn* for the <x>.[18] As Montaner (2004, 100, n. 2) accurately indicated, 'except for the limited influence of the Andalusi substrate on the phonology and the lexicon, the Aljamía never stopped being a Romance language, even in its most Arabized texts'.[19]

3.4 Dental consonants

The most common letters representing the dental consonants in Aljamía are د \ ذ, including the emphatic consonants ظ / ض. This is probably the result of a plausible loss of consonantal emphasis among Mudejars and Moriscos (Bouzineb 1986, 30–31). There is a distinction between the fricative and occlusive allophones [ð] and [d], but it is not so common. This distinction is more regular in later man-

cual, mediante la grafía árabe, reproduce el castellano hablado de una forma "acústica" (no "fonológica"), con sus asimilaciones, contracciones, etc.' Nevertheless, when reading her study, one has the impression that this 'vowel' will always be found in the Tarazona folio. However, this is actually not the case in graphic representation of the word *libras* in BRAH T19, where a *sukūn* appears between the /b/ and the /r/. On the other hand, this epenthetic vocalic element occurs in the other three words that have consonantal groups (*gruesa*, *tres y maestro*). For more information about the distribution of the epenthetic vowel and *sukūn*, see Martínez de Castilla 2010, 182–187.

14 BRAH T13, BRAH T18, BRAH T19, BNE 4953, BNE 5223, BNE 5377, BNE RES 245, RESC/62.

15 That is to say, Poema de Yúçuf, Urrea manuscript (particular collection), BNE 5267, BNE 5305, BNE 5313, BRAH T12, BRAH T13, RESC/13, RESC/30, RESC/33, RESC/37, RESC/52, RESC/64, RBME 1880, PAL 3226, BnF Arabe 774, BnF Arabe 1163.

16 Galmés de Fuentes 1970, 220–221, and several studies since then.

17 Steiger 1991, 53, n. 3.

18 More information in Martínez de Castilla 2010, 177–178.

19 'salvo la limitada acción del sustrato andalusí en la fonología y el léxico, la aljamía sigue siendo netamente una lengua romance, incluso en sus textos más arabizados'.

uscripts, and a possible phonological interference by the Andalusi seems less likely than the cases with the sibilants as in the example of /šarabe/ above). As described by Navarro Tomás (1918, 99), 'the occlusive articulation is produced when the dental consonant appears [...] after a "n" or "l"'. For example, in the manuscript BRAH T19, undated but copied at the beginning of the 17th century (Castilla 2019), we read on the one hand:

el día	ءالْدِيَ	(24r, 6)
al dueño	اَلْدُوَانُ	(112v, 3)
y-el dayuno	يَالْدَيُنُ	(182r, 8)
en desierto	ءَانْدَا شِيَارْتُ	(84r, 15)
en dineros	ءَانْدِنَارُشْ	(122r, 4)
en demandarlo	ءَانْدَامَنْدَرْلُ	(140v, 4)

And on the other hand:

a ḏonde	اَذُنْدَا	(*passim*)
conpañía ḏel mensajero	كُنْبَنِّيَ ذَالْ مَانْشَجَارُ	(1r, 2–3)
ḏelante ḏe los onraḏos	ذَالَنْتَا ذَا لُشْ اُنْرَ ذُشْ	(1r, 9)
a ḏoze ḏías ḏe la luna	اَذُ زَا ذِيَشْ ذَاللُنَ	(20r, 9–10)

4 Contents

As far as the contents of the manuscripts are concerned, most of these codices are miscellanies – or multiple-text manuscripts –, but there are also composite and unitary copies. Notably, we can find copies of excerpts of the Qur'an, *hadith* (traditions attributed to the Prophet Muhammad), fragments of juridical texts, magic, and so on. Exceptionally, there are some documents written in Aljamía (personal letters and legal documents), but unfortunately only few items have been preserved (see, for instance, Hoenerbach 1965, Labarta and García Cárcel 1981, Viguera 1982 or Viguera 1991).

The Aljamiado texts are markedly Islamic and this applies to both Mudejar and Morisco periods. The production was carried out during an earlier stage mainly in Aragon, in Castile, and later, as far as we can establish, in Algeria, Tunisia and in the heart of the Ottoman territories, in Salonica and Constantinople, during the Morisco period and after the expulsion. It can therefore be defined as an 'Islamic

literature made [...] by Muslims and for Muslims'[20] (Bernabé Pons 2010, 27). It is 'an Islamic variant of Spanish', as Ottmar Hegyi (1985) aptly defined it.

The manuscripts can be divided into three main groups: a) translations of earlier Arabic texts, the date of translation remaining unclear;[21] b) copies and adaptations of Western-European texts; and c) new works. A taxonomic distinction has been traditionally maintained between the Aljamiado manuscripts in Arabic letters and those in Latin script, the former being called 'Aljamiados' and the latter 'Moriscos'. For this reason, several mentions of '*manuscritos aljamiado-moriscos*' appear in specialist publications dealing with this production, both in articles and catalogues; this is a descriptive phrase that allows us to jointly treat copies written in Mudejar and Morisco communities where the two scripts are used.[22] However, such description seems confusing since when we find a manuscript in Arabic script and another one in Latin characters containing the same text(s), the differences between both are in most cases strictly graphic. In this way, neither the state of the language of a given Aljamiado text, nor its place of production can be defined on the basis of its script (Latin or Arabic characters).[23] For instance, 'The story of Abū Šaḥmah when his father, 'Umar, sent him to be whipped'[24] is found in various Aljamiado codices (Fig. 1):[25] they are all in Arabic script, with the exception of BNE 6016, written in Latin letters. BNE 6016 belongs to the same textual tradition as the three other witnesses of the text which have been preserved in the manuscript of Urrea de Jalón, BRAH T12 and BnF Arabe 774.

20 'Una literatura islámica, hecha [...] por musulmanes y para musulmanes'.

21 Although the knowledge of Arabic was almost completely lost in Castile and Aragon, research over the past few years have been throwing light on the possible use, albeit in a minority way, in the Aragonese area. Cf. Ferrando 1996, 177–195; Ferrando 2000, 195–200; García-Arenal 2010, 295–310.

22 As an example, it will be sufficient to mention publications like Galmés de Fuentes 1998; Galmés de Fuentes 1986; or Vespertino 2002–2004.

23 Although in initial studies of the Aljamiado literature it was assumed that there was a correlation between the use of the Arabic script and the Aljamiado texts written in the Iberian Peninsula on the one hand, and between the use of the Latin characters and the production in the exile after the expulsion, texts written in Latin script in the Iberian Peninsula and conversely others written in Arabic outside the Peninsula have been found. Cf. Suárez 2004, 20 and 21 and the bibliography cited there.

24 'El hadiz de Abu Xahma cuando lo mandó açotar su padre Omar'.

25 Taken from Martínez de Castilla 2010, 308.

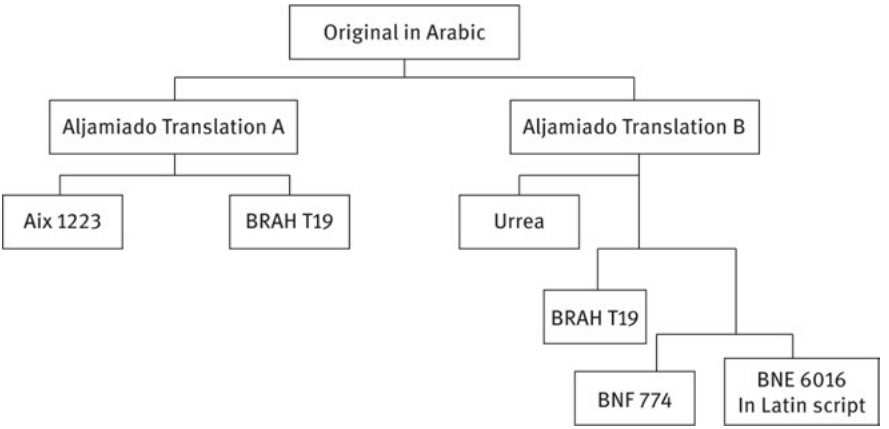


Fig. 1: Genealogical textual schema of 'The story of Abū Šaḥmah'.

In some cases, we know that an Aljamiado text in Latin script could have been transcribed from a copy in Arabic characters (Fig. 2). This is the case of the 'Prayer for the rite of the child's *fadas*' (i.e. the Muslim rite for giving a name to a newborn)[26] in BCM T232, in Latin letters, which was apparently copied from BRAH T19, unless both manuscripts stem from the same common exemplar (Martínez de Castilla 2010, 344).

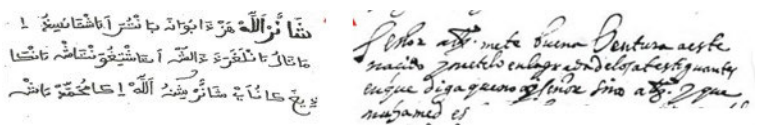


Fig. 2: Fragment of the 'Prayer for the rite of the child's *fadas*'. Left: BRAH T19, fol. 24v. Right: BCM T232, fol. 295r.

Also, in a few instances Spanish texts have been transliterated in Arabic script, either faithfully to the original or with changes. This way, a text produced in another culture and widely spread, serves as a basis either for a transliteration in Arabic script, as was done by the copyist of *The sayings of the seven wise men from Greece* (*Dichos de los siete sabios de Grecia*),[27] or for an adaptation to an

26 'Rogaria para las fadas de la criatura'.

27 Edited by Galmés de Fuentes 1991, 47–54. See also Castro (ed.) 1990, 407–410 and Ramos 2012, 848.

Islamic cultural context: this is the case with *The Handless Maiden* (*La Doncella de las manos cortadas*), a well-known legend in Middle Ages and still circulating in the oral European tradition (Thompson (1970, T411.1).[28] About the latter, Mary Elizabeth Perry (2005, 97) stated that 'the story of Carcayona reads as a Muslim version of the *Handless Maiden tale*'.

The language of the translated texts and of the new ones is therefore different. The new texts use a language closer to the "standard" use of the Christian Spanish texts than the copies of the translations, much more influenced by the Arabic; in other words, the use of the Arabic alphabet does not induce any special differentiation in the kind of language employed in these new Morisco works. Here is an example of a text written in Arabic script, produced in the 16th century, to help the Moriscos to leave Aragon in order to arrive safely to Salonica:

> [Itinerary from Spain]. [...] From there to Verona. Do not go through the city as you would pay a real per head. There you will ask for the road to Padua. There you will take a boat to Venice; from Venice to Bolonia [Bologna] or to Durazzo, or to Lesos or to Castelnou, that of these parts you will find first. You will enter an inn; you will pay half a real per day; and do not take anything from the inn as they will charge you triple.[29]

The language of the transliterated texts of Western European origin does not change either:

> [Sayings of Bias]. These are the sayings of Bias, and they are as follows: and in order to be well understood, the reader should think that each Sage is speaking with him: Look at yourself every day of your life in a mirror, take this advice from me.[30]

On the other hand, the texts resulting from a translation of an Arabic original exhibit a different state of the language (Fig. 3):

> The hadith of Abū Šaḥmah when his father, 'Umar, God may be pleased with him, sent him to be whipped. 'Abdallāh b. 'Umar said: 'Umar b. al-Khaṭṭāb had a son called Abū Šaḥmah

28 Valero Cuadra (ed.) 2000; Perry 2005. See also Martínez de Castilla 2006c.

29 BnF Arabe 774, fols 37v–38v. My own transliteration and translation. The text has been edited by Sánchez Álvarez 1982, 153–154. '[Itinerario de España]. [...] De allí a la Verona. No paséis por de dentro de la ciudad, que pagarés a real por cabeça. Allí demandarés el camino para Padua. Allí os embarcarés para Venecia; de Venecia para la Bolona [*sic*] o para Duracio, o para Lesos o para Castelnou, el que antes hallés d'estos puertos. [...] entrarés en una posada [...]; pagarés medio real por día; y no toméis nada de la posada, qu[e] os arán pagar de uno tres'.

30 BnF Arabe 1163, fol. 60v. Edited by Galmés de Fuentes 1970, 41. '[Dichos de Bías] Estos son los dichos de Bías, los cuales son los siguientes: I para ser bien entendidos, piense el le[c]tor que cada sabio habla con él: Mírate todos los días que vivieres al espejo; toma de mí esti consejo'.

who was a reciter of the Qur'an and when he was reciting it seemed that the Messenger of God was reciting. And Abū Šaḥmah fell ill, with a very bad illness, and the colleagues and friends of the Prophet Muḥammad were visiting him. One day, when he went there, they gathered in 'Umar's house and told him: 'O prince of the Faithful! If you made a promise, as did 'Alī b. Abū Talib for al-Ḥasan and al-Ḥusayn until God restored them to health, maybe God would restore to health your son Abū Šaḥmah and he would be healthy'.[31]

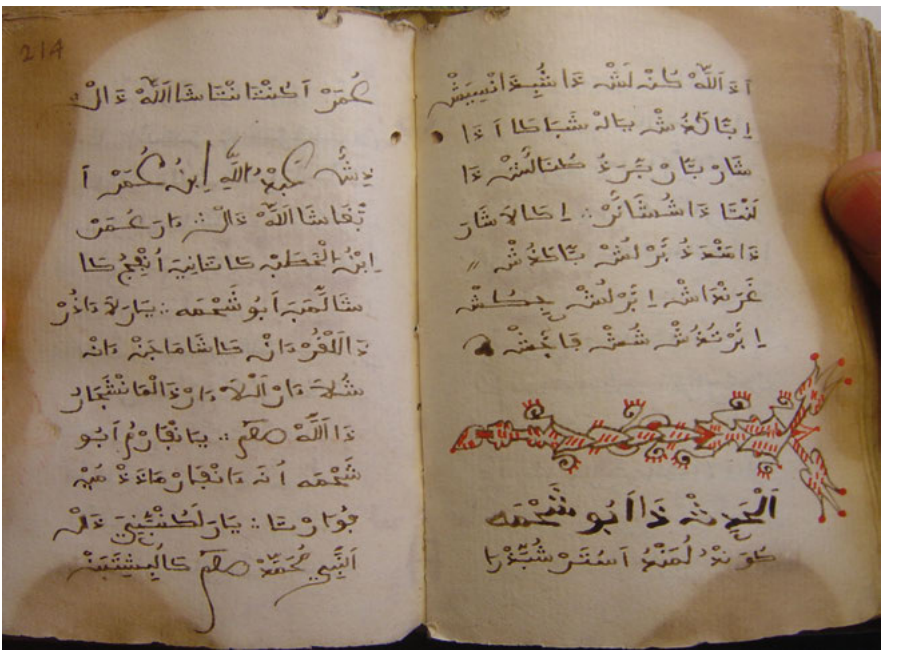


Fig 3: Bibliothèque Méjanes 1367, Aix-en-Provence, fol. 213v.

31 Aix 1367, fol. 213v. Edited by Martínez de Castilla 2005, 507. 'El hadiz de Abu Xahma cuando lo mandó açotar su padre Omar, aconténtese Allah d'él. Dixo Abullahi ibnu Omar, apáguese Allah d'él: Era Omar ibnu Elhatab que tenía un fijo que se llamaba Abu Xahma, y era lledor del Alcorán, que semejan en su leer al leer del mensajero de Allah (*ṣ'm*). Y enfermó Abu Xahma una enfermedad muy fuerte; y era la conpañía del anabí Muhamad (*ṣ'm*) que lo visitaba. Pues cuando fue un día, ajuntáronse en la casa de Omar y dixéronle: "Ye rey de los creyentes, si prometieses una promesa, así como hizo Ali ibnu Abi Talib por Alhaçan y Alhuçayni hasta que les invistió Allah la salud, por ventura que tu fijo Abu Xahma que le daría Allah salud y estaría luego sano"'.

We could thus conclude that the archaic state of the last text, as well as that of the other translations from the Arabic, can be at least partly explained by the conservative nature of the translations, some of them made probably in the 13th century, and copied again and again. However, even in the translated texts — the texts of Arabo-Islamic origin — we appreciate the influence exerted by the "standard" 16th century Spanish, used by the Christians. Then:

- Some copyists were aware of the linguistic trends of their times. This explains the presence of learned and literary words as well as Latinisms, related in almost all cases to specialised botanical vocabulary: *junqueruela* 'reed', *sisba* 'fruit of the jujube tree', *pollicios* 'sprouts' or *ixola* 'prickly and bitter plant' — in some cases they come from Latin through Arabic loans. I have also found a few hapax — words for which there is no documentary evidence before in any other place, in Christian or Muslim context, as *reviḍa* ('another life') or *boticaxear* ('put makeup on someone).[32]
- From a morphological point of view, some copyists used the synthetic superlative in *-ísimo*, in words like *noblísimo*, very common in Spain at the end of the 16th century but never found before that time. That period coincides with the last copies in Aljamía, which strongly suggests that the copyist had the knowledge of the linguistic uses of their time.

In these cases, it shows that the copyists were educated persons, attentive to the various linguistic innovations of their time and more used to Spanish texts in Latin script than in Arabic script. They probably knew Arabic quite well, without being well versed in the language since many mistakes are found in the Arabic fragments. In general, the texts in Arabic show more errors than the Aljamiado texts, and in many cases, like BnF Arabe 447, the orthography in Latin script is more accurate than in the Arabic one.

Apart from the use of a specialised vocabulary and application of the new morphological derivations found in the rest of the contemporary Spanish literature, other features, such as divisions of words and phrases, help us to detect some copyists' knowledge and familiarity with Spanish texts in Latin script. Thus, while in the Middle Ages every kind of grouping of words is found, in the 16th and 17th centuries the graphic agglutination is more limited, and we only find phrases like *enella o d'ellos* (= preposition + article) written as a single graphic group. On the other hand, the divisions of the words at the end of a line seem to be random in the Middle Ages. However, in modern times, the 16th and 17th centuries, the divi-

32 The Arabisms found in Aljamía often play the role of the Latin loan words in the contemporary Christian prose (Montaner 2004, 99–204). See Martínez de Castilla 2010, 210–218.

sion (when it occurs) tends to be syllabic: *des-/pués* 'afterwards', and is not only found at the end of a line, but at the end of a folio (recto or verso) too, even if in Arabic this segmentation is not possible.

Then, we have to take with caution some conclusions advanced by Hegyi (1981, 22), and widely accepted until now, that a) the Aljamiado manuscripts were produced by isolated individuals whose only concern was to preserve the religious tradition for the next generations; and b) due to their social exclusion,[33] they kept an archaic language, far from any linguistic innovation, because they 'do not share the linguistic ideal of the Renaissance, and they stay on the sidelines of the linguistic movement which heightens the Spanish language through the cultisms from Latin and Italian'.[34]

5 Conclusion

In the Aljamiado texts, we see a consistent system used for the transliteration of the Romance language in Arabic script. From the oldest copies preserved (15th century) until the last ones produced by the Moriscos at the beginning of the 17th century, a standardised graphic and orthographic system has been applied to these Spanish texts written in Arabic script. However, it is impossible to assert when this system started to be used, because of the lack of evidence from more ancient manuscripts. On the other hand, with the exception of a later Aljamiado copy of the end of the 17th or beginning of the 18th century, kept in a private collection in London, no hint of the possible use of other transliteration systems has been preserved, which impedes us from suggesting any hypothesis about the period of use and the selection criteria of one or another system.

Concerning the language, the standardisation process is double in Aljamía: on the one hand, the translations of the Arabic texts show a peculiar version of the Spanish language, characterised by a higher proportion of Arabisms (in lexicon, syntax and morphology), Aragonese layer and archaisms. These three features, very consistent throughout the Aljamiado production, show a conservative stage of the Spanish language – probably artificial – in the 16th century, the period when most of the preserved manuscripts were copied. However, this con-

33 'Las condiciones socioculturales no permiten la plena participación del grupo minoritario en la cultura oficial del grupo mayoritario, impidiendo que las novedades lingüísticas se filtren hasta ellos'.

34 'No comparten el ideal lingüístico del Renacimiento [y, por extensión, de conocimientos e inquietudes], y se quedan al margen del movimiento lingüístico que enriquece la lengua española por medio de cultismos procedentes del latín y del italiano'.

servative stage does not result from a deficient knowledge of the contemporary uses of Spanish. Moriscos were aware of the standardised use of the language of their time, and it is clearly demonstrated in the faithful and correct transliterations of Christian texts in Arabic script; or even more, they do use this normalised Spanish language to produce new texts (rather than copies of translated Arabic texts). In most of the cases, it is impossible to ascertain only on the basis of the language itself whether those texts were written by a Morisco.

With the exception of the exact transliteration of some fragments of literature that were very popular in contemporary Europe, Aljamiado manuscripts show a standardised kind of content: an Islamic one. The texts are either directly related to the religion, such as prayers, Qur'anic passages, Islamic law, exegetical commentaries, polemical texts about religion, etc., or indirectly, such as edifying literature, magic, or entertaining literature. The main part of this production shows a consistent common thread: it gives an easy access to the basic knowledge of what Islam means and what is a good Muslim is required to do. The information can be provided in Arabic letters (in most of the cases) or Latin script. Even if many hypotheses about the reasons of the use of one or the other alphabet have been argued (period, place, familiarity with the Arabic language and script, cultural identity, etc.), we still lack a definitive explanation (if this is possible at all) about the selection of one or the other script system. Although they are different, both are used in a consistent and standardised way.

Abbreviations

Aix	Bibliothèque Méjanes, Aix-en-Provence
BCM	Biblioteca de Castilla-La Mancha, Toledo
BNE	Biblioteca Nacional de España, Madrid
BnF	Bibliothèque nationale de France, Paris
BPal	Real Biblioteca. Palacio Real, Madrid
BRAH	Biblioteca de la Real Academia de la Historia, Madrid
RBME	Real Biblioteca del Monasterio de San Lorenzo de El Escorial, Madrid
RESC	Biblioteca Tomás Navarro Tomás, Consejo Superior de Investigaciones Científicas/ Centro de Ciencias Humanas y Sociales, Madrid
T	Biblioteca de la Real Academia de la Historia, Madrid (old call number)
Urrea	private collection (manuscript found in Urrea de Jalón), Saragossa

References

Bernabé Pons, Luis (2010), 'Los manuscritos aljamiados como textos islámicos', in Alfredo Mateos Paramio and Juan Carlos Villaverde (eds), *Memoria de los moriscos. Escritos y relatos de una diáspora cultural* [Memory of the Moriscos. Writings and stories of a cultural diaspora]. Madrid: SECC/BNE, 27–44.

Brisville-Fertin, Olivier (unpublished), 'El aljamiado como discurso de una comunidad'.

Bouzineb, Hossain (1986), 'El valor exacto de los signos gráficos', in Abdeljelil Temimi (ed.), *Actes de la première table ronde du CIEM: La littérature aljamiado-morisque: hybridisme linguistique et univers discursif*. Tunis: Publications du Centre de Recherches en Bibliothéconomie et Sciences de l'Information, 29–34.

Castilla, Nuria de (2005), *El manuscrito aljamiado T19 de la Real Academia de la Historia. Edición y estudio* [The Aljamiado Manuscript T19 of the Real Academia de la Historia. Edition and study]. Madrid: UCM.

Castilla, Nuria de (2006a), 'Cohesión y control: la oración de los moriscos a través del ms. T19 (BRAH). Documentos escogidos', in *Actas del IX Simposio internacional de Mudejarismo*, 165–179. Teruel: Instituto de Estudios Turolenses.

Castilla, Nuria de (2006b), 'Anduve mirando si parecía por allí algún morisco aljamiado', in Nuria de Castilla and Rodolfo Gil Benumeya (eds), *Cervantes y el islam* [Cervantes and Islam], 235–246. Madrid: SECC.

Castilla, Nuria de (2006c), Review of M. E. Perry, *The Handless Maiden*, in *Al-Qanṭara*, 27.2, 565–570.

Castilla, Nuria de (2010), *Una biblioteca morisca entre dos tapas* [A Morisco library between two covers]. Saragossa: IEIOP.

Castilla, Nuria de (2014), 'Qur'ānic Manuscripts from Late Muslim Spain: The Collection of Almonacid de la Sierra', in *Journal of Qur'anic Studies,* 16.2, 89–138.

Castilla, Nuria de (2016), 'Were the Moriscos in touch with Contemporary Ottoman developments? Twin Qur'ānic copies of the end of the sixteenth century', in *Intellectual History of the Islamicate World*, 4: 245–264.

Castilla, Nuria de (2019), 'An Aljamía Translation of the "Morisco Qur'an" and its Arabic Text', in C. Adang, M. García-Arenal and F. Rodríguez Mediano (eds), *Translating Sacred Texts*, Leiden: Brill.

Castilla, Nuria de (forthcoming), *Manuscrits mudéjars et morisques en France* [Mudejars and Morisco manuscripts in France]. Rome: Istituto per l'Oriente C. A. Nallino.

Castro, Antonio (ed.) (1990), Pedro Mexía, *Silva de varia lección*, II. Madrid: Cátedra.

Cervera Fras, María José (2010), *Manuscritos moriscos aragoneses* [Morisco manuscripts from Aragon]. Saragossa: IEIOP.

Corriente, Federico (2000–2001), 'El romandalusí reflejado por el glosario botánico de Abulxayr', in *Estudios de dialectología norteafricana y andalusí*, 5: 93–241.

Corriente, Federico (2002), *Gramática árabe* [Arabic Grammar]. Barcelona: Herder.

Ferrer Valls, Teresa (1993), *Nobleza y espectáculo teatral. 1535–1622* [Aristocracy and theatre show. 1535–1622]. Valence: University.

Ferrando, Ignacio (1996), '¿Un céjel andalusí-aragonés del siglo XVI?', in *Estudios de dialecto andalusí*, I: 177–195.

Ferrando, Ignacio (2000), 'Andalusi Arabic in Post-Islamic North of Spain: the language of Aragon's Mudejars and Moriscos', in Manwell Mifsud (ed.), *Proceedings of the Third*

International Conference of the Association International de Dialectologie Arabe. Malta: AIDA, 195–200.
Galmés de Fuentes, Álvaro (ed.) (1970), *Los amores de París y Viana* [Paris and Viana's love]. Madrid: Gredos.
Galmés de Fuentes, Álvaro (ed.) (1975), *El libro de las batallas. Narraciones* épico-caballerescas [The Book of the Battles. Epic-chivalresque tales]. Madrid: Gredos.
Galmés de Fuentes, Álvaro (ed.) (1991), *Dichos de los siete sabios de Grecia* [Sayings of the seven wisemen from Greece]. Madrid: Gredos.
Galmés de Fuentes, Álvaro (1998), *Los manuscritos aljamiado-moriscos de la Real Academia de la Historia* [The Aljamiado-Morisco manuscripts of the Real Academia de la Historia], Madrid: Real Academia de la Historia.
Galmés de Fuentes, Álvaro (1986), 'La lengua española de la literatura aljamiado-morisca como expresión de una minoría religiosa', in *Revista española de lingüística,* 16.1: 21–38.
García-Arenal, Mercedes (2010), 'Musulmanes arabófonos y musulmanes aljamiados', in *Al-Qanṭara*, 31.1: 295–310.
Hegyi, Ottmar (1981), *Cinco leyendas y otros relatos moriscos (ms. 4953 de la Biblioteca Nacional de Madrid)* [Five legends and some other Morisco stories (ms. BNM 4953)]. Madrid: Gredos.
Hegyi, Ottmar (1985), 'Una variante islámica del español', in *Homenaje a Álvaro Galmés de Fuentes*, vol. I. Madrid: Gredos, 647–656.
Hoenerbach, Wilhem (1965), *Spanisch-islamische Urkunden aus der Zeit der Nasriden und Moriscos.* Berkeley, Los Angeles: University of California Press.
Labarta, Ana (1982), 'Una página aljamiada hallada en Torrellas (Tarazona)', in *Turiaso* III, 227–233.
Labarta, Ana and Ricardo García Cárcel (1981), 'Un fondo de documentos árabes de moriscos valencianos', in *Actas de las Jornadas de Cultura árabe e islámica.* Madrid: Instituto Hispano-árabe de Cultura, 395–400.
Martínez de Castilla, Nuria. See Castilla, Nuria de.
Montaner, Alberto (2004), 'La aljamía: una voz islámica en Aragón', in José María Enguita Utrilla (coord.), *Jornadas sobre la variación lingüística en Aragón a través de los textos*. Saragossa: Institución 'Fernando el Católico', 99–204.
Morel-Fatio, Alfred (1892), *Catalogue des manuscrits espagnols et des manuscrits portugais* [Catalogue of Spanish and Portuguese manuscripts]. Paris: Imprimerie Nationale.
Navarro Tomás, Tomás (1918), *Manual de pronunciación española* [Manual of Spanish pronunciation]. Madrid: CSIC [reprint 1999].
Perry, Mary Elizabeth (2005), *The Handless Maiden. Moriscos and the Politics of Religion in Early Modern Spain*. Princeton, Oxford: Princeton University Press.
Quilis, Antonio (1981), *Fonética acústica* [Acoustic phonetic]. Madrid: Gredos.
Ramos, Rafael (2012), 'Para la tradición del *Libro de los doce sabios*', in Natalia Fernández and María Fernández (eds), *Literatura medieval y renacentista. Líneas y pautas* [Medieval and Renaissance Literature. Guidelines and standards]. Salamanca: SEMYR, 843–853.
Sánchez Álvarez, Mercedes (1982), *El manuscrito misceláneo 774 de la Biblioteca Nacional de París* [The miscellany manuscript BNP 774]. Madrid: Gredos.
Sánchez Álvarez, Mercedes (1995), 'Observaciones sobre el arcaísmo lingüístico de los textos aljamiado-moriscos', in *Sharq al-Andalus*, 12: 339–348.

Steiger, Arnold (1991), *Contribución a la fonética del hispano-árabe y de los arabismos en el ibero-románico y el siciliano* [Contribution to the Hispano-Arabic Phonetic and to the Arabisms in the Ibero-Roman and Sicilian]. Madrid: CSIC.

Suárez, Raquel (2004), *El tratado de materia religiosa de Mohanmad de Vera (Ms. Esp. 397 de la BNP)* [The treatise of religious matter by Mohanmed de Vera (Ms. Esp. BNP 397)]. Oviedo: unpublished dissertation.

Thompson, Stith (1970), *Index of folk-literature*. Bloomington, London: Indiana University Press.

Valero Cuadra, Pino (ed.) (2000), *La leyenda de la doncella Carcayona* [The legend of the Carcayona maiden]. Alicante: University.

Vespertino, Antonio (2002–2004), 'El aragonés de la literatura aljamiado-morisca', in *Archivo de Filología Aragonesa*, 59–60, n. 2: 1731–1756.

Viguera, María Jesús (1982), 'Partición de herencia entre una familia mudéjar de Medinaceli', in *Al-Qanṭara*, 3.1: 73–134.

Viguera, María Jesús (1991), 'Cuentas aljamiadas de Joan Meçod (Almonacid de la Sierra, siglo XVI)', in *Homenaje a Juan Martínez Ruiz*. Granada: University, 129–141.

Villaverde, Juan Carlos (2010), 'Los manuscritos aljamiado-moriscos: hallazgos, colecciones, inventarios y otras noticias', in Alfredo Mateos Paramio and Juan Carlos Villaverde (eds), *Memoria de los moriscos. Escritos y relatos de una diáspora cultural* [Memory of the Moriscos. Writings and stories of a cultural diaspora]. Madrid: SECC/BNE, 91–128.

Vincent, Bernard (2006), 'Reflexión documentada sobre el uso del árabe y de las lenguas románicas en la España de los moriscos (ss. XVI–XVII)', in *El río morisco* [The Morisco river]. Valence: University of Valence, 105–118.

Jan Schmidt

How to write Turkish? The Vagaries of the Arabo-Persian Script in Ottoman-Turkish Texts

Abstract: The chapter discusses the development of the orthography of Turkish texts in, what the Turks usually call 'the old script', i.e. the Arabo-Persian alphabet in its various forms as it appears in texts handwritten, printed and engraved, which were produced in the Ottoman Empire and the Republic of Turkey before 1928.

1 Turkish (Turkic) texts before the Ottomans

Turkish has been written since the 8th century, perhaps even earlier, although no texts have survived from this earlier period. The first surviving texts were written in a 'runic' alphabet found on inscriptions in stone and on paper in Central Asia. It was based on the Aramaic alphabet in its Iranian form, possibly together with some letters adapted from the Greek (Hephtalite) alphabet, as Sir Gerard Clauson has suggested (Clauson 1970). This alphabet was in use for two centuries and was mainly employed for monumental inscriptions by local rulers who were part of the Göktürk ('Celestial Turkish') confederation.

With the rise of the Uighur Khanate (principality) from 742, a second alphabet, or rather a group of alphabets, came into existence. This used a cursive script that had its origin in the cursive Sogdian script,[1] which, in turn, was based on the Aramaic script. The script was mostly used for Christian, Manichaean and Buddhist texts. In 763, Manichaeism, spread by Sogdian preachers, was declared to be the state religion by the Uighur Khan, Bögü Kagan; after the fall of the Uighur Empire in the 840s, the Uighur Turks moved to the oases of eastern Turkestan. Some of them converted to Buddhism (and later to Islam). In the 13th century, Uighur Turks still played an important role as clerks, teachers and even ministers in the Mongol Empire, where the Uighur script also began to be used for Mongol

I would like to take the opportunity to thank Barbara Flemming and Edith Ambros for their help and useful advice with this paper.

1 The Sogdians were an Iranian people from Central Asia, best known from their role as merchants on the Silk Road.

https://doi.org/10.1515/9783110639063-006

texts. Unlike the runic script, the Uighur script was not forgotten after the political power of the people who used the script had waned, but kept its prestige, even once also after most Turks had converted to Islam.

Surviving copies of an important Turkish text of the early Islamic period entitled *Kutadgu Bilig* ('Fortune-bringing Knowledge'), a work in the 'mirror of princes' genre from the 11th century written by Yusuf Khass Hacib of Balasagun for the Karakhanid prince of Kashgar, were not only written in Arabic, but also in the Uighur script. Later the script was still occasionally used in the chancelleries of the Timurids and Ottomans. The latest text in the script was perhaps the 'proclamation' meant as a political statement to other rulers of Turkish and Jingizid descent, in which the people of the *vilayet-i Rum* (Anatolia) were notified of the defeat of the Akkoyunlu prince Uzun Hasan at the hands of Sultan Mehmed the Conqueror in 1473 at the battle of Tercan (Otlukbeli). The text is preserved in the archives of Topkapı Palace in Istanbul.

With the adoption of Islam, the Arabic alphabet began to be used for texts in the Turkic languages of Central Asia. This was particularly stimulated by the voluntary conversion to Islam of the aforementioned Karakhanid dynasty in 960 (the dynasty ruled from 840 to 1212 in Turkestan). During their rule, Mahmud al-Kashgari, possibly a member of the ruling dynasty, wrote a work on the Turkish language in Arabic for the Abbasid Caliph in Baghdad, which was entitled *Dīwān Lughāt at-Turk* ('Compendium of Turkish Dialects'). This was in the 1070s.

In the period between the 11th and 13th century, Turkish (Turkic) hardly appeared at all as a written language in the Islamic domain. This changed in the 13th century, albeit hesitantly at first, when Turkish (Turkic) began to be written again in Anatolia, where the Ottoman state had its cradle during the later years of the Rum-Seljuk sultanate. The Seljuks, despite their Turkic background, had used Persian as their official language ever since it had begun to dominate the area in the 11th century. The power of the sultanate slowly crumbled and a number of small principalities arose, including the Ottoman realm, which ultimately took its place. Its rulers began to use Turkish – or, to be more specific, 'West Oghuz Turkic', the language spoken by the Turks who had migrated south-westwards into the area over the previous centuries – as their language of administration. 'There was then', in the words of Celia Kerslake (Kerslake 1998, 179), 'a concomitant upsurge of Turkish literary activity, much of it religious-didactic in character and often consisting of translation or adaption from Persian'. In the early period, before about 1300, this literature still showed evidence of linguistic influence from the older Karakhanid literary tradition, but that impact disappeared soon afterwards. What these early texts shared was a tendency to use – in the Uighur way – a spelling which employed very few vowels, whereas later Ottoman texts, appearing after approximately 1500, were written in a more explicit spelling.

2 Turkish texts in the Ottoman Empire

With the transformation of the Ottoman principality, which was founded around 1300, into a regional empire 'the Ottoman court and ruling class acquired a consciousness of imperial power [and] certain stylistic registers emerged – in which [written Turkish] was all but submerged beneath a heavy overlay of Arabic and Persian elements. These reflected the dominance of Arabic and Persian in the Islamic high culture and learning which the Ottoman elite sought to emulate and advance' (Kerslake 1998,179–180). These Arabic and Persian elements were not restricted to lexical features but included grammatical and syntactic ones as well. This went so far, in fact, that some works penned in this elevated literary style in the 17th century, when the development reached its peak, look like Persian texts with a Turkish conjunction or verb here and there. The spelling of these Arabic and Persian elements remained unchanged, but in some popular texts written by less well-educated people (and which often had their origins in oral literature), one sometimes finds orthographic adaptations based on the pronunciation of such foreign elements by Turks (see below for an example).

As for the orthography of the Turkish elements, no consensus was reached among writers and copyists about the right form, so it remained inconsistent. Only during the period of reform in the 19th century when manuscripts were replaced by printed texts were the first attempts made 'to define and regularize the language', in particular for didactic purposes. Reform also meant reform of public education, and Turkish became part of the curriculum in elementary and secondary schools for the very first time. The language was now specifically called 'Ottoman Turkish' (Kerslake 1998, 180) and developed into something resembling standard Turkish.

So, unlike Arabic, literary Turkish was not pressed into a 'classic' or 'standard' Turkish at an early stage. The spelling of the Turkish elements in the language in the older period was influenced by various complicating (and little-studied) factors, such as regional differences in the spoken language. The expansion of the Ottoman state and the migration of Turkish-speaking subjects within its expanding borders, which came to include North Africa, the Near East and southeastern Europe, meant that the kind of Turkish spoken in Niksar in 1500 must have been different from Turkish spoken in Sarajevo the same year. Another factor was the 'natural' development of Ottoman Turkish: Turkish spoken in Niksar in 1500 certainly was different from Turkish spoken in the same town in 1820. Change in the language, such as the gradual change in the system of vowel harmony, did not always lead to a change of spelling (a phenomenon also known in English and other languages), hence the orthography of late Ottoman Turkish contained a number of archaisms (see below).

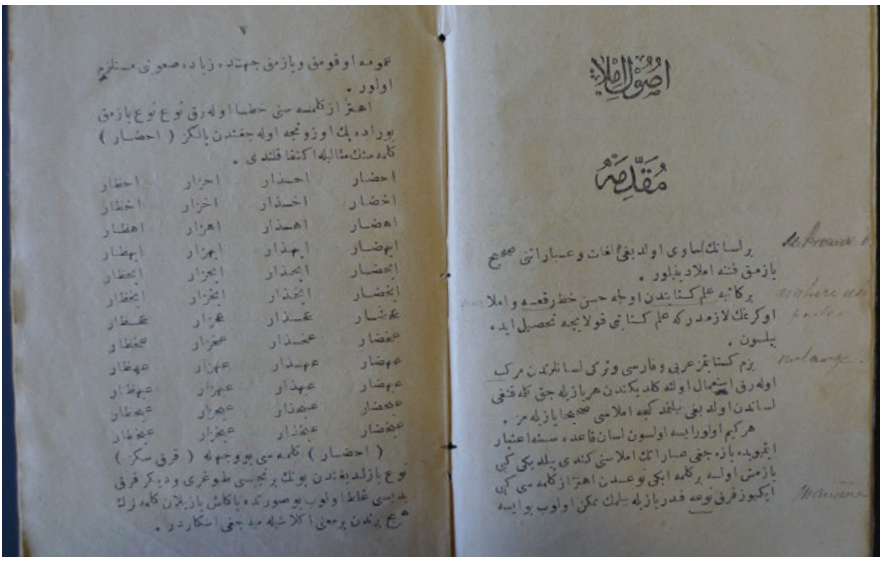
To a small extent, spelling was also influenced by the non-Ottoman Turkic language Chagatai (literary eastern Turkic), represented for example by the works of the Timurid scholar and poet Ali Shir Nawa'i (15th century), and literary Azeri Turkish also had an impact, represented first and foremost by the poet Fuzuli (16th century). These poets were admired by Ottoman literati and their work was even considered to be part of the Ottoman literary canon by some of them. A number of odd particularities in the spelling of certain Ottoman Turkish suffixes have been attributed to this influence.

These elements complicated the standardisation of Ottoman orthography, but that is not to say that chaos ruled. By approximately 1500, when old Ottoman Turkish had developed into early modern ('middle') Ottoman Turkish, a preferred way of writing the language adhered to by most authors and scribes emerged. In the 16th century, the court in Istanbul (established in 1453) became the dominant centre of literary culture (Hazai 2012). Perusing Ottoman manuscript texts written or copied between, say, the late 15th and early 20th centuries makes it abundantly clear that, with a few exceptions, the spelling of the Turkish lexical elements is quite consistent and variation is only marginal. The exceptions here, again, are mostly the few surviving manuscripts produced by less well-educated authors/scribes. As Mehmet Yastı has shown, in such literature, there could be a wide range of orthographic inconsistencies even within one manuscript. In a copy of a work entitled *Esrarü l-Arifin*, a didactic piece on mysticism in simple Turkish by a certain Seyfüllah Kasım el-Bağdadi, preserved in a copy made in 1498 and kept in the Süleymaniye Library in Istanbul, the word *uçmak* ('to fly') was spelt in four different ways and the word *zira* ('because') in as many as five. The letters *cim* (ج) and *çim* (چ) were mixed up – one finds both *gece* and *geçe* ('night'), for instance. Emphatic and non-emphatic consonants could be alternated: the word *su* ('water') appeared either with a *sin* (س) or a *sad* (ص), defective *benüm* (بنم) is found alongside *plene benūm* (بنوم), etc. (Yastı 2010). Perhaps the need for consistent spelling was felt less by the author, because he vocalised the text completely and therefore reading problems hardly existed. An interesting later example is provided by the Leiden manuscript Or. 1551, the memoirs of a mercenary soldier composed in 1834, which survives in a unique but incomplete copy in a codex format that is different from 'normal' Ottoman texts (see Fig. 1). There are strong clues in the text indicating that the author, who may have been illiterate, dictated it to a scribe.

Fig. 1: Title page of the Leiden MS Or. 1551. © Courtesy of the University Library Leiden.

The spelling of the Turkish words and the Arabic and Persian loanwords is highly erratic. Here are some examples (the column on the right gives the modern spelling in Latin script):

MS Or. 1551		**Normal spelling**		**Modern Turkish spelling**	
بك	<bk>	بيك	<byk>	*bin*	'thousand'
طفروج	<ṭfrwj>	تفرج (Ar.)	<tfrj>	*teferrüc*	'inspection'
ويلايتى	<wylʾyty>	ولايتى (Ar.+T.)	<wlʾyty>	*vilayeti*	'province'
تألف	<tʾʔlf>	تأليف (Ar.)	<tʾʔlyf>	*te'lif*	'author'



اصول املا

مقدمه

Fig. 2: Opening pages of *Usul-ı imla*. © Courtesy of the University Library Leiden.

During the period of reform, especially after 1908 when the Young Turks came to power, attempts were made to formalise the existing tendency to standardisation by publishing spelling manuals.[2] At least one of these was issued by the Ministry of Education on behalf of the Grammar and Orthography Committee (in 1917–19).[3] An earlier example, probably the first of its genre and intended for use by professional clerks, was Mehmed Raşid's *Usul-ı imla* ('Principles of Orthography') from 1886 (Fig. 2).[4] In the introduction, the author expresses his hope that the book will help the reader to avoid using idiosyncratic spellings of Turkish words and making spelling mistakes in Arabic and Persian loanwords; these would harm their professional reputation and even put them to shame ('*muharririnin fazl u kemaletine elbette nakise tiraş eder… belki hacaletini mucib olur*', pp. 3–4). To illustrate a particularly dramatic example of the way in which Arabic loanwords used to be misspelt, he presents a list of 48 orthographic variants of the word *ihzar* ('preparation') in the Preface, only one of which was correct (p. 7).

2 See under 'İmlâ' and 'Usul-ı imlâ' in Özege 1971–79.

3 Analysed in Gümüş 2008.

4 The year is mentioned in the Leiden University Library catalogue; Özege's bibliography does not mention a year of publication, nor is one found in the book itself. The author, as he explains in the introduction, was a teacher of calligraphy and orthography at the Naval Academy (Mekteb-

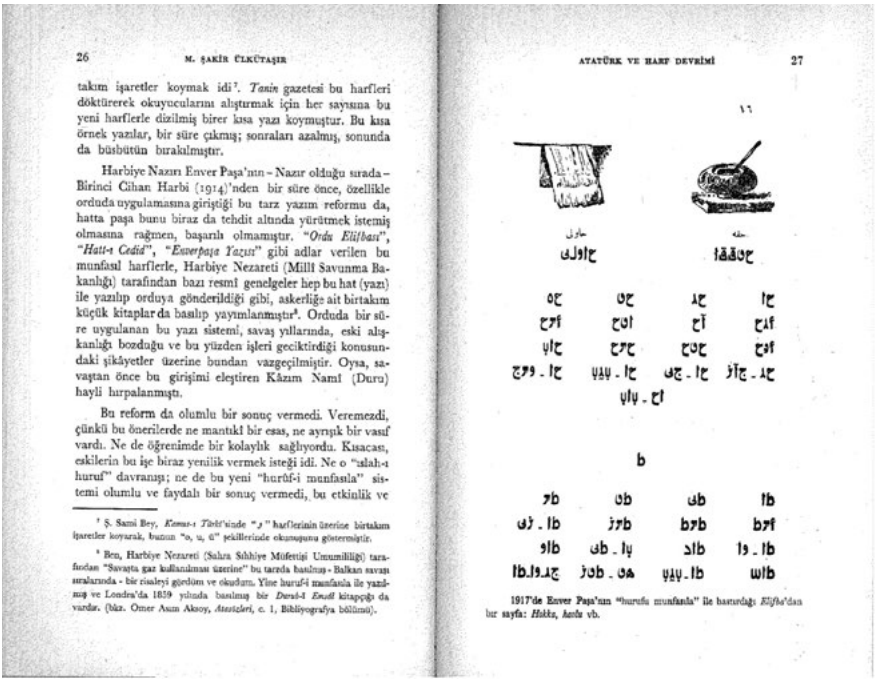
3 The orthography of Ottoman Turkish

So far, we have looked at some of the major tendencies in the development of Ottoman Turkish orthography. But how did the Turks adapt the Arabo-Persian alphabet to spell their own language? One problem here was that the Arabic script, like the Aramaic-based Uighur script, contains, to put it simply, too many consonants and too few vowels to do the job well. Basically (and ignoring historical, regional and other nuances for a moment), Ottoman Turkish contains twenty consonants and eight vowels, whereas the Arabo-Persian script has 31 consonantal graphemes and three vowel graphemes (the latter also functioning as consonants). Aesthetic value aside, in the words of Geoffrey Lewis (Lewis 2010, 27), 'there is nothing to be said in favour of the Arabo-Persian alphabet as a medium for writing Turkish'. There were various ways of dealing with the lack of vowels: (1) not to write any at all, (2) to use the three main (long) vowels in Arabic, *alif*, *wāw* and *yā'*, to indicate a/e, o/ö/u/ü and i/ay/ey respectively, or (3) to use superfluous 'emphatic' (velar) consonants for indicating unwritten 'back vowels' (a, o, u and ı) and non-emphatic consonants for indicating 'front vowels' (e, ö, ü, i). The matter was complicated further by the letter *kāf*, which could represent g, k, n, v or y (g and n in more *plene* texts could be specified by an extra slanting dash گ or three dots respectively). As a result, and in conjunction with the occurrence of a plethora of Arabic and Persian loanwords in most texts, many equivocal readings were possible: اولو may be read as Turkish *ulu* 'great', *ulu* (Arabic: 'possessors') or ölü 'dead', *evli* 'married', *avlu* 'courtyard' or *avlı* 'stocked with game'. دول can be *döl* 'progeny', *dul* 'widowed' or *düvel* (Arabic: 'states'). كل can be *gel* 'come', *gül* 'smile', *kel* 'scabby', *kel* (Arabic: 'lassitude'), *kül* 'ashes', *küll* (Arabic: 'all'), *gil* (Persian: 'clay') or *gül* (Persian: 'rose'. More reading problems arose because scribes and printers were not always careful about word divisions: بوسنه could stand for *bu sene* 'this year' or *Bosna* 'Bosnia', for instance.[5] Retention of archaic spellings or adoption of Chagatai usage caused words like ايو to be read as *iyi* 'good' long after ceasing to be pronounced '*eyü*' (the word *evli* just mentioned belonged in the same category, being spelt <evlw>) and كوپرى (*köpri*, 'bridge') was to be read as *köprü*.

In order to fix some of the existing drawbacks, from the middle of the 19th century onwards, ideas were launched for modifying the alphabet and/or changing the orthography of certain words, particularly by writing more *plene*, that is, using more letters per syllable, and more phonetically, e.g. by replacing *eyü* by *iyi* (ايى) and *köpri* by *köprü* (كوپرو).

i Bahriye) in Istanbul and had composed the treatise for his students. He claimed that it was the first publication of its kind.

5 These examples are taken from Lewis 2010, 27–28.



26 M. ŞAKİR ÜLKÜTAŞIR

takım işaretler koymak idi[1]. *Tanin* gazetesi bu harfleri döktürerek okuyucularını alıştırmak için her sayısına bu yeni harflerle dizilmiş birer kısa yazı koymuştur. Bu kısa örnek yazılar, bir süre çıkmış; sonraları azalmış, sonunda da büsbütün bırakılmıştır.

Harbiye Nazırı Enver Paşa'nın – Nazır olduğu sırada – Birinci Cihan Harbi (1914)'nden bir süre önce, özellikle orduda uygulamasına giriştiği bu tarz yazım reformu da, hatta paşa bunu biraz da tehdit altında yürütmek istemiş olmasına rağmen, başarılı olmamıştır. *"Ordu Elifbası"*, *"Hatt-ı Cedid"*, *"Enverpaşa Yazısı"* gibi adlar verilen bu munfasıl harflerle, Harbiye Nezareti (Millî Savunma Bakanlığı) tarafından bazı resmî genelgeler hep bu hat (yazı) ile yazılıp orduya gönderildiği gibi, askerliğe ait birtakım küçük kitaplar da basılıp yayımlanmıştır[2]. Orduda bir süre uygulanan bu yazı sistemi, savaş yıllarında, eski alışkanlığı bozduğu ve bu yüzden işleri geciktirdiği konusundaki şikâyetler üzerine bundan vazgeçilmiştir. Oysa, savaştan önce bu girişimi eleştiren Kâzım Namî (Duru) hayli hırpalanmıştı.

Bu reform da olumlu bir sonuç vermedi. Veremezdi, çünkü bu önerilerde ne mantıklı bir esas, ne ayrışık bir vasıf vardı. Ne de öğrenimde bir kolaylık sağlıyordu. Kısacası, eskilerin bu işe biraz yenilik vermek isteği idi. Ne o "ıslah-ı huruf" davranışı; ne de bu yeni "hurûf-i munfasıla" sistemi olumlu ve faydalı bir sonuç vermedi, bu etkinlik ve

[1] Ş. Sami Bey, *Kamus-ı Türkî*'sinde "ɔ" harflerinin üzerine birtakım işaretler koyarak, bunun "o, u, ü" şekillerinde okunuşunu göstermiştir.

[2] Ben, Harbiye Nezareti (Sıhhıye Müfettişi Umumîliği) tarafından "Savaşta gaz kullanılması üzerine" bu tarzda basılmış - Balkan savaşı sıralarında - bir risaleyi gördüm ve okudum. Yine hurûf-i munfasıla ile yazılmış ve Londra'da 1859 yılında basılmış bir *Durûs-ı Emsâl* kitapçığı da vardır. (bkz. Ömer Asım Aksoy, *Atatürkçü*, c. 1, Bibliyografya bölümü).

ATATÜRK VE HARF DEVRİMİ 27

1917'de Enver Paşa'nın "hurufu munfasıla" ile basturdığı *Elifba*'dan bir sayfa: *Hokka, kardı* vb.

Fig. 3: 'Simplified' spelling initiated by Enver Pasha, early 20th century. The two illustrated words on the top right-hand page are *havlu* 'towel' and *hokka* 'inkpot'. (From M. Şakir Ülkütaşır', *Atatürk ve harf devrimi*, Ankara 1974).

In order to overcome the shortcomings of the script, the statesman and intellectual Münif Pasha suggested in 1862 that diacritics – both existing and newly invented ones – should be used to compensate for the deficiencies in existing spelling or that the letters should be written separately, which would avoid the need for more than one form of a letter. A year later, the Azerbaijani scholar Feth-Ali Ahundzade held a public lecture in Istanbul in which he proposed to invent new letters to indicate vowels. None of these plans came to much, however, although in at least one major publication, Şemsüddin Sami's famous encyclopaedia, *Kamus*, from 1901, three diacritics were used to distinguish between o, ü and ö (normally all indicated by a ﻭ). During this late period, the notion of writing the letters separately resurfaced. A serious attempt to implement the idea was made by Enver Pasha, the politician and member of the ruling triumvirate, between 1913 and 1917; he especially intended to simplify the work of military telegraphists, and various texts produced by the Ministry of

War were consequently published in the new script (see Fig. 3).[6] None of these steps proved to be adequate or even acceptable, though, and the Arabo-Persian script was eventually given up altogether (in 1928).

As regards the various styles of scripts, in what follows, I will only discuss the types most commonly found and not indulge in calligraphic finesses and the nuances of sub-types, which are often difficult for the non-specialist to distinguish.[7]

4 Ottoman scripts: variation in the outward appearance of letters and their use

Along with the Arabic and Persian languages and literatures, the Ottomans inherited a manuscript culture from their predecessors, the Turkish principalities (emirates) in Anatolia, which in part were inheritors of the Seljuks. As far as I can tell, the Ottomans adopted the culture almost wholesale and, apart from making modifications to some of the inherited types of script and an additional language, Turkish, did not alter it or add to it much.

The most widely used type was, doubtless, *naskhī* (or *nesih* in Turkish) (see Fig. 4). This was chosen for works of prose, which also encompassed religious and scholarly texts including the Qur'an, and for histories, collections of stories and other such accounts. Legend has it that the script was invented by Ibn Muqla (d. 940). It was the most readable script of all at the time and was meant to furnish easy access to texts. Used in conjunction with diacritical marks, mostly employed in religious writings meant for a broader readership, it gave little room for error or misinterpretation. Its forms hardly changed over the centuries. It was also predominantly used in printed books, journals and newspapers from the mid-18th century onwards.

A larger and more artistic variant of *nesih* was called *thuluth* (*sülüs*). It was sometimes used for book or chapter titles or religious formulae, like the *bismillah* in Ottoman manuscripts, but is rarely encountered in archival documents (where *nesih*, *divani* and *rikʿa* predominate; see below). The script was mostly used for large inscriptions in and on public buildings like mosques and endowment libraries and is often found in calligraphic albums as well.

6 Ibid., 28–29.

7 In what follows, I rely mostly on Aktan 1995, 31–72.

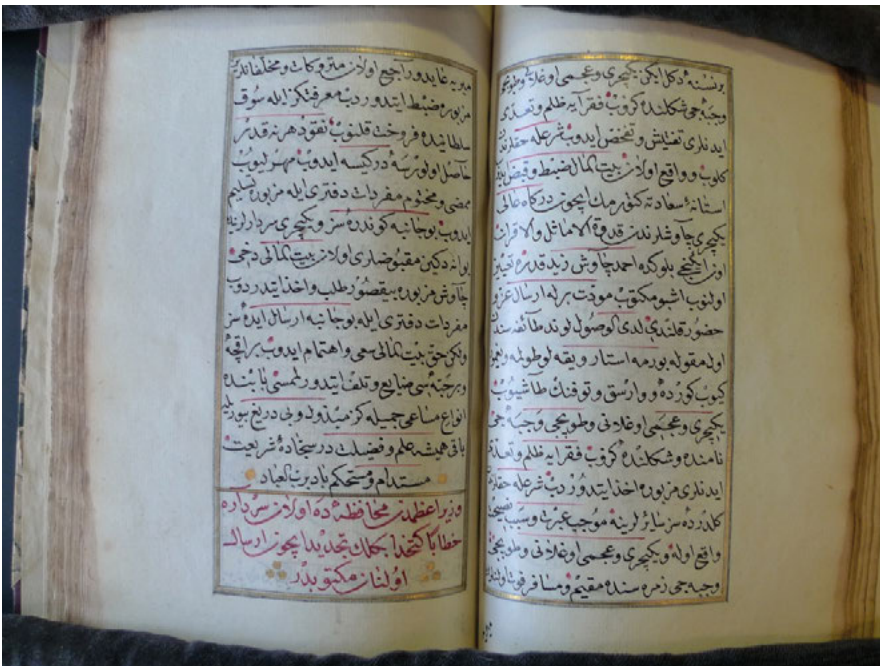


Fig. 4: Two pages from a multiple-text manuscript with text in ornamental *nesih*, 17th century. Leiden University Library, MS Or. 12.411. © Courtesy of the University Library Leiden.

A script of Persian origin – tradition has it that it was introduced into the Ottoman Empire by Persian calligraphers during the reign of Mehmed the Conqueror – also used what was known as *taʿlīq* (literally meaning 'suspension' in Arabic and called *nastaʿlīq* in Persian) (see Fig. 5). This became widely used in Ottoman manuscripts containing collections of poetry and other texts in the genre of *belles lettres*. It was also chosen for documents produced by the religious administration, such as charters of pious endowments and court registers. In its calligraphic (*celi*) variant, it is found on signboards and inscriptions on monuments and buildings which were part of pious endowments, such as mosques, religious schools (*medreses*), libraries and fountains; these texts are often in the artistic form of poems. The opposite of calligraphic *taʿlīq* is called *şikeste* (*shikastah* in Persian, literary meaning 'broken' and a 'sloppy', irregular version of the script), which was used for informal letters and notes by private individuals who did not want or need to spend time writing particularly neatly. It was not used much in the Ottoman Empire where more casual, idiosyncratic versions of *rikʿa*, to which I will return later, were preferred.

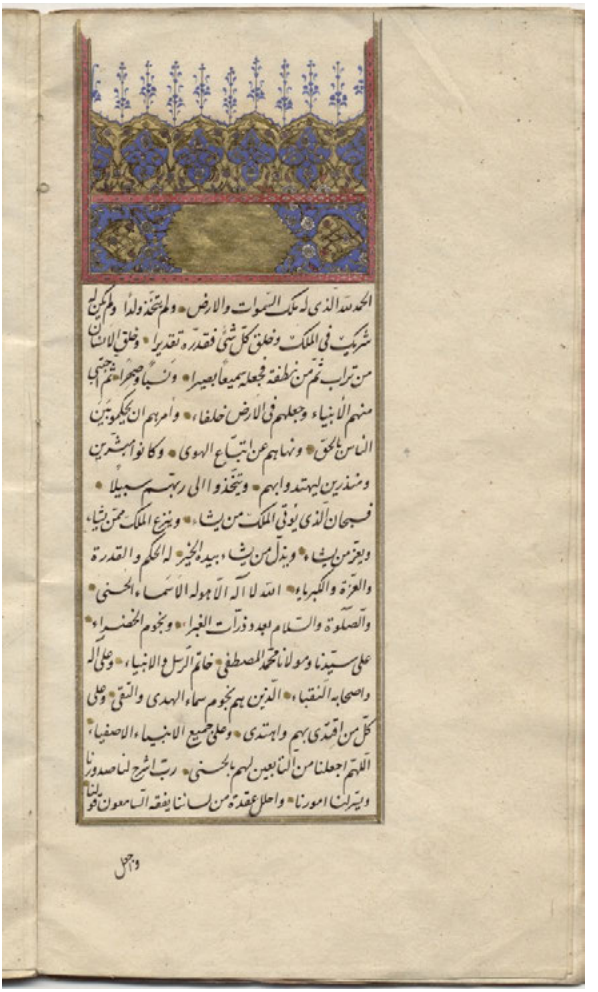


Fig. 5: Opening page of an Ottoman 'mirror for princes' with text in *taʿlīq*, 17th century. Leiden University Library, MS Or. 625. © Courtesy of the University Library Leiden.

There were two scripts which were almost exclusively used in the Ottoman bureaucracy. They shared the feature of irregular ligatures, meaning that in principle all the letters could be connected. Firstly, there was *siyakat* (*siyāqah* in Arabic), a defective script (dots are generally left out) expressing a specialised, encoded Arabo-Persian terminology used in the financial administration and inherited from the Seljuks (see Fig. 6).

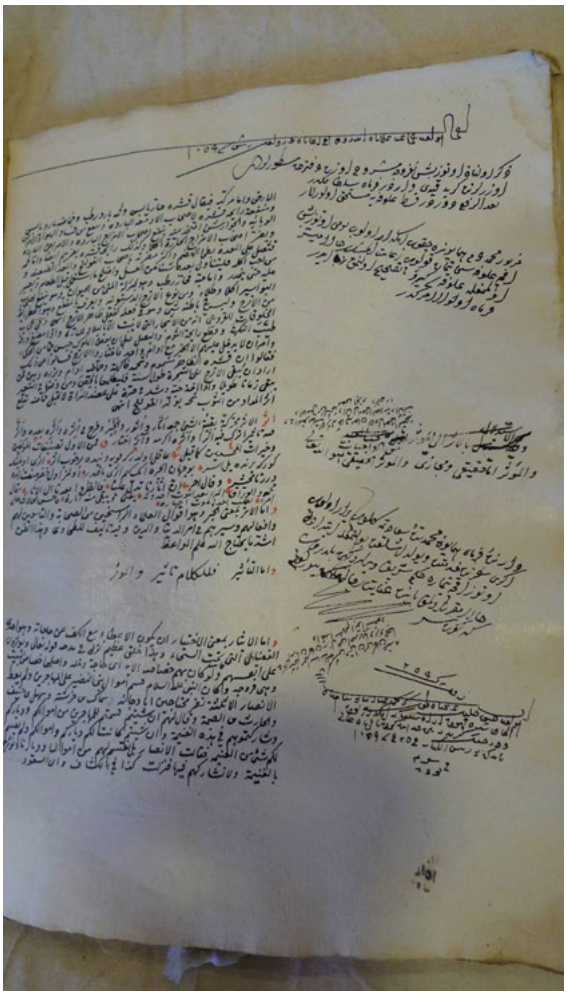


Fig. 6: A composite manuscript, 16th century, with marginal notes, partly in *siyakat* script. Leiden University Library, MS Or. 644, fol. 43b. © Courtesy of the University Library Leiden.

This was used for financial documents produced by various state institutions, including land registers, cadastral entries, parts of deeds of pious endowments and military records. To complicate matters (and increasing its illegibility for the uninitiated), the script changed greatly over time (which was hardly the case with the other Ottoman scripts – why this was so is unknown). The script also used special numbers consisting of Arabic letters meant to avoid falsification. *Siyakat* was abandoned during the last quarter of the 19th century.

Secondly, there was *divani* ('pertaining to the Council of State', also known as *dīwānī*), a calligraphic script used in official documents such as sultans' edicts (*fermans*), state letters and ministerial registers. This was invented by the Ottomans, although its origin is obscure (according to the polyhistor Mustafa Âli (d. 1600), it was developed from a *taʿlik*-like Persian script); it first appears in the 15th century in official documents and reached its full (classic) development in the 18th century. Like *siyakat*, the script was rarely printed. This may have had something to do with its illegibility for the lay reader. (Occasionally, *divani* was used for other genres as well (see Fig. 7); because the Dutch pastor and Orientalist Johannes Heyman (d. 1737) had texts copied by the Dutch consular secretary at Izmir, who normally used *divani*, some manuscripts in his library – now kept by the Leiden University Library – contain texts written in *divani*, a number of which are literary pieces.)

Finally, there was a widely used script called *rikʿa* (or *ruqʿa*, originally an Arabic term meaning a piece of paper or a brief message, among other things) (see Fig. 8). It was basically a simplified and more easily legible *divani* which had the practical advantage that it could be written quickly because of the large number of ligatures it contained. It was an Ottoman invention, appearing for the first time in 16th-century records, and eventually developed into a separate script in the 17th century. *Rikʿa* came to full fruition in the 19th century and was later adopted by Arabs and Persians alike. It was adopted widely in the Ottoman bureaucracy and became the predominant script for personal use in the late 18th century, when it also began to be used in manuscripts with texts in various genres. It was sometimes, albeit rarely, printed or used for inscriptions on seal rings or medallions, for instance. It is perhaps the only Ottoman script still used in Turkey today – by very old people who received their elementary education before 1928, which was when the Arabo-Persian scripts were abandoned by the government (suppressed, in fact) and the Latin alphabet was made compulsory for public use.

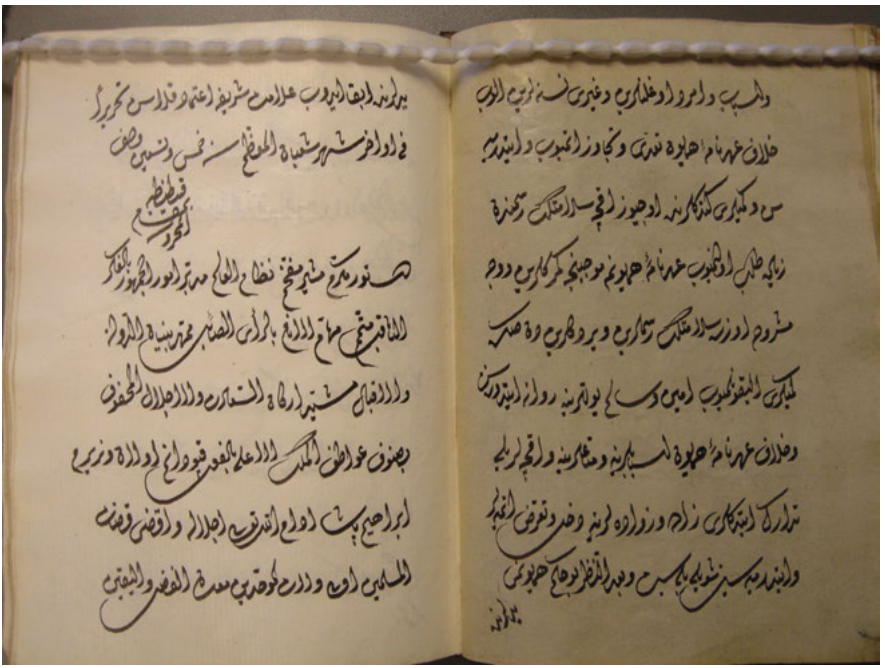


Fig. 7: Two pages from a multiple-text manuscript compiled by Johannes Heyman with texts in *divani* script, late 17th century. Leiden University Library, MS Or. 87. © Courtesy of the University Library Leiden.

5 Conclusion

So far, a few remarks have been made on the spelling of Turkish and the most common types of script found in Ottoman manuscripts, including archival documents. Little research has been done on aspects of Ottoman Turkish orthography as yet and we are far from having a comprehensive study on its history at our disposal. This is no surprise in view of the fact that our knowledge of the historical development of the Ottoman Turkish language is still rudimentary (not least, paradoxically, because linguists are mostly dependent on texts in a defective script produced by and for an urban elite). In this paper, I have largely based my findings on my own limited experience as a reader of Ottoman manuscripts and printed texts.

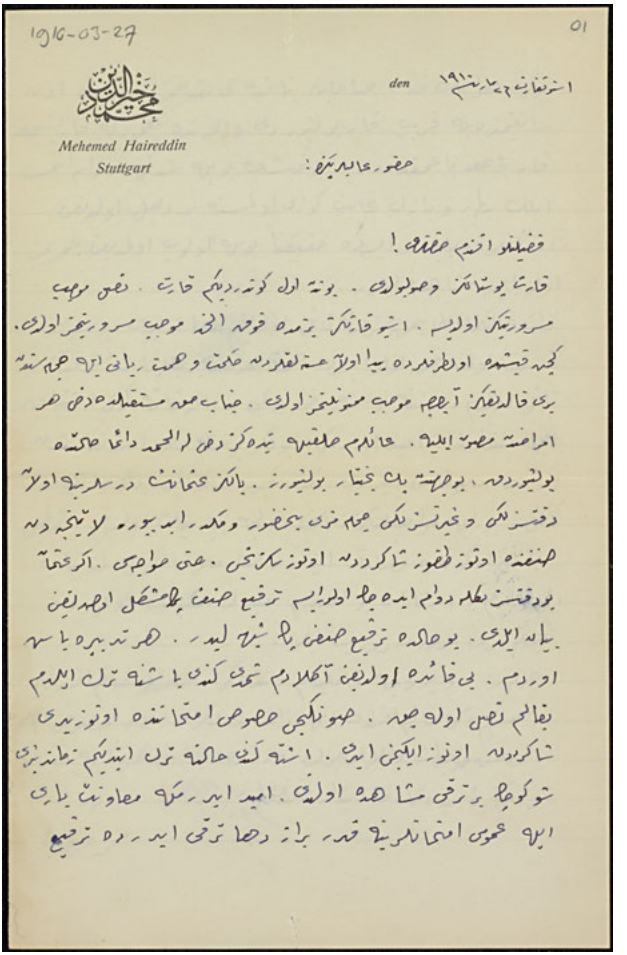


Fig. 8: A letter from Mehmed Hayrüddin in Stuttgart to Christiaan Snouck Hurgronje in Leiden, *rik'a* script, 1910. Leiden University Library, Or. 8952. © Courtesy of the University Library Leiden.

As for the central theme of this collective volume, standardisation, I would say that a tendency towards standardisation is visible in Ottoman texts written between 1300 and 1928, with the exception of texts produced outside the mainstream genres, but it was never actually formalised. Once the Ottoman Empire was securely established around the dynastic court in Istanbul in 1453 and early modern Turkish had replaced old Turkish around 1500, the spelling of texts – at least those serving the literary elite – began to become relatively stable. As far as I am aware, there were no religious or – before 1908 – political incentives, let alone

a state policy, that contributed to this development. Manuscript production was diffuse, that is, it was not concentrated in one or a few workshops, at least outside the palace and the state bureaucracy. As far as we can tell from colophons in manuscripts, the copying of texts was a thoroughly decentralised and individual activity. Publishing as a commercial activity only came with printing in the 19th century and will inevitably have contributed to standardisation even further. The formats used for manuscripts did not influence spelling, it seems.

For most Ottomans, formats and the aesthetic aspect of the script in its various styles seem to have been more important than the way in which words were spelt, and they continued to develop the shape of the letters and their ligatures in the footsteps of the Seljuks before them, in particular. Calligraphy was considered to be one of the highest forms of art, if not an esoteric science, compared to which mere orthography was insignificant.

References

Aktan, Ali (1995), *Osmanlı paleografyası ve siyasî yazışmaları*, Istanbul: Osmanlılar İlim ve İrfan Vakfı.

Clauson, Sir Gerard (1970), 'The origin of the Turkish "runic" alphabet', in *Acta Orientalia*, XXXII: 51–76.

Gümüş, Zehra (2008), 'Sarf ve Imlâ Encümeni'nin "Usûl-ı imlâ" isimli eseri ve imlâ meselerimiz', in *Turkish Studies*, 3/6: 289–305.

Hazai, György (2012), 'La place du XIVe siècle dans l'évolution de la langue turque', in Barbara Kellner-Heinkele, Simone-Christiane Rachmann et al. (eds), *Monumenta et Studia Turcologica: Ausgewählte Schriften von György Hazai*. Berlin: Klaus Schwartz, 250–266.

Kerslake, Celia (1998), 'Ottoman Turkish', in Lars Johanson and Éva Á. Csató (eds), *The Turkish Languages*. London, New York: Routledge, 179–202.

Lewis, Geoffrey (2010) (reprint), *The Turkish Language Reform. A Catastrophic Success*. Oxford: Oxford University Press.

Özege, M. Seyfettin (1971–1979), '*Eski harflerle basılmış Türkçe eserler kataloğu*', 5 vols, Istanbul.

Sâmî, Şemseddin (1317 [1901]), *Kamus-u türkı; kâffe-i lûgāt-i türkiye ile lisan-ı türkı'de müstamel kelimat ve ıistıila-hat-ı arabiye ve farisiye ve ecnebiyeyi cami olarak lisanımızın mükemmel lûgat kitabıdır.* Deri-Saadet, İkdam Matbaası.

Yastı, Mehmet (2010), 'Esrârü 'l-'Arifîn ve İmla Özellikleri Üzerine Bir Değlendirme', in *Türkiyat Araştırmaları Dergisi*, 28: 69–107.

Branka Ivušić

Developing Consistency in the Absence of Standards – A Manuscript as a Melting-Pot of Languages, Religions and Writing Systems

Abstract: This chapter deals with a late 16th-century multilingual Ottoman manuscript. It is a study of the anonymous scribe's attempt to write European languages in Arabic script and it tries to answer questions about possible orthographic models, the transfer of writing conventions from one language to another or from one script to another. The paper also looks at the scribe's own innovations and the consistency of the spelling system employed in this manuscript.

1 Introduction

1.1 An 'Oriental' manuscript and its 'European' texts

In the collection of Oriental manuscripts preserved in the Austrian National Library (Österreichische Nationalbibliothek, ÖNB) in Vienna we find a rather puzzling Ottoman multiple-text manuscript. A small, but not miniature volume of 162 folios[1], written in Arabic *naskhī* script and apparently by one hand only, it contains more than 300 texts in Ottoman Turkish, Persian and Arabic, but also in Croatian (a variety of Central Southern Slavonic that might be termed Croatian), Hungarian, German and Latin. The manuscript is assumed to be the work of an anonymous compiler, and, judging from its very regular appearance and the remarkably low number of corrections, it is likely to be a fair copy.

The number of languages found in the manuscript creates a hybrid impression which is only enhanced when we look at the contents. Most of the texts are

The present study is part of my PhD research at the Centre for the Study of Manuscript Cultures (CSMC) in Hamburg supported by the DFG.

1 Except for the last ten folios, which were left blank, and folio 107, which has only a catchword, all folios are covered by writing on both sides. The average page displays one justified block of 15 lines and, if it is a verso-page, a catchword in the left lower margin.

https://doi.org/10.1515/9783110639063-007

Ottoman Turkish. They range from a language guide to Persian, instructions in Islamic faith, astrology and mathematics to love poetry. The few Arabic texts are exclusively religious, mostly collections of *aḥādīṯ* while the Persian sample consists only of a few poems along with bits of Persian in the language guide. Such texts and this kind of content-specific 'division of labour' among the three most important languages of the empire are by no means uncommon for an Ottoman multiple-text manuscript, a so-called *mecmua*. The "European" samples on the other hand are not only remarkable because they show the use of Arabic script – and to some degree Ottoman writing conventions – for languages usually written in Latin characters (alongside Cyrillic and Glagolitic, in the case of Croatian) but also because of the content of their texts. All four languages appear first in a penta-lingual section of the manuscript introduced by a heading and a narrative paratext – both in Turkish – stating that what follows are the essentials of Christian and Jewish beliefs. These essentials turn out to be the Ten Commandments, The Lord's Prayer, and the Apostolic Creed in their Lutheran versions, split into smaller parts, which are usually numbered and presented first in Turkish, then in Croatian, Hungarian, German, and finally in Latin. The penta-lingual section is followed by one without any introductory paratexts comprising songs in all the European languages except Latin. With fourteen Christian religious songs, some of them composed by Martin Luther and by other prominent German reformers, as well as five secular songs, the German sample is by far the largest. Hungarian comes second with five Christian religious songs, one love song and one mixed text, Ottoman-Turkish with Hungarian insertions, which was also considered a love song by most previous researchers but which might be a *ṣufī* poem, describing the longing for knowledge rather than for another person (Sudár 2005, 75). The last text of this section is the only Croatian song in the manuscript. It has also traditionally been regarded as a love song – some authors even attributed Petrarcian features to it (Nametak 1981, 10) – but in fact it might also be a mystical religious song. The three texts mentioned last are important for the understanding of the whole manuscript because none of them could be identified in any other source and they might be works composed by the compiler. The only "European" text to be found outside these two sections is a German distich at the very end of an anthology of Turkish and Persian distiches.

1.2 The history of the manuscript

As mentioned above, the manuscript is among the holdings of the ÖNB (currently under the shelfmark Cod. A.F. 437, Fig. 1). It is not explicitly dated in a colophon or chronogram or the like, but folio 2r contains a note in the Latin language and

script: 'Miscellanea varia. Colloquium Persico Turcicum. Professio fidei Turcicae.' [Various miscellanies. Persian-Turkish phrase book. Profession of Turkish faith.], a rather understated and incomplete list of contents, but nevertheless useful, as it can clearly be attributed to Sebastian Tengnagel (Mittwoch and Mordtmann 1927, 76) who worked in the Imperial Court Library (later the Austrian National Library) from the 1590's until his death in 1636. Unfortunately, his handwriting did not change during this time (Unterkircher 1968, 133) and does not permit further dating, therefore the year of his death must serve as a *terminus ante quem*. Later cataloguers claimed that the manuscript once belonged to Tengnagel's own collection which he left to the library upon his death. Tengnagel elaborated two catalogues of this collection, one in 1612 (ÖNB Cod. 9539), the other (ÖNB Cod. 12650) at an unknown later date. No manuscript that could be identified with the one discussed here is mentioned in the first catalogue, and the second is not accessible for reasons of conservation. Thus, at present, it is not possible to determine whether Tengnagel wrote the note as owner or as custodian.

Obviously, the librarian, who was very learned and interested in Oriental languages, was not the first "European" to lay his hands on the manuscript. Someone who had no experience with Oriental books made a first foliation, starting at the end of the manuscript. The damage to this foliation on nearly every folio shows that the manuscript was trimmed, probably at the time it was rebound.

A *terminus post quem* is provided by a list of Ottoman Sultans (fols 69v–70r), which ends with Murad III, whose reign started in the year 1574 and ended in 1595. Moreover it was written on watermarked European paper. Curiously, there are eleven different watermarks.[2] Two of them are a so-called watermark pair, which means that they were produced with two different moulds in the same paper mill. All watermarks are located in the fold of the presently rather tightly bound codex, hence only a few fragments are visible, so that any identification with dated reference examples would be very tentative. Nevertheless, the watermark pair appears by shape, size and placement of the chain lines to be identical to one used in Vienna in 1587. Despite differences in detail some others also show great similarities to watermarks in manuscripts written in Austria from 1584 to 1604.

As with the other conditions of its production, there is no mention in the manuscript of its place of origin. The language samples found in the manuscript

2 Despite this high number of watermarks nothing in the manuscript suggests that it was written in more than one production process: Some quires indeed show just one watermark, and some watermarks can be found in one quire only, but, on the other hand, there are also quires with two watermarks and identical watermarks in different quires all over the manuscript.

indicate that it was most likely produced in the Kingdom of Hungary, large parts of which were occupied by Ottoman troops by the end of the 16th century.

The idea of the manuscript's having been produced by a European convert was proposed very early and is still maintained by some authors (Krstić 2011, 88f.; Römer 2014). The first cataloguer to describe the manuscript in more than a few words, Josef von Hammer, suggested this idea. Moreover he thought the compiler was an imprisoned crypto-protestant, who wrote Christian religious texts in Arabic characters in order to conceal them from his Muslim guards (Hammer 1812, 34 No. 281). The compiler's Hungarian origin was first promoted by Antal Gévay. Based on the colophon of the first Hungarian song – a Christian religious captive's lament – he believed he had discovered the compiler's identity even by name. The colophon says: '*ezer öt száz és hetven | esztendőben ‹ csonkatoronnak erős temlecében | palaszki györgy igen bízik istenben ‹ hogy nem || hadja sokáig az büntetésben*' [In the year one thousand five hundred and seventy, in the strong dungeon of Csonka tower, Palaszki György indeed trusts in God, so that He doesn't leave him in punishment for long] (fol. 60v, 13 – fol. 61r, 1). Gévay concluded that it was György Palatics (György Palaszki, according to the manuscript) who had compiled all the texts and written the manuscript. He even believed he could trace Palatics' alleged gradual estrangement from his Christian Hungarian origins and native tongue in the sequence of texts in the manuscript (Rexa 1901, 109). Later, most of his assumptions about György Palatics turned out to be wrong. The Hungarian nobleman was indeed the composer of the song – which can be found in Hungarian hymn books of that time – as well as a captive of the Ottomans for three years, but he never converted to Islam, nor did he die in captivity, as Gévay supposed, but held important positions thereafter. Although Hungarian researchers dismissed Gévay's idea about Palatics as being the compiler[3] as early as the late 19th century, the manuscript is still referred to as *Palatics-kódex* in Hungarian publications.

Later researchers had many intuitively appealing theories about the compiler's origin, religious affiliation and motives, but most of them relied more on impressions than on firm evidence. Thus we still know little about the person who produced the manuscript. The aim of the present paper is to show how he adapted the Arabic script when writing in Croatian, Hungarian, German and Latin, and, possibly, to shed some light on his identity; for example, his degree of familiarity with writing conventions in other writing systems might tell us something about his education.

3 See Sudár (2006, 9–12) for a more extensive overview of the discussion.

2 Arabic script in the European territories of the Ottoman Empire

Islam and Arabic script had been introduced into South Eastern Europe before the area was conquered by Ottoman troops and integrated into the Ottoman Empire. The process began in the late 14th century and was at its peak by the middle of the 16th century, and only then did Islam and Arabic script have any impact on the local population. In the late 16th century, Ottoman troops held fortresses located as far as today's borders between Hungary, Slovakia and Austria.

2.1 Writing the languages of the local populations in Arabic script

Although sporadic use of Arabic script is documented for a few languages in the region[4], only Albania and Bosnia, where large parts of the population converted to Islam, would develop a real manuscript culture in Arabic script in their respective local languages. In Bosnia such a manuscript culture can be traced back to the 17th century. First prints in Arabic characters appeared in the 19th century and continued into the 1940s. In the Bosnian research tradition, the phenomenon is known as *alhamijado književnost* 'aljamiado literature', a term which probably originated in an influential encyclopedia entry (Hadžijahić 1955, 144) where it was borrowed from Ibero-Romance studies. Bosnian researchers use the term *arebica* or *arabica* to refer to the Arabic script when used to write Central South Slavonic (CSS) varieties (the term is analogous to *ćirilica* 'Cyrillic' and *latinica* 'Latin script') that were (later) regarded as one language: Serbo-Croatian or Croato-Serbian (currently fragmented into the four national standards: Bosnian, Croatian, Montenegrin and Serbian). Manuscripts recording such varieties were produced in considerably smaller quantities outside of Bosnia and Hercegovina, respectively not by Bosnians or Hercegovinians. Perhaps the best known of these is one of two manuscripts containing a multilingual phrase book, probably from late 15th century Istanbul, with translations of Arabic sentences into Persian, Greek and a Serbian dialect (Aya-Sofya kütüphanesi 4750, see Lehfeldt 1989).

4 For more information on the spread of Arabic script in the Balkans see Zakhos-Papazahariou 1972 and Hegyi 1979, though e.g. the information they give on the existence of Arabic script manuscripts in Bulgarian / Macedonian could not be confirmed (Ivušić 2014, 94 n. 25).

Unfortunately most of the manuscripts in *arebica* were – if at all – studied only as texts, not as material objects in their own right; reliable editions are scarce. The only comparative study of writing conventions based on the manuscripts was published in 1969 by the German scholar Werner Lehfeldt (1969), although both earlier (Muftić 1964) and more recently (Drkić and Kalajdžija 2010), attempts were made to explore their graphematics.

Despite the fact that some Hungarian converts played important roles in Ottoman intellectual life, like the 16th century interpreter Murad or the 18th century printer İbrahim Müteferrika, there were never any mass conversions of native speakers of Hungarian. Thus Arabic script literacy and proficiency in Hungarian has never been a frequent combination. In fact the only known instances of Hungarian in Arabic script, except for the manuscript discussed here, are both attributed to Murad the interpreter (*Murad tercüman*), one of which is the only other known manuscript containing Latin in Arabic characters; the latter is an autograph of Murad's Islamic religious treatise in Ottoman Turkish and in Latin and Hungarian, all three languages being written in Arabic script with an interlinear Latin transcript (MS Marsh 179 in the Bodleian Library). The other consists of some marginal notes in the famous *Chronicon pictum* (Zsinka 1923), a 14th century illuminated chronicle. Pál Ács (2000, 312f.) assumed that these notes were also written by Murad's hand.

The only other known sample of German in Arabic script consists of a few words, idiomatic expressions and parts of a prayer in Evliya Çelebi's *Seyâhatnâme* (Kißling 1938; Römer 2009).

3 Adapting Arabic script for the European samples in the manuscript

Phonologically all four European languages (Croatian, Hungarian, German and Latin) differ substantially from Arabic. Whoever recorded them in the manuscript used the Persian and Ottoman Turkish varieties of writing, and did not have to adapt the Arabic script as used for Arabic. Thus he had additional graphemes at his disposal, and was familiar with the writing habits and conventions of Persian and Turkish. Nevertheless the writing system was inadequate for representing any of these European languages and our scribe was confronted with many problems especially in vowel transcription.

Evidently, the samples of the European languages in the manuscript are not mere transcriptions of a Latin script, or in the case of the Croatian sample, a Glagolitic or Cyrillic model. Each of them shows features absent in the non-Arabic

script written sources of that time. One such example is the use of *zayn* and *sīn* for a regular differentiation of voiced [z] and voiceless [s] in German (Ger.) and Latin (Lat.), in examples like Ger. اوُنسْ <ʾwʔuns^{0}> [ʔʊns] 'us' vs. اوُنْزَرْ <ʾwun^{0}zar^{0}> [ʔʊnzɐ] 'our' (fol. 34r, 6) or Lat. رَه زُوْرَّقْسيِتْ <rah zuwʔr^{2}aq^{0}syit^{0}> [rezu'r:e͡ksit] 'he arose' (fol. 37v, 15). This is a sub-phonemic distinction never found in the Latin script, in either manuscripts or in prints of the early modern period. Nevertheless, the scribe's attempts at representing spoken language does not mean that he based his notation solely on the phonetic principle; some morphological spellings are also found (cf. Section 3.3.2).

3.1 Consonants

Most of the consonant phonemes of these European languages are a subset of the Arabic phoneme inventory, hence their graphical representation was no challenge. The same holds for phonemes which are absent from Arabic but present in Persian and Turkish, and for which graphemes already existed. Nevertheless, each of the European languages has some consonants not found in the three Oriental languages. Table 1 shows what solutions the scribe found for their graphic representation. All the characters he used when writing consonants in the European samples are listed. Since, on the one hand, some of the encoded sounds are not phonemes, such as [ʤ] in German, and on the other hand ك <k> for instance cannot be regarded as a grapheme, not even in Ottoman Turkish, if we define *grapheme* as the smallest distinctive unit of written language, the two terms are not used in the table, although most of the character-sound correspondences displayed are also grapheme-phoneme correspondences.

Table 1: Consonants in the European samples

Character	Cro. sound	Hun. sound	Ger. sound	Lat. sound
ب <b>[5]	/b/	/b/	/b/	/b/
پ <b$_1$>	/p/	/p/	/p/	/p/
ت <t>	/t/	/t/	/t/	/t/
ج <ǧ>	-	-	[ʤ][6]	-
چ <ǧ$_1$>	/ʧ/, /ʦ/	/ʧ/, /ʦ/	/ʦ/	/ʧ/?, /ʦ/?
ح <ḥ>	/x/	/h/	/h/, /x/=[x]+[ç]	/h/
خ <ḫ>	/x/	-	/x/=[x]+[ç], /h/	/x/
د <d>	/d/	/d/	/d/	/d/
ر <r>	/r/	/r/	/r/	/r/
ز <z>	/z/	/z/	/z/	[z]
ژ <z$_2$>	/ʒ/	/ʒ/	-	[ʒ]
س <s>	/s/	/s/	/s/	/s/
ش <š>	/ʃ/	/ʃ/	/ʃ/	[ʃ]
ص <ṣ>	-	/s/	-	-
ط <ṭ>	/t/, /d/	/t/	/t/, /d/	-
غ <ġ>	/g/	/g/	/g/	/g/
ف <f>	/f/	/f/	/f/	/f/
ق <q>	/k/	/k/	/k/	/k/
ك <k>	/ʨ/, /ʤ/, /ɲ/	/c/, /ɟ/, /ɲ/	-	[ʤ]?, [ɟ]?
ڭ <k$_3$>	-	/ɟ/, /ɲ/	-	-
ل <l>	/l/	/l/	/l/	/l/
م <m>	/m/	/m/	/m/	/m/
ن <n>	/n/	/n/	/n/	/n/
ه <h>	([h])?	/h/	/h/, /x/ [ç]	-
ي <y>	/j/, [+pal]	/j/, [+pal]	/j/	/j/
و <w>	/v/	/v/	(/v/)	/v/

Even a cursory glance at the table reveals some important aspects of the scribe's adaption strategy: 1. Although the Arabic script is especially suited for adding characters through alternations of the number and position of dots – a method frequently employed when adapting it for other languages (Daniels 2014) – the scribe did not create any additional characters. 2. Arabic characters encoding sounds that are not present in European languages or which are not regularly used as a substitute, such as غ <ġ> for /g/, were not used. The scribe omitted the

5 The transliteration follows the proposal of Mumin and Versteegh 2014, 11–21.

6 The grapheme ج <ǧ> in the German sample appears in wordforms where the sound [ʤ] was realised due to regressive assimilation. There is no evidence for [ʤ] being phonemic in the sample, hence it is put in phonetic brackets in the table.

graphemes ث <t̲>, ذ <d̲>, ض <ḍ>, ظ <ḏ>, ع <ᶜ> for /θ/, /ð/, /dˤ/, /ðˤ/, /ʕ/ rather than linking them arbitrarily to sounds of the European samples regardless of phonetic proximity. Instead he chose the graphic representation of the Ottoman Turkish sound, which he perceived as phonetically closest even if that created new ambiguities. A closer look at some cases in which there is no obvious one-to-one character-sound correspondence will clarify further aspects of the adaption strategy.

3.1.1 ڒ <ğ$_1$>

The use of ڒ <ğ$_1$> for the affricate /ʦ/ is a good example of how a sound in the European languages is represented by the character encoding its closest match in Ottoman Turkish. The consonant ڒ <ğ$_1$> is not an original character of the Arabic script, it is a Persian innovation adopted in Ottoman Turkish writing. Both Persian and Turkish have a voiceless post-alveolar affricate /ʧ/, but lack an alveolar affricate /ʦ/. In Central South Slavonic, Hungarian and German both sounds are phonemic, while they are absent from Classical Latin. The Latin found in the sample contained in the manuscript is of course not Classical Latin, neither is it scholarly Neo-Latin, although the latter had some influence on the variety that clearly shows features of Medieval Latin. Therefore, it is unlikely that the ڒ <ğ$_1$> in words like تَرْڿِيَا <tar^{0}ğ$_1$iyʾa>, cf. Class. Lat. *tertia* 'third', represents a dental or alveolar stop /t/, or a velar stop /k/ in a word like اَنْڿِيلّاَمْ <ʾªn^{0}ğ$_1$yil^{2}ʾam^0>, cf. Class. Lat. *ancillam* 'maidservant (acc.)'. It is more probable that ڒ <ğ$_1$> represents /ʦ/ in تَرْڿِيَا <tar^{0}ğ$_1$iyʾa>, and /ʧ/ in in اَنْڿِيلّاَمْ <ʾªn^{0}ğ$_1$yil^{2}ʾam^0>. When ڒ <ğ$_1$> is written where Class. Lat. *ti* would appear, there is no reason to assume any other sound but /ʦ/, as the assibilation from /ti/_V to /ʦ/ was a sound change already documented in 2nd or 3rd century Latin; it was well established in Medieval Latin, and was believed to have been classical even by humanists (Stotz 1996, 219). The interpretation of those cases where ڒ <ğ$_1$> encodes a sound corresponding to Class. Lat. /k/_/i,e/ is much harder to justify. What can safely be ruled out is that ڒ <ğ$_1$> represents the velar /k/, as all attestations of /k/ followed by a back vowel or a consonant are encoded by ق <q> (see Section 3.1.2). In Romance languages, except Sardinian and Dalmatian, /k/ before front vowels was palatalised to /ʦ/ or /ʧ/ (a change which often continued to develop), and in Medieval Latin, it was also pronounced either as an affricate or as a sibilant (Stotz 1996, 183–185; Bonioli 1962, 73–78). The velar pronunciation was retained in Ireland, and Irish influenced parts of northern England well into the 12th century. As the character ڒ <ğ$_1$> is used to represent both sounds, /ʦ/ and /ʧ/, in the other language samples, and both pronunciations existed in Medieval Latin as well, we cannot decide which one to assume for the Latin variety.

3.1.2 ق <q>, ك <k> in Croatian, Hungarian, German and Latin

The employment of the graphemes *qāf* ق <q> and *kāf* ك <k> shows that, occasionally, a shared prominent distinctive feature like palatal articulation was enough to establish phonetic proximity for the scribe. Moreover it shows that he aimed at a unified transcription system for all four European languages. Like Ottoman Turkish, all of them have a voiceless velar stop /k/. In Ottoman Turkish both ق <q> and ك <k> are used to write /k/. As with the graphemes for the other pairs of emphatic and non-emphatic sounds in the Arabic source orthography, ق <q> and ك <k> serve to encode vowels in Ottoman Turkish: ق <q> is used in inherited words with back vowels as in *kurak* 'dry' written as قوراق <qūrāq>, and ك <k> for those with front vowels as in *kürek* 'oar' written as كورك <kūrk> (Weil 1917, 24; Buğday 1999, 13–15). As Ottoman Turkish is a language with vowel harmony based on the contrast front/back (front/not-front), such a system of vowel notation is quite effective. In the European samples, /k/ is represented by *qāf* only, both with back and front vowels, as can be seen in the examples in Table 2.

Table 2: ق <q> in the European samples

	Croatian	**Hungarian**	**German**	**Latin**
/k/ before back vowels	قوُ <qwu> *ko* 'who'	پُوقُولْرَاَ <$b_1$uwquwl^{0}r^a'a> *pokolra* 'to hell'	قوُمْ <qwum^0> *komm* 'come'	قوُنْترآ <qwun^{0}tr'^{ā}> *contra* 'against'
/k/ before and after front vowels	وَه قَه <wah qah> *veke* 'age'	قِيِوُولُمْ <qiywuwlam^0> *kivülem* 'besides me'	قِينْدَرْ <qyin^{0}dar^0> *kinder* 'children'	نَقْ <naq^0> *nec* 'neither'

Kāf, on the other hand, is used for different palatal sounds in Croatian and Hungarian, and very likely also in Latin. In the Croatian sample, ك <k> represents the voiceless palatal affricate /tɕ/ (<ć> in modern orthography) as in وَه كَه <wah kah> *veće* 'bigger', but also the voiced /dʑ/ (<đ>) as in سيكَه <sykah> *siđe* 'he descended' and the nasal /ɲ/ (<nj>) as in كَغُووُو <kaġuwwuw> *njegovo* 'his'.

In Ottoman Turkish, *kāf* also had various functions. It was employed for the stops /k/ and /g/, for the approximant /ɰ/, and for the velar nasal /ŋ/. In the Western Rumelian dialects (a group of present-day Turkish dialects in the Balkans), /k/ and /g/ before front vowels are not only articulated in pre-velar position, as in Standard Turkish (Stand. Trk.), but have been palatalised to palato-velars, palatal affricates or palatal stops (Friedman 2002, 613). We cannot say when this process took place, but already in early Turkish loanwords in

Serbo-Croatian, or loanwords transmitted via Turkish, /k/ and /g/ before front vowels are frequently, although not regularly (Németh 1970, 91), substituted by /ʧ/ and /ʤ/ respectively. Famous examples are *ćùprija* 'bridge', cf. Stand. Trk. *köprü* (Škaljić 1966, 200) or *đồn* 'sole of a shoe', cf. Stand. Trk. *gön* 'worked leather' (Škaljić 1966, 252). In Bosnian *alhamijado* manuscripts, *kāf* is regularly used for /ʧ/ and frequently for /ʤ/ as well (Lehfeldt 1968, 141–154). When Ottoman clerks wrote down Serbo-Croatian names in registers and other administrative documents, they rendered /ʧ/ by *kāf* (Aličić 2000, xxiii). In such registers *kāf* can occasionally represent the palatal nasal /ɲ/ and even an alveolar nasal /n/.

In the case of /ʧ/, there might also be another reason for using *kāf*. All writing systems employed to write CSS varieties in early modern times had difficulties representing this sound. In Latin script manuscripts and prints of the 16th century, various solutions, usually based on Italian orthography, can be observed. Typically <ch> is used, but <c> and <cch> occur as well, and none of them is satisfying as they often simultaneously represent the affricates /ʧ/ and /ʦ/. The two Slavonic scripts Glagolitic and Cyrillic were created to put Old Church Slavonic (OCS) into writing. Despite being closely related to CSS varieties, OCS has a rather different phoneme inventory. Since /ʧ/ is not part of it, a grapheme encoding the sound does not exist in either of the Slavonic scripts. In both, the sound /ʧ/ is usually represented by a character which was used for the OCS sound sequence /ʃt/, since OCS /ʃt/ and CSS /ʧ/ are corresponding reflexes of the same Proto-Slavonic sound. But as Trunte (2012, 186) points out, in Cyrillic manuscripts the grapheme к <k> for the velar stop /k/ is frequently used for /ʧ/. Moreover he found such a spelling, if only in one instance, in a Glagolitic print, a historiography published in 1531 in Rijeka (Ital. Fiume). It cannot be excluded that whoever wrote our manuscript, knew about these spellings in the Slavonic scripts.

Encoding the nasal /ɲ/ by *kāf* is almost unparalleled in Arabic script CSS. With one exception, and in only one instance, Bosnian *alhamijado* manuscripts do not have this spelling (Lehfeldt 1969, 156). In the Serbian sample of the mentioned multilingual phrase-book, /ɲ/ is represented by a Digraph نيـ <ny>, sometimes adding *hamza* above *yā'* (Lehfeldt 1989, 49). The rendering might be motivated by the Ottoman Turkish use of *kāf* for a nasal, although in Ottoman Turkish it is a velar nasal, and not palatal. But as we have seen so far, in the scribe's reanalysis of the character it was disassociated from the feature 'velar' and linked to the feature 'palatal'. Initially, this reanalysis was probably triggered by the phonetics of the sounds which corresponded to /k/ and /g/ before front vowels in Standard Turkish. But then the scribe took a further step: the nasal encoded by *kāf* in the Ottoman Turkish variety which served him as a source language for the script might not have been velar anymore but was certainly not palatal either, so when the scribe chose to encode /ɲ/ by *kāf* he picked the most logical

representation in the systematics of the writing system he devised, regardless of a phonetic justification, and thereby he emancipated the graphic representation from phonology.

3.1.2.1 Three-dotted *kāf* ڴ <k$_3$> and *kāf* ك <k>

In Ottoman manuscripts the alternative for encoding the velar nasal is the three-dotted *kāf* ڴ <k$_3$>. This character is also used in the manuscript for the palatal nasal, but only in the Hungarian sample, not in the Croatian one. In Hungarian, the three-dotted *kāf* encodes /ɲ/ as in آرڴق <ʾaʳ⁰k$_3$aq> [‘a:rɲe:k] *árnyék* ‘shadow’, but also the voiced palatal stop or affricate[7] /ɟ/ as in نَلَڴْ <nalak$_3$⁰> [‘nɛle:ɟ] *ne légy* ‘don’t be’. Both Hungarian sounds can be represented by *kāf* without dots as well, as in كيتِوُمْ <kyitwum⁰> [‘ɲitom] *nyitom* ‘I open’ and نَاكْ <nʾak⁰> [nɒɟ] *nagy* ‘big’. *Kāf* without dots is also used for the voiceless palatal stop or affricate /c/ in آكا <ʾaʾa> [‘ɒcɒ] *atya* ‘father’. But the spelling of Hungarian palatals is even more variable, because all of them can be represented by digraphs with *yā’*. Table 3 summarises the various graphic representations of Hungarian palatals, listing them in order of frequency.

Table 3: Graphic representations of Hungarian palatals

/c/	تي <ty>, ك <k>
/ɟ/	ك <k>, ڴ <k$_3$>, كي <ky>, ڴي <k$_3$y>, دي <dy>
/ɲ/	ك <k>, ڴ <k$_3$>, ني <ny>

There are two possible explanations for these spelling variations and, perhaps, both have played a part. Firstly, they might have been motivated by phonetic features of the Hungarian sounds. Unlike the Hungarian /c/ and /ɟ/, Central South Slavonic /tɕ/ and /dʑ/ are always realised as affricates, not as palatal stops. Palatal stops cannot be distinguished from palatalised velars regarding their distinctive features (Hall 1997, 70–76). If Hungarian /c/ and /ɟ/ are identical to /kʲ/ and /gʲ/, it is not surprising that the scribe used *yā’* as a marker of palatalisation to represent

7 There is an ongoing debate in Hungarian phonological theory as to whether /ɟ/ and /c/ should be classified as stops or affricates at the phonological level (Szende 1992, 119). Although it is not disputed that they are realised as affricates in certain contexts – and as stops in others – they pattern with stops in two important respects: First they can be realised by their unreleased allophones before stops, and second when they appear on both sides of a word boundary, they are merged into a geminate (Siptár and Törkenczy 2000, 83).

them. Central South Slavonic (CSS) /tɕ/ and /dʑ/, on the other hand, differ from /k^j/ and /g^j/ as they are [+strident]. The sounds /tɕ/, /dʑ/, /c/ and /ɟ/ are palatal but only the latter two show realisations which are not distinguishable from palatalised sounds. The presence of the digraph spellings with *yāʼ* for /c/ and /ɟ/, and the absence of that spelling for /tɕ/ and /dʑ/ might reflect the phonetic distance between these two sets of sounds. Secondly, the writing conventions of the Hungarian Latin script might have been a model for the digraph writings. As can be seen in the Hungarian examples quoted, digraphs are used in modern Hungarian orthography to represent each of the three sounds: <ty> encodes /c/, <gy> /ɟ/ and <ny> /ɲ/; these conventions were already well established in the late 16th century (Kniezsa 1959, 17f.). What remains unclear is why the three-dotted *kāf* does not appear in Central South Slavonic (CSS) at least as an allograph for /ɲ/. Today /ɲ/ is basically the same sound in all known varieties of Hungarian and Central South Slavonic but this may not have been the case when the manuscript was written.

The Latin sample shows only one instance of *kāf*, in وِيرْكِنَه <viyr^{0}kinah> *virgine* 'virgin (abl.)'. This is also the only attestation of what would be /g/ followed by a front vowel in Classical Latin. So instead of assuming that *kāf* represents a velar /g/, rendered by *ġayn* in all other attestations of the sound in the European samples, it seems much more likely that *kāf* represents a palatal, close to /ɟ/ or /dʑ/. That the sound in question might have been /dʒ/, the pronunciation of Latin /g/ followed by *i* or *e* in Italy (Bonioli 1962, 81f.), also seems unlikely. In that case the scribe would have used ج <ğ>, as he did when rendering the sound in Arabic, Turkish, Persian and German, where it is not phonemic but a product of assimilation.

German lacks this kind of palatal phoneme. Palato-velars, palatal affricates, stops and nasals, the domains of *kāf* in the other European samples, are all absent from the German phoneme inventory, consequently *kāf* does not occur in the German sample at all. We may assume that the scribe did not want to spoil the disambiguation he had devised for the other languages.

3.1.3 ح <ḥ>, خ <ḫ> and ه <h>

One case seems to contradict the idea that the scribe omitted Arabic graphemes representing sounds that are not part of the phoneme inventory of the European languages or their closest match: this is the use of Arabic characters for velar, pharyngeal and glottal fricatives. Although none of the European languages has a pharyngeal fricative /ħ/, all of the European language samples show the use of ح <ḥ>, the character which represents the sound in Arabic orthography. The explanation of this seeming exception lies in the use of such characters in Ottoman

Turkish, which served as a transmitter for the Arabic writing system. In Ottoman Turkish only the glottal fricative /h/ was phonemic, and yet the Arabic fricative graphemes which encode the velar and pharyngeal fricatives were used in Ottoman writing conventions. Their use was almost exclusively limited to words of Arabic origin but they are occasionally found in Persian and Turkish words. In most instances, however, the pronunciation was simply [h].

3.1.3.1 Croatian

In the Croatian sample ح <ḥ> is used alongside خ <ḫ> to represent the velar fricative /x/, but there seems to be a positional distribution: ح <ḥ> represents the sound in syllable onset only, as in حِيزُوْ <ḥyiz$_2$uwʔ> *hižu* 'house (acc.)', خ <ḫ> on the other hand is used in syllable coda (with one exception), as in ژِيوْيخ <z$_2$yiw^{0}yiḫ > *živih* 'the living (gen.pl.)'. The only occurrence of ه <h> as a consonant spelling in the Croatian sample is in the interjection آهْ <ʾāh^0> *ah* 'oh', and here it is impossible to tell whether the encoded sound was velar or glottal. Curiously, in Bosnian *aljamiado* manuscripts, خ <ḫ>, the original Arabic grapheme encoding /x/, is exceedingly rare and does not appear in the observed distribution (Lehfeldt 1969, 166–172). In the Serbian parts of the phrase book, ه <h> is written for /x/ in all positions, ح <ḥ> and خ <ḫ> appear only in one instance each, both in Oriental names: خُوَرَزمِسْقي <ḫuwarazmis^{0}qy> (Aya-Sofya kütüphanesi 4750, fol. 11r) 'Chorasmian' and حَسَنْ <ḥasan^0> (Aya-Sofya kütüphanesi 4750, fol. 14v) 'Hasan'. These examples of CSS in Arabic script seem to show that the transmission of the script via Turkish, as suggested above, obscured the function of خ <ḫ> in Arabic, so that the character could not simply be adopted for writing CSS /x/. Apparently the character ح <ḥ> became the default spelling for post-palatal fricatives in CSS written in Arabic script.

3.1.3.2 Hungarian and Latin

Hungarian has only one post-palatal fricative, traditionally assumed to be /h/, but in syllable final position the realisation is either [x] or zero.[8] However, in syllable initial position, the Hungarian sample shows 132 instances in which ح <ḥ> is used, as in حَازَاتْ <ḥaʾzaʾt> *házát* 'his house (acc.)', and there are four attestations

8 Siptár (1994, 213 and 265–268) still assumes /h/ to be the basic underlying phoneme of which the other realisations are derived. Siptár and Törkenczy (2000, 274–277) later stated that, at least in today's pronunciation of educated non-language-conservative speakers from the Budapest area (Educated Colloquial Hungarian), /x/ is the basic phoneme while [h] and zero have to be regarded as derived allophones.

of ه <h> (cf. Section 3.3.1). In the Latin sample the glottal fricative is also encoded by ح <ḥ>; the only use of خ <ḫ> appears in مَخَابَرِيسْ <maḫʾabaryis⁰> *mechaberis* 'you will commit adultery' and it reflects a velar fricative in the Greek loanword.

3.1.3.3 German

In German the situation is more complicated. The consonant phonemes of German include a glottal fricative and a velar fricative with two complementary variants: velar fricative [x] after back vowels and palatal fricative [ç] after front vowels and consonants. As can be seen in Table 1, ح <ḥ> is used for all three sounds; nevertheless, there is a definite preference. Of the 283 attestations of ح <ḥ> in the German sample only four represent the palatal fricative [ç], 22 represent the velar fricative [x] but 19 of the latter are instances of the same lexeme. Seven are special cases, because they represent various spelling pronunciations (cf. Section 3.3.3). The remaining 250 instances of ح <ḥ> indicate the glottal fricative. But the same sound is frequently encoded by ه <h>. However, the use of ه <h> seems to depend on position: 81 of 89 attestations of ه <h> (when used as a consonant character) in the German sample are syllable initial. Moreover, 53 of these spellings occur in words with the sound sequence *her*, most of them being forms or derivations of German *Herr* 'lord'. This spelling might have been influenced by the spelling of the frequent Ottoman Turkish هَر *her* 'every'. The other instances of the character ه <h> twice represent a syllable final fricative in the interjections *oh* and *ah*, which might be glottal as well as velar; in a further two examples, ه <h> reflects spelling pronunciations (cf. Section 3.3.3) and there are four instances of the palatal fricative, but the latter is always found in derivations of *mächtig* 'mighty'.

The default choice for representing both the velar and the palatal fricative was خ <ḫ>. Only eight out of 377 attestations render the glottal fricative and all eight are either instances of the word خُوخ <ḫuwḫ⁰> *hoch* 'high, up' or compounds with *hoch*, where the representation of the first fricative was probably influenced by the last.

3.2 Vowels

All European languages in the manuscript have more vowel phonemes than Arabic. The German and Hungarian vocalic systems are especially rich and, like Ottoman Turkish, they include rounded front vowels. Vowel quantity is phonemic in all of these European languages, and accent is also phonemic in both Central South Slavonic (pitch accent) and in German (stress accent), but neither quantity nor accent are encoded in the European samples. In each of them, the graphemes

representing long vowels in Arabic, *ʾalif* ا <ʾ>, *yāʾ* ي <y>, and *wāw* و <w> as well as the diacritics representing short vowels, *fatḥa* ◌َ <a>, *kasra* ◌ِ <i>, *ḍamma* ◌ُ <u>, are used regardless of quantity or of any feature other than vowel quality; this mirrors Ottoman Turkish writing habits. The same holds for the different representations of a vowel sound depending on its position in the word. It is impossible to display all the manifold vowel representations in all the languages (some of the polygraphic combinations occur in one instance only), but Table 4 (cf. Appendix II) shows the most frequent representations for each of them.

Some of the assumed vowel phonemes in Table 4 are questionable. For instance the diphthong /ɪɛ̯/ is the most plausible interpretation of the graphic sequence يَ◌ <iya> in Croatian تِيَلْوُ <tiyalwu> *tijelo* 'body', but this would be the only attestation in which Proto Slavonic **ě* (the sound known as *jat'*) is reflected as /ɪɛ̯/.

3.2.1 Implementing a marker of distinction

There is no simple one-to-one correspondence between graphemes and vowel phonemes. Nevertheless /i/ (/iː/, /ʏ/) can be distinguished from any other vowel. In most cases it is also possible to distinguish between renderings of /a/ (/aː/, /ɒ/) and those of /e/ (/eː/, /ɛ/, /ɛː/, /ə/). However, the graphic representation of rounded vowels is ambiguous in all the European languages. In the first commandment, which is also the first instance of the European languages in the manuscript, the Croatian, German and Hungarian texts contain three different rounded vowels, as shown in (1).

(1) Cro.	yʾa	sʾam^0	twuy^0	buwġ^0	dʾātyi		
	Ja	*sam*	*tv**o**j*	*b**o**g*	*da=ti*		
	I	am	your	god	COMP=2.SG.DAT		
	nabwudah			wah kah |	buwz$_2$yi		ʾyiʔ
	*ne=b**u**d-e*			*već-e*	|	*b**o**ž-i*	*i*
	NEG=be.PRS.PFV-3SG			greater-ACC.PL		god-?	and
	wyišah			ʾwud^0	manah		
	više (*viš-e*)			***o**d*	*mene*		
	more (greater-ACC.PL)			from	me		

'I am thy God, do not have greater and more/higher gods than me'.

Hun.	ʾan^{0} | wʾ^{ā}kwuq^{0}	ʾ^{ā}z^{0}	tah	ʾwurʾad^{0}	ʾyiš^{0}tanad^{0}
	*Én | vagy-**o**k*	*az*	*te*	***u**r-ad*	*isten-ed*
	I be.PRS -1SG	DET	you	lord-POSS.SG.-2SG	god-POSS.SG.2SG
	nalakanaq^{0} |	ʾan^{0}	qyiwuwlam^{0}	ʾiydah ġan^{0}	ʾyiš^{0}tanyid^{0}
	ne=legyenek |	*én*	*kív**ü**l-em*	*idegen*	*isten-id*
	NEG=be.IMP.3PL	I	except- POSS.SG.1SG	foreign god-	POSS.PL.2SG

'I am thy Lord, thy God, thou shalt not have foreign gods except me'.

Ger.	dwu suwl^{0}t	nyit^{0} ʾ^{ā}n^{0}dah rah	ġat^{2}ar^{0}	naban^{0}	| my$_{I}$^{0}r	
	*D**u** s**o**l-t*	*nit ander-e*	*gett-er*	*neben*	| *mir*	
	you shall-2SG.PRS	NEG other-ACC.PL	god/PL-ACC.PL	beside	me	
	ḥʾaban					
	haben					
	have.INF					

'Thou shalt not have other gods beside me'.

All rounded vowels are represented by the combination of *wāw* و <w> and *ḍamma* ُ <u> (in word initial position preceded by *ʾalif* ا <ʾ>). But already in the second commandment, the first attempts at disambiguation can be observed: a new combination involving *hamza* ء <ʔ> is introduced. The distinctive function of *hamza* is established most consistently in Latin: all instances of back mid /ɔ/ and /oː/ show *wāw+ḍamma* وُ <wu>, while *wāw+hamza+ḍamma* وُ <wʔu> renders back high (close) /ʊ/ and /uː/ only. In the second commandment in Croatian وُ <wʔu> occurs once representing a word final /u/ while the /u/ in mid-position in the same word is again written without *hamza*. In the Hungarian and German texts of the second commandment وُ <wu> remains the graphic representation of both back vowels, but *hamza* is used twice in the German text, representing /yː/ and /ʏ/. The fourth commandment shows the first use of *hamza* in Hungarian; as in German it was, originally, not used to differentiate the back vowels but to mark the rounded front vowels /y/ and /yː/ (or /øː/)[9]. Since the German sample is mainly Central Bavarian (Bichlmeier and Ivušić 2013), and because one feature of Central Bavarian is the delabialisation of Middle High German rounded front vowels, these vowels are present in a few instances only and *hamza* develops the same function of marking back high (close) vowels in German as in Latin and Croatian. The further the scribe

9 The spelling اوُدْنَبْرِوُلْ <ʾwʔud^{0}nab$_{I}$^{0}rwʔul^{0}> can be interpreted as *üdnepről* or as *üdneprűl* 'of the holiday'. The allomorph of the Modern Standard Hungarian delative ending, *-ről* is *-rűl* in various dialects; in late 16th century written Hungarian, both variants were used (Papp 1961, 44f. and 168).

proceeded in writing the manuscript, the more regular the use of *hamza* became. It spread to words which were written only with وُ in earlier attestations in the manuscript, and then to all positions in the word. In one of the songs following the penta-lingual section (fol. 42v, 10, Fig. 1) *hamza* has been added to represent the /u:/ in the German word نوُْنْ <nwʔun⁰> *nun* 'now'.[10] This gradual increase of consistency in the use of *hamza* can be observed in the Hungarian sample as well, but here it never becomes associated with one vowel only, rather it is used for both /u/ and /u:/ and for the rounded front vowels /ø/, /ø:/, /y/ and /y:/ as well as for /o/ and /o:/. Interestingly, no development in disambiguating back mid and back high vowels takes place in the Latin sample. Thus we find two distinct spellings in the first instances of the respective phonemes. As already mentioned, the manuscript is assumed to be a fair copy. It is possible that the first version of the Latin sample was written somewhat later than the samples of the other European languages, when the *hamza* spelling for *u* was already fixed.

3.3 Peculiarities of the graphic representation

3.3.1 Use of characters for Arabic emphatic consonants

As mentioned, no systematic use of the characters for Arabic emphatic consonants in vowel notation, as it existed in Ottoman Turkish, can be observed in any of the European samples. Nevertheless, a few traces of that use are present in each of the samples except for Latin. Croatian, Hungarian and German show some (Cro. 4, Hun. 3, Ger. 8) instances of the character ط <ṭ> representing /t/ and /d/ (only in Croatian and Hungarian) in words with back vowels. If the words are not *hapax legomena*, as three of the Croatian word forms are, they are attested with <t> and <d> spellings as well. In the case of Hungarian, adopting the Ottoman writing habit would have been possible without much effort since Hungarian has a type of vowel harmony close to Ottoman Turkish (front/back), and indeed the Hungarian sample shows more instances of the emphatic consonants compared to the other European samples. Thus, (a) it is the only European sample using ص

10 This spelling can be identified as a correction because the *hamza* is added in red ink. Red ink is used in the manuscript for certain elements as verse separators, ciphers (if part of an enumeration) and headings. In very rare cases diacritic vowels, dots and other script elements are written in red. Sometimes this was simply done for ornamental reasons, in which case, all diacritics in a verse would be red; more often, one of them was simply forgotten and added later, when the scribe had already switched to red ink, as is the case in the example cited.

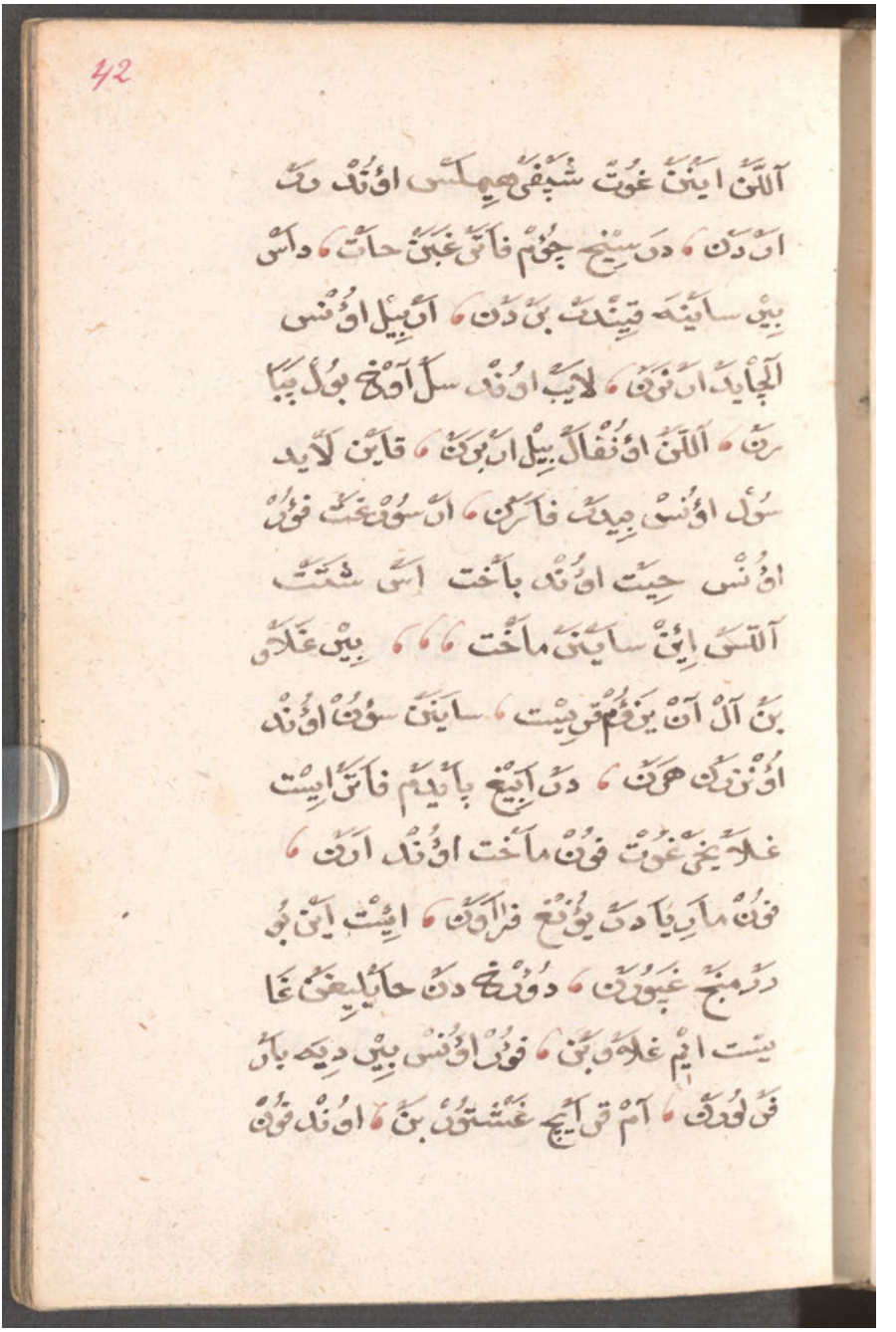


Fig. 1: Cod. A.F. 437, fol 42v. © Österreichische Nationalbibliothek Wien.

<ṣ> in three attestations, all in words with back vowels; (b) there are four word-forms which deviate from the usual encoding of fricative /h/ with ح <ḥ>, using ه <h> instead. All four words contain front vowels, three are instances of the palatal allomorph of the marker of situational possibility *-het* (denoting a meaning close to 'can'), which also appears with the usual ح <ḥ> spelling, but these three attestations with ه <h> are part of an enumeration, in which they stand in contrast with word-forms containing the velar allomorph of the marker *-hat* written with ح <ḥ>, as illustrated below (inverted commas are verse separators).

(2)

šah	ʾal^{0}hatam^{0} šah	ḥʾal^{0}ḥaʾtuwm^{0}ʿ	šah ʾuwʔl^{0}|hatam^{0}	šah
se	él-het-em se	hal-hat-om ʿ	se ül|-het-em	se
neither	live-can-1SG neither	die-can-1SG	neither sit-can-1SG	neither
yʾar^{0}ḥʾatwum^{0}ʿ	šah	ʾahatam^{0}	šah ʾyiḫʾatuwm^{0}ʿ |	
jár-hat-om ʿ	*se*	*ehet-em*	*se ihat-om*[11] ʿ	
walk-can-1SG	neither	eat.can-1SG	neither drink.can.-1SG	

'Neither can I live nor can I die, neither can I sit nor can I walk, neither can I eat nor can I drink'.

And finally, there is one exception to the general rule that the velar /k/ is represented by ق <q> in the European samples (cf. Section 3.1.2): In the mixed Hungarian/Turkish text, the final consonant of the Hungarian lexeme *lélek* 'soul' is rendered by ك <k> in all nine attestations. In other texts of the Hungarian sample, the same word is written with the usual ق <q>. It might be that in this mostly Turkish text Turkish writing habits spread to the Hungarian parts as well. However, it is not clear why only instances of one lexeme are affected, since the three other words in the text which have /k/ in the front vowel environment (*kegyelmed* 'Your Grace', *kegyelmes* 'gracious' and *szeretőmnek* 'to my beloved') show the ق <q> representation.

3.3.2 Morphological spellings

As has been argued so far, the notation of the European samples in the manuscript aims at representing spoken language and is not a transcript of a model in another

11 The initial /i/ of the verb *inni* 'to drink' is palatal in Modern Hungarian but it developed from a velar vowel /ɯ/ which merged with palatal /i/ in the first centuries of the Old Hungarian period, 896–1526, (Abaffy 2003, 320f.). Reflexes of the original velar quality can still be seen in Modern Hungarian, as the verb demands the velar allomorphs of endings and suffixes.

writing system. Nevertheless, the phonetic principle was not the only one employed in the writing of the European languages.

In Hungarian, the palatal stops or affricates /ɟ/ and /c/ (already discussed) are often created in a morphophonemic process when roots and stems ending in /d/ and /t/ merge with suffixes and endings with initial /j/. In the manuscript the resulting word-forms are usually written in a way that represents their morphological structure at the expense of their phonetics, as can be seen in examples like ويغاديۇنق <wyġˀdywʔnq> [‘vigɒ/aɟunk] *vigadjunk* ‘we shall celebrate’ (*vigad* + 1.pl. prs.imp.indf.).

In the German sample, final devoicing, a very prominent phonological process in most dialects since the Middle High German period and in Modern Standard German, is not represented graphically. The devoicing affects stops and fricatives in syllable final position, thereby obscuring paradigmatic relations, as in *Tag* [ta:k] ‘day’ vs. *Tage* [‘ta:gə] ‘days’. As in Modern German orthography, the graphic representation in the manuscript does not reflect devoicing. In the case of the manuscript, one might rather speak of *hardening*, since the consonants in question do not differ in voice but stand in a fortis-lenis opposition in Bavarian, but the same character is employed for both fortis and lenis as in تَاَغْ <tˀaġ⁰> *Tag* and in تَاَغَه <tˀaġah> *Tage*.

Both instances of morphological spellings were writing habits which, in the late 16th century, were developing into conventions in the Latin script orthographies of the respective languages; but since both reflect the analysis of rather transparent morphological processes, there is no need to assume an influence of Latin script writing.

3.3.3 Latin script convention or spelling pronunciation?

There is one further phenomenon in the German sample that, at first sight, is an adoption of a Latin script spelling convention. Several attestations in the German sample contain something that is termed *stummes h* ‘silent h’. The term refers to the character <h> when it is not pronounced as a fricative. In some words the spelling reflects a historical fricative pronunciation but in most cases *silent h* is a purely orthographical element, marking either vowel length or the syllable boundary. However, the presence of such spellings in word-forms like غَحَتْ <ġaḥat> *gehet* ‘goes’ in the manuscript is not necessarily an imitation of Latin script writing conventions. The realisation of the *silent h* as a fricative is a typical spelling pronunciation in German and as such often encountered in the speech of religious services and in many church songs. The distribution of the *silent h*

spellings in the manuscript suggests that, in fact, they might well be renderings of the spelling pronunciation as they occur in religious texts only.

3.3.4 Gemination and double characters

As in Arabic, but not in Ottoman Turkish, gemination is phonemic in Hungarian and Latin. In both languages it is indicated by doubling the consonant grapheme in Latin script writing. In Hungarian complex consonant graphemes, such as <sz> /s/ or <cs> /ʧ/, only the first component is written twice, hence /s:/ is represented by <ssz>. In Latin-script German, the doubling of consonant graphemes also exists, but, in contrast to Latin and Hungarian, it does not encode the quantity of consonants, which is not phonemic in German, but indicates that the vowel preceding the doubled consonant grapheme is short. In all three language samples, the Arabic diacritic *šadda* ّ <²> indicating gemination in Arabic is used, but all of them also show doubling of the consonant grapheme, sometimes even a combination of both is found, i.e., superscription of one of the two identical consonant graphemes with *šadda*.

Unsurprisingly, the last mentioned spelling is most frequently encountered with /l/ and /l:/, as for these consonants a prominent model, الله *allah*, existed. Writing the consonant grapheme twice occurs in Latin with /r:/, in Hungarian with /b:/, /l:/, /t:/ and /z:/, and in German, the following characters are doubled: لا <ll>, تت <tt> and سس <ss>. At morpheme boundaries, where the final consonant of the preceding morpheme and the initial consonant of the following morpheme are the same, the doubling of consonant graphemes seems natural and can be encountered in Ottoman Turkish as well. One might argue that the double spellings in Hungarian and Latin are due to a misconception of the morpheme boundaries by the scribe (although nothing in the Hungarian and Latin samples supports such an assumption), but in the German sample, where neither *šadda* nor the double graphemes can be phonetically motivated, the influence of Latin script writing habits seems to be the only explanation.

4 Conclusion

The inclusion of texts in four languages, not usually written in Arabic characters, into an Ottoman multiple-text manuscript, was an ambitious endeavour. The scribe had to adapt a single writing system for languages which differ substantially from each other and from those whose Arabic script writing habits

and conventions he knew. As shown, his adaption remained rather conservative, he did not devise any new characters nor did he use those which were part of the Ottoman Turkish inventory independently of their use in that language. Nevertheless, he developed original solutions for the graphic representation of phonemes which are not part of the phoneme inventories of Ottoman Turkish, Persian or Arabic; these solutions can be seen in his use of *kāf* ك <k> and *hamza* ء <ʔ>. His system was able to represent most of the consonant phonemes of the European languages, but the representation of vowels, especially rounded vowels, was very ambiguous and can be interpreted only by readers well familiar with the respective languages. The vowel phoneme inventory of Ottoman Turkish is also underspecified by vowel graphemes of Arabic script, but some of the shortcomings could be resolved by the Ottoman system of indicating the backness of the vowel by the graphemes for the uvular or pharingealised consonants in the environment of the vowel. In the European samples there is no systematic use of the Ottoman Turkish writing conventions, although some traces can be found.

This paper claims that the graphic representation of the European samples renders spoken rather than written language; nevertheless, these representations are not based only on phonetics. A few instances of morphological spellings which were also features of the respective Latin script orthographies can be observed. Some peculiarities of spelling could be influenced by writing conventions in other writing systems, but the only phenomenon where Latin script orthography is the most probable explanation is the use of *šadda* ّ <ʲ> and double graphemes in Hungarian, Latin and German. The fact that the scribe must have been familiar with German Latin script writing, at least to some extent, can be clearly seen in the German sample where *šadda* ّ <ʲ> and double graphemes are not used to represent gemination.

References

Abaffy, Erzsébet (2003), 'Hangtörténet (Az ómagyar kor)' [History of phonetics and phonology (In the Old Hungarian period)], in Jenő Kiss and Ferenc Pusztai (eds), *Magyar nyelvtörténet*. Budapest: Osiris kiadó, 301–351.

Ács, Pál (2000), 'Tarjumans Mahmud and Murad – Austrian and Hungarian Renegades as Sultan's Interpreters', in Bodo Guthmüller and Wilhelm Kühlmann (eds), *Europa und die Türken in der Renaissance*. Tübingen: Niemeyer, 307–330.

Aličić, Ahmed (2000), 'Uvod [Introduction]', in Adem Handžić, Snježana Buzov and Amina Kapusović, *Opširni popis Bosanskog sandžaka iz 1604. Godine*. Sarajevo: Bošnjački institut Zürich, Odjel Sarajevo & Orijentalni institute, iii–xxxvii.

Bichlmeier, Harald and Branka Ivušić (2013), 'Zur dialektologischen Einordnung der deutschen Texte einer osmanischen Sammelhandschrift vom Ende des 16. Jh.s', in Rüdiger Harnisch (ed.), *Strömungen in der Entwicklung der Dialekte und ihrer Erforschung. Beiträge zur 11. Bayerisch–Österreichischen Dialektologentagung in Passau September 2010* (Regensburger Dialektforum, 19). Regensburg: Edition Vulpes, 365–384.

Bonioli, Maria (1962), *La pronuncia del latino nelle scuole dall' antichità al rinascimento*. Turin: G. Giappichelli.

Buğday, Korkut (1999), *Osmanisch – Einführung in die Grundlagen der Literatursprache*. Wiesbaden: Harrassowitz.

Daniels, Peter (2014), 'The type and spread of Arabic script', in Meikal Mumin and Kees Versteegh (eds), *The Arabic Script in Africa – Studies in the Use of a Writing System*. Leiden, Boston: Brill, 25–39.

Drkić, Munir, and Alen Kalajdžija (2010), *Omer Hazim Humo: Sehletul-vusul – Grafija i leksika Sehletul-vusula* [Omer Hazim Humo: Sehletul-vusul – Writing and lexics of Sehletul-vusul], Mostar: Muzej Hercegovine.

Friedman, Victor (2002), 'Balkan Turkish in Macedonia and Adjacent Areas', in Hasan Celal Güzel (ed.), *The Turks – 6. Turkey*, Ankara: Yeni Türkiye, 609–616.

Hall, T[racy] Alan (1997), *The Phonology of Coronals*. Amsterdam, Philadelphia: Benjamins.

Hammer, Josephus de (1812), *Catalogus codicum arabicorum, persicorum, turcicorum bibliothecae palatinae vindobonensis* [Catalogue of the Arabic, Persian, Turkish codices of the Vienna court library], Vienna: no publisher mentioned.

Hadžijahić, Muhamed (1955), 'Književnost na arabici (aljamiado)' [Literature in arabica (aljamiado)], in *Enciklopedija Jugoslavije I A-Bosk*, 144–145.

Hegyi, Ottmar (1979), 'Minority and Restricted Uses of the Arabic Alphabet – The Aljamiado Phenomenon', in *Journal of the American Oriental Society*, 99 (2): 262–269.

Ivušić, Branka (2014), 'Ein „orientalisches" Manuskript und seine „europäischen" Texte', in Maria Kozianka and Bettina Bock (eds), *Schleichers Erben – 200 Jahre Forschung zum Baltischen und Slawischen*. Hamburg: Baar, 89–112.

Kálmán, Béla (1966), *Nyelvjárásaink* [Our dialects]. Budapest: Nemzeti tankönyvkiadó.

Kißling, Hans Joachim (1938), 'Einige deutsche Sprachproben bei Evlijā Čelebi', in *Leipziger Vierteljahresschrift für Südosteuropa*, 2: 212–220.

Kniezsa, István (1959), *A magyar helyesírás története. Második, javított kiadás* [History of the Hungarian orthography. Second, revised edition]. Budapest: Tankönyvkiadó.

Krstić, Tijana (2011), *Contested Conversions to Islam – Narratives of Religious Change in the Early Modern Ottoman Empire*. Stanford: Stanford University Press.

Lehfeldt, Werner (1969), *Das serbokroatische Aljamiado-Schrifttum der bosnisch-hercegovinischen Muslime – Transkriptionsprobleme*. Munich: Trofenik.

Lehfeldt, Werner (ed.) (1989), *Eine Sprachlehre von der Hohen Pforte*. Cologne, Vienna: Böhlau.

Mumin, Meikal, and Kees Versteegh (eds) (2014), *The Arabic script in Africa – Studies in the use of a writing system*. Leiden, Boston: Brill.

Mittwoch, Eugen and J[ohannes] H[einrich] Mordtmann (1927), 'Die Wiener Sammelhandschrift', in Franz Babinger, Robert Gragger, Eugen Mittwoch and J[ohannes] H[einrich] Mordtmann (eds), *Literaturdenkmäler aus Ungarns Türkenzeit – nach Handschriften in Oxford und Wien*. Berlin, Leipzig: Walter de Gruyter, 70–87.

Muftić, Teufik (1964), 'O arebici i njenom pravopisu' [About arebica and its orthography], in *Prilozi za orijentalnu filologiju*, 14: 101–121.

Németh, J[ulius] (1970), *Die türkische Sprache in Ungarn im siebzehnten Jahrhundert*. Budapest: Akadémiai kiadó.
Nametak, Abdurahman (1981), *Hrestomatija bosanske alhamijado književnosti* [Chrestomathy of Bosnian aljamiado literature]. Sarajevo: Svjetlost.
Papp, László (1961), *Nyelvjárás és nyelvi norma XVI. századi deákjaink gyakorlatában* [Dialect and language norm in the practice of our 16th century scribes]. Budapest: Akadémiai kiadó.
Rexa, Dezső (1901), Gévay Antal irodalmi munkásságához [About Antal Gévays literary production], in *Irodalomtörténeti Közlemények*, 11 (1): 105–117.
Römer, Claudia (2009), '16. yy. Arap harfleriyle yazılmış Almanca dinî ve dünyevi metinler ile Evliyâ Çelebi'nin Seyahatname'sindeki Almanca örnekler – bir karşılaştırma' [16th century German religious and worldly texts written in Arabic characters and the German examples in Evliya Çelebi's Seyahatname – A comparison], in Nuran Tezcan (ed). *Çağının Sıradışı Yazarı Evliyâ Çelebi*, Istanbul: Yapı Kredi Yayınları, 365–372.
Römer, Claudia (2014), 'Cultural Assimilation of a 16th-Century New Muslim – the mecmua ÖNB A.F. 437', in İlhan Şahin et al. (ed.), *CIÉPO 19, Osmanlı Öncesi ve Dönemi Tarihi Araştırmalar, İstanbul Esnaf ve Sanatkarlar Odaları Birliği Yayını*, cilt II, 607–620.
Siptár, Péter (1994), 'A mássalhangzók' [The consonants], in Ferenc Kiefer (ed.), *Strukturális magyar nyelvtan – 2. Kötet: Fonológia*. Budapest: Akadémiai Kiadó.
Siptár, Péter, and Miklós Törkenczy (2000), *The Phonology of Hungarian*. Oxford: Oxford University Press.
Sudár, Balázs (2005), *A Palatics-kódex török versgyűjteményei – Török költészet és zene a XVI. századi hódoltságban* [The Turkish poetry collections of the Palatics-codex – Turkish poetry and music in 16th century Ottoman Hungary]. Budapest: Balassi kiadó.
Škaljić, Abdulah (1966), *Turcizmi u srpskohrvatskom jeziku* [Turkish loanwords in the Serbo-Croatian language]. Sarajevo: Svjetlost.
Stotz, Peter (1996), *Handbuch zur lateinischen Sprache des Mittelalters – Band 3, Buch VII: Lautlehre*. Munich: Beck.
Szende, Tamás (1992), *Phonological Representation and Lenition Processes*, Budapest: Linguistics institute of the Hungarian academy of sciences.
Trunte, Nikolaos (2012), *Slavia Latina – Eine Einführung in die Geschichte der slavischen Sprachen und Kulturen Ostmitteleuropas*. Munich: Otto Sagner.
Unterkircher, Franz (1968), 'Hugo Blotius und seine ersten Nachfolger', in Alois Kisser and Josef Stummvoll (eds), *Geschichte der Österreichischen Nationalbibliothek. Teil 1 – Die Hofbibliothek: (1368-1922)*. Vienna: Prachner, 79–162.
Weil, Gotthold (1917), *Grammatik der Osmanisch-türkischen Sprache*. Berlin: Reimer.
Zakhos-Papazahariou, Emmanuel (1972), 'Babel Balkanique – Histoire politique des alphabets utilisés dans les Balkans', in *Cahiers du monde russe et soviétique*, 13 (2): 145–179.
Zsinka, Ferenc (1923), 'A Bécsi Képes Krónika török bejegyzései' [The Turkish notes in the Vienna Illuminated Chronicle], in *Magyar Könyvszemle*, 30 (3–4): 248–250.

Appendices

Appendix I: List of abbreviations

Languages

Class. Lat.	Classical Latin
Cro.	Croatian
CSS	Central South Slavonic
Ger.	German
Hun.	Hungarian
Lat.	Latin
OCS	Old Church Slavonic
Stand. Trk.	Standard Turkish

Other abbreviations

1.	first person
2.	second person
3.	third person
abl.	ablative
acc.	accusative
comp.	complementizer
dat.	dative
det.	determiner
gen.	genitive
imp.	imperative
indf.	indefinite
inf.	infinitive
neg.	negation
ÖNB	Österreichische Nationalbibliothek
pfv.	perfective
pl.	plural
poss.	possessive
prs.	present
r	recto
sg.	singular
v	verso

Appendix II:

Table 4: Most frequent graphic representations of vowel phonemes

Vowel	Position	Graphic representation	Language
	Initial	Ā <ʾā>	Cro., Hun., Ger.
		Á <ʾa>	Hun.?[12]
a	Mid	Á <ʾa>	Cro., Hun.?, Ger., Lat.
		Ā <ʾā>	Cro., Ger., Lat.
	Final	Á <ʾa>	Cro., Hun.?, Ger., Lat.
		Ā <ʾā>	Cro., Hun.?, Lat.
a:	Initial	Ā <ʾᵃ>	Hun., Lat.
		Á <ʾa>	Lat.
	Mid	Á <ʾa>	Cro., Hun., Ger., Lat.
		Ā <ʾᵃ>	Cro., Hun., Ger., Lat.
	Final	Á <ʾa>	Hun., Ger.
		Ā <ʾᵃ>	Hun., Ger.
ɒ	Initial	Ā <ʾᵃ>	Hun.?
		Á <ʾa>	Hun.?
	Mid	Á <ʾa>	Hun.?
		Ů <wu>	Ger.
	Final	Á <ʾa>	Hun.?
		Ā <ʾᵃ>	Hun.?
ɒ:	Initial	-	
	Mid	Ů <wu>	Ger.
	Final	-	
e, ɛ, ə	Initial	Á <ʾa>	Hun., Ger., Lat.
	Mid	´<a>	Cro., Hun., Ger., Lat.
		◌ʰ◌́ <ah>	Cro., Hun., Ger., Lat.
	Final	◌ʰ◌́ <ah>	Cro., Hun., Ger., Lat.
		◌̈◌́ <aẗ>	Ger.
		´<a>	Ger.

12 In Standard Hungarian and most Hungarian dialects, a short phoneme /ɒ/ contrasts with long /a:/. In the German sample, which is mainly Central Bavarian, words containing the vowels /ɒ/ and /ɒ:/ in Central Bavarian but /a/ and /a:/ in Standard German and other dialects show Ů <wu> spellings, suggesting that the sounds in question are /ɒ/ and /ɒ:/, not /a/ and /a:/. In the Hungarian sample, however, words with an assumed /ɒ/ are not written in a way that suggests any other sound but /a/. A short /a/ phoneme is known in northern Hungarian dialects (Palóc dialects) but there the correspondent of /a:/ in Standard Hungarian and the other dialects is /ɒ:/ (Kálmán 1966, 40), of which no traces can be found in the Hungarian sample.

e:, ɛ:	Initial	اَ <ʾa>	Hun., Ger., Lat.
	Mid	ـَ <a>	Cro., Hun., Ger., Lat.
		ـَه <ah>	Cro., Hun., Ger., Lat.
	Final	ـَه <ah>	Hun., Lat.
i, ı	Initial	اِي <ʾyi>	Cro., Hun., Ger., Lat.
		اِئ <ʾyʔi>	Ger.
	Mid	ـِي <yi>	Cro., Hun., Ger., Lat.
		ـِ <i>	Cro., Hun., Ger., Lat.
	Final	ـِي <yi>	Cro., Hun., Lat.
		ـِ <i>	Cro.
i:	Initial	اِي <ʾyi>	Hun., Ger., Lat.
		اِئ <ʾyʔi>	Ger.
	Mid	ـِي <yi>	Cro., Hun., Ger.
		ـِ <i>	Cro., Ger.
	Final	ـِي <yi>	Ger.
o, ɔ	Initial	اُو <ʾwu>	Cro., Hun., Ger., Lat.
	Mid	ـُو <wu>	Cro., Hun., Ger., Lat.
		ـُ <u>	Cro.
	Final	ـُو <wu>	Cro.
o:	Initial	اُو <ʾwu>	Hun.
	Mid	ـُو <wu>	Cro., Hun., Ger., Lat.
		ـُ <u>	Cro.
	Final	ـُو <wu>	Hun., Ger., Lat.
ø	Initial	اُؤ <ʾwʔu>	Hun.
		اُو <ʾwu>	Hun.
	Mid	ـُؤ <wʔu>	Hun.
		ـُو <wu>	Hun.
	Final	-	
ø:	Initial	اُؤ <ʾwʔu>	Hun.
	Mid	ـُؤ <wʔu>	Hun.
		ـُو <wu>	Hun., Ger.
		ـُ <u>	Ger.
	Final	ـُؤ <wʔu>	Hun.
u, ʊ	Initial	اُؤ <ʾwʔu>	Cro., Hun., Ger., Lat.
		اُو <ʾwu>	Cro., Hun., Ger.
	Mid	ـُؤ <wʔu>	Cro., Hun., Ger., Lat.
		ـُو <wu>	Cro., Hun.
	Final	ـُؤ <wʔu>	Cro., Lat.
		ـُو <wu>	Cro.
u:	Initial	اُؤ <ʾwʔu>	Hun.
	Mid	ـُؤ <wʔu>	Cro., Hun., Ger.
		ـُو <wu>	Cro., Hun., Ger.
	Final	ـُؤ <wʔu>	Hun., Ger.
		ـُو <wu>	Ger.

y, ʏ	Initial	اوُ <ˀwʔu>	Hun.
		اوُ <ˀwu>	Hun.
	Mid	وُ <wʔu>	Hun., Ger.
		وُ <wu>	Hun.
	Final	-	
yː	Initial	-	
	Mid	وُ <wʔu>	Hun., Ger.
	Final	وُ <wʔu>	Hun.
ɤ̞	Initial	-	
	Mid	ـِر <yir>	Cro.
		ـَر <ar>	Cro.
	Final	-	
a͡ɪ	Initial	آي <ˀᵃy>	Ger.
		اَي <ˀay>	Ger.
	Mid	اَي <ˀay>	Ger.
		ـَي <ay>	Ger.
	Final	اَي <ˀay>	Ger.
a͡ʊ	Initial	آو <ˀᵃw>	Ger.
		اَو <ˀaw>	Ger.
	Mid	آو <ˀᵃw>	Ger.
		اَو <ˀaw>	Ger.
	Final	اَو <ˀaw>	Ger.
ɪ̯ɐ̯	Initial	-	
	Mid	ـيَ<iya>	Ger.
	Final	ـيَه<iyah>	Ger.
		ـيَة<iyaẗ>	Ger.
ɪ̯ɛ̯	-		
	Mid.	ـيَ<iya>	Cro.
	-		
ɔ̯ʏ	Initial	اُوي <ˀuwy>	Ger.
	Mid	وُي <wuy>	Ger.
	Final	وُي <wuy>	Ger.
ʏ̯ɐ̯	Initial	-	
	Mid	وَ <wa>	Ger.
	Final	-	

Florian Sobieroj

Standardisation in Manuscripts written in Sino-Arabic Scripts and *xiaojing*

Abstract: Standardisation processes concerning orthography, handwriting and page layout can be observed in manuscripts written in Sino-Arabic scripts that may or may not include transliterated Chinese-language texts (*xiaojing)*.[1] Besides identifying some of these processes, it is the objective of the present paper to explore the *xiaojing* phenomenon with regard to name and script, earliest evidence as well as its function as a system of writing Chinese. The material used for this investigation are trilingual manuscripts written in Arabic, Persian and Chinese mostly produced in Northwest China in the context of Naqshbandiyya-based Sufism and higher education at the *madrasa*s. Accordingly, the texts inscribed in the manuscripts relate mainly to Islamic mysticism and dogma, to prayer and philology. In the presentation of this material, different page-layout formats and configurations of languages will be looked at and the conventions that have been followed in writing *xiaojing* will also be taken into consideration.

1 Introduction

In this paper, an attempt has been made to identify some standardisation processes concerning orthography, handwriting and page layout in manuscripts written in Sino-Arabic scripts (*khaṭṭ-i ṣīnī*), which include *xiaojing*, i.e. Chinese texts transliterated in the Arabic script. An effort has been made to grasp the *xiaojing* phenomenon by looking at its name, the script used for it, the earliest evidence of its use and by studying its function as a system of transcribing Chinese[2] while also pointing out its inherent deficiencies. The multilingual manuscripts in Arabic, Persian and Chinese used as material for this investigation were mostly produced in Northwest China in the cultural environments of the Naqshbandī Sufi and *madrasa* education. The material includes some of the most important Islamic texts which continue to be studied by Chinese Muslims, such as the mys-

1 Sino-Arabic scripts in this paper are defined as styles of Arabic writing which show a Chinese influence, irrespective of the language, be it Arabic, Persian or Chinese, while *xiaojing* signifies (dialect forms of) the Chinese language written in the Arabic script.

2 It will be shown that the Arabic writing system allows to record the spoken Chinese language in a way that the Chinese characters cannot.

https://doi.org/10.1515/9783110639063-008

tical *Kitāb al-Jahrī*, *Manāqib* and *Ashiʿʿat-i Lamaʿāt*, and additionally some grammars, catechisms, glossaries and prayer books. In this presentation of textual material, different page-layout formats and configurations of languages will be identified.

Besides covering these points, the *xiaojing* conventions generally observed by scribes will be outlined and exemplified. These include the following, among others: the choice of a small, unpretentious script (mostly); full (or nearly full) vocalisation; graph features serving as adaptations to the Chinese sound system, such as doubling of vowel signs within the *rasm* (consonantal skeleton) of a word; the letters *kāf* and *ṣād* with three dots above them; *fatḥa* added 'superfluously' to a consonant followed by *alif*; hybrid forms including Chinese characters (occasionally); *jīm* written in the initial form even when it occurs in isolation; and the lack of distinction made between the consonantal sounds x- (as in Chinese *xi*) and s- (as in *si*), although the sounds q- (as in *quan*) and j- (as in *jiang*) are distinguished. The *xiaojing* writing system also accurately represents final nasal consonants in accordance with dialectal differences, such as -n/-ŋ variation. In Mandarin variety spoken in Linxia (Lanyin 兰银 Mandarin, the dialect retaining some linguistic traits of Middle Chinese (MC) and which features prominently in the manuscripts discussed in this chapter) the final nasal is realised as [-n] and this sound is written as <n> in Arabic script. This differs from the *xiaojing* used for writing the dialects exhibiting the final velar nasal [-ŋ], which is represented by the letter *kāf* with three dots above.[3]

The Latin-script transliterations of the *xiaojing* transcriptions given in this chapter follow the pinyin system used in the People's Republic of China to transliterate Chinese, despite the fact that pinyin is a rather artificial system and far from ideal for recording the variety of the spoken languages. As the *xiaojing* texts are transliterations of Chinese dialects rather than of Mandarin (Putonghua), the pinyin transliterations can only be considered as approximations to the actual pronunciation.

3 The authors of the Chinese Wikipedia article '小儿经' (*xiaoer jing*) (Wikipedia '小儿经' *xiaoer jing* 2019) specify that *xiaojing* not only transliterates the language spoken in the Northwest (Lanyin) but also that of the Central Plain and of the Northeast. According to the charts in the same Wikipedia article, velarised /ŋ/ is represented by final *kāf* ڭ written with three dots above and by initial consonant with *sukūn* above it + vowel, e.g. يًا ڭ (= *yang* 羊, 'sheep'), وًا ڭ (= *wang* 忘, 'to forget'), يُو ڭ (= *yong* 用, 'use'); however, these combinations of letters have not been seen in the manuscripts studied, nor are they included in the Xining *xiaojing* syllabary (see below). It will be demonstrated (e.g. section 4.4.1.), that even the structure of the Chinese language transcribed in the manuscripts shows dialectal influence.

1.1 The name *'xiaojing'*

The name of the Arabic transliterations of Chinese language is used in a small number of variations. *Xiaojing* 小經 (lit. smaller canonical writings) is generally understood to refer to children because it was used in teaching young children Islamic texts in Chinese before they studied the Arabic language.[4] Corresponding to this understanding, the designation *xiaoer jing* 小兒經, meaning 'children's canonical writings' has also been used, and the name *xiaojing* may therefore also be considered an abbreviation of this designation. This didactic use of *xiaojing* is apparent in a small number of Chinese catechisms written in the Arabic script said to have been composed specifically for Muslim children, for instance those authored by Ma Tianmin 馬天民 in the middle of the 20th century, as discussed in section 4.3. However, Arabic transliterations of Chinese were not only employed in primary schools (*xiaoxue* 小學), which were financed by the Muslim communities (*jiaofang* 教坊)[5] and therefore enjoyed independence from the state, but they were also used widely in *jingtang jiaoyu* 經堂教育, or *madrasa* education, in Northwest China. This institution emerged in the middle of the 16th c. in Shaanxi province, from where it moved to southern Ningxia.[6]

4 Cf. Bakhtiyar (1994, vol. 4, 71).

5 Cf. Stöcker-Parnian 2002, 154.

6 More specifically, to the Tongxin 同心 area of Ningxia; cf. Zhou (2008, 47-48), who mentions the name of Hu Dengzhou 胡登洲 (d. in 1597) from Shaanxi province, the founder of *jingtang jiaoyu* (also Yang 1996, 79).

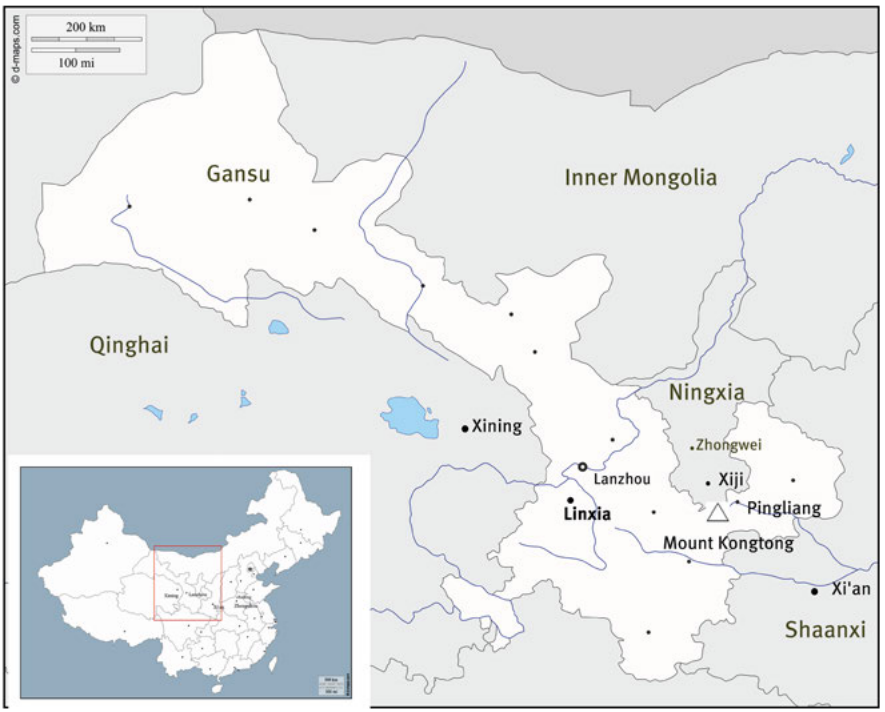


Fig. 1: The provinces Qinghai, Gansu, Ningxia, and Shaanxi.

Feng Zenglie 馮增烈 (2007, 618) has proposed an alternative explanation of the name *xiaojing* 小經, however. He explains the designation in relation to the complementary term *dajing* 大經: the 'larger *jing*' are the Islamic scriptures in the Arabic and Persian languages, whereas the 'smaller *jing*' (*xiao jing*) are those written in Chinese language, but in Arabic script. Besides this, Feng mentions the variant designation *xiaojing* 消經 (with nearly identical pronunciation [different tone of the first word], but written in different characters), which he explains as denoting the 'digesting of [Islamic] canonical literature' (*xiaohua jingwen* 消化經文).[7]

7 The authors of the article '小儿经' (*xiaoerjing*) (Wikipedia '小儿经' *xiaoer jing* 2019) claim, while referring to an internet publication as their source (footnote 2: 回族的语言和文字*Huizu de yuyan he wenzi*, published 2005 on the site '宁夏 旅游网' Ningxia lüyou wang, a tourism page), that in the Northwest the name variant *xiaoerjin* 小兒锦 is generally used, abbreviated as *xiaojin* 小锦, while in the Central plain and the Northeast the preferred form is *xiaoerjing* 小兒經.

1.2 The alphabet and scripts

The *xiaojing* writing system is based on the 28 letters of the Arabic alphabet, to which *xiaojing* adds some letter forms derived from Persian, such as *pa*, *čim* and *žay*, and a few relatively unfamiliar graphs such as *ṣād* with three dots above it (as used for the initial c-, as in *can* (Mandarin *cang*) 倉, 'storage').[8] The three dots above the Arabic letter *kāf* (to represent the phoneme j-, as in *jiu* 酒, 'wine' in the standard language [not identical with the one transliterated in the *xiaojing* texts])[9] are an adaptation to the Chinese sound system, whilst the doublings of *ḍamma* (nunation), *fatḥa* and *kasra*, which were added to the consonantal letters in the middle or final position, is a characteristic feature of *xiaojing*.[10] These orthographic features, with variations, are standard usage throughout the *xiaojing* manuscripts consulted. Although they are not written in one particular calligraphic style, the letters are mostly written in a very small script. Arabic texts such as the Qur'an, *Takhmīs al-Burda* (*Mukhammas*) and *Madāʾiḥ*, which are often accompanied by *xiaojing* glosses in the manuscripts, are written in the bold 'hieratical'[11] *khaṭṭ-i ṣīnī* style in many cases, which is understood to show the influence of Chinese calligraphy. The presence of a sacred text seems to function as a trigger, fostering a standard in handwriting which may or may not include *xiaojing*.

8 The letter *ṣād* with three dots is also used in the writings of the Tatars of Lithuania and Poland (as pointed out by Alessandro Gori); for this letter in Belorussian, Lithuanian and Polish Tatar manuscripts see Miškinene 2015, 66); a relevant image can be found in Sobieroj 2010, plate 3 (with three dots beneath [!] *ṣād*).

9 *Kāf* with three dots above was used for /ŋ/ in Ottoman Turkish, for /g/ in North African Arabic and Berber, for /g, ŋ, ng/ in Wolof, and in other sub-Saharan writing traditions as well (as indicated by the editors of this volume).

10 A useful set of charts of *xiaojing* letters and compounds accompanied by their Chinese character homologues in Hanzi (小經字母和拼音 ثِيَوْ كٍ زِ مُو خُوَ پٍ ىِ) is included in the Xining edition of *ʿAqīdat al-islām* / *Xinyang wenda* (no date), pp. 77–79. A list of the phonetic values of 36 consonants in the initial position and 73 syllables ending with a vowel can be found in the Wikipedia article '小儿经' (Wikipedia '小儿经' *xiaoer jing* 2019).

11 I have designated it this way because the classical Ṣīnī style was mainly employed to represent writings considered sacred by Northwest Chinese Muslims.

1.3 The significance of *xiaojing* as a system of transliteration

In his volume dedicated to the subject of Islamic education among the Hui, designated as a 'national minority' in China,[12] Zhou Chuanbin 周傳斌 highlights the unique role played by *xiaojing* as a system of phonetic transliteration of the Chinese language in use in the pre-modern era and as a direct (chronological) precursor of the Latin script.[13] However, as Zhou points out (2008, 60–61), outside China, in the former Soviet Union, the Cyrillic script has been used to transliterate Chinese in a version developed on the basis of *xiaojing*.[14] The Chinese in question are the so-called Donggan, Muslims from the provinces of Shaanxi and Gansu who were deported beyond the borders of the Qing Empire to Central Asia – Yang (1996, 71–72) adds that the destination was Kyrgyzstan, Uzbekistan and Kazakhstan. The deportees, referred to in the Russian Czarist empire as 'Dungan' but belonging to the same stock as the Hui in China, initially used *xiaojing* to transcribe their native language, but in the first half of the 20th century they were made to adopt the Cyrillic alphabet after having gone through a phase of using the Latin script.[15]

1.4 The earliest evidence of *xiaojing*

Zhou (2008, 57) mentions a stone monument erected in Xixiang 習巷 Mosque at Xi'an University as the earliest evidence of the usage of *xiaojing*. This monument bears an Arabic text on the construction of the building, including Chinese per-

12 Zhou (2008, 59) includes some pictures illustrating the usage of *xiaojing*, including a modern Arabic transcription of *Zhonghua renmin gonghe guo*, 'The People's Republic of China'; Zhou (2008, 79) offers a picture of *xiaojing* marginalia in a manuscript of the Arabic grammar *Zaowu misubaha* 遭五米素巴哈 [*Ḍaw' al-miṣbāḥ*].

13 Already in the late 16th century the Jesuits had developed a system of transliteration that has left traces in modern pinyin (for the wider context see e.g. Mungello 1989). An in-depth study of *xiaojing* transliterations carried out with the help of linguists while especially taking into account the unusual graph features may help to reconstruct the Chinese language spoken at the time under scrutiny. A comparative study of the *xiaojing* writing system and the Dunggan material from the former Soviet Union is also a desideratum.

14 It should be added that the Dongxiang ethnic group, who live in the Linxia region of Gansu province, also used a variant of *xiaojing* (cf. Yibulaxin 易卜拉欣 2007, 138) called Dongxiang wen 东乡文, 'writing of the Dongxiang' or Huihui wen 回回文, see (Wikipedia '小儿经' *xiaoer jing* 2019).

15 For the context see e.g. Dyer 1967.

sonal names and biographical notes written in the Arabic script in 1339–40.[16] However, *xiaojing* was only used more widely after the emergence of the *madrasa* system at the end of the Ming era (1368–1644). Feng (2007, 618b) mentions two manuscripts produced after the rise of the *madrasas* as the oldest evidence of the usage of *xiaojing* in Islamic writings which are designated as *jingji* 經籍) by Feng; one of these is an 18th-century copy of the Persian Sufi manual *Mirṣād al-ʿibād* written by the Kubrawī mystic Najm al-Dīn Rāzī (d. 654/1256; Okuyan 1988ff., vol. 32, 496–497); this includes Chinese glosses written in Arabic.[17] The manuscript was taken to France in 1909 and given to the Bibliothèque nationale in Paris as part of the exploits of the famous expedition to Gansu led by Henri d'Ollone (d. 1945). It was then described (albeit rather briefly) by Emile Blochet (1909, 290 [no. 8]).[18]

1.5 The shortcomings of *xiaojing*

One major shortcoming of this transliteration system was pointed out by Zhou (2008, 60), who says that a unified system of writing *xiaojing* has never existed; instead of adopting one standard, every Chinese dialect written in the Arabic script has been transcribed according to its own system, which means that certain conventions existed at the dialect level.[19] However, generally speaking, a lack of

16 A dating most likely referring to the manuscript rather than to the text. Zhou also fails to indicate whether or not the transcription of these names shows any characteristics specific to *xiaojing*. Bakhtiyar (1994, 4, 77; with an illustration) mentions the earliest evidence of *xiaojing* which he saw in a manuscript, namely a copy of Rashīd al-Dīn Faḍl Allāh's (d. 1318) work on medicine, *Ṭibb-i ahl-i Khitā*, apparently dated 1313 CE, i.e. during the author's lifetime (!).

17 In view of its importance as a foundational text for the Kubrawiyya Sufi order, the *Mirṣād* was translated into Chinese: 'Wu Zunqi 伍遵契 began translating it in 1672 and completed it six years later. He entitled it 歸真要道釋義 *Guizhen yaodao shiyi* 'Explanation of the Main Path of Returning to God', or *Guizhen yaodao* 歸真要道 'The Main Path of Returning to God', etc.' (Ma 1999, 34).

18 Feng (2007, 618b) confuses 'F. (sic!) Blochet' with Henri d'Ollone by making an explorer out of the cataloguer while also failing to mention d'Ollone's name. The *xiaojing* marginalia in this manuscript are also mentioned in Bausani (1968, 875).

19 *Xiaojing* transcriptions seem to have mainly been used in manuscript production in the Northwest (but see fn. 3). This impression is based on the perusal of some Arabic manuscripts that originated in the southerly province of Yunnan in which there are no traces of *xiaojing*. One example is a facsimile manuscript of the *Muttasiq al-naḥw*, which is an introduction to Arabic grammar written by Ma Fuchu 馬復初 (Na 2007, 330–331), i.e. Ma Dexin 馬德新 Yūsuf (d. 1874), printed in 1375/1955–56 and sold as a reprint from 1405 [= 1984-5] in Kunming (Chinese title on the front cover: *Jianming ayuxue* 簡明阿語學). The 102-page Arabic manuscript with a 12-page glossary is accompanied by extensive glosses mainly written in the margins, but not a single one is in *xiaojing*. The first page of the manuscript includes biographical notes on the author in

standardisation in its orthography is apparent. Feng (2007, 618b) and Yang (1996, 82–84) both highlight a lack of standardisation as well, which they trace back to the influence of dialects and the tendency of the authors of *xiaojing* (para-) texts to act independently of one another. Another reason is the different sound systems of Chinese and Arabic: the intonation and tone patterns of Chinese, i.e. the multiple tones (in the modern standard language four, in certain dialects more, in others less), are not represented graphically in the Arabic transcriptions, which often leads to confusion and mistakes. Besides that, no distinction was made in transcribing words consisting of one or two syllables respectively, such as *jiu* 'wine' and *ji you* 'chicken oil' (which probably sounded more or less the same; the examples are taken from Yang 1996, 84).

2 *Xiaojing* in *Kitāb al-Jahrī*

Ma Xuezhi 馬學智 Muḥammad Manṣūr, one of the authors of the sacred biography of the Naqshbandī (Jahrī) reformer of Guanchuan 關川 in Gansu, Ma Mingxin 馬明心 Wiqāyatullāh (d. 1781; Ma 2007, 345; Forbes 1960–2004, 5, 851a), says in the preface of his Arabic-language *Kitāb al-Jahrī* (Manṣūr 1933, 15)[20] that he is going to use *xiaojing* at certain points in the text, for which he gives the following reason: 'Most of the words of our shaykhs are in the Chinese language. Expressing them in Arabic is difficult, so some sayings are in Chinese, as it has been feared that the [meaning of the] original saying may be altered by the Arabic expression'.[21]

The *xiaojing* transcriptions scattered in the facsimile Arabic manuscript of the *Kitāb al-Jahrī* (429 pages of 15 lines written by at least two hands in *khaṭṭ-i ṣīnī* scripts) may be divided into personal names, toponyms, poetical verses, dialogues, proverbs, glosses and compounds.

Arabic, while the glossary has translations of well over a thousand words (*mufradāt*) occurring in the text, which are written in Chinese characters.

20 The *K. al-Jahrī* on the history of eight generations of masters of the Jahriyya branch of the Naqshbandiyya order was authored by Ma Xuezhi 馬學智 Manṣūr (d. 1923) and was published in 1933 as a facsimile manuscript written in a *khaṭṭ-i ṣīnī* script by Ibrāhīm, Manṣūr's servant (a partial Chinese translation was published in 1997). The text starts with the vita of the 'Pathfounder', Ma Mingxin, and ends with that of Ma Zhenwu 馬震武, who died in 1960 (cf. Sobieroj 2016a, 140).

21 The passage is translated in the *Daotong shizhuan* (Manṣūr 1997, Introduction, 10): '[...] in order to maintain the original meaning of the Mawlā's [i.e. Mingxin's] words (*baochi yuanhua*), I have used a *xiaoerjin* 小兒錦 (!) transcription [in] many places and hope that the reader will understand [...]'.

2.1 Personal names

Laojun 老君, which literally means 'old man', designates the deified Laozi. It occurs in a question uttered reproachfully by Ma Mingxin ('Why, then, are there people [in my ritual circle] who still believe that the "Old Man" is God' [*man iʿtaqada al-Laojun ilāhan*]) and was directed at an erstwhile Taoist monk who had converted to the Islam of the Jahriyya. The name *laojun* written with the Arabic definite article has three dots above the letter *kāf* and nunation above the final *yāʾ*: اللوكّيٌ .

The letter *kāf* representing 'Jie' in the name of the famous 18th-century mystic Liu Jielian 劉介廉[22] (Manṣūr 1933, 40) is likewise written with three dots above the horizontal line.

The names of the (eras of the) Qing emperors Qianlong 乾隆 and Kangxi 康熙, which are mentioned in *Kitāb al-Jahrī* (Manṣūr 1933, 69) to specify the year of the 'martyrdom' (*shahādatuhū*), i.e. death, and birth of Mingxin respectively, are also transcribed in the Arabic script, viz. as للسّلطان الكِيَنْلُنْ and للسّلطان الكانسي. Only the first name, *al-Sulṭān al-Kianlun* (= *al-Qianlong*), is provided with vowel signs; the second one was left void of any vocalisation. The lack of vocalisation in the writing of al-Kansi is indicative of the absence of standardisation – unless it has been omitted because the *matres lectionis* make the *rasm* unambiguous.[23] To sum up, then, there seems to have been a tendency to standardise the transcription of names by prefixing the Arabic article *al-* to Chinese names.

2.2 Toponyms

In accordance with *xiaojing* conventions, topographical names are usually provided with full vocalisation in the Arabic text (e.g. in Manṣūr 1933, 32, *al-Lianhuachen* اللِيَنْخُواچِن 蓮花城; *chen* is used instead of *cheng* 'town' for standard /ŋ/). However, in the second occurrence of the place name four lines below that, the vocalisation is dispensed with, perhaps because of having been considered redundant. On p. 32, the names of the town of *Fuqiang* 伏羌 and the mountain called *Liujiapo* 劉家破 are also stated in the Arabic script, the latter name with full vocalisation.

In the toponym *Qin'an xian* 秦安縣, written as الكِنْغَنْسِيَن (Mansūr 1933, 37), the non-phonemic initial consonant sound of *an* 安 is irregularly rendered by *ghayn*

22 I.e. the mystical philosopher Liu Zhi 劉智 of Nanjing (d. 1745; Luo 2007, 321–322), the 'Ibn ʿArabī of China'.

23 In contrast, the name of the chieftain of the village of Didianzi, Ma Laoye 馬老爺, is only provided in the Arabic translation: *al-Amīr al-Farasī* (Manṣūr 1933, 76, penultimate line).

as opposed to *ʿayn*. (This is the so called 'empty initial', or 'zero onset', which have a range of variation such as [ʔ], [ɣ], [ŋ] and [ɦ], Duanmu 2007).

The name of the river called *Miaoer he* 廟兒河 المِيَوْعَخُوَ is transcribed with full vocalisation, including the word *he* for 'river', with *ḍamma* above the letter *khāʾ* and *fatḥa* above the letter *waw*. The name of the *Gaoshan* 高山or 'High Hill' mountain mentioned three lines above, in contrast, is given in the Arabic translation, viz. *al-Jabal al-ʿālī* (Manṣūr 1933, 45).

The toponym *Guanmenkou* 関門口 (Manṣūr 1933, 46), to which the definite article is added, has been transcribed with vocalisation. The vocalisation of *men* 門 'gate' with *ḍamma* مُن and *kou* 口 'mouth'/'entrance' with *kasra* beneath the letter *kāf* كِو is also noteworthy.

In the place name *Puer zhuang* 普爾莊 (Manṣūr 1933, 46), consistently, the final [-ŋ]sound of *zhuang* has not been reproduced in the Arabic script either. Conspicuously, the syllable *er* has merely been transcribed by the letter *ʿayn* with the *fatḥa* vowel sign.

Only rarely has a place name been left without any vocalisation at all. In one such instance, the unvocalised *xiaojing* is explained in an Arabic translation: on p. 76, the location called *Didianzi* 底店子, which was populated by people who honoured Mingxin, but hated his successor, the Shaykh of Pingliang, is transcribed as ديدينز, but the name is preceded by an Arabic translation and note, viz. *ribāṭ al-qaʿr, wa-huwa bi-lisān al-ṣīn Didianzi* 'hospice of the depression, which is called D. in Chinese'. It appears that toponyms are regularly provided with vocalisations in the manuscript, while less often a place name is given in an Arabic translation.

2.3 Chinese poetry

The lines of a *duilian* 对聯 or 'antithetical couplet of parallel sentences'[24] which Ma Mingxin recited extemporaneously in reply to a request by some dignitaries at the governor's court in Xi'an, are transcribed in the Arabic script. As for the features of the script , the *khaṭṭ-i ṣīnī* used for transcribing the *duilian* has been

24 The couplet quoted by Mingxin is designated as poetry both in *Kitāb al-Jahrī* and the Chinese translation: The daozu 道祖, 'Pathfounder', was asked to compose poetry (*inshād al-shiʿr*, *zuo shou shi* 作首詩) by his detractors and he accepted the challenge: 'You want me to compose verses, while you cannot compose the first line (*shang lian* 上聯). Well, in that case I will compose the first and you do the second line' (*dui xia lian* 对下聯 ; Manṣūr 1997, 14-15; Manṣūr 1933, 28). The *duilian* includes rhyme words, viz. 'Shaanxi' 陝西 and 'ma ti' 馬蹄 ('horses' hoofs').

provided with full vowel signs, and in accordance with the traditional techniques of writing poetry in Arabic and Persian manuscripts, the endings of the verses have been marked by clusters of dots. The script of the *xiaojing* does not differ from that of its textual environment apart from its vocalisation.

As an inclusion of the material in the manuscript of Zhanye's 亶爺 *Manāqib* (e.g. p. 21) shows (for the text see below [p. 191]), it has been standard procedure to transcribe the poetical verses composed and recited by the masters of the Jahriyya in *xiaojing* rather than translating them into Arabic.[25] The verses are made to stand out by using full vocalisation, and verse markers have been inserted between the hemistichs and at the end of the lines as well.

2.4 Dialogues

A second example of a *xiaojing* passage integrated within the main text of the *Kitāb al-Jahrī* can be found in the second chapter, which is dedicated to the life of Wiqāyatullāh's successor, the 'Pingliang taiye' 平凉太爺 (*Ḥaḍrat Maulānā al-Aʿẓam al-Pinliānfūwī*; i.e. shaykh of Pingliang prefecture), also known as Mu Xianzhang 穆憲章 (cf. Ma 1999 [1985], 93). A somewhat obscure dialogue conducted between the Sufi Shaykh and the Taoist abbot (*sayyid al-ruhbān* 'lord of the monks') of the nearby monastery on Mount Kongtong[26] is given in the original Chinese wording without a translation.[27] The dialogue included in the narrative seems to constitute an exercise in *mufākhara*, in which two disputants try to outwit one another. The four (unrhymed) lines of *xiaojing* stand out against their textual environment because of the vowel signs and the circles added as

25 The recitations – again following the conventions of Arabic and Persian manuscript culture – are regularly introduced by the formula *fa-anshada [shiʿran bi-lughat al-Ṣīn]*, 'he recited (a poem in Chinese)' (e.g. *Manāqib*, p. 71; also pp. 211, 212, 216, 226 and 264).

26 The greatest scenic attraction of the Pingliang area is the Taoist monastery on the summit of Kongtongshan 崆铜山 (Mount Kongtong) which is also known as *Jiu gong ba tai shier yuan* 九宫八台十二院, 'Nine palaces, eight terraces, twelve courtyards'.

27 'One day [...], the Pingliang Taiye went to the famous Mt Kongtong west of the city to gather some medicinal herbs. Once he had climbed up to the summit of the mountain, he encountered the head of the monks there, who was called Chen Banxian 陈半仙, the "half-immortal Chen" [a bad pun or mockery, FS]. When the latter saw the Pingliang Taiye, he said: "You are wearing a dress of coarse material (ءِ بُـ تَوْ *tubu yi* 土布衣) [...]". The Taiye said: "I can move the pillar which borders on the sky ([...] غَـ شِوْ بًا تِيـَ *wo shou ban tian bian zhu* 我手搬天边柱) with my hand while my foot is on the ground" (ءِ كِّيُوَ تًـ دَوْ دِ *yi jiao tan dao di* 一脚探到底). The monk could not think of anything to respond to that. The Taiye picked the herbs [he wanted], went back down the mountain and returned' (Manṣūr 1997, 79; Manṣūr 1933, 91).

'verse markers'. The traditional standards of Arabic poetry quotation have been followed by using these devices, although the four lines in themselves do not constitute poetry (Fig. 2).

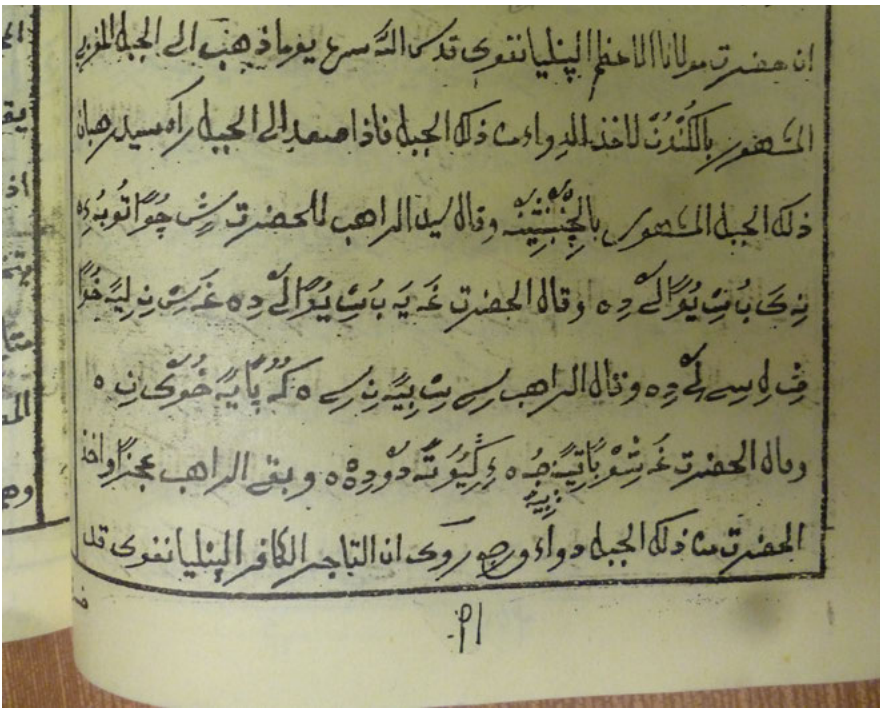


Fig. 2: *Kitāb al-Jahrī* (a dialogue between the Taoist prior and the Shaykh).

2.5 Proverbs

Proverbs, too, have been left devoid of any Arabic translation in *Kitāb al-Jahrī*; they are provided in *xiaojing*, obviously resulting from the author's apprehension that the meanings of the idiomatic expressions would not be adequately conveyed by a translation.

The two parts of the proverb, constituting a *chengyu* (成語) of four words,[28] in Manṣūr 1933, 113 (Manṣūr 1997, 101), which is written in Arabic script with vocali-

28 The idiomatic phrase is uttered, in the third chapter on the life of the master Chuanchang

sation in the manuscript, are separated by a verse marker in the form of a circle. There is an overline at the beginning, the middle and the end of the proverb. The placing of three dots above the letter ص in the word *cang* 倉 (the last word of the phrase) is a regular feature. However, *cang* is written with the homophone character 滄 rather than 倉 in the *Daotong shizhuan*. The consonant x- in *xiu* 修 'to build' is represented by *sīn* as in other cases too.

The description above serves to show that proverbs and words exchanged in dialogues have been treated in the same way as poetry in terms of their graphic representation.

2.6 Glosses (interlinear and marginal)

Besides the *xiaojing* included in the main textual body of the *Kitāb al-Jahrī*, numerous glosses mostly written in Chinese characters have been added to the manuscript. These were written down around 2008 by a user who at that time was a student and a novice of Honglefu 鴻樂府 Convent and he owned the manuscript. There are also a few cases of Chinese words in *xiaojing* transliteration between the lines and occasionally in the margins as well.

An example of an interlinear gloss can be found in Manṣūr (1933, 37, l. 4 from below) where the phrase *māta min dhālik majnūnan*, 'he died as a result of this, as a madman', has been explained by the expression *tiba* 提拔 'promoted by rank',[29] written underneath the line; the letter *tāʾ* has been provided with *kasra*, and the letters *bāʾ-alif* have been written with *fatḥa*.

On p. 29, the uncommon word *qaswara* apparently denoting a lion is explained beneath the line by the Chinese term *shizi* 獅子 written in Arabic script: شِ ذ.

A lexical gloss has been added in the outer margin on p. 78: قُوافُ. This entry transcribes Chinese *guafu* 寡婦 as the Arabic word *armala*, i.e. 'widow'. It seems that the scribe initially wrote قُافُ and only added the letter *waw* between *qāf* and

Taiye 船廠太爺 (Manṣūr 1939, 105-133), by the ʿĀlim (ṣaghīr) sulṭānī, 'little kingly scholar', called Xiaoshan Wangye (小山王爺) in the Chinese translation (Manṣūr 1997, 101). The ʿĀlim sulṭānī criticised the Sufi master whom he charged of insincerity, by uttering a proverb which is replete with historical connotations (*qāla bi-lisān al-ṣīn*, 'he said in the Chinese language'; Manṣūr 1933, 113): Mingxiu zhandao / andu chencang 明修棧道 暗渡陳倉, غَـدُو چِه ضَنان * مِ سِيُو جَّا دَوْ [building a covered way along a precipice openly – secretly taking the old path to Chencang] (Manṣūr 1933,113; Manṣūr 1997, 101). Peiqi Yan of Munich University (LMU) has kindly guided me to the literary source of this proverb (cf. Cai and Sun 蔡 2008, 545).

29 In Sufism, the *majnūn*, turned mad through passionate love of God, is considered seized by divine attraction (*jadhba*), hence 'promoted'.

alif afterwards or else he superscribed the letter *waw* so it would be interpreted as the vowel sign *ḍamma*.

The expressions *ibtilāʾan wa-ʿkhtibāran* 'trials and tests' encountered on p. 66, l. 6 are accompanied by two words in *xiaojing* transcription written beneath the base line, the first of which is illegible. The second word transcribed with full vocalisation is clear enough, however, and can be deciphered as كَوْيَاً or *kaoyan* 考驗 'test, testing'.

As we have seen, a few marginal and interlinear lexical glosses in *xiaojing* transcription have been added to the manuscript of *Kitāb al-Jahrī* and they reflect practices of annotation typical in the Arabic manuscript culture in their layout and functionality.

2.7 Hybrid compounds

There are also some examples of a mixture of the two, Chinese characters and Chinese in Arabic transcription. An example of this phenomenon may be found in *Kitāb al-Jahrī*, p. 54, line 5 from below, where the Arabic word for 'nose', *khayshūmī*, is explained underneath the line by the hybrid expression 子بِ, i.e. the letter *bāʾ* with the vowel sign *kasra* followed by the character *zi* to render the Chinese word *bizi* 鼻子 for 'nose'. The reason for this strange combination of Arabic and Chinese may be the difficulty the student had in finding an Arabic letter or compound to adequately represent the Chinese syllable *zi* (he may also have had trouble remembering the Chinese character for 'nose' or found it cumbersome to write the numerous strokes it consists of). The Chinese word for 'parrot' (*yingwu* 鸚鵡)[30] on p. 56 is given in an analogous format, i.e. the first word of the compound is written in Arabic letters (*ḥamza* with two *kasra*'s). A further example is the expression *maocao* (the translation [p. 62] correctly has *caomao* 草帽), or 'straw hat', (*qalansuwwat al-ḥashīsh*) written as a hybrid form (草مَوْ) between the lines of p. 74. The expression serves as a means of describing the shape of a fragrant flower which grew on the roof of the mosque in Pingliang and symbolised the Pingliang Taiye as the head of the order. The hybrid marginal gloss on p. 78 consists of a combination of لُو (= *lü* 綠 'green') and the character 翠 (*cui*) to render the Arabic word *zakhārif* 'splendour, decoration' (the meaning of the compound *cuilü* 翠綠 is 'emerald green').

30 The *Kitāb al-Jahrī* uses the Persian word *ṭūṭī* instead of Arabic *babbaghāʾ*.

3 *Manāqib-i Awliyā* 'The virtues of saints'

Another literary version of the sacred history of the Jahriyya Menhuan is the *Manāqib* (*M.-i Awliyā*), and it seems to show a higher proportion of *xiaojing* writing than does the *Kitāb al-Jahrī*. The Arabic text in question was composed by Zhanye 瞻爺 ʿAbd al-Aḥad in the twentieth year of the Republic of China (= 1931). The manuscript published as a facsimile of a *khaṭṭ-i ṣīnī* manuscript with a total of 361 pages was copied by Ma Lugou 馬麓溝 Ṣadīqullāh of Xiji in Ningxia province at the order of his shaykh. There is a colophon (p. 361) including the name of the scribe (*min yad* 'by [the] hand [of]' etc.), Ṣadīqullāh ibn Qamar, but neither a date is stated nor the place of publication of the facsimile. The *Manāqib*, which is composed of five chapters, was designed to serve as a supplement to the *Rashaḥāt* 'percolations' by ʿAbd al-Qādir. In contrast to the latter work, which only provides biographical information on Ma Mingxin and the Pingliang Taiye, Zhanye added the lives of three more successors of the Pathfounder, namely Chuanchang Taiye 船廠太爺 Quṭb ul-ʿālam (p. 82ff.), Siyueba Taiye 四月八太爺 (p. 112ff.) and Shisan Taiye 十三太爺 (p. 194ff.).[31] The Arabic work was translated into Chinese by Ma Siren 馬思仁 and the Chinese version entitled *Mannageibu* was published in Zhongwei, Ningxia province, in 2012.

Not only are there numerous explanatory glosses between the lines (e.g. the unfamiliar Arabic toponym *yamm al-timsāḥ al-aswad*, 'Sea of the Black Crocodile', a literal, albeit slightly inaccurate translation of the name of the province of *Heilongjiang*, 'Black Dragon River', which has been transcribed above the line as خِـ لٌ كِيَانْ (p. 83). There are also some *xiaojing* glosses written in the margins (pp. 194 and 291; a few other glosses are in Chinese characters, e.g. on p. 250).

There is a relatively high number of Chinese poems written in the Arabic script in all five chapters (pp. 71, 206, 207, 211, 212, 216, 226, 241 and 264–265).[32] Some lexical glosses in *xiaojing* between the lines of poetry quotations have been written obliquely above the words that are explained (e.g. on p. 28 where Ma Mingxin quotes verses of the *Mukhammas* while exhorting his disciples to be assiduous in their spiritual practices).

31 Another somewhat shorter text on the history of the Jahriyya entitled *Rashf* has also been added on pages 362–395. This was written in year 8 of the Republic (= 1919), apparently by the same scribe.

32 Zhanye's text also includes some poems in Persian (on pp. 279, 346, 349 and 351) and Arabic (pp. 166–167).

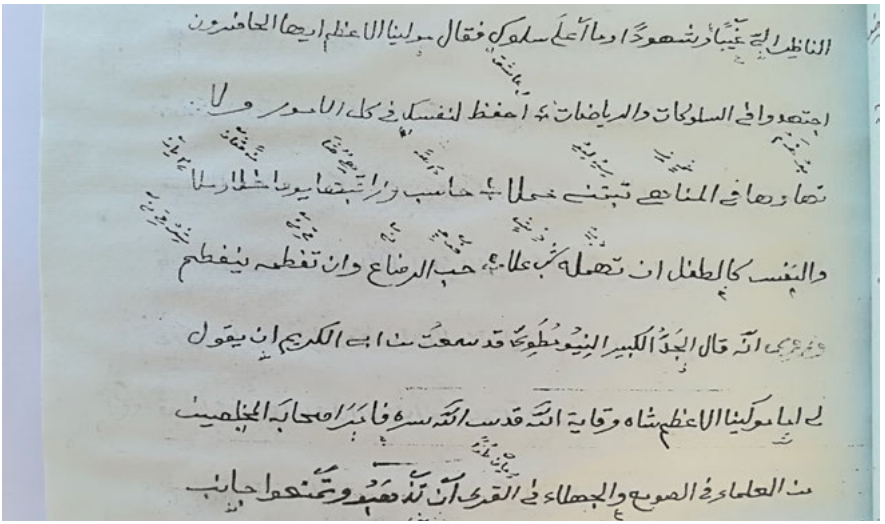


Fig. 3: Oblique annotations, Zhan Ye's *Manāqib al-Jahriyya*, p. 28.

The layout of the poems is a little inconsistent. On one hand, some of the poems are written with spaces between the Chinese verses, which are treated like hemistichs of Arabic poetry, and they are separated by graphical elements as well (e.g. on pp. 71 and 211). Other poetical quotations, on the other hand, are written continuously with a raised inverted comma serving as a verse marker (as on pp. 226 and 264–265).

Besides this, there are a few mysterious Chinese prose words uttered by the shaykhs of the Jahriyya and transcribed in the Arabic script, which answer questions such as 'how will the Islamic religion be fulfilled?' (literally 'sealed'; p. 170f.). As an answer, some Chinese nouns are quoted in juxtaposition and the number of words is mentioned. On p. 171 (= Ma 2012, 73), for instance, Shaykh Quṭb ul-ʿālam answers the above question as follows: 'When [the meanings of] five Chinese characters are fully manifested (*tammat*), Islam will return to its origin, namely دَو وُ سِ لٌ بان (*dao, wu* [?], *sin, lun, ban*; 道務曾 [!, i.e. 僧]龍幫)'; the transliterated Chinese words are identified, in a parallel transmission of the same text, through use of Chinese characters in the translation (Ma 2012, 50 = *Manāqib*, p. 119), but unfortunately an explanation of the relevance of these terms is not provided in the respective passages (the five nouns may be translated literally as 'way', 'duty', 'monk', 'dragon' and 'assistance').

Another relevant 'transmission' (the passages are introduced by *ruwiya*, 'it has been transmitted') is *Manāqib* on p. 95 (= Ma 2012, 41): 'When nine Chinese characters are realised, Jahrī Islam will spread' (*idhā taḥaqqaqa tisʿat aḥruf al-ṣīniyya yantashiru l-islām al-jahrī*).

4 *Xiaojing* in other Sino-Arabic manuscripts

4.1 *Ashiʿʿat-i Lamaʿāt*

Another *khaṭṭ-i ṣīnī* manuscript with content that is distinctively Sufi in nature also includes glosses in *xiaojing* and was published in facsimile (no place, no date; the Chinese preface is dated 1987). It is an Arabic translation of a Persian commentary on the famous Persian mystical tract entitled *Kitāb al-Lamaʿāt* 'Book of [divine] flares' by Fakhr ul-Dīn ʿIraqī (d. 686/1287 or later). In Northwest China, the *Lamaʿāt* has traditionally been studied together with a specific derivative text called the *Ashiʿʿat-i Lamaʿāt* 'Gleams from the flares' by Nūr ul-Dīn Jāmī (d. 898/1492),[33] and notwithstanding the fact that the (bilingual) title on the book cover is '*Kitāb al-Lamaʿāt bi-lisān ʿarabī* 'Book of [divine] flares in the Arabic language' – *Guangdian xueli* 光電學理', the text is none other than that of Jāmī's commentary, albeit in an Arabic translation. The copying of the text was completed on 2 Dhū al-Ḥijja 1410 (26 June 1990), as indicated on p. 255. Besides the Arabic translation, the manuscript also contains the poetical lines of ʿIraqī (or Jāmī) in the Persian text, which are overlined. Wide spaces have been left between the lines on a number of pages, which were meant to be filled with glosses of Chinese characters usually accompanied by *xiaojing* transcriptions.[34] In a few exceptional cases, as on pp. 65, 73, 102 and 106, characters have been added in the margins along with *xiaojing*. The broad margins have largely been filled with explanatory glosses in Arabic written obliquely against the frame of the text panel, mostly starting with *qīl*, 'it was said', *qawluhū*, 'his word' or *wa-l-murād*, 'this means'.

The poems (which are often quatrains) are quoted within the prose text in the original Persian language, followed by an Arabic translation.[35] They are also accompanied by Chinese translations written in Chinese characters and *xiaojing*. As for the format, two lines of characters and two lines of *xiaojing* are fitted in the wide space between two lines of the main text. For instance, on p. 48 (Fig. 4) line

33 Jāmī's commentary was translated under the title of *Zhaoyuan mijue* 昭元秘訣 by Po Nachi 破衲痴 (She Qiling 舍起靈), who died in 1710; cf. Ma 1999 [1985], 31; Luo 2007a, 500. We still need to find out whether or not the Chinese translations added to the Persian poems scattered throughout the text are those of the *Zhaoyuan mijue*. Zhou (2008, 78) includes a picture of the title page of the translation by She Qiling.

34 E.g. on pp. 7, 8, 9, 43–45, 46 (no Chinese characters), 48, 49, 60–62, 69, 70, 107–113 and 119.

35 At the beginning of the text (p. 9), the Arabic translation is denoted as *taʿrīb* in a gloss written above the line.

1 is a Persian verse[36] and line 2 is a Chinese translation of the last line on the previous page (= the first verse of the Persian Rubāʿī) written in characters. This line is followed by line 3 which is a *xiaojing* transcription, followed by line 4 written in characters, providing translation for (Persian) line 1. Line 4 is followed by line 5, which again is in *xiaojing*. The first line on p. 48 is this:

خواهم که بخوانمش بصد نام اما او برتر از آن است که کنجد در نام

'I would like to call him by a hundred names. He is too sublime to fit into a name, though'.

Lines 4 and 5, which are awkward Chinese translations of the above verses, run as follows:

ڤ كِيِك تِيًا خَوْ چِه ظٌّ تِيُوَ وُ نَا تِيُوًا قَوْ بُو كَه تِيُو

[37]求可不高玄那無 闕尊稱號千將欲

'I shall praise the honoured one in a thousand names / there is nothing as high and mysterious (as him) who is not to be found (by any human means)'.

The Arabic translation is in line 7:

تمنّيتُ أن أسمّيه بألف اسم لكّنه أعلى ممّا وسع في اسم

Thus, the structure of such complex translational technic is as follows: Persian is translated into Chinese, Chinese is transliterated in *xiaojing* which is followed by translation into Arabic.

However, not all the translations of Persian verses are consistently represented in this way: on p. 46 a quatrain quoted is followed by two lines of *xiaojing* translation, but a character version is missing.

36 This corresponds to p. 53 of the *Lamaʿāt* edition by Hadi Rastegar Moqaddm Govhari (thus transcribed on the English title page), Qum: Bustān-e Ketāb Publishers, 1390/2011.

37 The Chinese characters written from right to left may be transliterated in pinyin as follows: *yu jiang qian hao cheng zun que / wu na xuan gao bu ke qiu*, corresponding to the tentative reading of the *xiaojing* line: *wo jian qian hao chen zun que (?) / wu na xuan gao bu ke qiu*.

Besides that, *xiaojing* glosses are included between the lines, explaining Arabic expressions, as on p. 46 where the central expression *lamaʿān* in the phrase *wa-bāna lamaʿānu kamālihī*, 'the glow of His perfection has appeared', is explained as ثِيَان كِيَان قُوَانْ *xian jian guan* = *xian dian guang* 顯電光. As for the orthography, it is worth noting that the phoneme x- (as in *xian* 顯) has been transcribed as ثِ and d- (in *dian*) as كِ.

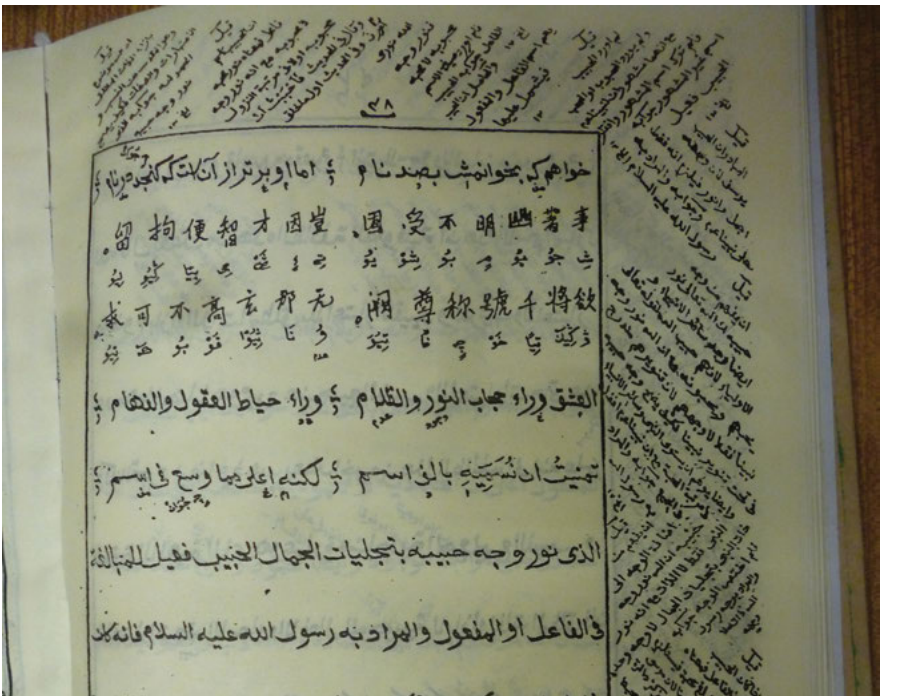


Fig. 4: The *Kitāb al-Lamaʿāt* (divine flares), facsimile manuscript, no date, no publisher's information.

4.2 *Xiaojing* glosses between the lines and in the margins: a Persian grammar

Chinese glosses written in the Arabic script have also been written in the margins and between the lines of the main text in a facsimile manuscript of the Persian grammar entitled *Minhāj* 米乃哈知, which was published recently (there is no indication of the date and place, though) and sold in a Muslim bookshop in Ping-

liang in 2014.[38] The Persian work is one of the *madrasa* books used in Northwest China and listed by Ma (1999 [1985], 29): 'The author of the *Hawā-i minhāj* is the Chinese Muslim scholar Chang Zhimei 常志美. The book has been written in Persian and gives a summary of Persian grammar. It was designed to help students in mastering the Persian language'.

The incipit of the facsimile manuscript written in a distinctive *khaṭṭ-i ṣīnī* style[39] goes بدان الهمك الله تعالي كه سخنهاء بارسي بر سه كون است, i.e. 'Know – may God inspire you – that Persian words are of three types'). There are just 12 lines to each page. The booklet contains a total of 66 pages altogether and includes a colophon which mentions the name of the scribe and date of completion. According to this, it was copied in Rajab 1070 (هزار و هفاتادم) (= March–April 1660) by Muḥammad ibn al-Ḥakīm al-Zīnamī (?) al-Shandunī al-Ṣīnī, الزينمي الشندوني الصيني, i.e. a scribe originating from Jinan in Shandong province. Beneath the place name, 'al-Zaynamī', the explanatory word 'Zinan-fu' has been written in a minuscule script with a vocalisation which is standard for *xiaojing*, i.e. doubling of vowel signs within the *rasm* of a word – in this case, doubled *fatḥa* above *nūn*, namely زِنَّافُ (Zinan-fu = 'Jinan prefecture).

The glosses in this manuscript are mostly in Persian, but sometimes in Arabic and occasionally in Chinese, which is written in the same style of the Arabic script as the glosses in the two other languages. As a rule, the explanations have been added *above* the word that is explained, as in the first sentence on p. 1 where it says that Persian words are of three types, namely nouns, verbs and adjectives (*ism*, *fiʿl* and *ḥarf*). Two examples are given for each type. The category of the verb is exemplified by *zad wa gasht*, 'he hit and turned away', the first example being translated by Chinese دَا (= *da* 打 'to hit'). As elsewhere (a mark of orthographic standardisation), *fatḥa* has been added to the letter *dāl*, although it is followed by *alif* and is therefore unnecessary phonetically. A *xiaojing* gloss written beneath the line can be seen on p. 54: in the overlined phrase توانم كه ترا از ناداني رها كنم 'I can free you from your ignorance', the noun *nādānī*, 'ignorance', is explained as بُ جِ دَوْ *bu zhi dao*, '[I] don't know' (or: 'not knowing' which is equal to 'ignorance'). It may be noted that in this transcription, the letter *jīm*, although occurring in isolation, is written in the initial and not the final form, which is contrary to the conventions of the Arabic script, but standard in *xiaojing* orthography.

38 The book cover only contains the four characters (*Mi nai ha zhi*), which transcribe the Arabic/Persian title.

39 The tapering ends of the letters *waw* and *rāʾ* are conspicuous here. Besides this, no distinction has been made between the other consonants (*bāʾ*, *kāf* etc.) and their equivalents in ʿAjamī.

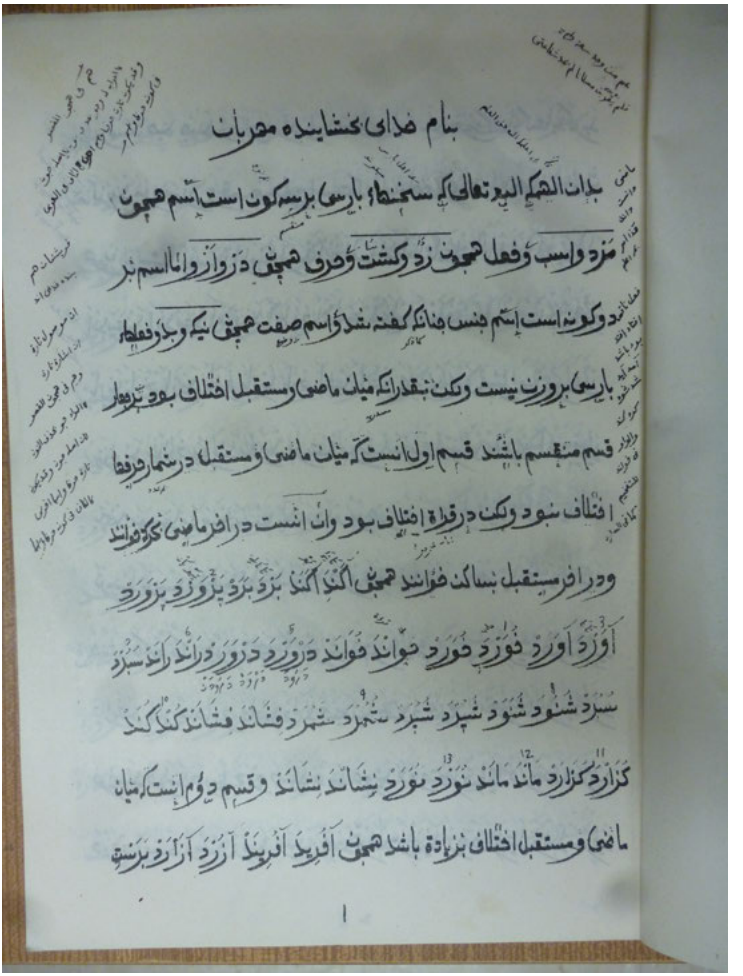


Fig. 5: The *Minhāj* (a manual on Persian grammar), facsimile manuscript, no date, no publisher's information.

4.3 Independent *xiaojing* texts: Ma Tianmin's catechisms

4.3.1 The catechism *Taḥāwur al-kalām fī masā'il al-ṣalāt wa-l-ṣiyām*

There are also manuscripts where *xiaojing* does not appear in the form of glosses in the margins or elsewhere, but where it figures in its own right. This is the case with some texts written by the 20th-century author Ma Tianmin 馬天民 Ibn

Maʿṣūm ʿAbd al-Raḥmān, as in the *Taḥāwur al-kalām fī masāʾil al-ṣalāt wa-l-ṣiyām*, a catechism on issues of ritual prayer and fasting.[40] In the facsimile manuscript published in 1406 (= 1985–6) in Baozi Mosque 堡子清真寺 in Linxia 臨夏 – the title on both covers of the booklet is only in Chinese and runs *Yisilanjiao libai fengzhai wenda* 伊斯蘭教禮拜封齋問答, 'Questions and Answers on Islamic Prayer and Fasting' – six lines of Arabic text are accompanied by ten lines of a corresponding Chinese text written in the Arabic script covering about two thirds of each page. Rather than constituting a literal translation of the Arabic text, the *xiaojing* [41] text is an independent version of Ma's catechism, which is made up of questions and answers. Unlike the Arabic text, the *xiaojing* text is written in a slightly larger script with vocalisation and punctuation marks (commas, colons and full stops), which bears witness to the recent (20th-century) date of the copy. The recurrent expressions *wun* وُ and *da* ذَا, i.e. *wen* 問, *da* 答 'question/answer', have been put in brackets.

Starting at the opposite end of the booklet (pp. 2–22), there is a Chinese version of the text printed (in stark contrast to the manuscript character of the Arabic/*xiaojing* text!) in Chinese characters, which can be used as a key to understanding the *xiaojing* text. The Chinese translation written in characters was made in 1993 by one Ma Xiqing 馬希慶 in or near Baozi Mosque (p. 22).

---

40 Another work – an Islamic catechism entitled *Taḥāwur al-kalām fī ʿaqāʾid al-islām*, 'On the articles of Islamic belief', which was written in 1952 – was composed in the same format (and published in Linxia as a facsimile manuscript: Baozi Mosque, Shawwāl 2004 [reprint]). I have included a short discussion of the text in Sobieroj 2016b, 61–62. The introduction to the catechism states on p. 3 that the author has also written various other works: (i) only in Arabic (five Chinese [!] titles are listed), (ii) in Arabic accompanied by a *xiaojing* translation (the catechism and two more Chinese titles) as well as (iii) some Chinese texts written exclusively in an Arabic transcription (five titles). The author's *Tuḥfat al-ṭullāb fī maʿrifat al-ṣiyagh wa-l-iʿrāb* on grammar seems to represent the type of works written in Arabic without *xiaojing*. This impression emerges from the study of the reprint of a 41-page booklet first published in 1376/1956 in Hanjiasi Mosque in Linxia ('[...] printed again after completion of a Chinese trans.' [p. 1]; Baozi Mosque 1410/1990). The grammar designated on the title page, *Qawāʿid al-nahw - alaboyu jianming yufa jingtang yu* 阿拉伯語簡明語法經堂語, 'Succinct Arabic Grammar (*madrasa* language)', was designed to be used by beginners after mastering morphology (*ʿilm al-ṣarf)* at the *madrasa al-thanawiyya al-ʿarabiyya* (p. 1 or second title page) and it includes the printed Arabic text with Chinese translations printed in characters on the same line. This seems to have become the standard format of modern editions of Arabic grammar texts such as the *Kāfiya*, published in 2005 in Zhoupo 周坡 Eastern Mosque in Tai´an City 泰安市, Shandong, for the use of *madrasa* students (*manlia jing jianben* 滿倆經簡本).

41 The term '*xiaojing*' is used in the Chinese preface of the *Taḥāwur al-kalām fī ʿaqāʾid al-islām* and is probably also referred to in the Arabic introduction, viz., حرر ها الحقير بتْيَوْكِنْ (a minuscule letter *ṭāʾ* ط is placed above *kāf* ك of تْيَوْكِنْ – a very uncommon orthographical addition).

The section on ablution, headed *ābdast* – *xiaodin* (= *xiao jìng* 小淨 'minor ablution'), includes the following *xiaojing* passage: (وُ) ثيو دٍ دِ فرض شيِ شما ، دُوَ شو سي كياً : سي نياً ، سي ليان شو ، دَو ليان جو، مو طِو دِ سيِ فْ جـ ءِ ، سي ليان كيو، دو الخ (دا) . An attempt to transcribe the Chinese passage above using a system of transcription that closely depends on Pinyin but also takes account of the added vowel signs yields the following: (*Wun) xiaodin di fard shi shima, duwa shao? (Da) si jian: si nian (!), si lian shou, dao lian zhou, mu tou di si fun zhi i, si lian jiou, dao lian huai gu* 'Q: What duties are involved in the minor ablution? How many are there? A: There are four duties: to wash your face, to wash both your hands up to the elbow, to wipe one of the four sides of your head [and] to wash both of your feet up to the ankles' (Fig. 6).

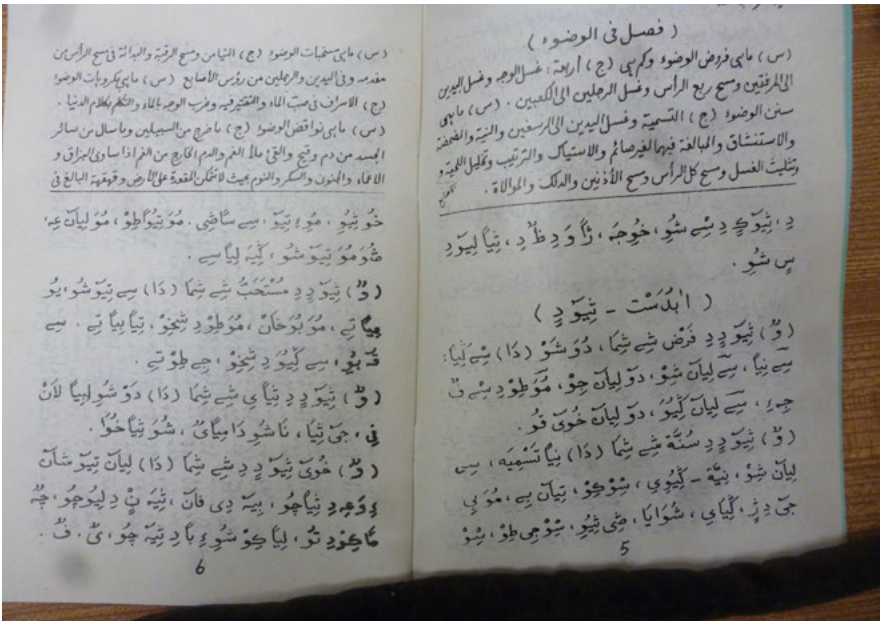


Fig. 6: The catechism of Ma Tianmin *Taḥāwur al-kalām fī ʿaqāʾid al-islām* (On the articles of Islamic belief).

In the above section, the author does not explicitly say how to pronounce the letter *ḍād* in the Arabic noun *farḍ*, 'duty'. In the Hanzi translation, *farḍ* is rendered by three characters with the phonetic value *folezuo* 佛勒作, which may be interpreted in that the noun is pronounced closer to the Persian pronunciation than to the Arabic. Mandarin *lian* for 'face' is transcribed as *nian*. Mandarin

jing for 'ablution' is rendered as *din* دِ. No distinction is made in the transcription between the consonants x- and s- in *xi*, 'wash', and *si*, 'four', both of which are rendered by the letter *sīn* – in standard Chinese, the two sounds are clearly distinguished in terms of their pronunciation. The consonant j- in *jian*, 'condition', for instance, is transcribed by the letter ك with three dots above it and one *kasra* beneath it. The letter x- in *xiaodin* (= *xiaojing*, 'minor ablution') is rendered as ڭِ with *kasra*, while j- in *jian* is rendered as *kāf* with three dots, which seems to be inconsistent; in *jing* in the expression *xiaojing* (= *xiaodin*), the initial consonant is written as d-. Given that the initial /j/ in MC is a merger of Middle Chinese *kj/gj and *dz/ts, it is plausible that this conservative dialect spoken in Linxia and represented in the manuscript has preserved some features of Middle Chinese.

The author of the credo must have spoken this dialect because he was active as a *madrasa* teacher in Linxia. According to the Chinese introduction of the booklet on the credo, the author was teaching children at Hanjiasi 韓家寺 Mosque in Linxia لينشيا, where he was known as 'Jingkou si shifu' 井口四師傅 ('the fourth master of a place called Jingkou'; الدنكوي).

4.3.2 Xining version of Ma Tianmin's catechism (*ʿAqīda*)

Another version of Ma Tianmin's catechism paralleling the Linxia version[42] was published at Dongguan 東関 Mosque[43] in Xining 西寧, the capital of Qinghai province,[44] which borders on Gansu. The Xining version is entitled *ʿAqīdat al-is-*

42 Cf. n. 41.

43 The largest mosque in Xining and religious centre for at least 100,000 Muslims living in the town, it was founded in the Hongwu era (1328–98) of the Ming dynasty. The mosque was destroyed in reaction to an armed uprising by Hui and Salar Muslims led by the Jahrī Ahong Su Sishisan 蘇四十三 and others against the Qing in 1781. It was rebuilt later, however, and was closed temporarily during the 'Cultural Revolution', then re-opened in 1979 (cf. Wu 1995, 459–460).

44 The undated booklet, which I acquired in Xining around 2010, has a photograph of the *Dongguan qingzhen dasi xuanli ta* 東関清眞大寺宣禮塔, or 'Great Eastern Bar Mosque with its minaret' on the cover, which may be taken as an indication of the place of publication. The same mosque (and picture) also appears as the place of publication of an undated booklet on prayer (*al-Adʿiya al-lāzima wa-awrād al-ṣalāh / Changyong duwa libai nianci* 常用都哇禮拜念詞), which contains the printed Arabic text of the prayers, but – contrary to what one should expect in light of analogous works – without any *xiaojing*. A transliteration of the text is provided in Chinese characters printed beneath the Arabic words, but according to the phonetic value of the individual characters (= Chinese 'Ajami!) and without offering any clues as to the meaning. Occasionally an explanation in Arabic is provided in Chinese characters, explaining how and when to apply a prayer text, as on p. 23: 'a prayer to recite while washing the right foot' (*xi you jiao nian de duwa;*

lām Yisilan xinyang wenda 伊斯蘭信仰問答, 'Questions and Answers on Islamic Belief', and features a text in *xiaojing* transcription with simplified Chinese characters written between the lines in the same direction as the Arabic script. In contrast to the manuscript nature of the *risāla*, the Arabic is printed (!) in a bold script, whereas the Chinese characters have been added in smaller types (seemingly also printed), from which it follows that the reader's eye is immediately attracted to the Arabic text rather than the Chinese. Unlike the Linxia publication of Ma Tianmin's text, the Xining *ʿAqīda* neither includes the Arabic words of the catechism nor a Mandarin Chinese version printed as a separate text. Comparing the interlinear Chinese characters of the Xining version (XN) with the Chinese version in the Linxia publication (LX), it emerges that there are discrepancies in the words and even in the structure of the sentences. In the section on disbelief, *kufr* is defined as negating what is obligatory to believe in and to look down upon what is exalted by the law. Whereas the Chinese has the words *fouren bixu xinyang de* 否認必須信仰的, 'to negate what one must believe in', the interlinear text of XN (p. 55) and the *xiaojing* text of LX (p. 40) written on the lower part of the pages has *yinmei guixin ta biding de nage* 隱昧歸信它必定的那個 or يِ مُوِ قوِ سٍ تَا بِي دٍ دِ نَقْ, the main differences being in terms of syntax, and the lexical variants *fouren/yinmei* (مُوِ of 隱昧, as indicated by the vowel sign, may have to be pronounced with a diphthong analogous to *gui* 歸); this results in two texts that appear to be almost completely different.

The above differences may be due to the effort made by the editors of the *ʿAqīda* to give Chinese character equivalents to the vernacular of Xining in the one publication, whereas the translation in the Linxia edition is Mandarin and does not give any equivalents of the *xiaojing* transliteration of the spoken language (of the Linxia region).

Then there are differences in the two publications relating to the orthography rather than to pronunciation. Here are some examples of the orthographical differences: *ta shi* 它是, i.e. 'it [*kufr*] is', written as تا شى in LX and تا ش in XN; *qingshi zai she-ri-er libian zunzhong* 輕視在舍日兒裏邊尊重, 'it is to belittle what is honoured in the law', is written as تٍ شِى ذَى شرع لبيًا ظٌ جهٌ in LX and كٍ ش زَى شرع لبيًا ظٌ جٌ in XN; *nage* 那個, 'that one', is written نَقَه in LX and نَ قَ in XN; *zhenzhu* 真主, 'true lord', i.e. God, is written جٍه جُ in LX and جٍ جُو in XN. The differences in spelling are indicative of a lack of shared standards (or of a discrepancy between the standard language and the dialect) despite the mutual geographical proximity of Xining and Linxia.

---

the prayer text in characters runs: *anla hunmai sanbiti gaidaimanye* 安拉混麥 etc., i.e. *Allāhumma thabbit qadamayya* 'God, make my feet firm!').

4.4 *Xiaojing* translations following in the same line

4.4.1 *K. Fiqh Kaydānī* and two related texts

In a facsimile copy of the famous Arabic work on the Islamic ritual prayer composed by Luṭfallāh al-Kaydānī (lived *c.*750/1349; GAL, vol. 2: 253) and hence entitled *K. Fiqh Kaydānī* (pp. 43–108), the *xiaojing* translations are not added as glosses, but follow in the line immediately after the Arabic wording, which they translate into Chinese. The *xiaojing* is written in the same *khaṭṭ-i ṣīnī* style of the Arabic script as the words of the Arabic text, which is overlined as a rule, but it usually takes up 2–3 times the space of the Arabic expressions translated. The *xiaojing* is vocalised, and the individual sentences are mostly introduced by the expression *ta yao [shuo]de shi* or *ye shi*, 'he means to say (also)'.[45]

The text copied in 1316/1898 – the colophon on p. 108 says 'the translation of this book was completed in Rabīʿ II 1316/Aug.–Sept. 1898' – forms the second of three parts all written by the same hand and in the same format with Arabic and *xiaojing* following in the same line, and was published on 30 September 2008 (no place mentioned) as a booklet bearing a bilingual title on the cover page: *Kaydān Kitāb* and 小經開達尼 (*Xiaojing Kaidani*).

The third text in the collection (on pp. 109–132) is the well-known and much-copied tract – in the Ottoman lands – on the condition of the ritual prayer *Shurūṭ al-ṣalāh* (cf. Sobieroj 2010, no. 191, part 3), while the first text (pp. 2–42) is a less well-known anonymous text on dogma and ethics whose title is given in the preface as *Risāla tataʿallaqu bi-bayān al-iʿtiqad wa-l-akhlāq wa-l-afʿāl (ʿalā l-tartīb wa-l-ijmāl)* ('A tract relating to the exposition of the creed, ethics and works, arranged logically in a summarising way').

While a colophon has been added to the second text, the third one is followed by a list of contents written in Arabic and headed by the Chinese word *mulu* 目錄 for 'table of contents' and dated 30 September 2008 (again in characters).

45 An attempt to transcribe the beginning of the *xiaojing* text yields the following: *ta yao de shi quan zan yu zhongshi weidu tiaoyang shiba tian yang khalq de naga zhu de* (他要的是全讚于眾世唯獨調養十八天養 etc.) [...] (الحمد لله رب العالمين تايودِ شِ ، كُواً زًا ءِ ظٌ شِ وِي دُو تيَوْ يَانْ شِبَا تيًا يَانْ ، خَلَقِ دِ نَفَه جوُ دِ والعاقبة للمتقين). Whereas in two accessible Mandarin translations of the Qur'anic expression *al-ʿālamīn* (the worlds) occurring in Sura 1:2, for instance, *yuzhou* 宇宙 or *quan shijie* 全世界 is used, the *xiaojing* employs the Arabic noun *khalq* plus prefixed *shiba tian* 十八天, i.e. 'the one lord who nurtures' the eighteen heavens – the latter classification being common in Buddhism.

4.4.2 *Xiaojing* Qur'an translation

The format of *xiaojing* translations following immediately after a reference text in the same line as has been chosen in the above Chinese edition of the *Fiqh Kaydānī* has also been adopted for a full translation of the Qur'an.[46] A combination of printed Arabic texts and handwritten *xiaojing* translations that is reminiscent of an incunabulum is inscribed in the spaces between the individual Qur'anic verses initially left empty. This work was published on *ǧumʿa* (Friday), 23 August 2002 (no place mentioned). The Arabic text in the publication is a transcription of the readings of *Ḥafṣ ʿan ʿĀṣim*,[47] which corresponds to the *riwāya* of the official Cairo edition from 1924. The Qur'anic verses are overlined and separated from the translations by verse markers with inscribed Arabic numbers. The Chinese translations are written in a fully vocalised Arabic script which follows *xiaojing* conventions as described in the introductory section. The language of the translation is a Chinese vernacular, which emerges from the pronunciation of individual words as well as from the syntax. *Sūrat al-ikhlāṣ*, one of the shorter chapters at the end of the *muṣḥaf* (no. 113), will be looked at by way of example. The *xiaojing* transcription of the first and the last verses (v. 1, 3–4) runs as follows:

اَي جِـ شِـ آ نِ شُوَآو تا شِـ الله دُو ءِ دِ [...]،[48]

تَا مَوْ يُـ سِ زِ ، زِ مَوْ يُـ سِ تَا ، وُ ءِ قَ ُزِ شِـ ، ءِ وِ تَا جَ بُ شِـ ءِ بَا دِ

The unfamiliar orthography of what is pronounced as *shuo* 說 'say' in Mandarin is indicative of dialect,[49] as is *sin* for *sheng* 生 'beget' [children] (or for *xing* 性 'gender, sex'?). The negation transcribed as *maw you* in *xiaojing* in v. 3, finally, is a dialect variant of Mandarin *meiyou* 沒有.[50]

46 A copy of this work was offered for sale in 2014 in a small mosque bookshop in the centre of Wuzhong City.

47 I.e. the readings of ʿĀṣim (d. 744) as transmitted by his student Ḥafṣ.

48 An attempt to transpose the above in pinyin yields the following: *ay zhi sheng a ni shuo ta shi Allah duyi de* 'Prophet, say: he is God, the one' *[...] ta mei you xing ren / ren mei you xing ta / wu yige ren shi / yin wei ta zhe bu shi yi ban de*.

49 On the dialects of Northwest China, see Yang 1996, 3.

50 With the aim of highlighting the specificities of the Chinese *madrasa* language (*jingtang yu* 經堂語) of Ningxia/Gansu, Ding Shiren 丁士仁 (Ding 2008, 48–53) contrasts a *jingtang yu* translation of the first five verses of Sūrat al-Baqara which he has transcribed in Chinese characters with a modern Chinese Qur'an translation (by Ma Jian 馬堅) of the verses in question (reproduced in Zhou 2008, 56–57).

4.5 Arabic-Chinese glossaries

Another type of Islamic manuscripts contains Arabic-Chinese glossaries into which *xiaojing* transcriptions have been incorporated. These were compiled to help readers understand a number of texts traditionally taught at the *madrasa*s.[51] The reference texts to which the glossaries refer that I was able to consult are (1) Saʿd al-Dīn al-Taftāzānī (d. 791/1389; GAL, vol. 2: 278–280), *Mukhtaṣar al-maʿānī*, on rhetorics, and (2) 'Mullā ʿIṣām al-Dīn' [al-Isfarāʾinī; Sobieroj 2010: no. 185] (*Manlia zidian* 滿俐字典) – which, in fact, is Jāmī's *Sharḥ* – on the *Kāfiya* of Ibn al-Ḥājib on grammar. I found manuscripts of two different glossaries relating to Mullā ʿIṣām al-Dīn.

4.5.1 *Lughāt Mukhtaṣar al-maʿānī*

The beginning of the glossary entitled *Lughāt Mukhtaṣar al-maʿānī al-mutarjama bi-lisān al-ṣīnī* 伯亞尼字典 (*Boyani zidian*, 'Dictionary of the Bayān'), six lines of which are included in the copy, has been written on the second page in Chinese characters and *xiaojing* . The incipit runs: نحمدك يا من شرح صدورنا لتلخيص البيان, 'we praise you who have enabled us [lit. have widened our chest(s)/breast(s)] to summarise the explication'. The Arabic words of the incipit are written with blank spaces separating them, which are filled with translations (*women zanni huokai xiongtang zhaijian mingyan* 我們讚你辖开胸膛摘簡明言, 'we praise you for opening up [our] chest/breast so that we may select words of wisdom' etc.'). The *xiaojing* transcriptions with vocalisation, on the other hand, are written above the Arabic expressions, i.e. وَمٌ زَانِي خُوَ كَيْ ثِيٌ تَان جَي كِيَاً مٍ ياً. The three parts, Arabic, characters and *xiaojing*, have all been written by one and the same hand. Nunation is added at the end of some words, as in *wamun* (= *women*, 'we') and in *thiun* (= *xiong*, 'breast/chest'; the consonant x- is transcribed by ـثـ).

The glossary itself starts on the second half of this page. It is divided into sections which follow the structure of the reference text and are headed by captions such as *muqaddima*, *ʿilm al-maʿānī*, and *aḥwāl al-isnād al-khabarī*. The format is four columns in which the individual lemmas are overlined and followed by Chinese translations in characters and *xiaojing* (the texts written in the two scripts are exactly the same). The first lemma, *al-Taftāzānī*, is explained in Arabic as [derived from] a place name (*ism mauḍiʿ*, without a transcription). The second is *al-faqīr*, 'the poor', followed by a Chinese translation in characters (written

51 Zhou (2008, 58) includes a picture of a manuscript folio of an Arabic–*xiaojing* glossary.

from right to left: 求濟 *qiuji*) and *xiaojing*: كِيُو كْي. Thirdly, *al-ghanī*, 'the rich', 無求濟 كِيُو كْي ءُ, etc. The lemmas are not arranged alphabetically, but according to the order of their occurrence in the text.

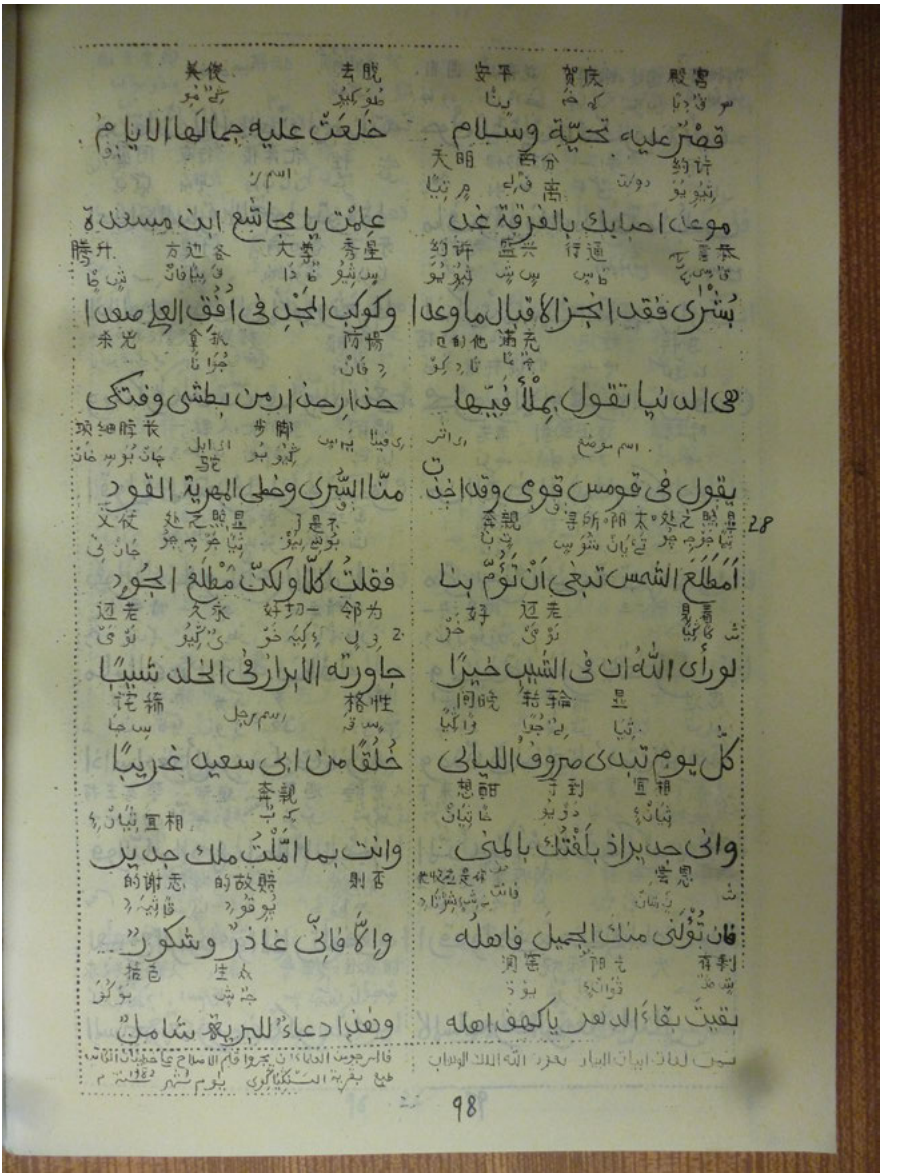


Fig. 7: The *Bayan zidian* (Dictionary of the Bayān).

The glossary ends on p. 70 with the date 28 March 1983 and adds a list of the poetic verses taken from a different manuscript, which are included in the *Bayān* with translations in characters and *xiaojing* written above the lines. The 26 pages of verses were published (*ṭubiʿa* [lit.: printed]) in 1982 at a place whose *xiaojing* transcription may be construed as the (topographically meaningful) name *Sanjiaqiao*, i.e. 'three-family bridge'. The last verse, only five words of which are represented in Chinese in the manuscript (marked by an asterisk here), runs: *baqīta* baqāʾa l-dahri* yā kahfa* ahlihī / wa-hādhā duʿāʾun li-l-bariyyati* shāmilun**, 'May you remain as long as the world/time remains, cave of his people / This is a prayer which encompasses the whole creation'. The five words translated as well as transcribed are 剩存 شٍ ظٌ, 光陰 ءٍ قوانْ, 匈洞 يو دٌ, 衆生 شٍ جهٌ and 包括 بَوُ كُوَ (Fig. 7).

Owing to the atomistic nature of the *xiaojing* glosses, which only explain individual words, the tendency to translate text into a Chinese dialect seems to have been checked here. This contrasts with the passage-wise dialectal *xiaojing* translation, as in the texts authored by Ma Tianmin, which were discussed in section 4.3.

4.5.2 *Lughāt Mullā ʿIṣām al-Dīn*

4.5.2.1 *Qaryat al-Sanjiaqiao* version

Like the format of the *Lughāt Mukhtaṣar al-maʿānī*, the version of the *Lughāt Mullā ʿIṣām al-Dīn*, which was published in facsimile form at the same place, i.e. *Qaryat al-Sanjiaqiao*, starts with an incipit of the text, only three lines of which are provided. The interspaces between the Arabic words are filled with Chinese characters, whose *xiaojing* transcriptions are written above (most of) the Arabic expressions, i.e. الحمد لوليه والصلوة على نبيه وعلى آله واصحابه المتادبين بادابه, 'Praise to Him who deserves it and prayers for His prophet and his (i.e. the latter's) family and companions who have trained themselves in his manners', 全讚唯獨應受他的主 كِيواً زاً وِي دو ءٍ شِو تَا دِ [...] 一切隨同他的人 一切受學禮他的人 將他的一切禮儀者, جو [...] ءٍ كِيَه صُو طً تَادِ زٍ ءٍ كِيَه شِوْ ثِيُوَ لي تَى دِ زٍ كْيانْ تَا دِ ءٍ كِيَه لِي ئِي جه.

In contrast to the *Lughāt Mukhtaṣar*, the lemmas of this glossary are not arranged in columns, but are written from right to left in horizontal lines. The Arabic expressions are followed by translations in characters, above which *xiaojing* is written. The booklet consists of 26 unnumbered pages of 17 lines ending with a colophon, which states the date of copying (20.4.1402/10.6.1983) and place of publication. In the gap between the explicit and colophon, the compiler has inserted a request in Arabic asking the scholars (*al-ʿulamāʾ*; i.e. the competent readers) to correct any mistakes he may have made in writing and translating.

The system of transcription seems coherent internally, although it partly deviates from other authors' systems. The scribe conventionally differenti-

ates between the consonants q- (as in *quan* 全), which he transcribes as ڭ, and between j- (as in *jiang* 將, transcribed by ڨ); the [-ŋ] sound of Mandarin Chinese, consistently rendered by n (*jiang* = *jian* ڮيان); the consonant r- (as in *ren* 人), which is transcribed by *rāʾ* with three dots above it (ڑ – a graph which corresponds to Persian *žay*); and the consonant x- (as in *xue* ڛيۆ學), which goes back to a h-sound and is rendered by the letter *thāʾ* (Fig. 8).

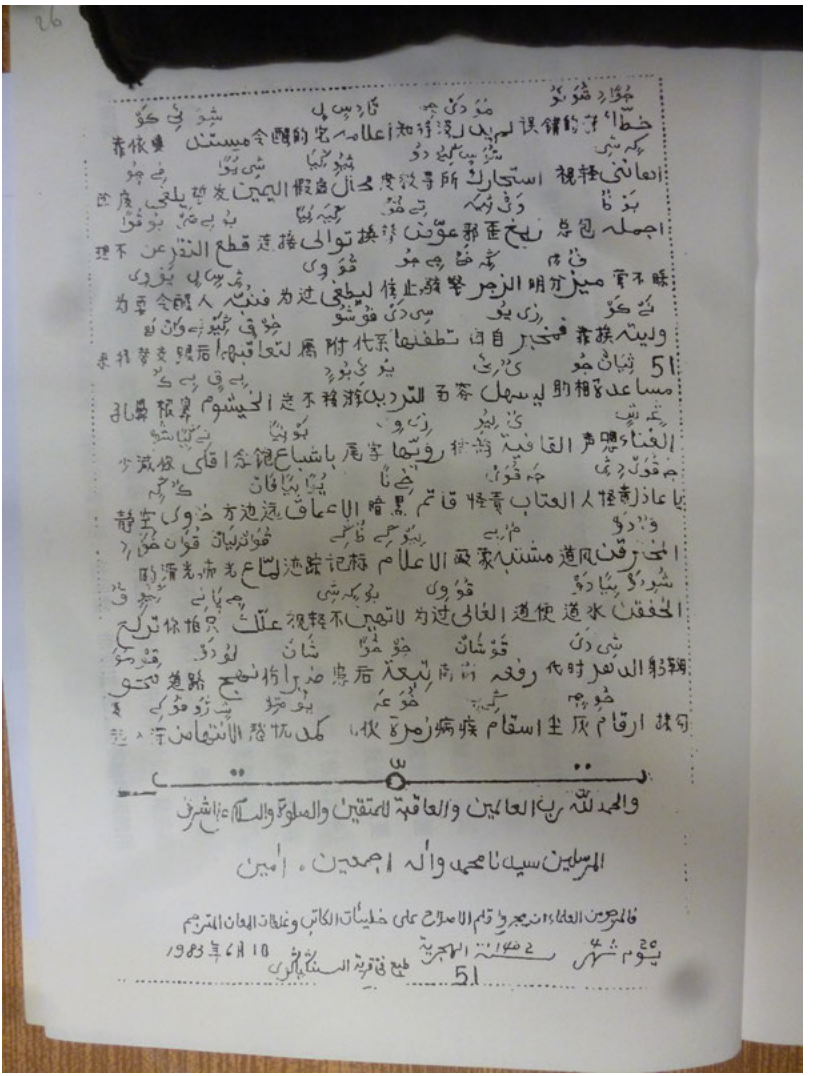


Fig. 8: The *Lughāt Mullā ʿIṣām al-Dīn* (A glossary on Islamic faith).

4.5.2.2 Kaifeng version

The second facsimile manuscript of the *Lughāt al-Mullā* (*Manlia zidian*) was published in 1356/1937–38, or the 26th year of the Republic of China جُنْهُوَا مِنْقُوَيْ 中華民國 ('Chinese republic' with the phoneme h / kh in *hua* [China] 'irregularly' written as ـھـ <h>, and final velar nasal [-ŋ] [in *zhong*], 'regularly' written as <n>) in the city of Kaifeng الكَيْفُنْني in Henan province under the auspices of Shams al-Din جِنْزِيَانِني Zhenjiang, and it was purchased in Pingliang in 2014.

In this booklet, the lemmas are listed in two columns (mostly as two keywords with their translations, which have been fitted into one line in each column) and they are arranged in order of their occurrence in the glossary. The Arabic expressions written in a bold *khaṭṭ-i ṣīnī* style (fat feet, tapering upstrokes)[52] are followed by *xiaojing* without any characters (there are, exceptionally, a few characters written above the *xiaojing* words), but characters re-emerge in the Chinese-language colophon where the date of publication is stated as well as the name of the editor (*bianjizhe*), Wang Hongxiang from Zhenjiang 鎮江王洪祥.[53]

The translation of the first lemma الحمد لوليه, 'praise to Him who deserves it', starts with the same words as in the *Sanjiaqiao* manuscript (plus the copula *shi* 是), which are transcribed differently, however, namely as follows: قَا زًا شِ وِ دُ ءِ الخ (in the *Sanjiaqiao* manuscript, it reads كِيواً زاً شِ وِي دو for Pinyin Mandarin: *quan zan shi wei du ying*). The difference could be attributed to the influence of the regional dialect of Henan where the text was published, although it is far from certain that there was any such influence. The text may have been transcribed in the Northwest and only been published in Kaifeng (as a reprint, *zai ban* 再版). If that were the case, it would be evidence of an extreme inconsistency in writing *xiaojing* and is therefore highly unlikely. However, there is also a chronological aspect that needs to be considered: the Kaifeng publication is half a century older than the modern Sanjiaqiao manuscript, which is also apparent in the usage of Persian words explaining some lemmas that are not found in the modern manuscript (e.g. *zamān*, 'time', for *yawm*, p. 77; *siyāh*, 'black', for *as-sawād*, *sebīd* (sic!), 'white', for *al-bayāḍ*, p. 78; and *ʿeshq*, 'passionate love', for *al-hawā*, p. 70).

52 For these characteristics see Sobieroj 2014, 103–106.

53 The last page opposite the explicit contains a pricelist for '[Islamic] canonical books' (*jingshu* 經書) which were available for sale at the Wenhua street mosque in Kaifeng. Apart from the present text, they include a glossary entitled *Lughāt al-Wiqāya* relating to *al-Wiqāya*, the textbook on Ḥanafī law.

4.6 *Munājāt*

Xiaojing is found in two places in a facsimile *khaṭṭ-i ṣīnī* manuscript which mainly includes an assembly of Qur'anic verses designed to be used as prayers for Muslims in Gansu:[54] firstly, on the title page inscribed beneath the title *Munājāt*, i.e. 'prayers of intimate discourse', and secondly, in the colophon. The Chinese words written in the Arabic script inform the reader (albeit inaccurately) about the contents of the booklet, which spans 54 pages of five lines per page. The Chinese subtitle runs سَا شِي سَا دُوَ آيت, i.e. 33 Qur'anic verses/passages (*duan* transcribes the character 段). The numbering of the *āyāt* is written in Persian (*seyyom āyat*, *čahārom āyat*, etc.) in a different script. This indicates that the Persian-language glosses in *khaṭṭ-i ṣīnī* manuscripts serve a similar explanatory or text-structuring function to *xiaojing*. As a matter of fact, the number of Qur'anic verses included in the *Munājāt* far exceeds 33, hence the noun *duan* should be understood as denoting passages or even chapters – the first of which is Sūrat al-Fātiḥa – rather than verses.

The *khaṭṭ-i ṣīnī* style in which the Qur'anic verses have been written has a calligraphic quality and features 'fat feet and slender ankles', which are refinements usually found in copies of the sacred Islamic texts in China. The prayer beginning with *Allāhumma ´ftaḥ lanā abwāb al-raḥma*, 'God, open to us the gates of mercy!', which follows Sura 114, i.e. the last of the Qur'an quotations, in contrast, is written in a different, unpretentious script which does not feature any of the above characteristics.

The colophon on the last page includes the name of the scribe, Maḥmūd ibn al-Khabīr ibn al-Shaykh al-Rasūl al-Watouwī. On p. 55, the main text borders on two rectangular panels which are each filled with one line of *xiaojing* written vertically from top to bottom in the same style as the Persian text dividers, and they contain a note on the person who is to recite the 33rd section.

5 Conclusion

The influence of Arabic manuscript culture has fostered processes of standardisation in *khaṭṭ-i ṣīnī* manuscripts. This can be seen in the adoption of familiar page-

54 The prayer booklet was sold in 2012, more or less illegally, in one of the small shops on the main road in Guanghe 廣河, a village and centre of Islamic culture in Gansu situated south of Lanzhou. Like many other publications, it lay hidden under some prayer scarves, carpets and other day-to-day objects used by Muslims.

layout formats by the Chinese Muslim scribes and artisans. Wide spaces have been left between the lines of the main text, which is also the case in Ottoman manuscripts where Turkish or Persian glosses have been added to explain certain Arabic expressions. And yet the division of the page into a section including an Arabic reference text and a (larger) section reserved for translation into a non-Arabic language, as is the case in Ma Tianmin's *xiaojing* catechisms, is a format hardly ever seen in manuscripts originating in the core Islamic countries.

Among the many different kinds of page layout and combinations of languages and scripts adopted in the *khaṭṭ-i ṣīnī* manuscripts including *xiaojing* texts (or glosses), we can find five main types:

1. *xiaojing* in conjunction with translations in Chinese characters added between lines of Arabic poetry quotations (*Lamaʿāt*);
2. translations written obliquely as isolated glosses, up against selected expressions of a reference text (the grammar book called *Minhāj*);
3. *xiaojing* figuring on a par with an Arabic reference text by covering the larger part of the page (*Taḥāwur al-kalām)*;
4. substitution of Arabic text by *xiaojing*, which is accompanied by Chinese characters written between the lines that reproduce the *xiaojing* vernacular (ʿ*Aqīda*); and
5. a widely used format adopted in Qur'an translations (and other texts), which features Arabic followed by *xiaojing* translations in the same line.

Three different types can be noticed in Arabic-Chinese glossaries: *xiaojing* translations written above the word to be explained (similar to type 2) or inscribed in the line following the lemmas which are listed in two or more columns (similar to type 5), or *xiaojing* translations additionally accompanied by characters, which is what the scribe who penned a copy of the *Lughat Mukhtaṣar al-maʿānī* did.

On a micro-level, it has been observed that *xiaojing* poems included in the Arabic manuscripts imitate the classical models of Arabic poetry quotation in terms of their conventions: the Chinese poems are introduced by the conventional Arabic formula *anshada*, 'he recited', and the verses that are fully vocalised are separated by graphical elements such as circles inserted individually or in clusters (see above pp. 187 and 192).

Classical Arabic has influenced the Chinese vernacular inscribed in *khaṭṭ-i ṣīnī* manuscripts directly (one example being the use of the word *khalq* in a *xiaojing* catechism where the translation written in characters offers a Chinese equivalent) and through mediation of Persian, which Hui Muslims consider their

ancestral language.[55] The contribution of the Persian writing system to Sino-Arabic manuscript culture can be seen in many ways: letters such as *čīm* and *žay* used in *xiaojing* to transcribe the Chinese consonant r- as in *ren* 人 ['man'] have been borrowed from the Persian alphabet. Persian literary works are studied at the *madrasa*s (although nowadays the influence of Persian is receding due to the increasing weight of the Arabic language) and they are quoted in the texts introduced in this study. The oldest of the group of texts narrating the hagiography of Ma Mingxin, the *Rashaḥāt*, was partly written in Persian. *Xiaojing* translations have not only been added to Arabic texts, but also to works in the Persian language such as the *Minhāj*. Persian insertions have even served to structure the text of some Sino-Arabic works (e.g. *Munājāt)* and, finally, Persian words have been used to explain Arabic lemmas in the manuscripts of the glossaries.

Traditional cultural practices like dictation of text by the *mudarris* during a lesson were followed by the Chinese *madrasa*s, but the amenability of *xiaojing* to misinterpretation (as shown in the funny story of *fiqh* related among Hui students about the confusion of wine and chicken oil due to shortcomings in the Arabic transcription of Chinese) may have impeded its full adoption. The hybrid forms of *xiaojing* graphemes and characters written in manuscripts by students (probably during lectures) also indicates that people found it awkward to write *xiaojing*. Chinese students were constantly confronted with having to make a choice between characters and *xiaojing*, and the aversion to writing characters because of religious reservations ascribed to earlier generations of Muslims had long been overcome, hence the presence of characters in the margins of *khaṭṭ-i ṣīnī* manuscripts and the ongoing publication of Chinese translations of Arabic texts.

As for the role of religion, specific sacred Islamic texts such as the Qur'an and the liturgical texts recited daily within the Jahriyya Sufi order have all played a role in the standardisation of the script adopted for literary transmission. Unlike works of *tafsīr* or biographies (even including Jahrī hagiography), these texts – *Mukhammas (Qaṣīdat al-Burda), Madāʾiḥ, Mawlid al-nabī* and *Munājāt* – have required the pious scribe to choose a calligraphically valuable, traditional *khaṭṭ-i ṣīnī* style showing features which are mostly absent from copies of manuscripts from other genres.

In the liturgical texts just mentioned, standardisation in handwriting and choice of format has been achieved in mutual dependency. The texts have all been written in the bold, hieratical variant of *khaṭṭ-i ṣīnī*, mostly in a layout of five lines a page (the two-volume *Mawlid* edition has seven lines),[56] which, at least in

55 The numerous borrowings from Arabic (and Persian) have been listed in various glossaries (e.g. Yang 1996, 87–122).

56 A manuscript of the *Kitāb al-Mawlid* on the celebration of the Prophet's birth, copied by

China, has also been the preferred format of the partial Qur'an texts (*juzʾ*) published in Yunnan province as wood or stone prints. However, the five-line format is also a natural choice because in the *Mukhammas* the text consists of units of five hemistichs. The liturgical texts are left clear of Chinese glosses despite the large empty spaces separating the lines, but publications of *Mukhammas* facsimiles with a Chinese translation printed or handwritten on opposite pages are very popular among Muslims in the region. Part of the traditional format of the *Mukhammas* is also the addition of a Persian poetical translation written in two-and-a-half lines at the bottom of each page. In a modern facsimile publication of the *Madāʾiḥ* (Hong Kong 2005), however, a Chinese translation written in characters with a five-line stanza is provided at the bottom of each page.

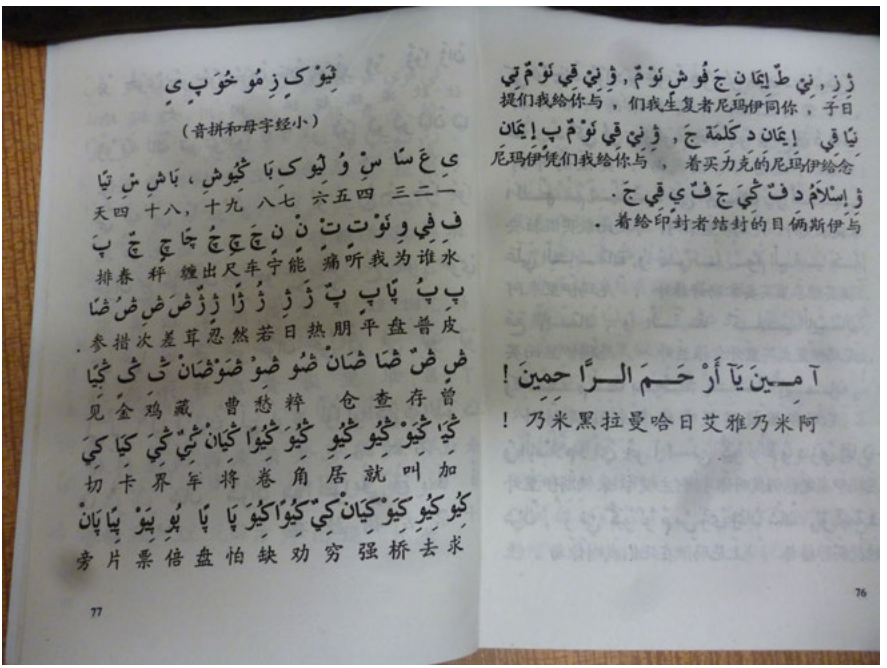


Fig. 9: A printed text of the *ʿAqīdat al-islām* which includes the transliteration tables.

the above-mentioned Maḥmūd ibn al-Khabīr and published in facsimile in 1999. It is stated on p. 127 of the *tiwen* 提文 ('treatise'?) part of the book that the *Kitāb al-Mawlid* was bequeathed by Ahmad ʿAqīla al-Makkī [in Arabia] to the Sufi Abū al-Futūḥ Ma Laichi 馬來遲 (d. 1766), Ma Mingxin's colleague.

Attempts have been made to standardise the orthography of *xiaojing* texts and paratexts in manuscript production. A case in point is the *xiaojing* syllabary appended, with a didactic motive, to the Xining edition of Ma Tianmin's much-studied *ʿAqīda* (Fig. 9).

Processes of standardisation have been at work, as is indicated by the regular presence of vowel signs added to the transcriptions in the manuscripts consulted and by the doubling of *fatḥa*, *kasra* and *ḍamma* or the redundant *fatḥa*, all of which are characteristic features of *xiaojing*. The rule that the letter *jīm* is written in the initial form even when it occurs in isolation is generally observed.

Even so, there are still some major inconsistencies in the graphic representation of a number of Chinese sounds:[57]

a. the representation of the initial non-phonemic consonant sound in words such as *an*, 'peace', or *er*, 'child', is inconsistent, for example (*ghayn* and *ʿayn* are used interchangeably; in the proverb 'building the covered way...' in *K. al-Jahrī*, p. 113, *an* is rendered by *ghayn* + doubled *fatḥa*);
b. consonant x- (a fusion of the two MC initials s- and h-): in the above proverb, *xiu*, 'build', is written with *sīn* as سيو (as it may be traced back to an older s-); *sīn* also appears in *xi*, 'wash'; however, *sīn* has also been used to transcribe s- in (unvoiced) *si* , 'four'; in *xian*, 'appear', x- is transcribed by ﺙِ , the same as in *xiong*, 'breast'.
c. Finally, in the toponym 'Linxia', x- has been written using the letter *shīn*. The j- consonant (again representing two MC initials) is usually written as *kāf* with three dots above it, ﻛٟ, but some aberrations have been observed as well: j- in *xiaojing* has been written with a minuscule letter *ṭāʾ* placed above *kāf*; conversely, *kāf* with three dots, vocalised with *kasra*, ﻛٟ , has also been used to represent d- in *dian*, 'lightning', and j- in *jing* (apparently pronounced as *din* in the Linxia dialect) has also been written as *dāl* (ﺪٟ).

57 Many such inconsistences may be accounted for by the fact that *xiaojing* represents a dialect variant which retains features of MC.

Nowadays, Arabic text editions accompanied by *xiaojing* in any of the variations shown above compete with formats where translations in characters are printed in the same line as the Arabic wording, and the impression is that at least in printed publications Arabic followed by Chinese translations this will eventually become the standard. The fact that Arabic sacred text has occasionally been transcribed by Chinese characters used exclusively for their phonetic value indicates that even transliterations written in characters, considered more elegant by some, can prevail over *xiaojing* .[58]

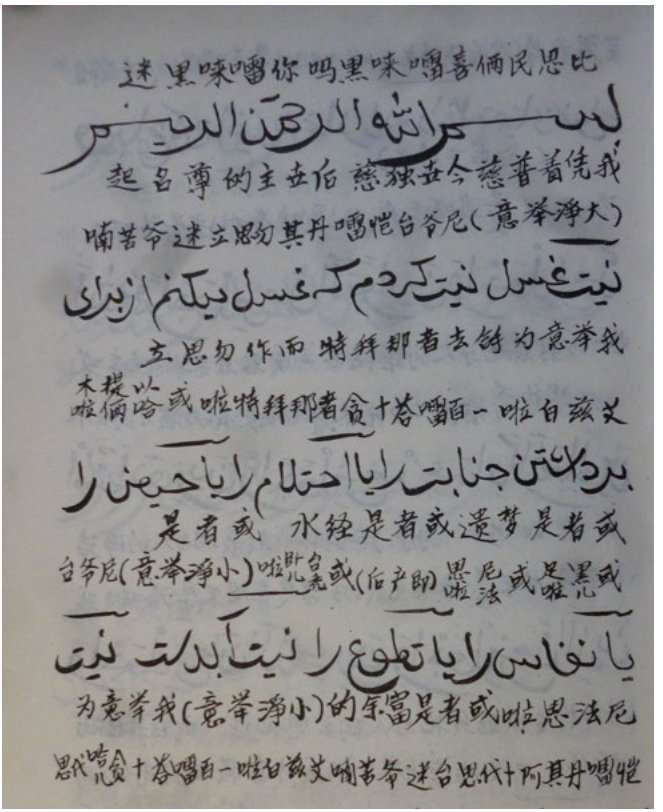


Fig. 10: The *Daʿawāt al-muslimīn* (a collection of prayers).

58 One example is a collection of prayers entitled *Daʿawāt al-muslimīn*, published as a facsimile manuscript in 1982 by the Islamic Association of Beijing (Beijing shi yisilan jiao xiehui 北京市伊斯蘭教協會) (Fig. 10). The collection also includes a Chinese translation written beneath the line, from the left to the right (the phonetic transcription written above the line runs from right to left).

References

Bakhtyar, Mozafar (1994), 'China', in Geoffrey Roper (ed.), *World Survey of Islamic Manuscripts*, Vol. 4, London: Al-Furqān Islamic Heritage Foundation, 63–116.

Bausani, Alessandro (1968), 'Un caso estremo di diffusione della scrittura araba: il "sino-araba"', in *Oriente moderno*, 48: 857–876.

Blochet, Emile (1909), 'Notice sur onze manuscrits persans rapporteés du Kan-sou', in *RMM,* 19: 285–293.

Brockelmann, Carl (1937–49), *Geschichte der arabischen Litteratur* [= GAL]. 2., den Suppl.-Bänden angepasste Auflage. Vol. 1–2. Suppl. vols 1–3, Leiden: Brill.

Cai Xiangyang 蔡向阳and Sun Dong孙栋 et al. (2008), *Hanyu chengyu fenlei da cidian* 漢語成語分類大辭典. Hubei: Chongwen shuju.

Daotong shizhuan 道統史專. See Manṣūr 1997.

Ding Shiren 丁士仁 (2008), *Jingtang yu de jiben tezheng he guanjian yuqi* 經堂語的基本特征和關鍵語氣.

Duanmu, San 端木三 (2007), *The phonology of Standard Chinese.* Second Edition. Oxford: Oxford University Press.

Dyer, Svetlana Rimsky-Korsakoff (1967), *Soviet Dungan: The Chinese language of central Asia: alphabet, phonology, morphology*. Bloomington: Asian Studies Research Institute.

EICh. – See *Zhongguo yisilan baike quanshu.*

Encyclopaedia of Islam (1960–2004), Second edition. 12 vols, Leiden: Brill.

Feng Zenglie 馮增烈 (2007), 'Xiaojing 小經'. In EICh, 618. Chengdu.

Forbes, A. D. (1960–2004), 'Ma Ming-Hsin', in *Encyclopedia of Islam*, vol. 5, 851a. Leiden.

GAL. – See Brockelmann.

Guanli Ye 閔裏爺 Abū l-Amān ʿAbdalqādir (1993), *Reshihaer* 熱什哈爾 (= *Rashaḥāt*). Zhang Chengzhi 張承志 (ed.), Yang Wanbao 楊万寶 and Ma Xuekai 馬學凱 (trs.). Beijing: Yinshua yichang.

Lipman, Jonathan (1999), 'Sufism in the Chinese Courts: Islam and Qing Law in the Eighteenth and Nineteenth Centuries', in Frederick de Jong and Bernd Radtke (eds), *Islamic Mysticism Contested. Thirteen Centuries of Controversies & Polemics*. Leiden: Brill, 553–575.

Luo Wanshou 籮万壽 (2007a), 'Liu Zhi 劉智', in EICh, 321–322. Chengdu.

Luo Wanshou. 籮万壽 (2007b), 'She Qiling 舍起靈', in EICh, 500. Chengdu.

Ma Siren 馬思仁 (2012), *Mannageibu* 曼納給布 (= Chin. trans. of Zhanye, *Manāqib*). Zhongwei shi Xuanhezhen Dongtai Dongdasi (Ningxia).

Ma Tong 馬通 (1999), *Zhongguo yisilan jiaopai menhuan suyuan* 中國伊斯蘭教派門宦溯源, Yinchuan: Ningxia renmin chubanshe. Reprint of the 1985 edition.

Ma Xuezhi 馬學智 Manṣūr Burhānaddīn Muḥammad (1997), *Daotong shizhuan* 道統史传 (Chinese trans. of Manṣūr, *K. al-Jahrī*). Ma Yi 馬義; Wang Shengjun 王生俊; Ke Fuyin 柯福銀; Hai Wenxue 海文學 (trs.). Xiji: Xijixian Beidasi.

Manṣūr Burhānaddīn Muḥammad (Ma Xuezhi 馬學智) (1352/1933) ['23rd year of the Republic'], *K. al-Jahrī*. Facsimile edition. No place.

Miškinene, Galina (2015), 'K istorii transliteratsii slavjanojazičnykh arabografičnykh rukopiseĭ litovskikh Tatar: na primere vil'njusskoĭ školy kitabistiki', in Joanna Kulwicka-Kamińska et al. (eds.), *The Tafsir of the Tatars of the grand duchy of Lithuania. Theory and research*. Toruń, 61–72.

Mungello, D. E. (1989), *Curious Land. Jesuit Accomodation and the Origins of Sinology*. Honolulu.

Na Guochang 納国昌 and Ma Chenghua 馬成化 (2007), 'Ma Dexin 馬德新', in EICh, 330–331. Chengdu.
Okuyan, Mehmet (1988–), Necmeddîn-i Dâye. In *Türk Diyanet Vakfı İslam Ansiklopedisi*, vol. 32, 496–97. Istanbul.
Sobieroj, Florian (2010), *Arabische Handschriften*. T. 9 (Verzeichnis der Orientalischen Handschriften in Deutschland [VOHD]). H.-O. Feistel (ed.). Stuttgart: Franz Steiner Verlag.
Sobieroj, Florian (2014), 'Arabic manuscripts on the periphery: Northwest Africa, Yemen and China', in Jörg B. Quenzer, Dmitry Bondarev and Jan-Ulrich Sobisch (eds), *Manuscript Cultures: Mapping the Field* (Studies in Manuscript Cultures, 1). Berlin: De Gruyter, 79–112.
Sobieroj, Florian (2016a), 'The Chinese Sufi Wiqāyatullāh Ma Mingxin and the Construction of His Sanctity in *Kitāb al-Jahrī*', in *Asiatische Studien*, 70 (1): 133–169.
Sobieroj, Florian (2016b), *Variance in Arabic Manuscripts. Arabic Didactic Poems from the Eleventh to the Seventeenth Centuries: Analysis of Textual Variance and Its Control in the Manuscripts* (Studies in Manuscript Cultures, 5). Berlin: De Gruyter.
Stöcker-Parnian, Barbara (2002), *Jingtang Jiaoyu – die Bücherhallen Erziehung. Entstehung und Entwicklung der islamischen Erziehung* etc. Frankfurt a. M.: Peter Lang.
Wu Jianwei 吳建偉 (1995), *Zhongguo qingzhensi zonglan* 中國清真寺綜覽 , Yinchuan: Ningxia renmin chubanshe.
Wikipedia "小儿经" *xiaoer jing* (2019), Chinese Wikipedia article '小儿经' https://zh.m.wikipedia.org/zh-hans/小儿经 (last accessed 06.02.2019).
Yang Zhanwu 楊佔武 (1996), *Huizu yuyan wenhua*. 回族語言文化, Yinchuan: Ningxia renmin chubanshe.
Yibulaxin 易卜拉欣 (2007), 'Dongxiang yu xiaojing', in EICh, 138. Chengdu.
Zhanye 氈爺 ʿAbd al-Aḥad (n. d.). *Manāqib-i Awliyā*. Facsimile edition. No place.
Zhongguo yisilan baike quanshu 中國伊斯蘭百科全書 (EICh) (2007), *Encyclopaedia of Islam*, Chengdu: Sichuan cishu chubanshe.
Zhou Chuanbin 周傳斌 (2008), *Jiaoyu. Xinhuo xiangchuande Huizu jiaoyu* 教育. 薪火相傳的回族教育. Yinchuan: Yinchuan chubanshe.

Jan van der Putten

A Collection of Unstandardised Consistencies? The Use of Jawi Script in a Few Early Malay Manuscripts from the Moluccas

Abstract: Jawi, a form of Arabic script in use in insular Southeast Asia, has been employed for writing Malay manuscripts for at least eight centuries (13th–20th centuries). Introduced and disseminated through Islamic canonical texts such as the Qur'an, it became the principal means of communication and knowledge production over the years and across the region. Studies of Malay palaeography and orthography have been rare and far between, and are mainly limited to colonial scholars designing the rules to be taught in the educational system and a number of surveys that record the orthography of certain texts.

Inspired by a recent publication which concisely discusses the palaeography and orthography of 60 dated or datable Malay manuscripts, the following chapter will expand on some of the observations made by the scholars. I will discuss a number of 17th–19th-century manuscripts originating from the Moluccas and compare some characteristics of their orthography with those of the *Hikayat Tanah Hitu*, a Malay historical narrative compiled in the mid-17th century. In this exploration of orthographic consistencies and peculiarities in Malay manuscripts originating from one region, I argue that the ambiguity of Jawi was a welcome characteristic that could accommodate the needs of multilingual communities found in the Malay world.

1 Introduction

Throughout history, the writing of Malay has made use of a variety of scripts in concordance with the prevailing influence of a major belief system or hegemonic colonial regime: Hindu-Buddhist traditions made use of a southern Brahmi script that spawned a number of vernacular forms; Islam brought Semitic languages and the Arabic script; and European imperialism and 'modernity' brought Latin script, which would eventually dominate the written and printed communications in insular Southeast Asia from the 19th century onwards, even though Christianity never became the dominating religion of the region. Insular Southeast Asia, where the Malay language was developed, spread and is now widely used,

https://doi.org/10.1515/9783110639063-009

is predominantly Islamic except for the northern part of the Philippines and some areas in eastern Indonesia. It is generally assumed that Islam gained its first firm footing in the 13th century with the conversion of the ruler of Samudra-Pasai at the northern tip of Sumatra. Other polities in the islands soon followed, and a Muslim trading network in the region gradually developed, renewing ties and connecting with larger networks across the Indian Ocean.

Islam not only introduced its canonical texts with their Arabic language and script, it also informed and enhanced contacts with other Muslim traditions such as Gujarati and Persian. These newly introduced traditions met with existing ways of disseminating texts and knowledge, with which they mixed in different ways according to local circumstances in very complex and multifarious processes that took place over a long period of time. In the multi-linguistic environment of the Indonesian archipelago, Malay was one of the vernaculars that had been used for a long time to communicate official matters of state and religious affairs and was apparently sufficiently known in 13th-century North Sumatra to become the vehicle of transmission for the new religion. Although there are a few examples of Islamic texts represented in locally derived Brahmi scripts, the Arabic script seems to have soon taken over the dominant role in disseminating religious knowledge and also other, more secular topics such as business affairs.

Expanding Muslim trading networks greatly stimulated the dissemination of Islam, its texts and the languages of transmission, Arabic, Persian and, most extensively, Malay. Due to geopolitical circumstances and reasons, the nodes of this network dispersed over a myriad of islands, all with their own local interests and traditions. Later, other Asian and European trading partners joined, and eventually disturbed the balance within the network. Therefore, it is difficult if not impossible to discern a single and strong Arabic or Malay Islamic tradition that would encompass the whole region and spur further developments. Throughout history, there have been powerful centres that have extended great influence over a longer period of time in political and cultural terms, such as Melaka in the 15th century, Aceh and Banten in the 16th until the mid-17th centuries, and Makassar in the 17th century. However, the impact these power centres impressed upon a more generally defined Malay culture was relatively short-lived and therefore limited in scope. Of course, the presence and gradual expansion of the influence of the Dutch East Indies Trading Company (Verenigde Oost-Indische Compagnie or VOC) also thwarted the establishment and continuation of a firm Malay Islamic tradition.

Although we perhaps cannot speak of a strong, single Malay tradition, the Arabic script which seems to have become current in insular Southeast Asia in the 16th and 17th centuries made it possible for narratives to be circulated within a cultural realm heterogeneously and temporarily delineated by a combination of trade agreements, religion and language. The Arabic script, developed to represent the

Semitic language in close connection with the writing down and distribution of the holy Qur'an and other liturgical and theological texts, was adjusted to enable the translation of these texts into languages of other language families. Probably based on a familiarity with similar adjustments made for Persian and Indian languages, five consonants were added for writing Malay (<p>: ڤ, <c>: چ, <ng>: ڠ, <ny>: ڽ and <g>: ݢ), while in the Javanese system an extra two characters were added to represent the retroflex stops /ʈ/and /ɖ/. These modifications were realised by adding diacritical dots to the existing forms of the Arabic characters, e.g. the Malay voiceless palatal stop /ʧ/ is represented by the addition of three dots below the Arabic ح <ḥ>. The set of Arabic characters to write Javanese is called Pégon. It is used next to locally developed Brahmi-derived as well as Latin scripts, while Jawi is the common name to refer to the alphabet used to write texts in Malay and several other regional languages. Palaeographical and orthographical studies of Jawi as used in manuscripts are rare, and were mostly carried out by colonial officials and missionaries who made it their business to disparage Jawi spelling. They argued that it was defective in terms of not representing the full range of vowels and, probably more importantly for most of them, that it was connected to Islam. Inspired by such considerations, they made efforts to 'force Jawi spelling into a Procrustean bed of presumed rules and standards' (Kratz 2002, 22).

On the other hand, Ulrich Kratz (2002) also contends that it may in fact be this 'defective' nature of the Jawi spelling not fully indicating the vowels of a word which allows for a variation in pronunciation and writing, a characteristic well appreciated in a multilingual environment and conducive to joining people together in a common cultural realm. In a recent publication, issued as a tribute to Ulrich Kratz's scholarship on Malay manuscript studies, Annabel Teh Gallop refers to this particularity in Malay spelling as a 'tradition of inconsistency' which may provide a good way of describing the impressions one gets upon reading Malay manuscripts (Gallop 2015, 29).

In this chapter I would like to follow up on some of the points Kratz and Gallop have indicated in their recent publications and explore a few characteristics of the Jawi spelling as found in a small number of manuscripts originating from the Moluccas in the 17th century. I must emphasise here that this is a first preliminary exploration of this kind, building on what has been displayed in *A Jawi Sourcebook for the Study of Malay Palaeography and Orthography* (Gallop 2015), which contains a concise survey of the writing in 60 dated Malay manuscripts spanning almost 350 years. In the next section, I will concisely discuss some earlier studies about Jawi spelling before going on to introduce and contextualise the manuscripts I explore in the remainder of this paper. As will become clear, the state of the art does not allow me to make any sweeping conclusions about standardisation of the Jawi script or the development of the orthography in Malay manuscripts.

2 Studies in Malay orthography

With reference to the recent studies by the above-mentioned scholars of Jawi script, earlier studies about Jawi orthography in general have put much effort into determining how ill-equipped the modified Arabic script was for the representation of Malay sounds. Moreover, they sought to prescribe certain rules that would improve the deplorable state of affairs and make the Malays write their language in a proper manner (see Kratz 2002 and Gallop 2015). In an elaborate survey of the history of the scripts used for writing Malay, Hashim Musa (1997) understandably focuses on a description of the Jawi script which to a certain extent is still being used and propagated in certain Malaysian quarters. His book presents long lists of spelling particularities that were gleaned from a wide array of Malay texts, ranging from a number of early inscriptions on tombstones and the earliest evidence of Jawi script in the region on the Trengganu Stone (dated 1303 CE) to a selection of manuscripts from the 16th up to the 19th century, to the spelling of a few books, treatises and news items in periodicals published in Malaysia in the 20th century. In a concise discussion of earlier studies on this topic, the author simply reverberates a well-known proposition made by the British missionary-cum-scholar William Girdlestone Shellabear (1901) that the relatively fixed standard of the writing in early Malay manuscripts was due to the recent transfer of the script from Arab teachers to Malay pupils. This standard for the writing of Malay is supposed to have gradually faded in subsequent periods as more local particularities seeped into the writing system. These local particularities mainly pertain to the representation of a few consonants, such as the opposition between *f* and *p, g* and glottal stop, and especially the writing of vowels in the words. Hashim Musa also refers to a study by the Korean Kang Kyoung Seok, who proposed the thesis that the Jawi spelling had gone through three stages in its development for the purpose of representing the Malay language. The first phase is characterised by a faithful rendering of Malay sounds by applying an Arabic standard with a full set of diacritics, including vowel points. In the second phase, this standard was gradually left behind and vowel points were replaced by a set of semi-vowels comprising *alif*, *wāw* and *yā'* . These represent the vowel pairs a/ə, o/u and i/e (è, é), which were inserted in the penultimate syllable of normal two- or three-syllable words in Malay which contains the word stress.[1] The third phase was characterised by the inclusion of these semi-vowels also in the ultimate syllable, a feature which is visible in the standard Jawi spelling systems developed in modern-day Malaysia (Hashim 1997, 78–80, Kratz 2002, 23).

1 For a discussion about the *alif* used to represent /ə/, see below.

As a general rule of thumb, the most obvious characteristic of the second phase often found in Malay manuscripts produced between the 16th and 19th centuries is the spelling of common words such as *jadi* 'to be, become' as جاد <jad>, and *orang* 'human being' as اورڠ <awrng>.[2] However, this feature of not indicating vowels in certain positions and the use of a set of ambivalent indicators (i.e. the use of semi-vowels), or even the total absence of any indication of vowels, which regularly occurs in the spelling of certain words,[3] charges the written texts with a certain degree of ambiguity which has led to severe criticism of past and present scholars and other observers imbued by a print-literate-inspired longing for unambiguousness of written texts (cf. Milroy and Milroy 1991, 64–69). Such criticism, of course, tacitly ignores similar inconsistencies in spelling and other linguistic peculiarities which were common in European written texts from an earlier period when mass education was not yet established and the standardisation of language usages was much less developed than it is now. Written communication is usually considered superior to speech in Western civilisations and by consequence applied as a yardstick to measure cultures in other parts of the world. However, two caveats must be brought forward with respect to 'measuring' Malay writing in accordance with such a 'universal' truism. Besides the fact that we are dealing with texts which were written down during a time when virtually no structured formal education existed in the Malay world, it may also be conceived that there are languages in which a certain degree of ambiguity in texts may be more appreciated than a formal, unambiguous form of communication. Such a consideration was proposed by Geoffrey Benjamin with regard to the informality and situatedness of colloquial Malay, which cherishes a high degree of ambiguity and invites interlocutors to join the 'language play'. This informal or condensed variety of the language shares certain important characteristics with the Malay used in narratives recited for local audiences, such as a-historicity, a lack of interest in the relations between events, and a downplaying of individuality. Benjamin describes this condensed form of Malay as a system in constant flux, where we are faced with contextual and fleeting shifts in cultural meanings projected by native speakers on the linguistic forms being used (Benjamin 1993, 355–56).

2 Words starting with the vowels i/e or u/o are introduced with the *alif-yā'* and *alif-wāw* respectively.

3 This commonly leads to consonant frameworks which can accommodate different vocalisations. The framework سنتق (s-n-t-q), for instance, is interpreted in Klinkert's Malay-Dutch dictionary (1947, 603–4) as *santak* ('to thump'), *səntak* ('to pull'), *santuk* ('sleepy; to accidentally hit something'), *suntuk* ('being obstructed'), *sintok* ('tree species: cinnamomum *sintoc*'), *sintik* ('type of small oyster').

A similar proposition is made by Ulrich Kratz, who calls the major perceived shortcoming of not indicating all vowels of the words in what is to a certain extent a received Jawi spelling a 'major strength'. Part of his attractive contention is that Jawi had established a 'shared written link' among regions and dialects (Kratz 2002, 23). This kind of written link seems to have existed in particular between Malay and Minangkabau cultural realms in Sumatra, whose languages share a large number of lexical items but differ in their pronunciation. In the case of such similar languages it seems quite straightforward to conceive that the inherent ambiguity of linguistic expressions in the writing system of these languages was considered to be a positive feature, as it marks a more open venue for communication than a form rigidly controlled by non-negotiable rules for the spelling of particular words. The consonant framework سنج (s-n-j), for instance, allows speakers of these languages to realise a different pronunciation, [sənja] and [sanjo] for Malay and Minangkabau respectively, which in the oral/aural mode of communication may not be readily recognised by them as one and the same word. However, its visual representation in written form is identical in Arabic script and therefore the meaning is readily understandable for speakers of both languages as 'dusk'.

A second caveat we need to make in this respect concerns the disregard of a differentiation of certain textual genres and the possible existence of more stable local traditions that produced Malay texts in which a higher degree of consistency was applied than general statements allow for. As the accuracy of Malay scribes' copying of doctrinal Islamic texts is generally considered higher than in more secular texts (see e.g. Voorhoeve 1964), it seems likely that the former textual genre will show a higher degree of consistency in spelling in comparison to the latter. We should also be aware that, although Malay was the language of state and religious learning from the 16th century onwards in many parts of the archipelago and beyond, people used different local languages in daily conversation in polities such as Aceh, Makassar, Bima and the Moluccas. These distinct polities were nodes in Islamic trading and linked scholarly networks; here local traditions may have developed their own particularities in terms of textual layout, shape and spelling in manuscripts produced at the courts of the rulers and other centres of learning. The study of these kinds of regional traditions found in clusters of related and datable documents may be a more fruitful approach, leading to a certain understanding about developments in the palaeography and orthography of Jawi writing, as Annabel Teh Gallop convincingly proposes in her recent publication. As preliminary examples of such traditions she mentions a few characteristics of an Acehnese religious book hand which was current in religious writings in the 18th and 19th century, and a Moluccan chancery hand found in documents from the royal courts in the Moluccas (Gallop 2015, 34–36). It is this notion of a

local tradition in the production of manuscripts I will engage in here and explore to what extent it can be applied to a number of manuscripts that originate from the Moluccas and the royal court of Makassar on the island of Sulawesi.

3 Malay documents from the Moluccas and Makassar

Although it was already widely used as a trading language in eastern parts of the archipelago at an early stage, Malay was not the only or the most dominating language of the region. In the 17th century, Malay had to compete with Portuguese, Spanish and Arabic as language of diplomacy; it was used in combination with Arabic in the dissemination of religious knowledge; and it also served as contact language next to a welter of local languages in daily conversations and trade negotiations. Because of the invaluable spices grown in the Moluccan islands, for centuries the region had been a trading hub which attracted a host of merchants, priests and soldiers from the archipelago, such as Javanese, Makassarese and Butonese, and from other parts of the world, including Arabs, Chinese, Indians and Europeans. Quite a few minor local rulers tried to establish their own polity to surf the prosperous wave of increasingly globalised trading networks, which inevitably led to harsh competition and full-scale wars between local contenders who frequently called in the assistance of outside forces to decide the conflicts with their neighbours and other contenders. In the 16th century the islands were alternately dominated by the northern polities of Ternate and Tidore. The former first secured Portuguese support for their political agenda, and later the Dutch helped the rulers to consolidate their power and prosperity into the 17th century, ironically also making Islam the dominant religion in the islands, despite Portuguese and Dutch proselytising activities. Not long after its arrival, the Dutch VOC implemented a gruesome monopoly on the cultivation and trade of cloves and nutmeg by forcefully expelling all foreign competitors and playing regional powers off against each other, or killing them if they failed to comply with the new rules.

It is in this highly volatile context that the documents I will discuss are set. The first and most important document is the *Tale of Hitu* (*Hikayat Tanah Hitu*). Hitu was a polity which had united 30 settlements on the north coast of Ambon, and it was jointly ruled by four prominent families. Imam Sifar Rijali, a learned member of one of these four ruling families, is reported to have written *Hikayat Tanah Hitu* while he was staying at the viceroy's court of Gowa in Makassar. Because of severe reductions in the clove production imposed by the Dutch VOC, the govern-

ment of Hitu had declared a war that would rage in the central Moluccan islands with intervals during the 1640s and 1650s. Rijali had taken refuge in Makassar to ask the assistance of the Makassarese at Gowa as the main competitors of the Ternatean overlords of Hitu. The rulers of Ternate were too much inclined to go along with Dutch rule. Subsequently, in the 1650s one of the factions of the Ternatean ruling family rebelled against this pro-Dutch stance of the sultan and open fighting broke out on the islands under the leadership of Majira, a distant member of the ruling family of Ternate. As one of the Hituese leaders, Rijali intended to restore the authority of his government. It was perhaps to provide documentation of the ongoing war to support his appeal or at the personal request of the enlightened viceroy of Makassar, Karaeng Pattingalloang, that Rijali compiled a prose narrative about the events and their historical context.[4]

Hikayat Tanah Hitu more or less chronologically deals with the early state formation in Hitu (1500–38), wars the Hituese waged on the Portuguese (1538–1605), the monopolisation of the clove trade by the VOC (1605–43), and Hituese armed opposition against the abolition of the Hituese government (1643–46). After writing the tale during his exile in Makassar, Rijali went back to the Moluccan islands in 1653, where a copy of the work was made in the 1650s for another member of his family. It is this copy that was passed on into the hands of the well-known German-born merchant, botanist and historian Georg Eberhard Rumphius, who spent most of his life in Ambon. This copy then, possibly by way of the Dutch Reverend Valentijn, eventually ended up in Leiden. Both Rumphius and Valentijn, well-known contemporaneous commentators on Moluccan affairs and important contributors to its historiography, used the tale to mine information for the writing of their own histories of the region. In his monumental *Oud en Nieuw Oost Indiën* (1724–26) Valentijn mentioned the tale as one of the texts being circulated within Muslim quarters. He interpreted this as an indication of the scholarship of the Muslim part of the Moluccan population, whose knowledge and ability to speak and write Malay he considered much better developed than the level of proficiency of the Christians under his tutelage. He was convinced that this better command of Malay by Muslims was due to the fact that they possessed a number of Malay writings, which they lent to each other to read or copy. Among the other texts he found in Ambon, Valentijn listed works which are now considered as classics in Malay traditional writing, including *Sulalat al-Salatin* or *Sejarah Melayu*, containing historical tales focusing on Malacca; *Hikayat Amir*

4 For a full account of the historical context and a description of the manuscript containing this text, see Stravers, Van Fraassen, and van der Putten 2004. The unique manuscript is preserved under Cod. Or. 5448 in Leiden University Library.

Hamzah, a tale about Prophet Muhammad's uncle; and *Hikayat Nabi Muhammad* and *Hikayat Nur Muhammad*, two tales about the Prophet Muhammad.

Hikayat Tanah Hitu is one of the very few relatively old text examples we have from eastern Indonesia, which adds to its importance for manuscript studies in general and the study of the development of the Jawi script in particular. The manuscript originally consisted of 53 folded folio pages which seem to have been bound with thread in quires. Since most of these pages are torn in the fold, the manuscript now almost exclusively comprises loose pages, measuring about 32 by 20 centimeters. The beginning of the tale is missing and the pages have been numbered using Arabic numerals from 2 up to 107 by one of the later owners or users. The first 80 pages each contain 17 lines in orderly Arabic characters, whereas the script in the remaining pages is less regular and these pages contain 15 to 17 lines. The manuscript held at Leiden University Library is still clearly legible, but the paper is rather worn and tattered at the edges. Even though the manuscript was probably originally bound, there is no sign of any covers or end-papers. The first and last pages are provisionally repaired with Japanese tissue and quite a few pages have greasy stains.

The extant text starts with the end of what must be the initial episode of the tale, which indicates that not many pages have gone missing from this manuscript. A few doodles that embellish the pages which now serve as first and last page, indicating the beginning and end of the extant manuscript, suggest that any covers and the beginning of the text might have been missing already by the time the copy was made or passed on into Dutch hands. The remaining part is complete, since the last page of the manuscript contains the end of the text.

The paper used for the copy of Rijali's text reveals some indications about the age of the manuscript, as the last 14 pages contain a watermark (a fool's cap) and a countermark (consisting of the initials for VOC). The typical fool's cap watermark indicates that the paper dates from the second half of the 17th century, whereas the countermark shows that the paper was ordered by the VOC.[5] Therefore the manuscript may be dated to the second half of the 17th century and is among the oldest Indonesian manuscripts extant. This is in contrast with most of the preserved Malay manuscripts which contain copies of texts made during the 19th century for successive generations of owners, as climatic conditions in the tropics render paper a highly perishable medium for the distribution of texts.

5 The paper shows a fool's cap with a seven-pointed collar of a kind which is likely to have been manufactured in the second half of the 17th century, while fool's caps with five-pointed collars are generally from the first half of that century (see Laurentius and Laurentius 2008, 2: vii).

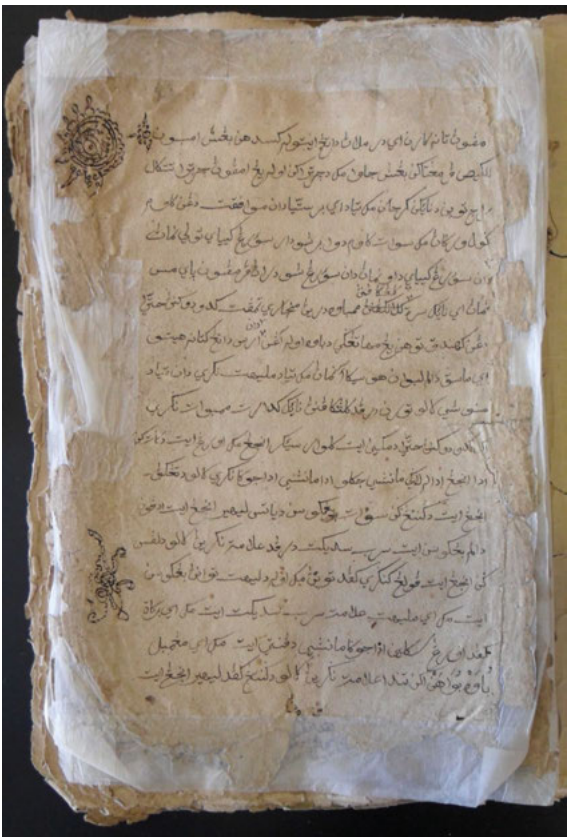


Fig. 1: First page of the *Tale of Hitu* (*Hikayat Tanah Hitu*), Leiden University Library, Cod. Or. 5448. © Courtesy of the University Library Leiden.

The *Hikayat Tanah Hitu* manuscript not only contains indications about its age, but also about its origin. In this respect a piece of calligraphy that appears twice in the manuscript deserves special attention: in the top margin of pages 77 and 78 of the manuscript (shown in Fig. 2), we find the phrase *Min Bulan Nustapi* (written as من بلن نسطفي (m-n-b-l-n n-s-ṭ-f-y) and meaning 'belonging to Bulan Nusatapi', the sobriquet of one of Rijali's cousins). This may be interpreted as the hallmark or inscription of the original owner(s) of the manuscript; Nustapi or Nusatapi was the name of Rijali's lineage. The manuscript also offers indications that it was a copy of an older one, perhaps even of the original text by Rijali. At certain intervals in the manuscript we find a word in the margin that has served as a sign for someone to indicate that the reading during a certain session had come to that

particular point in the text. Although quite common in the Islamic tradition of the Middle East, this Arabic word, *balagh* (بلغ; 'reach, transmit, report'), is not a mark that appears frequently in Malay manuscripts. In this manuscript it seems to indicate the reading by an authoritative reader who compared the copy with the original and in the end gave his approval to the copy. This approval is found on page 92, where the word *saḥḥ* (صح; 'authentic, acknowledged, legal') is written in the margin. These marginal notes of *balagh* and *saḥḥ* suggest that a member of the Nusatapi family carefully checked the manuscript, gave his approval to the copy and eventually inscribed the name of his family as an indication of ownership in the manuscript.

Fig. 2: Inscription of the name of the original custodians of the manuscript; pages 77 and 78 of the *Hikayat Tanah Hitu*, Leiden University Library, Cod. Or. 5448. © Courtesy of the University Library Leiden.

As indicated before, there are not many Malay documents that have been preserved from this period and region, but some scattered manuscripts may serve as extant material with which the orthography and some palaeographic characteristics of the handwritten *Tale of Hitu* can be usefully compared. These documents include a manuscript preserved in the Staatsbibliothek in Berlin which contains (1) a (fragment of an) undated letter from one of the Hituese leaders probably compiled in connection with one of the other texts in the manuscript, namely (2) an agreement between the Ternatean Sultan Mandar Syah and the Dutch Governor-General of 1652; and (3) the text of an agreement between the Dutch and Sultan Hamza in Ternate from 1638 (Staatsbibliothek, Ms. Or. Fol 409, 1–3).[6] The other documents comprise a letter from a Ternatean leader in exile in Makassar, Kime-

6 The text of the letter seems closely related to the problems surrounding the visit the Ternatean Sultan Hamza paid to Hitu and the consequences of the new political relations as imposed by the VOC in 1637 (see Stravers, Van Fraassen and van der Putten 2004, 61–62; 178–85). I surmise that the letter originally may have been an attachment of the agreement between Sultan Hamza and the VOC.

laha Salahak Abdul Kadir ibn Syahbuddin, to the British East India Company dated 23 May 1658 (Gallop and Arps 1991, 38), and two handwritten narratives copied in Ambon at the beginning of the 18th century as presented in *A Jawi Sourcebook* by Vladimir Braginsky (2015, 54–55; 62–63). As Annabel Teh Gallop notes, even a cursory comparison between these documents and two early-19th-century letters from Ternate and Tidore she scrutinises for the same publication (Gallop 2015, 82–85) already shows that the handwriting of these documents differs markedly from the 'wispy and spidery' hand and specific form of some letters that are characteristic for the Moluccan chancery hand (Gallop 2015, 36). Below I will discuss a few palaeographic and orthographic characteristics of the writing in the copy of the *Hikayat Tanah Hitu* and indicate similarities with the other extant Malay documents mentioned above.

4 Characteristics in the writing of the *Tale of Hitu*

Taking a glance at the mise-en-page of the first page of the *Hikayat Tanah Hitu* text (Fig. 1), it is obvious that the text block takes a central position on the page with three wider and one smaller margin, the narrowest margin being on the side of the paper where the manuscript is bound. This layout is consistent throughout the manuscript, and we can see that the first page of the manuscript with the doodles in the left-hand, wider margin is the recto side while a possible title page (first recto) and start of the text (first verso) are missing (see Fig. 1).

The writing is regular but not refined. The most obvious characteristic is the rather elongated, slightly slanted top stroke of the letter *kāf* (*k/g*) above its straight upright 'trunk'.[7] While the tails of certain letters, particularly the *rā'* and *wāw*, are often nicely rounded and elongated, the handwriting overall does seem to be quite common and does not show any distinct characteristics which would set the manuscript apart from the bulk of other Malay manuscripts.

The orthography of some of the words is arguably more specific to this text or copyist, and to a certain degree also represents the conventions of the period and

7 See Fig. 1, but it may be even more obvious on other pages, such as the page of the manuscript which was included as an illustration in van der Putten 2015, 51. The <g> can be distinguished from the <k> in the Jawi script by placing one or three dots above or below the character *kāf* (ک). In Malay texts, these dots quite often are omitted, especially in common words such as *juga* ('also'; spelled جوك [j-w-k]), but in *Hikayat Tanah Hituthe* opposite idiosyncrasy is applied by adding dots to the *kāf* in some words, even though these are normally spelled with a <k>, as for instance in the word kəluarga ('siblings, family') spelled كڤلوركا [g-w-l-w-r-g-a].

place where the copy was made. What must be mentioned first and foremost here is the use of the *tashdid* or *shaddah*, a diacritical mark to strengthen and normally double the consonant in Arabic, which takes the form of small 'w' on top of the enhanced character. The borrowing of this diacritic in Jawi orthography has triggered quite a number of comments from scholars' earlier studies of Malay texts, which are described and discussed by Russell Jones (2005). The most common explanation for the use of the *tashdid* in Jawi orthography is that it would indicate a schwa (a mid central vowel [ə]) in the preceding syllable. As there is no specific diacritic in Arabic script for the schwa, most frequently referred to with the Javanese name *pĕpĕt* in Malay studies, an *alif* may possibly be used to indicate this vowel. However, as mentioned above, the half-vowels *alif*, *wāw* and *yā'*, used in Malay spelling to indicate the vowel, are normally included in the penultimate syllable which contains the word stress, whereas syllables containing a *pĕpĕt* most often are not stressed. The convention of the use of the *tashdid* in these cases seems logical, because geminating a consonant does seem to have an effect on the quality of the preceding vowel (cf. Khattab and Al-Tamimi 2008). As has been noted by several scholars, a similar method of geminating the following consonants to indicate a schwa is also found in Old Malay and especially Javanese inscriptions from South Sumatra and Central Java, which were written in a script derived from the southern Brahmi script (Jones 2005, 281–2). Mahdi notes that for the Brahmi script used in Old Malay inscriptions, consonants are geminated following a prefixal *–r*, such as in *marvvaṅun* ('to rise'), while in root words it is only once found in the Old Malay inscriptions found in South Sumatra in the cognate for Malay *bətuṃ* ('bamboo', *pattuṃ* in the inscription), but occurs more frequently in later inscriptions (see Mahdi 2005, 187–8).

Although logically we tend to look at Arabic phonology to describe the *tashdid*'s function and effect with regard to Malay orthography, most interestingly the indicated function of the diacritical mark for Malay is in agreement with James Collins's comments about the historical development of a Malay dialect in eastern Kalimantan. In a concise overview about these dialects, he notes that the Proto-Malay *pĕpĕt* in Berau Malay has merged with /a/, and that the gemination of consonants following a penultimate syllable which originally contained **ə* is historically related to this merger (Collins 2006, 39).

Although he states that these characteristics are not shared by Kutai Malay, the major other dialect in the region, the examples Collins gives for the Malay dialect of Berau mirror the system of Jawi spelling as we find it in the *Hikayat Tanah Hitu* and other texts and may indicate a common characteristic in Austronesian languages in which consonants are geminated under influence of a schwa in a preceding syllable.

In discussions about the inclusion of the *tashdid* in Jawi manuscripts, its use is usually considered as being an indication of the relatively old age of the manuscript, but as Russell Jones (2005, 289) has noted, this interpretation is not absolute, as we can also find *tashdid*s in 19th-century manuscripts. This being the case, however, I think that a frequent use of this diacritical mark in original Malay words does represent a somewhat archaic tradition which can tell us something about the age of the particular manuscript. I would argue that this is also the case with regard to the manuscript under discussion here and that, to a certain degree, the use of the *tashdid* indicates a certain convention which was current in the Malay world, more specifically in the eastern archipelago, in the 17th century. In the *Hikayat Tanah Hitu*, then, we find a rather frequent use of the *tashdid* on the following consonant after a *pĕpĕt* in a penultimate syllable of words such as *bənar* ('right, correct', بنّر spelled [b-nʷ-r]), *bərkəlahi* ('to fight', spelled بركلّا [b-r-k-lʷ-a])[8], *tətak* or *mənətak* ('to slash', spelled منتّع [m-n-tʷ-']), *mənang* (to win, spelled منّغ [m-nʷ-ng]), *sri* (honorific title in combination with sultan, spelled سرّي [s-rʷ-y]). This use of the *tashdid* occurs a few times in the other documents originating from the same time frame and region, for example, *tətak* and *bəsar* with a *tashdid* on the second consonant is found in the fragment of an undated letter preserved in the Staatsbibliothek in Berlin (MS Or. Fol. 409-1, see Fig. 3), while in the Malay contract from the same collection of manuscripts (MS Or. Fol. 409-2, 3), the *tashdid* only occurs systematically in the word *dia* (third person pronoun singular, spelled ديّ [d-yʷ]), which was possibly done to distinguish the word from the preposition *di* ('in, at'). The page of the manuscript copy of the *Tale of Isma Yatim* (*Hikayat Isma Yatim*) which illustrates Vladimir Braginsky's concise notes on the same (Braginsky 2015, 55) yields a better crop of *tashdid*s: in the first seven lines we find these diacritical marks used in the words *majəlis* ('council', spelled مجلّس [m-j-lʷ-s]), *sekalian* ('all', spelled سكلّين [s-k-lʷ-y-n]) and *duli* ('dust', part of a formula referring to the king, spelled دلّ [d-lʷ]). The latter spelling, which also occurs regularly in the *Hikayat Tanah Hitu*, may be connected to the Sanskrit origin of the word, while the other two both may indicate the schwa in the preceding syllable.

8 This spelling occurs a few times in the *Hikayat Tanah Hitu*, sometimes without a *tashdid* and all of them without a final *yā*'. The undated letter in the Staatsbibliothek contains twice the same word spelled with the final *yā*' and without a *tashdid* (بكلاي [b-k-l-a-y], see Fig. 3, beginning of line 11 and 12).

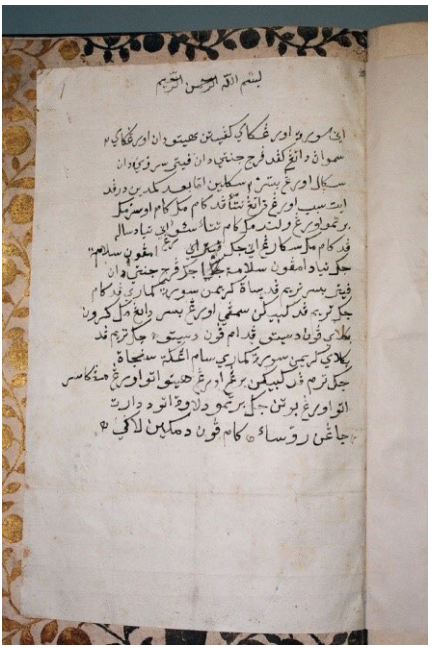


Fig. 3: Fragment of the undated letter from Hitu, probably late 1630s, Staatsbibliothek Berlin, MS Or. fol. 409-1[10]. © Staatsbibliothek zu Berlin - Preußischer Kulturbesitz, Orientabteilung.

The same page of the *Tale of Isma Yatim* also shows another characteristic which may be described as a spelling convention it shares with the writing of the *Tale of Hitu*. In line 4 of the page displayed in Braginsky's description (Braginsky 2015, 55), we find the word *jənis* ('sort') spelled as جينس (*jinis*), where the insertion of the half-vowel in the penultimate syllable may be observed as indication for the vowel of the final syllable. The same spelling method of writing the vowel of the ultimate syllable in the penultimate syllable which contains a schwa occur in a few instances in *Hikayat Tanah Hitu*, in words such as *kəris* ('dagger', spelled كيرس [k-y-r-s]), *bəlum* ('not yet', spelled بولم [b-w-l-m]) and *Jəpun* ('Japan', spelled جوڤن [j-w-p-n]). In a few instances in the *Tale of Hitu* we find *pəti* ('trunk, coffin') spelled as ڤيتّي [p-y-t^{w}-y], in which the convention of writing the assimilated half-vowel is found in combination with the inclusion of the *tashdid* on the following consonant. This spelling convention is restricted to a few words only, and it is obscure what might have triggered it. It seems clear, however, that representing the schwa in Malay texts in the Moluccas caused some problems, possibly related

to specific traits of local languages, and Malay dialects of the region generally lack this phoneme in their phonological system.[9]

Besides these regularities which are to a certain extent shared with some other texts from the early period of Malay writings, *Hikayat Tanah Hitu* also shows some particularities which may be considered idiosyncrasies of the scribe, possibly induced by the spoken vernacular, while others are rather commonly found in Malay manuscripts. These spelling particularities concern a somewhat regular, albeit certainly not consistent, omission of the voiceless glottal fricative /h/ in initial, medial and final positions, while in other instances the *h* is added in words which do not originally contain it. For instance, on the first page of the text in lines 6, 7 and 8, we find respectively *məmbawah*, *dibawah* ('to bring') and *labuan* ('mooring'), which in a perhaps more common orthography would be spelled *məmbawa*, *dibawa* and *labuhan* (see Fig. 1). Apart from this very common characteristic which is also found outside Old Malay manuscripts and the Malay language, we find a particular confusion of certain nasals and the omission of a glottal stop, most frequently before the suffix *–kan* is attached to the words. Some examples in the *Tale of Hitu* comprise *buankan* ('to discard, exile', instead of *buangkan*), *sampang* ('small boat', instead of *sampan*), *Seran* (name of an island, instead of *Seram*), *dinaikan* ('to rise, install', instead of *dinaikkan*), *enda* ('to want, will', instead of *hendak*) and *anaku* ('my child', instead of *anakku*). Both these characteristics may have been influenced by a vernacular language or a local dialect of Malay.

The use of vowel points in the Arabic script in the *Tale of Hitu* is mainly limited to the spelling of Arabic words and a few proper names, such as فَرْدِيْرِيكْ حوُتْمَنْ [fᵃ-rº-dyº-ryⁱ-kº ḥwᵘ-tº-mᵃ-nº] (*firdirik hutman*: *Frederik de Houtman*). In a few other instances diacritics are provided to indicate the exact spelling of the word, which may be due to a lack of familiarity with the word on the part of the writer or copyist or due to his intention to highlight the word for another reason.

If the spelling may be considered rather inconsistent at times in this manuscript, the opposite can be said about the morphology, which quite closely follows common practice of most of the texts comprising the bulk of extant classical Malay narratives.[11] One of the basic characteristics of Malay morphology of transitive verbs is the assignment of roles to agent and object through the use of

9 The local Malay dialect, Ambonese Malay, does not have a schwa (Collins 1980, 18), while in the local Austronesian languages, the schwa was changed into other vowel sounds (see Stresemann 1927, 95–100).

10 I am very grateful to Annabel Teh Gallop for providing me with the photographs of the manuscript held in Berlin. For a description and illustrations of these documents, see Wieringa and Hanstein 2015, 62–65.

11 In this category of traditional Malay narratives I do not include doctrinal Islamic treatises,

a specific form of the verb: when the action is viewed from the agent's perspective, the verb will be prefixed with *mə-* and a nasal which is homorganic with the initial sound of the verbal root. The nasal will precede the initial voiced sound of the root or will replace the initial sound if it is voiceless. For instance, the root *buang* will change into *məmbuang* ('throw away'), while *panggil* will transform into *məmanggil* ('to call'). It has been noticed that this general rule of prefixing voiced sounds is relatively new and that the occurrence of deviating forms in texts is an indication of their age (see Jones 2005). In older texts, we occasionally find forms such as *məmunuh* instead of *məmbunuh* ('to kill') and *məmuat* instead of *məmbuat* ('to make'), but the form *mənəngar* (and not *məndəngar*, 'to hear') is quite persistent and commonly found in texts from the 19th century as well (for an early example see Braginsky 2015, 55, line 2).

The text of *Hikayat Tanah Hitu* somewhat consistently follows this general rule, which forms an indication that the author and/or copyist was well acquainted with the rules of the formal written register of this language. In prefixing the active verb marker *mə*N-, hardly any deviating forms emerge compared to common practices. The text furthermore frequently and consistently uses the morphological possibilities available in Malay grammar to indicate a reciprocal action. In traditional Malay texts, the most obvious instance for this form is in the frequent war or fighting scenes in which the adversaries shoot, stab, hack or curse at each other. This is also the case in *Hikayat Tanah Hitu*, which contains forms such as *sərang-mənyərang* ('to attack each other'), *alah-mengalah* ('to defeat each other') and *tembak-mənembak* ('shoot at each other') as examples in which the second part of the reduplication is affixed while the first part consists of the root of the verb.[12] Another way to indicate reciprocity is by using the circumfix *bər-...-an* with possible reduplication of the root, which we find in examples such as *bərjanji-janjian* ('promise each other'), *bərtikam-tikaman* ('to stab at each other') and *bərsumpah-sumpahan* ('make a vow to each other').

Only in a few single instances do we find examples of more archaic grammatical forms, such as *dipəpatutan* ('to put in order', modern Indonesian *dipatutkan*) and *dibəbohonkan* ('to lie about something', modern Indonesian *dibohongi*).[13] These forms contain a partly reduplicated root form, which is extinct in modern standardised Malay.

---

which seem to follow other standards heavily influenced by the Arabic originals they were translated from or based on.

12 *Sarang-manyarang* is spelled with *alif* in the penultimate syllable, while *tembak-menembak* is spelled with a final –h (ه) instead of a normal *qaf* (ق).

13 Again in the *Hikayat Isma Yatim*, copied in Ambon, we find a similar irregular form, *bəpərsəmbahkan* ('to present'; Braginsky 2015, 55, line 11), while a few lines further down we can find the regular passive form of the word in *dipərsəmbahkan* (ibid., line 17).

5 Conclusion

I need to emphasise that the limited scope of this paper only allows for a perfunctory and preliminary discussion of a topic as broad as the use of Jawi spelling in Malay manuscripts during a period of over three centuries and a distribution over such a vast region. I have given a short survey of points brought forward in previous studies and described some palaeographic and orthographic characteristics of the writing found in a manuscript copied in the mid-17th century in Ambon, which I compared with the writing in a few documents originating from the same period and region.

Older discussions about Jawi orthography mainly deal with the ways in which vowels are represented, since Arabic usually indicates only three long vowels. I have given specific attention to the *tashdid*, frequently occurring in the text of the *Hikayat Tanah Hitu* and in some of the other extant Malay documents, demonstrating a usage which may be in agreement with certain historical developments in geminating consonants and a merger of the schwa into /a/, as noted by James Collins for Berau Malay.

Furthermore, the inconsistencies in the spelling of words I have touched on may reflect phonological characteristics of vernaculars and local dialects which are accommodated by the Jawi script, thereby providing a written link bridging the different repertoires of communication. It does not seem too surprising that the majority of correspondences in the use of *tashdid* in the *Tale of Hitu* were found in texts of the same genre of extended narratives which were also circulating in the Malay world by way of mouth, next to their written form preserved in the extant manuscripts. In the vast cultural realm of the Malay world it may be difficult to find a homogeneous body of texts that provide many indications of standardised traditions, but collections of unstandardised consistencies as accommodated by the Jawi script have certainly proven to be sufficient to communicate through time and space and provide the ambiguity appreciated by interlocutors with different linguistic backgrounds. This shared cultural feature of different literary traditions from such a vast geographical space has certainly assisted in considering their ensemble as one 'Malay' tradition, however heterogeneous this may be.

References

Benjamin, Geoffrey (1993), 'Grammar and polity: The cultural and political background to Standard Malay', in William A. Foley (ed.), *The role of theory in language description*. Berlin, New York: Mouton de Gruyter, 341–92.

Braginsky, Vladimir (2010), 'Newly found' manuscripts that were never lost: three François Valentijn manuscripts in the collection of Muzium Seni Asia (MS UM 81.163)', in *Indonesia and the Malay World*, Vol. 38, no. 112: 419–58.

Braginsky, Vladimir (2015), 'Ambon, 1702' and 'Ambon, between 1707 and 1710', in *Indonesia and the Malay World*, Vol. 43, no. 125: 54–55; 62–63.

Collins, James T. (1980), *Ambonese Malay and Creolization Theory*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Collins, James T. (2006), 'The Malayic variants of eastern Borneo', in Fritz Schulze and Holger Warnk (eds), *Insular Southeast Asia. Linguistic and Cultural Studies in Honour of Bernd Nothofer*, Wiesbaden: Harrassowitz Verlag, 37–52.

Gallop, Annabel Teh (2015), 'A Jawi sourcebook for the study of Malay paleography and orthography: Introduction', in *Indonesia and the Malay World*, Vol. 43, no. 125: 13–39.

Gallop, Annabel Teh and Bernard Arps (eds) (1991), *Golden letters: writing traditions of Indonesia/ Surat emas: budaya tulis di Indonesia.* Londan: The British Library; Jakarta: Yayasan Lontar.

Hashim Musa (1997), *Epigrafi Melayu: Sejarah Sistem Tulisan dalam Bahasa Melayu*, Kuala Lumpur: Dewan Bahasa dan Pustaka.

Jones, Russell (2005), 'On the use of the Arabic tashdid to represent the pĕpĕt (ĕ) in Malay script', in *Indonesia and the Malay World*, Vol. 33, no. 97: 281–92.

Khattab, Ghada, and Jalal Al-Tamimi (2008), 'Phonetic Cues to Gemination in Lebanese Arabic', in *Laboratory Phonology*, 11: 75–76.

Klinkert, H. C. (1947), *Nieuw maleisch-Nederlandsch Woordenboek met Arabisch Karakter* [5th impression]. Leiden: Brill.

Kratz, E. Ulrich (2002), 'Jawi spelling and orthography: a brief review', in *Indonesia and the Malay World*, Vol. 30, no. 86: 21–26.

Laurentius, Theo, and Frans Laurentius (2008), Watermarks, 1650–1700, found in the Zeeland Archives, 't Goy-Houten: Hes and De Graaf Publishers.

Mahdi, Waruno (2005), 'Old Malay', in A. Adelaar and N.P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*. London: Routledge, 182–201.

Milroy, James, and Lesley Milroy (1991), *Authority in Language. Investigating Language Prescription and Standardization*. London, New York: Routledge.

Putten, Jan van der (2015), 'Hila, Ambon, between 1653 and 1662', *Indonesia and the Malay World*, Vol. 43, no. 125: 50–51.

Shellabear, W.G. (1901), 'The Evolution of Malay Spelling', in *Journal of the Straits Branch of the Royal Asiatic Society*, 36: 75–135.

Stravers, Hans, Chris Van Fraassen and Jan Van der Putten (2004), *Ridjali: Historie van Hitu. Een Ambonse geschiedenis uit de zeventiende eeuw*. Utrecht: Landelijk Steunpunt Educatie Molukkers.

Stresemann, Erwin (1927), *Die Lauterscheinungen in den Ambonischen Sprachen*, Berlin: Dietrich Reimer.

Valentijn, François (1724–26), *Oud en Nieuw Oost-Indiën, vervattende een naaukeurige en uitvoerige verhandelinge van Nederlands mogentheyd in die gewesten*. Amsterdam: Joannes van Braam.

Voorhoeve, P. (1964), 'A Malay scriptorium', in John Bastin and R. Roolvink (eds), *Malayan and Indonesian studies. Essays presented to Sir Richard Winstedt on his eighty-fifth birthday*. Oxford: Clarendon, 256–66.

Wieringa, E. P., and Thoralf Hanstein (eds) (2015), *Schrift Sprache/Aksara dan Bahasa. Indonesische Handschriften der Staatsbibliothek zu Berlin Preußischer Kulturbesitz. Ausstellungskatalog*, Berlin, Jakarta: Staatsbibliothek zu Berlin – Preussischer Kulturbesitz; Museum Nasional Indonesia.

Dmitry Bondarev and Nikolay Dobronravin

Standardisation Tendencies in Kanuri and Hausa Ajami Writings

Abstract: The Kanuri and Hausa manuscript cultures had a standard model of the Arabic writing system whose uniformity is grounded in Arabic literacy in its Qur'anic and Classical forms. The scribes of both cultures (Kanuri starting from the mid-17th century and Hausa from the early 19th century) adhered to the principle of total orthographic uniformity in writing the Arabic texts but they used variable orthographic systems for writing in local languages (Ajami). Having been in contact for a long time, Kanuri and Hausa manuscript cultures share a similar type of Arabic script and belong to the Central Sudanic area of Ajami writing characterised by specific graphemic choices for some sounds.

Both Kanuri and Hausa orthographies developed from conservative simplified systems strictly modelled on Arabic letters to a more elaborated encoding of sounds. However, unlike Hausa writing of a later period, Kanuri tradition remained largely unaffected by graphemic innovation. Both cultures have identifiable sets of grapheme-phoneme combinations which were stable within a restrictive range of the phonemic and graphemic inventories. From a diachronic perspective, Hausa writing shows a tendency towards a closer match between the number of phonemes and graphemes, whereas in Kanuri there is a tendency of the retention of the spelling of some high-frequency lexical and grammatical items.

1 Introduction

The expression 'standardisation tendencies' in the title of our chapter might trigger unintended connotations, both positive and negative. On the positive side, 'standard' is usually equated with 'stability' and 'norm', and standardisation is customarily seen as a process of reaching an orderly state out of the chaos of variation. The development of a standard orthography is seen as progress, and as an explicit expression of widespread assumptions 'about the empowerment and liberating effects of literacy' (Blommaert 2004, 645). Various attempts by UNESCO to standardise orthographies for the world's languages speak vividly for the assumed high value of a standard.

On the other hand, the same title may be read negatively by scholars who challenge and problematise the ideologies of 'graphocentrism' which dominate

modern society whereby 'a language is not seen as "complete" unless it has acquired a standard orthography' (Blommaert 2004, 645). The expression 'standardisation tendencies' may thus be understood as an attempt to highlight the 'evolutionary positive' side of otherwise chaotic spelling, and to present writings in Kanuri and Hausa as systems aspiring to improvement by developing proper literacy.

Discussing standardisation tendencies in this chapter, we neither defend nor deconstruct the assumed evolutionary unidirectionality of writing systems towards a standard form (albeit, as will become clear, our approach lies nowhere near any evolutionary model of orthography development).

Our task is to document ranges of orthographic variation and uniformity and, where possible, identify their underlying causes in two writing cultures largely unaffected by European-driven standardisation policies. In doing so, we consider orthography as a system of graphemic combinations conditioned by various factors rather than as a fixed system of spelling rules. Thus, we avoid the more restrictive notion of orthography commonly defined as 'the standardized variety of a given, language-specific writing system' (Coulmas 2003, 35) or as 'the set of conventions for writing words of the language' (Sebba 2007, 10–11). When dealing with diachronic and synchronic variation in a given orthographic system, we first outline a graphemic set for each phoneme of the language and then analyse which linguistic and extra-linguistic factors licence the selection of graphemes for such a set. For example, a phoneme /b/ can be written as <b, f, m, p>, but not as <s> or <k>. Or /g/ can be written as <g> or <k> but not as <b>, or /l/ as <ḍ> or <l, r> but not as < g> or <b>, and so on. The sets of phoneme–grapheme correspondences may have different degrees of stability conditioned by linguistic domains, such as phonology, morphology, syntax, lexicon, prosody, genre, etc. For example, in various Kanuri Arabic-based orthographies, monosegmental and bisegmental consonantal sequences [nasal] – [plosive] (are typically written as single letters used for plosive consonants be it word-initial or word-medial (/**nd**árá/ = <**d**ara> 'where', /ʧasu**nd**ogi/ = <ṯasu**d**ūgī> 'they know') but in a specific morphological environment the nasal consonant may be written as a separate grapheme (/ʧu**nd**ogi/ = <ṯu**nd**ūgi> 'he knows'). The nasal segment can also be explicitly written in the etymological spelling of words borrowed from Arabic (Bondarev 2014b, 128–132).

Part of this approach involves tracing the development of such graphemic sets over time, and examining the components of an identified set as to whether they changed or were reduced to a one-to-one correspondence. Such chronological variations reflect possible phonological changes and/or cultural exchanges between different writing systems, be it in contact areas or under the hegemony

of a prestige orthographic system. Chronological graphemic variations may also account for the emergence of fixed restrictive conventions, or of a standard.

From this point of view, we consider orthography not in terms of the 'standard versus nonstandard' dichotomy, but rather as a combination of tendencies, not necessarily unidirectional. Thus, we can talk about patches of standard spelling within a system of internally organised sets conditioned by a multiplicity of linguistic and extra-linguistic factors. One set may be leaning towards a one-to-one representation, while the other may expand the scope of graphemes and thus become more variable.

A good example of such changing orthographic practices in contact manuscript cultures involves the Arabic-based orthographies used in various historical periods for writing Kanuri and Hausa. This chapter is an attempt to compare the variation and consistency of general orthographic tendencies in Kanuri and Hausa manuscript cultures and to identify the graphemic sets particularly prone to standardisation.

2 Kanuri and Hausa: shared and different history

The Kanuri and Hausa manuscript cultures have long been in contact with each other. Both were situated in the Sahelian region in what is now northern Nigeria and its neighbours to the north and east, and both grew out of early sub-Saharan Islamic polities. The Kanuri manuscript culture is rooted in the ancient Kanem-Borno whose ruling dynasty was one of the earliest in the Sahel to adopt Islam. The early Muslim elite society of Kanem can be traced back to the 11th century, and from the late 12th century onwards we find accounts of 'considerable development of Islamic learning in Kanem' (Hunwick 1995, 16). The Arab written sources from the 12th to the 16th century as well as local 16th century accounts about scholarly activity in Kanem-Borno can be taken as indirect evidence of the antiquity of the Kanuri manuscript tradition. However, the earliest manuscripts belonging to the Kanuri manuscript culture only go back to the 17th century with one manuscript dated 1669 (Bivar 1960, Bondarev 2006, 2014a).

The legendary history of Islam in Hausaland starts with the arrival of foreign scholars such as al-Fāzāzī (d.626/1230) or al-Maghīlī (d.909/1503–4 or 910/1504–5) travelling through Borno or from ancient Mali (Lippert 1900; Palmer 1908; Starratt 1993). More Islamic books were reportedly brought to Kano by the Fulani from Mali in the reign of Yakubu, son of Abdulahi Burja (dated by Palmer as 856–867/1452–1463 [Palmer 1908, 76–77]). Local written culture in Arabic, also linked with migrant scholars, already existed in Kano and Katsina in the

17th century (for more details see Hunwick 1995). However, the earliest substantial manuscript data related to Hausa goes back only to the first decades of the 19th century (see e.g. Hiskett 1975, 18). Thus, the development of the Kanuri manuscript culture can be observed over a time span of about 350 years while Hausa manuscript culture is only observable over the last 200 years. This makes any comparative study of these two cultures slightly anachronistic.

Such comparison is also complicated by historical interference created by the Western and Central Sudanic cultures between which the Hausa manuscript tradition developed. In the 15th century, the Hausa cities Kano and Katsina came into the sphere of activity of the Wangara (Wangarawa) – the merchants and clerics who came from the west, from ancient Mali and Songhay polities. It may be argued that the Wangarawa in Hausaland belonged to the Jahkanke scholarly lineages, ultimately stemming from the Soninke-speaking communities in what is now western Mali; however, their linguistic identities changed over time.[1] According to one hypothesis, the Wangarawa who reached the Hausa city-states in the 15th century spoke a Mande language (possibly Jula, see Al-Hajj 1968, Akinwumi and Raji 1990); according to another, they spoke Songhay (Lovejoy 1978). Later, with the expansion of Borno in the 16th century, the Hausa became increasingly exposed to Kanuri culture, including Kanuri administrative systems and their manuscript tradition, especially discernible in the script style. Sometime in the 18th century, Hausa states again came under western influence, this time from the Toronkawa Fulani Muslim scholars among whose circles reformist ideas developed into the jihad movements, ultimately resulting in the creation of the Sokoto Caliphate in the early 19th century (Hiskett 1957, Levtzion 2000, 83–86).

Given these connections with various scholarly Muslim communities, it is safe to suggest that the Hausa manuscript tradition developed in contact with cultures both to the west (Wangarawa, Fulani) and to the east (Kanuri), probably with varying influence from each side. For example, the most common Hausa script style (defined as 'Kanāwī' or 'Hausāwī') derives from the Borno calligraphic tradition (Brigaglia and Nobili 2013), whereas some Sokoto scribes were apparently more familiar with the Western Sudanic and Saharan styles (for a tentative classification see Nobili 2011). The techniques of glossing the Arabic texts in Hausa, e.g. the marking of glosses as *ʿjm* (Ajami), were possibly influenced by the Western Sudanic tradition, but this remains an open question.

1 According to Lamin Sanneh, the family name (*jamu*) of 'Abd al-Raḥmān Zaiti or Zagaiti, the head of the Wangarawa who came to Kano, should be read as Jakhite. Sanneh saw the Wangarawa as a Manding-speaking community of the Jakhanke (Sanneh 1989, 32–35).

Interestingly, most of the surviving manuscripts recognised as originating in the Kanuri and Hausa cultures coincide with the earliest attestations of writing in these languages. Therefore, the material side of Kanuri and Hausa manuscript cultures is validated both on linguistic grounds and by the evidence of codicological and palaeographic features. Thus (if we disregard the question of the application of a literary language to wider cultural and regional domains), when we see Hausa or Kanuri in manuscripts, it is typically the most solid evidence for the existence of these manuscript cultures, whereas script style is much less indicative as a feature of attribution, and layout even less. This is still more relevant for Hausa, because extant Hausa manuscripts come from a vast array of cultural and sociolinguistic contexts ranging from the 19th-century Caribbean and Brazil to West and North Africa.

3 Kanuri literacy in Arabic script

There are several sociolinguistic dimensions to the study of Kanuri texts in Arabic script, or Kanuri Ajami. The earliest written evidence of literacy in the area northwest of Lake Chad comes from annotated Qur'an manuscripts dating to the 17th century and produced in the Borno Sultanate, what is today northeast Nigeria and southeast Niger. The annotations were written in a distinct linguistic variety called Old Kanembu or Tarjumo, closely related to modern-day Kanuri (Bondarev 2013a, b, 2014a, b). Kanembu, as part of the term Old Kanembu, refers to a prominent dialect cluster of Kanuri spoken around northern, northeastern and eastern areas of Lake Chad. This was the region of the Kanem Sultanate, or Kanem – an influential Islamic polity established around the 12th century. Kanem is considered to be a political and religious predecessor to the Borno Sultanate and the annotated Qur'anic manuscripts are the earliest surviving witnesses to literacy practices in the domain of Islamic education spread between the region of Kanem and Borno. The linguistic variety used in the annotations of the Borno Qur'ans took on the wider role of a specialised language used exclusively for the purpose of translating Arabic texts at various intermediate and higher stages of Islamic education. This tradition still survives in northeast Nigeria, albeit diminishing under the pressure of a growing influence of non-Kanuri speakers and of reformist-oriented religious leaders.

There is some inconclusive evidence that Old Kanembu had developed into a classical language used for composition and unrelated to the translation function. Thus, a five-line verse poem (visually resembling a classical *takhmīs* style) in a variety of Kanembu found by Bondarev in Mao, Chad, in 2011 (Fig. 1) shows features which are typical of Old Kanembu in the Borno Qur'ans. However, there

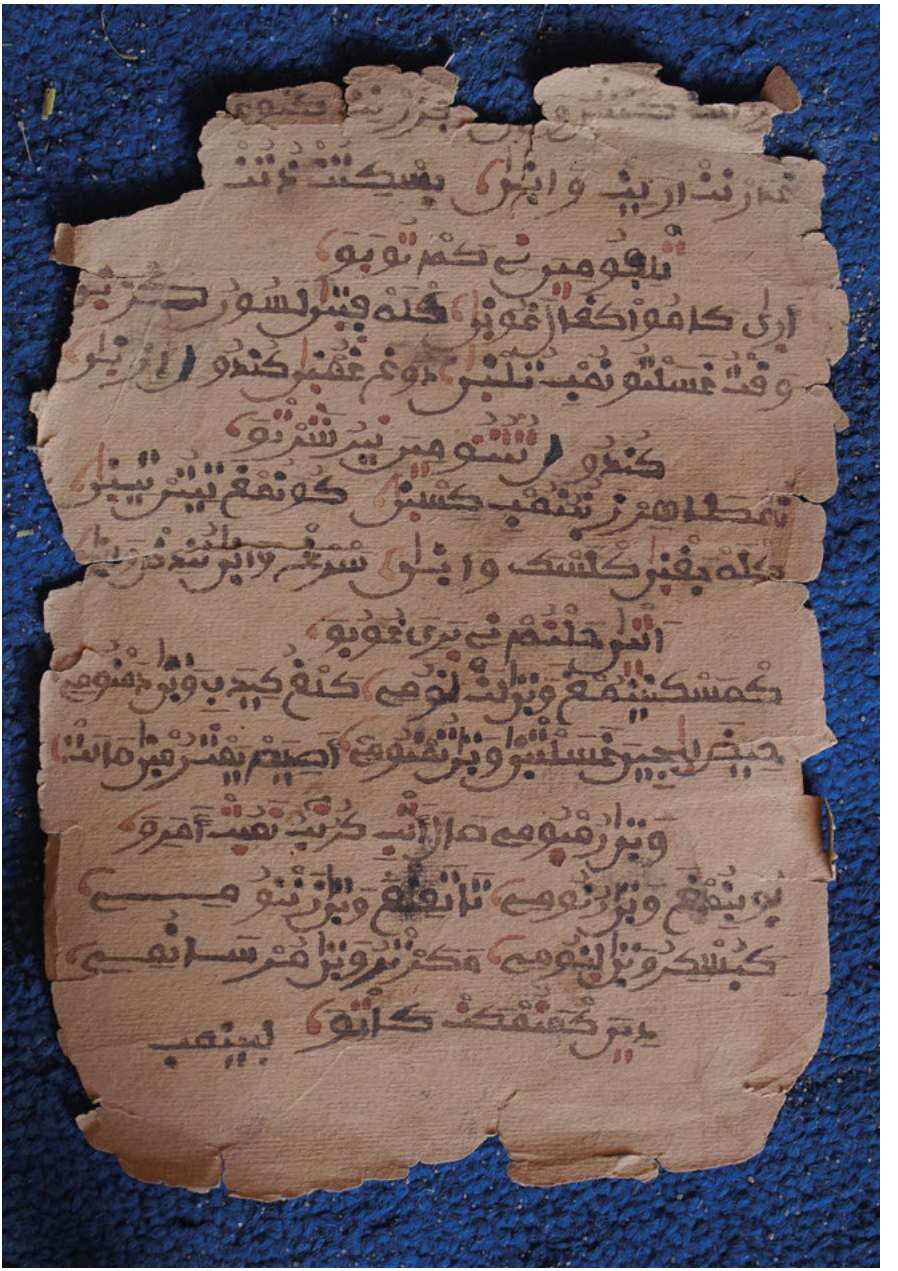


Fig. 1: Poem in Kanembu on religious and secular duties. Mao, Chad, Mala Abbakar Abdala's collection. Courtesy of SOAS University of London, SOAS Digital Collections, Borno and Old Kanembu Islamic Manuscripts, *qasida* in Kanembu in *wuʿāẓ* genre.

are many features in the language of the poem which set this variety apart from both Old Kanembu and Kanembu dialects, giving the impression of a pan-dialect idiom that could only have been used for literary compositions.

Another dimension of literacy in Kanuri Ajami has to do with writing in the western (non-Kanembu) varieties of Kanuri dialects. The earliest evidence for writing in the western dialects of Kanuri dates to the beginning of the 20th century and it comes from historical records, mostly dealing with the lists of the Kanem and Borno rulers and a summary of their deeds. This genealogical and historical genre known in Kanuri as *gargam/girgam* might have existed in written form as early as the 1790s (Bondarev 2014a, 120–121). All known copies of *gargams* originate from the area of Borno, situated west of Lake Chad, and account for the use of (western) Kanuri dialects. To date, no written *gargams* from the east of the historical Kanem-Borno (in what is now Kanem province in the Republic of Chad) have come to light, and it is impossible to say whether the same *gargam* genre was written in a variety of Kanembu.

A distinct corpus of writing in Kanuri consists of Ajami manuscripts commissioned by pre-colonial European explorers and colonial administrators. In terms of their relevance to the study of spelling conventions in Kanuri and Kanembu manuscripts, the most interesting are those Ajami texts written by those Kanuri speakers schooled in the traditional system of literacy based on Arabic and Islamic learning. This type of Ajami is represented in the short folktales published in 1911 by Philip Askell Benton, a colonial officer (formally called Assistant Resident) who collected the texts some time at the beginning of the 20th century (Benton 1911; treated in Bondarev 2014b, 132, 139–140). Ajami writings commissioned by Europeans and transcribed by scribes who did not belong to the Kanuri-speaking cultural area are of lesser relevance to the present study because these scribes had their own spelling conventions which were idiosyncratic or influenced by other cultural areas. The earliest such examples are Kanuri texts and word lists collected by James Richardson, an English traveller, during his expeditions to sub-Saharan Africa between 1848 and 1851 (Bondarev 2007, 67–68). Specimens of Kanuri Ajami written outside the Kanuri and Kanembu speaking cultural areas were also published by the German scholar Rudolf Prietze (1914, 1930).

In more recent times, Kanuri writing in Arabic script has remained visible in two distinct genres: translations of Arabic texts, and religious didactic poems. Translations are written in a variety of Old Kanembu, a continuation of the practices found in the commentaries of the ancient Borno Qur'anic manuscripts. The Islamic didactic poems are written in a literary variety of Kanuri, possibly similar to the registers used in the *gargams* and folktales. The texts of both genres are circulated in printed market editions – cheap facsimile reproductions of handwritten originals.

3.1 Kanuri orthographic conventions

Tendencies in the orthographic conventions of the Old Kanembu Ajami writing used in the Borno Qur'ans have been described in Bondarev 2006, 2007, 2014b. The present section gives a short summary of previous findings in the study of the orthography of Old Kanembu: consonantal graphemes and a quite detailed analysis of the development of vowel encoding, followed by a presentation of general graphemic features that withstood the test of time and regional differences.

The scribes who wrote in earlier varieties of Kanuri and Kanembu were conservative users of the Arabic script adapted for the languages of West Africa. Unlike many other Ajami graphic systems which gradually developed special signs for their non-Arabic phonology (see a comprehensive list in Mumin and Versteegh 2014, 1–22), Ajami writing in Kanuri and Kanembu remained graphemically minimalistic even in the more recent manuscripts and facsimile editions. Thus, there have been no graphemic innovations throughout the whole history of Kanuri and Kanembu Ajami writing outlined in the previous section.[2]

3.2 Consonants in Kanuri Ajami

One possible reason for graphemic conservatism is veneration of the Arabic script as the written emanation of God's word – a factor difficult to prove for earlier manuscripts (due to the absence of written testimonies left by the Borno scribes) but impossible to rule out in a Muslim society. The question of veneration aside, the lack of graphemic innovations might plausibly be accounted for by the phonological system of Kanuri, especially in the domain of consonants. In the different analyses, and depending on the Kanuri dialect, the number of consonants ranges from 19 (Cyffer 1998, 19) to 21 (Hutchison 1981, 17) to 25 (Lukas 1937), and 24 in Mowar (Bulakarima 2001, 42). In Kanembu dialects, Jouannet (1982) postulates from 20 to 26 consonants for Ngaldoukou Kanembu and Lukas (1931) 26

2 There may be slight evidence to the contrary. In 1885, Jean Marie Le Roux met a Kanuri-speaker (from Kukawa) in Algeria who wrote several pages at his request. According to Le Roux, 'He wrote from top to bottom. His writing bears a lot of resemblance to Arabic writing; but it should be noted that it employs a greater number of letters than does the Arabic alphabet'. (Le Roux, 1886, x, fn1) (Ce nègre écrivit, en ma présence, quelques pages que j'ai conservés. Il écrivait de haut en bas. Son écriture a beaucoup de ressemblance avec l'écriture arabe ; mais il est à remarquer qu'il emploie un plus grand nombre de lettres que ne comporte l'alphabet arabe.) It is however possible that the French author was just confused by the peculiarities of the Central Sudanic writing style.

consonants for Kaidi Kanembu. For Old Kanembu – the earliest written variety of Kanuri and Kanembu – 23 consonants have been reconstructed (Bondarev 2014b). Compared to these numbers, Arabic has 28 consonants represented in writing by 28 graphemes. Of the lesser numbers of Kanuri / Kanembu consonants, only two are potentially problematic for choosing an appropriate Arabic letter. One is a bilabial plosive /p/ (occurring only in syllable final position) and the other is a voiceless alveolar affricate /ʧ/ (or in earlier text /ts/). The consonant /b/ is written either as <b> or <f> and the consonant /ʧ/ as <t̲>.

Other choices of Arabic consonantal letters are based on almost straightforward correspondences between the remaining Kanuri consonants and homorganically similar Arabic consonants. The fact that Kanuri has four prenasalised stops *mb*, *nd*, *n̲dʒ* (**n̲dz* in Old Kanembu) and ŋ*g* (analysed differently as monosegmental or bisegmental) which are absent in Arabic does not complicate orthographic choices, because these consonants can be written with the existing Arabic letters either as monographemic non-sonorant consonants or as digraphs. Monographemic encoding of prenasalised stops is the most typical convention so that *mb*, for example, is written as <b>. Digraphic orthography is usually applied for nasal-plosive sequences when they occur at syllable boundaries (Bondarev 2014b, 128–32). Such sequences are perfectly normal in Arabic and writing them in Old Kanembu or Kanuri does not require any graphemic innovation. Thus, the sequence /n-d/ is written as <nd> both in Old Kanembu/Kanuri and Arabic, for example, *andī* 'we' and *ʿinda* 'at' respectively.

Given that the choices of appropriate consonant letters were not complicated by any implosive or glottalised consonants (nonexistent in Kanuri, unlike Hausa) and given that the number of Arabic graphemes is higher than the number of consonants in Kanuri or Kanembu, there was no apparent need for Kanuri and Kanembu scribes to invent new graphemes.

The consonant letters which stand for the Arabic phonemes absent in Old Kanembu were not redundant, rather they were utilised in two different ways. One was the retention of Arabic spelling for Arabic loanwords, and the other was free spelling variation. These nine letters are *ṣād*, *ḍād*, *ṭāʾ*, *ẓāʾ*, *ʿain*, *qāf*, *xāʾ*, *ḏāʾ*, and *šīn*.

Retention of Arabic spelling in Old Kanembu writing is illustrated here by the letter *ḍād*. One of the most frequent words in Old Kanembu is <larḍ> /lárdə/ from Arabic *al-ʾarḍ* 'earth'. This is typically written above or in close proximity to the Arabic source word. Another example is the verb <ḍaro> 'to harm' from Arabic *ḍarra*. Similar to <larḍ>, <ḍaro> occurs in visual proximity to the Arabic *ḍarra*. However, it is also often written above the Arabic *ẓalama* 'to do wrong; to harm'. This means that the spelling of Old Kanembu <ḍaro> was not simply dependent on visual correspondence with the Arabic graphemic source word but rather was

an abstract orthographic rule which could be reconstructed as prescribing that 'Arabic loanwords be written with the relevant Arabic consonants' (an etymological spelling comparable to the orthography of the English loanwords *cliché*, *haute*, *kitsch*, or *oeuvre*).

These same nine letters are sporadically used in free variation with the letters encoding homorganic phonemes or phonemes which have a similar manner of articulation. For example, in Old Kanembu writing, *šin ~ ṯāʾ ~ sīn* are used for /s/, *ẓāʾ ~ ḏāl ~ zāl* for /z/ and *qāf ~ kāf* for /k/.

In the earlier Old Kanembu manuscripts these spelling variants are much less common than in the later ones. The Borno Qur'an manuscript MS.1YM (most probably written before the end of the 18th century) is a good example, showing more frequent variation than the Borno Qur'ans of the 17th-century.

One of these nine letters, *šīn*, becomes more stable in the manuscripts of the 19th and 20th centuries, in the sense that it is not used as a variant of *sīn* to encode the phoneme /s/ but is only used to represent /š/. The emergence of *šīn* as a letter in its own right was most probably conditioned by emergence of palatal consonants in Kanuri (e.g. **sin* > *šin* 'eye', see Bondarev 2014b).

Approximately around the same post-18th-century period, the letter *ḍāʾ* – hitherto a rare variant of *zāl* for /z/ – became much more prominent in the Old Kanembu of the Qur'an manuscripts, in the Kanuri of the kings' lists (*gargam*), in Kanuri writings commissioned by colonial officers and scholars, and in the Tarjumo and Kanuri didactic poems of the late 19th century. It is still unclear what conditioned this increase in frequency.

3.3 Vowels in Kanuri Ajami

As is typical for other sub-Saharan Ajami traditions, the vowels in most Kanuri and (Old) Kanembu texts are written with the vocalic diacritics borrowed from Arabic vocalised writing. Some rare examples of writing vowels in *plene* – by means of the three Arabic letters used for long vowels – have been found in an Old Kanembu manuscript in Leeds University Library (MS 357, Bondarev 2007, 69). At present, we do not know whether this spelling convention existed alongside the more typical diacritic-based spelling of vowels or whether it was a one-off innovation.

The number of vowel phonemes in Kanuri and Kanembu ranges from 6 to 11 depending on dialect and/or analysis: 6 vowels in Mowar Kanuri (Bulakarima 2001); 7 vowels in Yerwa Kanuri (Hutchison 1981, Cyffer 1998); 11 in Ngaldoukou Kanembu (Jouannet 1982). The number of vowels in Old Kanembu cannot be established conclusively, but there were at least six (i, e, ə, a, o, u). Other possible numbers of vowels may have been seven, nine or eleven. Given that the six pho-

nemes are clearly distinguishable in Kanuri and (most of the) Kanembu Ajami texts, we will only be dealing here with orthographic conventions used in writing these vowels. Of the six, the most ambiguous are /o/ and /u/ since they are underspecified, being represented by the same diacritic sign *ḍamma*, used in Arabic for the short vowel /u/. However, even in the earliest extant manuscripts, the /o/ in word final position was often differentiated from /u/.

The most typical graphemic solution for encoding /o/ in final position was a combination of the *ḍamma* with the letter *wāw* followed by the letter *alif* surmounted by a *sukūn* sign. The earliest spelling of this type occurs in the Qur'an manuscript MS.3ImI in the annotations written before 1669. For example, in *kasikō* 'it will be' (fol. 13r; Q.2:137). 

Another manuscript of a similar age (written before 1689) has the same orthography for final /o/. One can see a *sukūn* above *alif* in *tatarṭō* 'he will enter' (Q.84:12; MS.Arabe 402, fol. 250r). 

The same spelling is found in the Qur'anic manuscript MS.2ShK (e.g. Q.2:19, and is very consistent throughout the manuscript) and in three other early manuscripts such as MS.4MM, MS.Konduga, and MS.Kaduna.AR.33-1. The historical *gargam* texts written in Kanuri show the same *ḍamma-waw-alif*-sukun encoding of final /o/ (MS.H 279–282, Bondarev 2014a).

Nevertheless, variations in spelling do occur in most of the manuscripts. One common spelling is marking /o/ with the same combination of letters and diacritics but without a *sukūn* above *alif* as in MS.Arabe 402 *sasikō* 'they will be (in hell)' Q.82:14, fol. 249r. 

A less frequent variant of the previous spelling of final /o/ is using only *ḍamma* and *wāw*. Although potentially ambiguous with /u/, it is very consistent in MS.Kano.Tahir.

The combination '*sukūn* above *alif*' is undoubtedly modelled on an orthographic rule in the Qur'an manuscripts, namely, marking the *alif* as a silent letter in the ending of the perfect forms of the verbs in the 3rd person plural, e.g. *qālū* 'they said'. The frequency of this form in the Qur'an is very high and many Old Kanembu glosses are written next to such verb forms, creating an orthographic harmony between the spelling of the Qur'an and that of the commentary in Old Kanembu. For example, in the dated Qur'an MS.3ImI, the Qur'anic phrase Q.2:71 *wa mā kādū yafʿalūn* 'they almost did not do it' or 'they were on the point of not doing it', where the underlined verb *kādū* is written with the final *alif* surmounted by a *sukūn*, is translated into Old Kanembu *bukiyā kisadiro tadikibū*, with the final vowel in the last word written exactly as in the Qur'anic *kādū* (fol. 8r; Q.2:71).[3] 

3 It is unclear whether *kisadiro* is written with *ḍamma* above *rā'* or with a *sukūn*.

This particular case can be ambiguous because the negative suffix *bo* might have been pronounced [*bu:*] as in Tarjumo, but other instances of such spelling clearly correspond to final /o/ as in the examples above.

A peculiar type of marking /o/ in final position is found in MS.Konduga, where the letter *hāʾ* is vocalised by a *ḍamma* (e.g. fol. 18r; Q.2:199) <walsikin-ruhu> corresponding to what can be reconstructed as **walsikinr*ō 'from where they emerged', the final *-ro* possibly being an adverbial marker).

Some lexical items with final /o/ (such as in *agō* 'thing') were written with reduced notation (using only *ḍamma* + *wāw*) and are found as early as the 17th century in MS.3ImI. The spelling of this word has been consistent throughout the history of Ajami in Old Kanembu and Kanuri, starting from the Borno Qur'ans of the 17th and 18th centuries to the early 20th century (as seen in Benton 1911, VII). However, some items with final /o/, for example the indirect marker *-ro*, were written in an underspecified manner, that is by using only *ḍamma* – the sign for either /o/ or /u/. What is remarkable is that, in all known manuscripts, *-ro* was never written differently, which shows that the underspecified spelling of *-ro* was highly conventionalised across time and manuscripts.

In general, the final vowels in graphic words (not necessarily matching the lexical units) had the highest tendency for standardisation, on condition that these words occur frequently. One such item is the 3rd person singular suffix *-i* in perfective verb forms which tends to be written with the letter *yāʾ* (*gulji* 'he said'). Another item is the adverb *afī* (reconstructed as *abí*, modern form *awí*) 'what' having already been standardised in the 17th-century manuscripts ( Q.82:18, Arabe 402, fol. 249r).

It is interesting that, starting from the earliest manuscripts, most of the orthographic conventions were already in place. Indeed, the spelling principles of many frequent items did not change over time, including *agō* 'thing', *afī* 'what' / 'which', *sikī* 'there is' (e.g., MS.3ImI Q.3:78; MS.2ShK Q.2:157), demonstratives *tī* 'that' and *nī* 'those', independent pronouns *hū* 'I', *nī* 'you', *tī* 'he/she', *handī* 'we', *nadī* 'you (pl.)' and *tandī* 'they'.

One significant exception to the spelling of the final /i/ in the pronouns is MS.3ImI, where this segment is written in four different ways. The first (and infrequent) spelling uses a single vocalic sign *kasra* (e.g. *tandi* 'they') and the three other spelling combinations are based on the letter *hāʾ*; these *hāʾ*-based spellings are as follows: (1) *kasra* followed by the letters *yāʾ* and *hāʾ* (*tandīh*); (2) *kasra* followed by *yāʾ* and the letter *hāʾ* vocalised by kasra (*nandīhi*); (3) *kasra* followed by *h*āʾ and *kasra* (*tandihi* 'they', *nandihi* 'you pl.').

1.  *tandīh* (Q.3:20–21)

2.  *nandīhi* (Q.3:28)

3.  *nandihi* (Q.2:233)  *tandihi* (Q.3:91)

The word final *hā'* deserves special attention because of its several usages in Old Kanembu and Kanuri writings. Beside the above examples where the use of *hā'* could have been motivated by phonological features such as high tone or length of the vowel, the same (earliest) manuscript MS.3ImI has the letter *hā'* for demarcating the terminus of the graphic (and possibly also prosodic) unit *isnuyi-ka* 'to the deceased' written as <isnuyikah>: *isnuyika* (deceased.DO) '(Allah will bring life) to the deceased ones' Q2:73.

This 'otiose' use of the letter *hā'* is less common in some manuscripts and more common in others. Some pages of MS.2ShK give the impression that there was a strong tendency for *hā'* to become a standard convention for marking the final demarcation of lexical and phrasal units. The idea that something was going on in the scribes' minds about the usage of this letter in final position can be inferred from the variant spellings of the final graphemic segments. Firstly, *hā'* is most frequently marked with a *sukūn* (a zero-vowel sign), but sometimes it is left unmarked as in later manuscripts. Secondly, the items which do not have *hā'* at the end are all function words, grammatical markers and pronouns, such as the direct object clitic *ka*, indirect/adverbialiser clitic *ro*, ablative postposition *kan*, definitive determiners *tí* <tī> 'this' and *aní* <anī> 'those'. The only grammatical items written with the final *hā'* (rarely with a *sukūn* above) are the subject marker *-yi* <yih> and the genitive *-bi* <bih>. Thus, the tendency in the manuscript MS.2ShK is as follows: the final *hā'* is used in content words. Exceptions are the content words with final /o/ which are written in "full" orthography (*ḍamma-wāw-alif-sukūn*).[4]

Two grammatical elements – the subject marker *yi* and genitive marker *bi* – are written with *hā'* at the end. No other function word ends with the letter *hā'*. One representative example of this spelling convention in MS.2ShK is the recto of folio 13 corresponding to Q.2:120–125.

4 The use of final *hā'* with or without *sukūn* may be influenced by the various ways of writing wordfinal /-a/ in Arabic. The same explanation may apply to such uses of *hā'* in Hausa manuscripts from the 19th century up to the present (see below).

This (leaning towards an) orthographic rule of writing the letter *hāʾ* in MS.2ShK does not feature in other manuscripts. For example, in the Borno Qur'an manuscript from Konduga (MS.Konduga), *hāʾ* is used for tagging the end of a content word, except for the content words with final back vowels *u* and *o*. But *hāʾ* is also used to signal the end of most of the grammatical and functional items, including the subject marker *-yi* <yih> and the genitive *-bi* <bih>, but excluding the partitive clitic *kami* 'from within'.

In the genealogical *gargam* manuscripts of the late 19th century, *hāʾ* is not used at all, with the exception of MS.H 282, where the diphthong *ai* in the word *mai* 'ruler' is written with a final *hāʾ*<mayh> (the same encoding of the diphthong in the word *mai* is also attested in the earlier Qur'an manuscript MS.1YM).

In more recent manuscripts of the qasida genre there seems to be a tendency to write the final *hāʾ* when there is phonetic (non-phonological) aspiration after *a* and *o* at the end of each hemistich. This is shown in the following line of a didactic poem on the fundamentals of Islam.

<Allah mejī din-ro noʿatah̲ gabtuma t̲amun numbud̲nah̲>

Alla méjī dînro noáta, ngabtáma sámun nəmbuzána

'The eternal existence of God is known, (it is) like (something) remaining endlessly'

(al-Barnawi 1997, 2)

The hemistich final *e*, *a* and *u* that have prosodic lengthening (i.e. phonetically long vowels) motivated by poetic intonation, and also final *i*, are typically not marked by the letter *hāʾ*.

Writings in Kanuri, Kanembu and Old Kanembu belong to what we call the '*imāla* type' of Ajami; this is manifested in a specific representation of /e/. *Imāla* is a dot below the letter indicating a fronted /a/ pronounced closer to [e] in the Qur'anic Arabic of the Warsh variety, i.e. the variety used in the Qur'anic reading tradition transmitted by ʿUṯmān ibn Saʿīd al-Quṭbī.

In the Old Kanembu of MS.2ShK, *imāla* indicates a vowel /e/ (e.g. Q.2:137). The combination *imāla* + *alif* in the form of *yāʾ* (*alif maqṣūra*) + a short *alif* above may have been used for either prosodic long [ē] or a high tone /é/ (Q.2:140). Given the temporal proximity of this manuscript to the dated (1669) MS.3ImI (Bivar 1960, Bondarev 2014a), this is probably the earliest attestation of the use of *imāla* for this purpose. However, in many other cases what corresponds to /e/ in the later manuscripts (e.g. MS.1YM) or in modern Kanuri, Kanembu and Tarjumo is written using *kasra*, which is the default grapheme for /i/. For example, the agentive marker *ye* and genitive *be* are always written as <*yi*> and <*bi*>.

In contrast, the scribe of the dated MS.3ImI writes *kasra* for what corresponds to *imāla* in MS.2ShK or to /e/ in all known modern varieties. Nor does the 17th century manuscript (MS.Arabe 402) have *imāla*. This is interesting because fronted /a/ in the Qur'anic Arabic transmitted in the Warsh version is written with *imāla* and it is unclear why the scribes of these two early manuscripts did not copy the diacritic sign for a similar sound in Old Kanembu (as many other scribes did). It would be tempting to assume that, historically, there was no /e/ in Old Kanembu and thus it was absent in the language that served as the source for this learned variety. This is however hardly plausible. Firstly, all known Saharan languages to which Old Kanembu belongs have /e/. Secondly, the scribes of the manuscripts contemporary with these two do use *imāla* for the same lexical and grammatical items.

Out of the six vowels that can be reconstructed for Old Kanembu, the mid-central schwa /ə/ has the most unambiguous orthography. It is typically written with *sukūn* (zero vowel). This is found in all later manuscripts and types of Kanuri Ajami.

3.4 General tendencies in Kanuri orthography

As can be seen in the examples described above, orthographic conventions tend to stabilise in one single manuscript and differ from manuscript to manuscript (see also Bondarev 2014a, 145–6, 2014b, 113). One possible factor for the internal regulation (or levelling) of a manuscript's orthography is the shared space in terms of both writing support and the physical location of the manuscript within the same group of scribes. As suggested previously:

> In the past, the manuscripts served a surrogate role of public platforms in the competition of the auditory-vocal and written forms, i.e. when paper was more expensive the written form was visually shared by more people – teachers and students – and so conventions in spelling were due to a collective writing space. (Bondarev and Tijani 2013, 133)

However, there are a number of orthographic conventions which remain the same across manuscripts, irrespective of the time and place of their production. For Kanuri, six such stable tendencies may be identified.

1. Underrepresentation of the nasal segment in prenasalised stops, such as *mb*, *nd*, *n͡dʒ* (**n͡dz* in Old Kanembu) and ŋg written as <b>, <d>, <j> and <g> respectively.
2. Retention of Arabic spelling for Arabic loanwords.
3. The letter *kāf* used for *k* and the letter *ghain* for *g* (as in Hausa Ajami discussed in Section 4 below).
4. Writing schwa ə with a *sukūn* sign.

5. Tendency to distinguish between *o* and *u* at the end of graphic words or graphic phrases (not necessarily matching lexical or phrasal units).
6. Orthographic uniformity of higher frequency lexical and grammatical items.

Some of the conventions listed above are more widespread and found in several writing traditions (for example, underrepresentation of the nasal segment in prenasalised stops), while others are more region-specific and only shared within a literacy contact area (like the letter *kāf* for *k* and *ghain* for *g*, or the Arabic "weak" letters used for suprasegmental features such as the combination of stress and tone); yet other conventions are language-specific (like schwa written with the *sukūn* in Kanuri and Old Kanembu).

Some high frequency items in Old Kanembu could also have been encoded in a more standardised way due to their prosodic prominence at the end of an intonational phrase, as is evidenced in modern day practices of commenting the Qur'an in Tarjumo. Many phrase-final and sentence-final items (e.g. [ʧígí:] 'there is', [gə̀n:] 'in') have such a prominent pitch and length in Tarjumo recitations that the whole practice of commenting the Qur'an is sometimes referred to by laymen as '*cigi* and *gen* recitation'.

One factor accounting for the orthographic uniformity of many lexical and grammatical items in the Kanuri Ajami writings of the later period (e.g. the *gargams* and texts commissioned by Europeans) might have been the scholarly background shared by the scribes:

> the gargams [the late nineteenth century] are written in a careful, sometimes calligraphic hand comparable to [...] the Borno Qur'anic manuscripts [...] The [...] conventions in Benton's texts [1910s] [...] together with the confident calligraphic hands [...] betray their authors' affiliation with the ʿulamāʾ circles and suggest that they were used to write commentaries in a more codified Old Kanembu (Bondarev 2014b, 139–40).

4 Hausa literacy in Arabic script

The development of Hausa literacy in Arabic script, or the Ajami tradition, was, until the 20th century, both successful and marginalised. The marginal character of Hausa Ajami was primarily motivated by the position of Hausa as a secondary written language (after Arabic). Marginalisation of written Hausa was both conceptual and literal, as we find many Hausa glosses on the margins of texts written in Arabic. Such glosses are rarely taken into consideration in the descriptions of the manuscripts.

It is thus not surprising that many scholars, starting with such authorities as Mervyn Hiskett, largely ignored the Hausa glosses when describing West African

Ajami writings. This is in strong contrast to the approach dominating Irish or Slavonic studies. In the latter, the early cases of 'native' marginal usage have been extensively studied and generally interpreted as the beginning of the respective local literacies.

Very little is known about Hausa or related Chadic languages before the 18th century, so one has to rely on linguistic reconstructions. In the 16th century, Leo Africanus stressed the role of the so-called Gobir language in the region between the Niger and Lake Chad, that is, between the Songhay Empire and Borno. Gobir was one of the Hausa states, and it is generally assumed, with some degree of certainty, that the Gobir language was an early form of Hausa. Later on, the Hausa people and language were mentioned as Afnu, apparently a Kanuri word (Afunó).[5]

Before the early 19th century, Hausa as a written language was in no comparable to Soninke, Fula or Berber (in any script). In Western Sudanic Africa, the spread of Ajami in scholarly varieties of Soninke, Fula and Wolof might have been influenced by the peculiarities of historical development, especially in the coastal regions where Islamisation encountered an ever growing European presence (with written contracts, treaties, etc.).[6] In Central Sudanic Africa, Old Kanembu (Tarjumo) was commonly used in Qur'anic glosses at that time. The written rendering of words and sentences in Old Kanembu was already a norm rather than an exception. In all such cases, a certain tendency towards standardisation was attested.

Not the slightest evidence of written Hausa can be found for that period. There are a few words (personal and place-names) in Arabic texts, but this in no way compares to the written tradition of the Middle Niger region. Later, there are a few references to the Hausa literature produced in Katsina as early as the 17th century by Wali ɗan Masani (ʿUmar b. Muḥammad b. ʿAbd Allāh b. Nūh al-Barnāwī al-Kashnāwī, or Ɗan Masani(h), b. *c.*1003/1594–5, d. 2 Rajab 1078/18 December 1667).[7] Hausa sources mention only one by Wali ɗan Masani by Yahaya (1988, 38) and this particular work is interesting, in as much as it was recited from memory by a member of Wali ɗan Masani's extended family three centuries after his death, and not from a manuscript. This means that the Katsina scholar could indeed have authored the text, but there is no proof as of yet that the work was ever written down.

5 Tomasz Habraszewski (1967, 63), in his study of a 17th-century vocabulary of Kanuri collected by Evliya Celebi, a Turkish traveler, suggested that one word, *gurasa* 'bread' written as *kurasa*, might be Hausa: 'This is a Hausa word, not known to a Kanuri dictionary'. The word in question was a borrowing from Arabic, and is also used in the Sudanese and Chadian/Shuwa Arabic dialects with the same meaning, so it was not necessarily Hausa.

6 Tal Tamari, personal communication. Also see, e.g., Brooks and Mouser 1987.

7 On Ɗan Masani and his works see John O. Hunwick et al. 1995, 29–20.

It is quite possible that this poem and some other Hausa texts had been transmitted orally since the 17th centuries, but the dates for written Hausa (Dan Masani's lifetime, born *c.*1003/1594–5, died in 1078/1667 according to Hunwick et al. 1995, 29) suggested by Yahaya (1988, 31–42) have proved incorrect. The manuscripts referred to by Yahaya are much more recent, although we cannot exclude the possibility that some texts might have been copied from earlier works.

Yahaya's view of Hausa as a written language before the early 19th-century jihad and the creation of the Sokoto Caliphate is similar to that of Gottlob Adolf Krause, a German scholar and pioneer of Ajami studies in the 19th century.[8]

The first dated example of Hausa Ajami (in a multilingual text) was written in the Caribbean diaspora in 1817, although there are a few non-dated poems in Hausa which have been cited as examples of writing before the 19th century. The list of such poems is extremely short, and we do not know if any of these had been transmitted in written form prior to the foundation of the Sokoto Caliphate. In one such poem, known as 'Billahi arumu' and attributed to Muhammadu na Birnin Gwari, Hausa is referred to as *Baubauci*, literally, 'pagan language'. The unknown author also tried to explain why he dared to use this language instead of Arabic, and this type of explanation – even 'defense' – is often found in non-Arabic writings, both in Sudanic Africa and elsewhere in the Islamic world. Such a work, if dated, would mark an important step in the growth of Hausa Ajami, from simple glosses to an authored literary production. However, there is a significant discrepancy in the dates given for Muhammadu na Birnin Gwari's life. According to Hunwick et al. (1995, 233) he flourished in 1850. Yahaya refers to a much earlier date, giving 1178/1758 as the date of his birth.[9]

Turning to the standards of Hausa as a written language since the early 19th century, we can deduce a few general rules. First, until the end of the pre-colonial period there was no uniform tradition of Hausa literacy. Geographically speaking (as mentioned in Section 2), there were two poles of attraction. One variety of Hausa Ajami may be described as Western or Sokoto-centred and close to the written traditions of the Central Niger region. Another variety (mainly in Kano) looks like an offshoot of the venerable centuries-old tradition of Borno (Kanuri) Ajami. Both were reflected not just in their specific conventions of transcription, but also in their respective styles of handwriting. Even today, when buying an Ajami book in Northern Nigeria, one can easily refer to the style used in Kano as opposed to that of the printed Middle Eastern publications in Arabic and, interestingly, that of Zaria Ajami (Sokoto is now only a minor centre of Ajami publishing).

8 Krause 1884, 29, see footnote 10 below.

9 Yahaya 1988, 45.

The east-west division in Hausa Ajami standards is not surprising, given the role of "Western" Songhay-speaking and "Eastern" Kanuri-speaking Islamic scholars and schools in the development of learning in the Hausa states. Looking to the Western tradition, Gottlob Adolf Krause wrote:

> Prior to the Ful (Fulbe) the Hausa possessed a script, also Arabic – with some slight variations, and, if my inquiries are confirmed, the Songhai had written their native language before the Hausa.[10]

This hypothesis has only partially been confirmed, as there are many Songhay words in the historical chronicles written in Arabic, such as *Ta'rīkh al-Sūdān*. The practice of advanced studies in the "East", that is, in Borno and adjacent Hausa cities such as Hadejia, continued well into the 20th century. At the same time, scholars coming from the "French" territory (mainly from Niger and often speaking the Songhay Zarma languages) continued to resettle in Nigeria, establishing their schools and bringing with them the 'western' style of writing and transcription of non-Arabic texts. On the other hand, many Songhay-Zarma speakers came to the centres of Islamic learning in Nigeria to study, so the migration was not unidirectional.[11]

4.1 Hausa orthographic traits and conventions

Except for a very small number of publications, the marginal usage of Hausa in glosses remains largely unexplored in Ajami studies. In the present paper, we are looking at the glosses found in the Arabic manuscripts produced in the 19th century. The results of our research may be summarised as follows:

1. Vocalisation typical for the Central Sudanic cultural area, i.e. a specific vowel-sign for the /e/ vowel;
2. the 'emphatic' Arabic letters are either obsolete or used as symbols for those Hausa consonants which do not exist in Arabic;
3. in a few cases, there is a tendency toward new symbol-creation. However, with the exception of the letter *ṭā'* with three dots, none of these attempts can be seen as really successful;

10 'Vor den Fulen besassen die Haussaner eine Schrift, ebenfalls die arabische mit einigen geringen Abweichungen, und wenn meine Erkundigungen sich bestätigen sollten, so hätten vor den Haussanern die Songhai ihre Muttersprache geschrieben'. (Krause 1884, 29).

11 On the Zarma in Nigeria see e.g. Dobronravin 2000, 91–101; Gulbi and Bunza 2014.

4. the difference between the Western and Eastern conventions is diminishing over time, so that the only survival of it now is the difference in the styles of writing, but not in the transcription of Hausa phonemes;
5. Some conventions, or graphic strategies, do exist, but they are fluid, being more an approximation to a standard rather than a rigid set of norms. This is especially true in the case of length and tone representation, as well as that of labialised and palatalised Hausa consonants.

Each of these tendencies may now be explored in more detail.

4.2 Vowels in Hausa Ajami

In this section, we follow Philip Jaggar's analysis of the Standard Hausa vowel inventory. According to Jaggar,

> Hausa has a 10 vowel system, comprising five basic vowels /i, e, a, o, u/ with phonemic vowel length, in addition to two diphthongs /ai/ and /au/ [...] In medial position in native words, only long /ē/ and /ō/ occur. If the syllable becomes closed by a coda as a result of a morphophonological rule, /ē/ and /ō/ automatically shorten to /e/ and /o/ and merge with centralized /a/ (only short vowels occur in closed CVC syllables).[12]

In the Hausa dialects, the picture is different. Ahmadu Bello Zaria wrote:

> It seems to us quite difficult, if not impossible, to provide a 'comprehensive' study of vowel differentiation across the numerous Hausa dialects examined. This is because vowel oppositions may be quite unstable from dialect to dialect, from speaker to speaker and from utterance to utterance within the same idiolect.[13]

As for more specific variations, according to Zaria, 'final /ee/ in the standard dialect corresponds to /ii/ in the areas around Zaria and Bauchi. This kind of correspondence generally happens after a nasal' (Zaria 1982, 52). Thus, it is not surprising that in many Hausa manuscripts the same word may occur with either /i/ or /e/.

Writings in Sudanic African languages have a strong tendency to retain the vowel-signs. This may be explained by the syllable structure of African languages in the region. Nonvocalised texts are sometimes almost undecipherable; a Hausa

12 Jaggar, 2001, 9–11. Vowel length is not marked in modern Latin-script orthography.
13 Zaria 1982, 183.

text is not easily understood if the vowel signs are absent in words with a CVCV or CVCVCV structure. A few exceptions with nonvocalised Hausa texts are known. However, this is highly unusual in the marginal glosses, where the very purpose of their usage is the clarification of the main Arabic text. Having studied a great number of glossed Arabic manuscripts, we can argue that a Hausa gloss is normally vocalised. As for the main text, if it is written in Hausa, it can be predicted, with a great degree of certainty, that the text will be vocalised. When vowel-signs are not marked, it may be the result of hasty (incomplete) writing. Otherwise, one could suggest that nonvocalised texts served as written supports for the oral transmission of knowledge. In any case, lack of vocalisation is not a rule, and various explanations may be found for it.

With the exception of the diasporas outside Sudanic Africa, the front vowels /i/ and /e/ are marked with different signs, known respectively as *wasali bisa* (Arabic *kasra*) and *imāla* (a loanword from Arabic). As is the case in Kanuri Ajami, *imāla* is a dot below a letter borrowed from the Warsh tradition of Qur'anic spelling. By the 19th century, *imāla* had already been used in Sudanic Africa, in Fula and Kanuri written conventions.

In contrast to the systematic transcription of front vowels, the back vowels /o/ and /u/ were usually marked with the same sign, known as *rufu'a* (in Arabic, *ḍamma*). This is interesting, as the same rule applies to local writings in Fula. On the other hand, Western Sudanic literacy in Fula and a few other languages have an additional symbol for /o/, an inverted *rufu'a*. The question remains as to whether the Central Sudanic usage is archaic, preceding that of Western Sudanic Africa, or whether this is a case of the independent development of two divergent norms of transcription. In any case, Hausa literacy may be described in areal terms as a part of the "non-*o*" literacies in Sudanic Africa. Although the graphic combination of the *ḍamma* with the letter *wāw* followed by the letter *alif* surmounted by a *sukūn* sign occurs both in Hausa manuscripts and in those of the Kanuri/Kanembu area, its use, unlike in Kanembu/Kanuri, is not restricted to encoding /o/. It might be argued that this feature was borrowed from the Borno tradition, but it became more decorative than graphemic.

There were a few attempts to create symbols for *o* in Hausa Ajami in the late 19th and early 20th centuries, but these were apparently rare and definitely unsuccessful. One such example is found in a Hausa manuscript from the Barne collection at Special Collections, SOAS Library, University of London (MS 380271). In this manuscript, the *o*-sound is rendered with a combination of *rufu'a* above the letter and a small *rufu'a* below the same letter. This way of writing was certainly borrowed from the Warsh tradition of Qur'anic reading, but the meaning of the combination in the Warsh was different, being a marker of long vowels in the affixed personal pronoun *-hū*. In the 20th century, a new way of writing the

o-sound was invented, most probably in Zaria. As the Hausa *o* is almost always long, that is /ō/, the combination used includes a supporting *wāw* with a zero-sign (*sukūn*) above it, to differentiate /ō/ from /ū/, which does not require an additional *sukūn* above the *wāw*. However, even now the /o/ is not marked in many modern publications in Hausa Ajami.

An extreme case of fluidity is demonstrated in the marking of vowel-length and tones. The latter were most probably ignored or only recognisable in combination with the length of Hausa vowels. It seems that long final vowels were left unmarked if they were combined with low tone, as in <doki> *(dōkȉ)* 'horse'. The long vowels were usually marked, as in the Arabic written tradition, with an additional *wāw*, *yāʾ* or *alif* respectively for /ū/, /ī/ and /ā/. In the case of long /ē/, two variants of symbol combination were initially applied, either an additional *yāʾ* or a *yāʾ* without dots (e.g. a stump in the medial position) with a so-called 'red *alif*' above the additional letter. The same 'red *alif*' could also be used without any supporting letter. Nowadays, the standard Ajami rule for /ē/ is the use of a *yāʾ* without dots supplemented with a 'red *alif*', which is no longer red in colour.

In a few pre-colonial Hausa manuscripts, usually from present-day Ghana and Togo, one more method of length-marking was used before colonisation. Thus, a combination of *hamza* and *sukūn* could be found above the supporting letters (*wāw*, *yāʾ*, *alif*). Adam Mischlich thought this was how the 'mid-range vowels' were marked.[14] This view has not been proven by any research, and in modern Hausa Ajami no such vowels are known. It cannot be excluded that the now forgotten combination of *hamza* and *sukūn* was in fact used to mark a certain tonal pattern, but this is far from clear.[15]

Until the 20th century, the letter *alif* could also be combined with an additional diagonal line across the main letter (*alif mai suka*, 'alif with spear'). In the glosses, '*alif* with spear' was apparently not used. This variety of *alif* is uncommon in modern Hausa Ajami.

4.3 Consonants in Hausa Ajami

Standard Hausa has 32 consonant phonemes. Among them, the glottalised set is represented with so-called "hooked" letters in modern Latin-script orthography. According to Jaggar, '*ɓ* and *ɗ* are laryngealised (often implosive) bilabial and

14 'Hamza da dámri (Hamza in Verbindung mit dámri) über einem Konsonanten deutet an, daß die Silbe weder lang noch kurz, also mittellang ist.' Mischlich 1906, xxxiii.

15 See also Dobronravin, 2006, 139, n. 36.

alveolar stops, hooked *ƙ* is a glottalised velar ejective, the digraph *ts* [sˀ] is an ejective alveolar sibilant, and the digraph *'y* is a laryngealised palatal glide (derived via reduction of a /dĩy/ sequence)'. Labialised and palatalised consonants are written as digraphs. With few exceptions, as noted by Jaggar, 'the four palatalized /fy, ky, gy, ƙy/ and three labialised /kw, gw, ƙw/ unit phonemes all contrast with the corresponding plain segments before /a(a)/ (/fy/ is a marginal phoneme)'. There are also two 'R-phonemes', the alveolar tap/roll /ř/ and the retroflex native flap r [ɽ], not marked in modern Latin-script orthography (Jaggar 2001, 5–8).

The picture is different in the Hausa dialects, especially in the Western dialectal cluster and in the diaspora. The glottalised consonants in the Western dialects also include. e.g. [ʧˀ], [sʷˀ], [bʸˀ], [bʷ ˀ], [dʷˀ] (Gouffé 1969, Zaria 1982, 50). Most of these additional consonants occur and are phonemic only in word-initial position before /a/. On the other hand, the consonant [ʧˀ] regularly corresponds to the standard [sˀ]; in word-initial position both [ʧˀ] and [sˀ] occur, and the difference between them is phonemic. In the Hausa diaspora, there is a tendency towards the reduction of the glottalised set. In Ghana, the glottalised consonants are replaced with non-glottalised ones, such as /s/ instead of [sˀ] (Zaria 1982, 178). Moreover, there are local differences in the treatment of loanwords. According to Zaria, 'in the area around Daura one finds standard /dˀ/ corresponding to /sˀ/. This kind of correspondence is limited to names which originate from Arabic' (Zaria, 1982, 50).

Turning to the transcription of Hausa consonants, it may be said that, together with the Kanuri writing tradition, Hausa literacy belongs to the '*g-ghayn* area'. It means that, with a few exceptions in the diasporas, the consonant /g/ is invariably written with a *ghayn* (Hausa *angai*). The same graphemic choice is also found in Eastern Fula Ajami. To a lesser extent, the *g-ghayn* area also includes parts of the Songhay region and a few languages in the Northern part of present-day Ghana such as Mamprule and Dagbane.[16] The *g-ghayn* area is opposed to the *g-kāf* area in Western Sudanic and North Africa. In the literacies of the latter regions, /g/ is represented with the letter *kāf*, either as such or with three additional dots to mark the voiced counterpart to /k/.

The 'emphatic' Arabic letters have been used in different ways throughout the history of Hausa literacy. This variety of usage may reflect dialectal variation.

16 This is witnessed in some manuscripts from the collection of the Institute of African Studies, University of Ghana, such as IASAR/28, *ʿAbd Allāh b. al-Ḥājj al-Ḥasan's Niṣāb al-ḏahab*, composed in Dagbane in 1361/1942–43 (Hunwick 2003, 597); or IASAR/54, Alfa Muntaga's poem in Mamprule in praise of Shaykh Aḥmad al-Tijānī (*qaṣīda fī madḥ li-Shaykh Aḥmad al-Tijānī*). The xerox copies of these unpublished manuscripts were consulted by Nikolay Dobronravin in the Herskovits Library of African Studies of the Northwestern University, Evanston, USA.

In Western Hausa dialects there are two different consonants, /ʧˀ/ and /sˀ/, both corresponding to /sˀ/ in modern Standard Hausa and close Hausa dialects (the so-called Kano dialect or dialect cluster). As there were two different types of Hausa Ajami literacy, one of them was probably influenced by dialectal pronunciation.

The closest example of a standard can be seen in the use of *qāf* as opposed to the letter *kāf*. By the end of the 19th century this pair had already been used to differentiate two consonants, plosive /k/ and glottalised /ƙ/ in Hausa.[17] However, before this standard developed, there existed another set of rules for the transcription of Hausa consonants. In earlier manuscripts, both *kāf* and *qāf* are invariably found to denote both /k/ and /ƙ/. The use of *qāf* for Hausa /k/ was then retained in some manuscripts, but only in loanwords from Arabic, such as <*arziqi*> (to be read as *arziki* 'wealth, prosperity', from Arabic *al-rizq*, with assimilated article *al-*) or <*loqaci*> (to be read as *lokaci* 'time', from Arabic *al-waqt*). In modern Hausa Ajami, the use of *qāf* is restricted to the transcription of glottalised /ƙ/, and the historical spelling of Arabic loanwords is normally ignored.

If the differentiation between *kāf* and *qāf* was made relatively early in Hausa Ajami, this was not the case of the other 'emphatic' letters. The letters *dāl* and *ṭaʾ* are nowadays a regular pair representing respectively plosive /d/ and glottalised /ɗ/. In the precolonial Hausa manuscripts and marginal glosses, this opposition was less common. Both consonants could be written as *dāl*. Comparable to the case of *kāf* and *qāf*, the choice between *dāl* and *ṭāʾ* was apparently of little importance for the Hausa scribes, as long as they could recognise the words. In borrowings from Arabic the use of *ṭāʾ* was more consistent than in other strata of Hausa lexicon.

The letter *ṭāʾ* could also be used to denote another glottalised Hausa consonant, namely /sˀ/. As mentioned above, in modern Latin orthography this phoneme is transcribed with the digraph ts. In Hausa Ajami, there was no visible preference for the reading of *ṭāʾ* as /ɗ/ or /sˀ/ before the 20th century, and it seems that such preferences were more individual than regional. Until the 20th century, the letters *ṭāʾ*, *ẓāʾ*, *sīn* and *ṯāʾ* were all used to transcribe the sound which corresponds to glottalised /sˀ/ in modern standard Hausa.

17 Such use of *qāf* for Hausa /k/ in loanwords is found, e.g., in the Hausa manuscript (dated 1879) written in Ghat, in today's Libya, where one can see <qabīla> (to be read as *kabila* 'tribe', from Arabic *qabīla*) and <yākī> (to be read as *yaƙi* 'war'). The Ghat manuscript does not have a special sign for /e/; the final /o/ in verbs is systematically marked with a combination of *wāw* and *alif* with *sukūn*, as in <yā-ḏō>, to be read as *ya zo* 'he came'. As a 'Middle Eastern' diaspora feature, the *qāf* in this manuscript has two dots above, while *fā* is written with a dot above or below the letter (Ghānim 1998).

In the 20th century, with the standardisation of Hausa in Latin script, glottalised /sˀ/ was transcribed with an extra letter, the *ṭāʾ* with three additional dots. Since the 1950s this usage has been strengthened by the publication of Boko(Latin)-Ajami transcription tables.[18] The same extra letter is used in modern books printed in Ajami in Nigeria. This symbol has found its way into the set of additional Arabic symbols in the Unicode of the 2000s. The creation of *ṭāʾ* with three dots is a rare case of a successful introduction of a new symbol in Hausa Ajami literacy. Unlike *ṭāʾ*, the Arabic letter *ẓāʾ* (with one diacritical dot) was only used for /z/; this letter is fairly rare in modern Hausa Ajami.

Two centuries of Hausa Ajami development have resulted in the fixation of symbols for glottalised /ƙ/, /ɗ/ and /sˀ/. The rendering of /ɓ/ and /ˀy/ is another story, or perhaps two other stories. In the pre-colonial manuscripts and marginal glosses, these two consonants were not marked separately from /b/ and /y/, that is, the letters *bāʾ* and *yāʾ* were used to denote them. This lack of differentiation may be explained, first of all, by the absence of corresponding 'emphatic' letters in the Arabic written tradition. No single Arabic letter could be easily selected for the transcription of one of these Hausa consonants. Besides, certain dialectal influence cannot be ruled out. In Western Hausa dialects, /ˀy/ is non-existent, corresponding to the sequence /dĩy-/, as in the pair *dĩyauci*, modern standard Hausa *'yanci* 'freedom' and *dĩya* (singular feminine), corresponding to modern standard Hausa *'ya* 'daughter; (historically), free woman'; (plural) – *'ya'ya* 'children'.

With the development of the Hausa Latin-based orthography, the letters *ɓ* and *'y* (a graphical combination of ' and y) were created, and they are now used in most printed Hausa texts. This standard has also influenced Hausa Ajami. As a result, glottalised /ɓ/ has variously been marked with one, two or even three extra dots below the *bāʾ*. Today's prevalent form is a *bāʾ*-shaped letter with three dots below. This form is still far from convenient, as /ɓ/ may be found in combination with the vowel /e/, resulting in an awkward combination of three plus one dot below the letter. Three dots below were also added to the letter *yāʾ*. In some manuscripts even today neither /ɓ/ nor /ˀy/ is marked with any additional symbols.

Interestingly, in a few manuscripts and printed publications, the glottalised /ˀy/ is rendered with the letter *ʿayn* and three additional dots above it. This is not surprising, if we take into consideration the way the letter *ʿayn* has been read by most Hausa speakers. In both cases the standard pronunciation is that of a glottal stop, marked with a *hamza* in standard Arabic spelling. Less educated and diaspora scribes even wrote 'Allah' with an *ʿayn*, as <*ʿallā*>. Thus, the rendering of

18 See, e.g. 'Ajami Boko' 1961; Wali/Binji 1969.

/ʔy/ with an *ʿayn* plus additional dots may be interpreted as the transcription of a glottal stop with /y/ as an extra feature.

Beside the transcription of glottalised consonants, the development of a Hausa Ajami standard included the transcription of the affricates. In most Hausa dialects, these are /ʤ/ and /ʧ/, corresponding respectively to the letters *j* and *c* in modern Hausa Boko spelling. The use of *c* has long been the norm, but before it was developed, and sometimes even nowadays, the digraph *ch* is also used to denote the same Hausa phoneme.

In Ajami, /ʤ/ has always been represented with the letter *jīm*. Only in one rare diaspora document of 1817 is this letter also used to denote the voiceless affricate /ʧ/, as in <luqaji> *lokaci*. In the same rare document from Trinidad, the phoneme /ʤ/ is also represented with the letter *ḏāl*. It is not clear to what extent this usage was influenced by the Western Sudanic tradition, where there was a more marked confusion in the rendering of affricates. In the great majority of Hausa manuscripts and glosses the letters *ḏāl* and *zāʾ* (also known in Hausa as *zayra*) are used indiscriminately even now.

As for /ʧ/, there were initially two major traditions of its transcription. In Sokoto-centered manuscript practice, the phoneme was written with the letter *šīn*. This usage may have been influenced by the peculiarities of Hausa pronunciation in an environment dominated by Fula speakers. In Kano, as in the manuscripts of Borno, the same consonant was written with the letter *ṯāʾ*. This usage has been retained and has nowadays become the norm.

It is worth mentioning that, in Hausa Ajami, the letter *ṯāʾ* is also found with two more variants of reading. In borrowings from Arabic, and especially in personal names, *ṯāʾ* may be read as /s/, as in <ʿuṯmān> (to be read as Usman). In a few Hausa manuscripts and marginal glosses in the 19th century the same letter was also used to denote glottalised /sʔ/, as in <ṯōrō> (to be read as *tsoro* 'fear'). This usage seems to be totally forgotten now.

A further pair of letters in Hausa Ajami are *rāʾ* and *ḍād*. The latter turned out to be convenient for representing both /l/ and the retroflex *r*-phoneme in Hausa manuscripts and marginal glosses. The use of *ḍād* for /l/ is common in Sudanic Africa and may be explained by the same reading in some parts of the Arab world. As for the retroflex *r*, there were a few Hausa manuscripts where this phoneme was consistently marked with *ḍād*, e.g. *saṛkī* 'king, chief'. In modern Hausa Boko spelling there is no difference between the two *r*-s, and some Hausa dialects do not have the retroflex *r* at all. Thus, in modern Hausa Ajami, the letter *rāʾ* corresponds to both phonemes. As for the letter *ḍād*, it is only found in a few loanwords from Arabic, mainly in personal names.

The flexibility of Hausa Ajami standards is still visible in the treatment of labialised and palatalised consonants as well as for vowel-length and tone. As

mentioned above, in modern Hausa Boko, labialised and palatalised consonants, mainly /kʷ/, /ƙʷ/, /gʷ/, /kʲ/, /ƙʲ/ and /gʲ/, are marked with the respective digraphs, kw, ƙw, gw, ky, ƙy and gy. Other cases of labialisation and palatalisation are much more rare. There were no 'spare' letters in the Arabic script to denote such consonants, and digraphs are not common in the Arabic written tradition. So it is understandable that there was little incentive to underline labialisation or palatalisation of consonants in pre-colonial Hausa writings. In fact, such consonants were marked through the use of 'similar' vowel-signs. In most manuscripts and marginal glosses, the corresponding signs were *rufu'a*, for the labialised consonants, as in <gubru> (to be read as *gwabro*), modern Hausa Boko *gwauro* 'bachelor', and *imāla* for the palatalised ones, as in *gēra* (to be read as *gyara*), modern Hausa Boko *gyara* 'to fix'. In a single manuscript, already mentioned above (*Littafin tuba*, from Barne's collection at SOAS), the scribe did try to mark the labialised consonants with an additional small *rufu'a* below the main letter. This attempt was, however, unsuccessful, and no other manuscripts in this style have been found. In modern Hausa Ajami, influenced by the Hausa Boko standards, both labialised and palatalised consonants have been marked. This was done either by adding three dots to the original letter (*kāf, qāf, ghayn*) or through a combination of symbols. In the latter case, the letter *wāw* is supplemented with a *madda* above it to denote labialisation, as in *kwaɗo* 'frog'. To mark a palatal consonant, the letter *yāʾ* with *madda* is added to it. This approach is now prevalent, but not yet uniform in publications printed in Nigeria.

The use of final *hāʾ* (in some manuscripts both with or without *sukūn*) may be influenced by the practices of writing the word final /-a/ in Arabic. This usage is also found in modern Ajami publications where the same words occur with or without final *hāʾ*, like <ṣallā> and <ṣallāh> for *salla* 'prayer'.

The general tendency in the development of Hausa Ajami may be described as a slow growth of a set of rules.[19] The trend toward the standardisation of Hausa in Arabic script went through significant outside influences in the last century. The British in colonial West Africa introduced an additional letter, *fāʾ* with three dots below to denote /p/ in 'penny'. This symbol was only used on coins and has never been adopted as a part of Hausa literacy standards. A much more sig-

19 The topic outside the scope of this chapter is word breaks. In the early stage of Hausa writing in Arabic script, it was common to see the hyphenation (breaking the word) and disjointed writing of single words. The most consistent feature of this style was the disjoined writing of the words with a medial long /ē/. In such words, the supporting *yāʾ* was left unconnected with the next letter of the word, as in <barē wā> (*barewa*), 'gazelle'. In modern Hausa Ajami such incongruences are very rare.

nificant outside initiative took place in the 1980s, when the ISESCO created an all-embracing system of Arabic-script transcriptions. The ISESCO model completely ignored the traditions of Hausa Ajami, e.g. in the use of *kāf* with additional dots for /g/ instead of the common *ghayn*. As a result, the new transcription was practically ignored by the users of Ajami in Nigeria. Elsewhere in West Africa, the story was slightly, but not significantly, different. A few Islamic and Christian publications appeared in the ISESCO orthography, mostly in Fula in (or for) Guinea, Mali and Cameroon. The current position of the Unicode consortium allows for the further use of Hausa Ajami both offline and on the Internet.

5 Conclusions

In this comparative overview of orthographic tendencies in the Kanuri and Hausa manuscript cultures we have tried to document ranges of orthographic variation and uniformity, identify some of their underlying factors and outline the graphemic sets particularly prone to standardisation.

Both cultures had a standard model of the Arabic writing system whose uniformity is grounded in Arabic literacy in its Qur'anic and Classical forms. Therefore, the concept of regulated uniform orthography was not alien to the scribes who added annotations in the vernacular to the standardised Arabic texts, often written by the same scribe. Nonetheless, the scribes of both cultures (and the claim can easily be extended to most of the Arabic-based writings discussed in this book) did not show any particular adherence to the principle of total orthographic uniformity. The only noticeable restriction on variation is manifested in a conservative use of the Arabic letters, rather than in particular choices of letters and their combinations for representing Kanuri and Hausa sounds. The shared disfavour of invented graphemic designs is particularly observable at the earliest point in time when both cultures can be synchronically compared, that is, in the first part of the 19th century – the period for which we have the earliest extant Hausa manuscripts.

This conservatism is even more striking for Hausa because, unlike Kanuri which has fewer consonantal sounds than the number of consonantal letters in Arabic, the consonantal inventory in Hausa is larger than the graphemic inventory in Arabic (28 Arabic characters vs 23–26 consonants in Kanuri vs 32 consonants in Hausa (Newman 2000, 392)).

We cannot ascertain whether it was only pressure from the phonological system or some other (extra)linguistic factors which shaped later orthographic conventions in Hausa. What is clear in the history of Hausa Ajami writing is that

it developed more consonants to better suit its phonological system, and sets of standard rules gradually increased with time.

The economical approach to the use of Arabic letters in the history of Kanuri writing is comparable to that of Hausa with the only difference being that, throughout the whole history of Kanuri Ajami writing, no new letters were invented. The main reason for that seems to be the smaller number of Kanuri consonantal phonemes. The veneration of the Arabic script – an attitude which cannot be ruled out but which is difficult to prove due to a lack of written evidence – may only partially explain the uninventiveness of the Kanuri scribes. The same respectful attitude to the letters of the Scripture did not inhibit the invention of new characters in Hausa in later manuscripts, or in many other written traditions of Islamic sub-Saharan Africa.

Some similarities between Kanuri and Hausa Ajami in their graphemic choices for writing certain sounds visibly identify both cultures as belonging to the same larger literacy area. Thus, both use <gh> for /g/, unlike the manuscript cultures to the west (for example Soninke) where the most common choice for /g/ is <k>. We call the former the '*g-ghayn*' type of orthography as opposed to the '*g-kaf*' type. Another feature setting Kanuri and Hausa apart from the western cultures is the absence of *tanwīn* signs used to indicate a word-final sequence of a vowel and a nasal consonant, such as *-an*, which, in Kanuri and Hausa writing, would be spelled with the diacritic sign *fatḥa* for [a] and the letter *nūn* for [n]. Since we do not have manuscript evidence of Hausa writing before the early 19th century, it is impossible to say whether Hausa had a different 'western' selection of orthographic choices before having closer contact with Kanuri.

Apart from the similarities, our comparative analysis reveals some contrasting features summarised as follows. One of the most striking differences between Kanuri and Hausa orthographies is the encoding of /o/. For Kanuri, a special spelling of the final /o/ is already attested in the early Old Kanembu manuscripts which is a combination of the *ḍamma* + *wāw* + *alif* surmounted by a *sukūn* sign. Contrastively, the earliest known Hausa Ajami writing belonged to the 'non-*o*' type of orthography and specific spelling solutions for /o/ were only sporadically tried by the scribes. In view of the fact that Hausa manuscript culture was significantly influenced by that of the Kanuri (the similar script style being the most salient feature), the reason why it was not until the 20th century that Hausa scribes introduced a graphemic combination for /o/ similar to that used in the Kanuri manuscripts still needs to be explained.

Both cultures have identifiable specific sets of grapheme-phoneme correspondences, which changed over time. From the diachronic perspective, such sets can be considered stable since it was only the mapping of a limited number of graphemes on a limited number of phonemes that changed.

In Old Kanembu writing, the phoneme /s/ was written as *sīn*, then as *sīn* and *ṯāʾ* and then as *sīn*, *ṯāʾ* and *šin*, and later only as *sīn*. The additional letters were also used for homorganically similar phonemes. Thus, the letter *ṯāʾ* was also used for /ts/ and in the later manuscripts for /ʧ/, while the letter *šin* was also used for /š/. What is important here is that this graphemic-phonemic subset was quite stable over time within a given number of available choices, with only the mapping changing. Another such stable set with a restricted range of graphemic choices is observable for the phoneme /z/ which was written with the letters *ẓāʾ*, *ḏāl* and *zāl*, or for the phoneme /k/ written as *qāf ~ kāf*.

In Hausa Ajami, the salient stable sets can be identified as follows. At different times and in different areas, the phoneme /ʧ/ was written with the letters *šin* (in earlier manuscripts in the area of Sokoto) and *ṯāʾ* (in Kano manuscripts and later everywhere). The letter *ṯāʾ* was also used for homorganically similar /s/ in Arabic loanwords which originally have /ṯ/ (and are written in Arabic with *ṯāʾ*) and for the glottalised /sˀ/. Another set consists of the letters *rāʾ* and *ḍād* and the phonemes /l/ and retroflex /ɽ/. In earlier times, *ḍād* was used for /l/ and occasionally for /ɽ/. In modern Hausa Ajami, both rhotics are written with the letter *rāʾ*, whereas *ḍād* is only used in the etymological spelling of Arabic loanwords.

Thus, we can identify sets of grapheme-phoneme combinations which were stable within a restricted range of the phonemic and graphemic inventories. The choices within such sets were conditioned by phonological, morphological, and mimetic factors, the latter being the etymological spelling of Arabic loan words and shared scribal practices manifested on the pages of the same manuscript.

In more general historical terms, Hausa writing shows a tendency towards a closer match between the number of phonemes and graphemes, especially in more recent manuscripts. In the history of Kanuri writing, there is also a noticeable change in the choice of Arabic letters but there are two parallel tendencies. One is the diachronic rearrangement of grapheme-phoneme correspondences within specific grapheme-phoneme sets, while the other is the retention of the spelling of some high-frequency lexical and grammatical items.

References

Ajami Boko (no author) (1961), Zaria: Gaskiya Corporation.

Akinwumi, Olayemi and Adesina Y. Raji (1990), 'The Wangarawa Factor in the History of Nigerian Islam: the Example of Kano and Borgu', in *Islamic Studies* 29, 4: 375–385.

Al-Barnāwī al-Yarwāwī, Abī Bakar Gonīmī (1418/1997), *Al-daʿwa bi-l-ḥikma*. *Al-ʾūlā*. *Qawāʿid al-islām al-musammā bi-munqiḏu al-ʿawāmmi min al-malāmi* (*Invitation to Wisdom*.

Part One. The foundations of Islam, known as the redemption of the masses from evil), Maiduguri, Shimāl Naijīriyā (in Arabic and Kanuri. Mss facsimile).

Al-Hajj, Muhammad (1968), 'A 17th Century Chronicle on the Origins and Missionary Activities of the Wangarawa', in *Kano Studies,* 1.4: 7–16.

Benton, Philip Askel (1911), *Kanuri readings including facsimiles of mss., transliteration, interlinear translation and notes, also a complete English-Kanuri vocabulary and a partial Kanuri-English vocabulary*, London: Oxford University Press.

Bivar, Adrian David H. (1960), 'A Dated Kuran from Borno', in *Nigeria Magazin*, 65: 199–205.

Blommaert, Jan (2004), 'Writing as a Problem: African Grassroots Writing, Economies of Literacy, and Globalization', in *Language and Society*, 33: 643–671.

Bondarev, Dmitry (2006), 'The Language of the Glosses in the Bornu Qur'anic Manuscripts', in *Bulletin of the School of Oriental and African Studies,* 69, 1: 113–140.

Bondarev, Dmitry (2007), 'The Nigerian Qur'anic Manuscript Project: Retrieving a Unique Resource on the Kanuri Language and Culture', in *African Research and Documentation*, 103: 65–82.

Bondarev, Dmitry (2013a), 'Language change induced by written codes: a case of Old Kanembu and Kanuri dialects', in Esther-Miriam Wagner, Ben Outhwaite, Bettina Beinhoff (eds), *Scribes as Agents of Language Change* (Studies in language change), Boston, Berlin: De Gruyter Mouton, 291–323.

Bondarev, Dmitry (2013b), 'Qur'anic Exegesis in Old Kanembu: Linguistic precision for better interpretation', in Tal Tamari and Dmitry Bondarev (eds), *Qur'anic Exegesis in African Languages*. Special Issue of the *Journal of Qur'anic Studies*, 15(3): 56–83.

Bondarev, Dmitry (2014a), 'Multiglossia in West African Manuscripts: a Case of Borno, Nigeria', in Jörg B. Quenzer, Dmitry Bondarev and Jan-Ulrich Sobisch (eds), *Manuscript Cultures: Mapping the Field* (Studies in Manuscript Cultures, 1), Berlin, Boston: De Gruyter, 113–155.

Bondarev, Dmitry (2014b), 'Old Kanembu and Kanuri in Arabic Script: Phonology through the Graphic System', in Meikal Mumin and Kees Versteegh (eds), *The Arabic Script in Africa: Studies in the Use of a Writing System*. Leiden, Boston: Brill, 107–142.

Bondarev, Dmitry and Abba Tijani (2013), 'Performance of multilayered literacy: Tarjumo of the Kanuri Muslim scholars', in Kasper Juffermans, Yonas M. Asfaha and Ashraf K. Abdelhay (eds), *African Literacies: Ideologies, Texts, Education*, Newcastle upon Tyne: Cambridge Scholars Publishing, 118–146.

Brigaglia, Andrea and Mauro Nobili (2013), 'Central Sudanic Arabic Scripts (Part 2): the Barnāwī', in *Islamic Africa*, 4(2): 195–223.

Brooks, George E. and Bruce L. Mouser (1987), 'An 1804 Slaving Contract Signed in Arabic Script from the Upper Guinea Coast', in *History in Africa*, 14: 341–348.

Bulakarima, Shettima Umara (2001), *A Study in Kanuri Dialectology: Phonology and Dialectal Distribution in Mowar*. Maiduguri: Awwal Printing & Publications.

Coulmas, Florian (2003), *Writing Systems: An Introduction to their Linguistic Analysis* (Cambridge Textbook in Linguistics). Cambridge: Cambridge University Press.

Cyffer, Norbert (1998), *A Sketch of Kanuri*. Cologne: Rüdiger Köppe.

Dobronravin, Nikolay (2000), 'Hausa, Songhay and Mande Languages in Nigeria: Multilingualism in Kebbi and Sokoto', in H. Ekkehard Wolff, and ORIN D. Gensler (eds), *Proceedings of the 2nd World Congress of African linguistics, Leipzig. 1997*. Cologne: Rüdiger Köppe, 91–101.

Dobronravin, Nikolai (2006), 'TAM markers in 'Littafen nahwowin hausance': interpreted as verbs?', in Bernard Caron and Petr Zima (eds), *Sprachbund in the West African Sahel*. Louvain: Peeters, 129–143.

Ghānim, ʿImād al-Dīn (1998). *Al-Dawākhil al-Lībīyah fī majmūʿat dirāsāt lil-Raḥḥālah al-Almānī Ghūtlūb Ādūlf Krāwzah*. Ṭarābulus: Markaz Jihād al-Lībīyīn li'l-Dirāsāt al-Tārīkhīyah.

Gulbi, Abdullahi Sarkin and Dano Balarabe Bunza (2014), 'The Occupational Crafts of Zarma in Argungu Emirate', (a paper presented at the International Conference on the Relations between the Regions of Dosso (Niger), Sokoto, and Kebbi (Nigeria), organized by Abdou Moumouni University Niamey (Niger), Usmanu Danfodiyo University Sokoto, and Kebbi State University Aliero, on 14th–17th December 2014. https://webcache.googleusercontent.com/search?q=cache:JXrq2SWE1B8J:https://www.amsoshi.com/2017/07/19/occupational-crafts-zarma-argungu-emirate/+&cd=1&hl=en&ct=clnk&gl=de

Gouffé, Claude (1969), 'A propos de haoussa ['y]', in *Word*, 25 (1–3): 131–139.

Habraszewski, Tomasz (1967), 'Kanuri — language and people — in the 'Travel-Book' (Siyahatname) of Evliya Celebi', in *Africana Bulletin*, 6: 59–66.

Hiskett, Mervyn (1957), 'Material Relating to the State of Learning among the Fulani before Their Jihād', in *Bulletin of the School of Oriental and African Studies*, 19 (3): 550–578.

Hiskett, Mervyn (1975), *A History of Hausa Islamic Verse*. London: School of Oriental and African Studies, University of London.

Hunwick, John O., et al. (1995), *Arabic Literature in Africa. V. 2: The Writing of Central Sudanic Africa*. Leiden, New York, Cologne: Brill.

Hunwick, John O. et al. (2003), *Arabic Literature in Africa. V. 2: The Writing of Western Sudanic Africa*. Leiden, Boston: Brill.

Hutchison, John P. (1981), *The Kanuri Language. A Reference Grammar*. Madison: University of Wisconsin.

Jaggar, Philip J. (2001), *Hausa*. Amsterdam; Philadelphia: John Benjamins.

Jouannet, Francis (1982), *Le kanembu des Ngaldoukou: langue saharienne parlée sur les rives septentrionales du Lac Tchad. Phonématique et prosodie*. Paris: SELAF.

Krause, Gottlob Adolf (1884), *Ein Beitrag zur Kenntniss der Fulischen Sprache in Afrika*. Leipzig: Brockhaus.

Levtzion, Nehemia (2000), 'Islam in the Bilad al-Sudan to 1800', in Levtzion, Nehemia and Randall L. Pouwels (eds), *The History of Islam in Africa*. Athens, Ohio: Ohio University Press, 63–92.

Le Roux, Jean Marie (1886), *Essai de dictionnaire français-haoussa et haoussa-français : précédé d'un Essai de grammaire de la langue haoussa, Maganan haoussa ... renfermant les éléments du langage parlé par les nègres du Soudan, accompagné d'une carte de l'Afrique septentrionale*, Alger: A. Jourdan.

Lippert, Julius (1900), 'Sudanica', in *Mitteilungen des Seminars für orientalische Sprachen*, 3 (3): 198–207.

Lovejoy, Paul E. (1978), 'The Role of the Wangara in the Economic Transformation of the Central Sudan in the Fifteenth and Sixteenth Centuries', in *Journal of African History*, 19 (2): 173–193.

Lukas, Johannes (1931), *Die Sprache der Káidi-Kanembú in Kanem*. Berlin: Dietrich Reimer.

Lukas, Johannes (1937), *A Study of the Kanuri Language, Grammar and Vocabulary*, London, New York: Published for the International African Institute by the Oxford University Press.

Mischlich, Adam (1906), *Wörterbuch der Hausasprache*. Berlin: Georg Reimer.

Mumin, Meikal and Kees Versteegh (2014), 'Introduction', in Mumin, Meikal and Kees Versteegh (eds), *The Arabic Script in Africa: Studies in the Use of a Writing System*. Leiden, Boston: Brill.

Newman, Paul (2000), *The Hausa Language: An Encyclopedic Reference Grammar.* (Yale Language Series). New Haven: Yale University Press.

Nobili, Mauro (2011), 'Arabic Scripts in West African Manuscripts: A Tentative Classification from the de Gironcourt Manuscript Collection', in *Islamic Africa*, 2 (1): 105–133.

Palmer, Herbert Richmond (1908), 'The Kano Chronicle', in *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 38: 58–98.

Prietze, Rudolf (1914), 'Bornu-Lieder', in *Mitteilungen des Seminars für orientalische Sprachen*, 17: 134–260.

Prietze, Rudolf (1930), 'Bornu-Texte', in *Mitteilungen des Seminars für orientalische Sprachen*, 33: 82–159, i–xxxii.

Sanneh, Lamin O (1989), The Jakhanke Muslim clerics: A religious and historical study of Islam in Senegambia. Lanham, MD: University Press of America.

Sebba, Mark (2007), *Spellings and Society: the Culture and Politics of Orthography around the World*. Cambridge: Cambridge University Press.

Starratt, Priscilla Ellen (1993), *Oral History in Muslim Africa: Al-Maghili Legends in Kano*. Ph.D. dissertation, Ann Arbor: University of Michigan.

Stolbova, Olga (2016), *Chadic Etymological Dictionary*. Moscow: Institute of Oriental Studies.

Wali, Na'ibi Sulaimanu and Haliru Binji (1969), *Mu koyi Ajami da Larabci* (We learn Ajami and Arabic). Zaria: Northern Nigerian Publishing. Co. (in Hausa).

Yahaya, Ibrahim Yaro (1988), *Hausa a rubuce: tarihin rubuce rubuce cikin hausa*, (Written Hausa: the history of writing in Hausa), Zaria: Kamfanin Buga Littattafai na Nigeria ta Arewa (in Hausa).

Zaria, Ahmadu Bello (1982), *Issues in Hausa Dialectology*. PhD dissertation, Indiana University.

Lameen Souag

Kabyle in Arabic Script: A History without Standardisation

Abstract: The history of writing Kabyle, the largest Berber variety of Algeria, is often reduced simply to the history of the development of its currently dominant Latin orthography; Kabyle before the 20th century is often viewed as essentially an unwritten language. This oversimplification obscures an extensive, though somewhat peripheral, corpus of Kabyle written in Arabic script starting before the 19th century and continuing in some contexts up to the present. Unlike Berber varieties in some other regions, however, Kabyle shows few signs of ever having developed a standard Arabic-script orthography, and the transcription conventions used show little or no continuity between manuscripts from different periods. This lack of standardisation reflects discontinuities in the region's educational and political history, in combination with certain differences between Kabyle and Arabic phonology.

1 Introduction

As might be expected from Arabic's long history as the dominant literary language of North Africa, Berber languages have frequently been written in Arabic script. In at least two regions during the pre-colonial era, such writing went beyond the sporadic quoting of words or poems and came to be used for a significant corpus of largely religiously motivated manuscripts: the Sous Valley of Morocco in the west, and the Ibāḍī areas of Libya and southern Tunisia in the east (Boogert 1997; Brugnatelli 2011). This contrasts strikingly with the most populous Berber-speaking region between them, Kabylie in north-central Algeria. Relatively few Berber manuscripts in Arabic script have been reported from the area, even though it is known to have contained many *zaouias* (religious colleges) with substantial manuscript collections, and the colonial-era research tradition that provided Western academics with their first glimpse of the Sous and Ibāḍī manuscript traditions (Luciani 1893; Calassanti-Motylinski 1905) treated Kabyle as essentially an unwritten language. In present-day Kabylie too, there is little awareness of any history of writing Kabyle in Arabic script; Haddadou (2004), of the University of Tizi-Ouzou, devotes a whole section of his *L'alphabet berbère* to the Arabic alphabet, alluding to Moroccan, Libyan, and Mozabite examples, yet mentions no examples of its use for Kabyle.

https://doi.org/10.1515/9783110639063-011

Contrary to the impressions of colonial-era researchers and modern Algerians alike, however, several Kabyle manuscripts in the Arabic script survive. They are strikingly similar in their content and purposes to the better-known Sous and Ibāḍī materials, although rather less numerous. Moreover, the practice of writing Kabyle using Arabic script has continued in some contexts even up to the present day, despite being rejected by most activists. Nevertheless, a close examination of the orthographic conventions used reveals sharp discontinuities. While individual materials produced in similar circumstances show some signs of standardisation, such embryonic conventions have rarely been maintained for long. This lack of standardisation, and the rarity of Kabyle manuscripts in Arabic script relative even to other Berber varieties, seem to be the result of two factors: the sharp discontinuities in the region's educational and political history over the past two centuries, and the specific differences between Arabic and Kabyle phonology.

2 Contexts of literacy

By the start of the 19th century, the region of Kabylie, largely independent of the nominally Ottoman central government in Algiers, was covered by an extensive network of zaouias (*timɛemmert* in Kabyle) – religious colleges providing a brief primary education to large numbers of children, and a higher education to the minority who stayed longer. Their curriculum included literacy, memorisation of the Qur'an, Arabic grammar, Sufism, Maliki law, hadith, interpretation of the Qur'an, arithmetic, geometry, astronomy, and poetry (Daumas 1847, 61–63). Colloquial languages were of little concern to this educational system, whose core texts were in Classical Arabic and whose scholars in turn normally wrote in Classical Arabic. However, an important function of graduates was to provide religious guidance in Kabyle to non-Arabic-speaking lay people, and in some cases this motivated them to write down Kabyle texts of their own composition.

The French conquest, starting in 1830 with the occupation of Algiers but not encompassing the Kabyle heartland until 1857, had a disruptive effect on this system from the start. Parts of the elite fled to Syria, preferring exile to French rule, and villages and mosques were destroyed in the process of conquest (cf. Hanoteau 1867, 144). In order to govern the newly conquered areas, the French state found it necessary to study Kabyle, resulting in a flurry of publications incorporating Kabyle text to varying degrees; these were not, however, aimed at a Kabyle audience.

In 1871, Shaykh El-Mokrani and Shaykh Aheddad led a major rebellion against France, which at its peak involved not only their native region, eastern Kabylie, but

most of central and eastern Algeria. When France finally defeated them in 1872, the consequences for the region were dire. A heavy punitive tax was imposed, beggaring many families, and lands owned by tribes and foundations – including zaouias – were confiscated by the state. Leading rebels, including major religious scholars, were exiled or executed, and restrictions were imposed on Islamic education. In such a situation, the resources available to the zaouias, and the prospects for their graduates, were a fraction of what they had formerly been.

Meanwhile, a new educational system was beginning to provide more attractive prospects. Between 1873 and 1883, the Pères Blancs opened eight Catholic schools in Kabylie; the French government followed suit in 1882, opening another eight state schools (Dirèche 2007). In these schools, the language of instruction, and more generally of literacy, was French; together with their modern vocational focus, this qualified their graduates for subaltern positions within the colonial economy. These schools were attended by only a small minority, but their graduates' relative prosperity gave them an importance out of proportion to their numbers. Starting in the early 20th century, the importance of French was increased by labour migration to France, as remittances became a major economic resource for Kabyle villages. Somewhat later, the Association of Muslim Ulama, founded by Ben Badis in 1931, attempted to redress the balance by expanding Islamic education and shifting it towards a modernised curriculum taught in Arabic (Heggoy 1973).

In 1948, the Algerian literacy rate, excluding French citizens, was only 6% (United Nations 1955, 436). It was left to the War of Independence (1954–1962) to change this situation. The outbreak of the revolution was itself enough to provoke belated reforms; with the Constantine Plan of 1958, France attempted to appease and remould Algerian public opinion by, among other things, building new schools and increasing enrolment. After independence in 1962, the pace of change accelerated: the new Algerian government extended education throughout the country, making literacy one of its highest priorities. The language of instruction was initially usually French, since, at independence, most of the small minority of graduates in subjects other than Arabic and Islamic studies had studied in French and had little command of Standard Arabic. However, the government pushed hard to change the language of instruction to Arabic, despite opposition from some of those educated in French. By the end of the 1980s, all levels up to the end of secondary school, as well as many university subjects, were taught in Arabic.

Berber, however, was treated rather differently. The post-independence government viewed the idea of Berber identity as divisive, and saw no place for the language in education; the only appropriate arenas for its public use were oral contexts such as radio broadcasting or music, and even there it was put under

pressure. In combination with other factors, this helped provoke increasing hostility in Kabylie to the government's education policy; one of the era's first large-scale protests, in 1980, was triggered by the government's banning a university lecture on Berber poetry. Massive emigration from rural areas to the largely Arabic-speaking cities throughout this period led to extensive language shift among emigrants' children, compounding the feeling that Kabyle was under threat.

In 1988, the government was forced by budgetary problems and popular protests to open up the political arena to independent parties and newspapers, creating new options for the use of Kabyle in writing. In 1995, after a large-scale school boycott in Kabylie compounded the multiple crises then facing the state, it conceded the gradual introduction of Tamazight teaching into selected schools (Zeroual 1995). Over the early 21st century, it went further, declaring Tamazight first a national language (Bouteflika 2002), and finally an official one (Bouteflika 2016). In theory, 'Tamazight' refers to a language of which all Berber varieties are considered to be dialects. In practice, the Tamazight of most government textbooks and of books by Kabyle speakers – which make up the large majority of Berber publishing in Algeria – is simply Kabyle with a number of Arabic or French loans replaced by neologisms. Even now, Tamazight occupies a very limited place in the educational system; it is taught as an optional subject, not available in all schools, and has not been adopted as a language of instruction for any other subject. However, Tamazight-language publishing, made possible by the post-1988 liberalisation, continues.

3 Sources

No list of Kabyle works in the Arabic script can be guaranteed or even expected to be complete. Many potentially relevant collections remain uncatalogued, and the observed domains of Arabic script usage include contexts, such as private letters or political posters, which are rarely preserved systematically. Nevertheless, enough material is available from the past two or three centuries to yield a general picture. The distribution of this material can only be understood in the light of its educational and political contexts.

3.1 Indigenous efforts in the pre-colonial and early colonial periods

3.1.1 The zaouia tradition

The manuscript libraries of the zaouias of eastern ('Petite') Kabylie have yielded a small but significant number of manuscripts in Kabyle. None include dates, unfortunately; some can be shown to date to the 19th century, while others are thought to predate this period, but in most cases no solid evidence is available. The best documented ones are the religiously focused materials from the collection of Shaykh Lmuhub Ulaḥbib of Tala Uzrar (near Bejaia), born in 1822 (Aïssani 1998), some of which may predate the 19th century. 476 manuscripts in this collection have been catalogued (Mechehed 2007), although this does not reflect the full picture, since some manuscripts have passed into other hands, and many were burned by French troops in 1957. Most of this collection is in Arabic, but it includes five Kabyle manuscripts: three copies of translations of al-Sanūsī's Creed (KA 21, KA 22, TIA 08), of which the oldest (KA 21, Fig. 1) has been addressed in detail by Gutova (2011); an anonymous poem about the virtue of knowledge by Said Ali Cherif, born 1820 (LIT 21/2); and a child's note (DVS 09). Luciani (1893, 159) attests to the existence at his time of a Kabyle translation of the same work at the zaouia of Sidi Yahia ben Hamoudi, among the Ait Ourthilane (which he describes as the only example of a Berber Islamic manuscript in Algeria). GEHIMAB (2006, 5) mentions another Kabyle creed in manuscript found at the zaouia of Tilmiouine (Ouzellaguen), and the Ministry of Religious Affairs (2009, 24) gives a barely legible photograph of another theological manuscript in Kabyle, cited as belonging to the family of Shaykh Oubelkassem of the *zaouia* at Boudjelil (bombed by the French army in 1956). Another religious poem, from the library of Shaykh Aheddad, is mentioned in Aïssani (2012). Manuscript no. 3066 of the Bibliothèque Nationale d'Algérie is described by GEHIMAB (2006, 2) as a 10-page verse text on *fiqh* in the Kabyle dialect of the Soummam valley; the photo is illegible.

A rather different genre, equally motivated by the needs of zaouia students, is represented by the Algerian Arabic-Kabyle wordlist from the library of Cheikh Aheddad, partly reproduced by GEHIMAB (2006, 5) and transcribed and analysed by Aïssani (2012). Zaouias often attracted students from outside the region, who would have needed to learn Kabyle for practical purposes.

Many early North African Arabic-language texts, such as the early 18th century pharmacopoeia of Ibn Ḥamādūsh (2001, 57, 83), contain scattered words in Berber, commonly proper names or plant names. While such transcriptions are of limited

Fig. 1: First page of manuscript KA 21 from the Lmuhub Ulahbib Collection (Tala Uzrar, Ain-Legradj, Bejaia, Algeria), a Kabyle translation of al-Sanūsī's creed. © Djamel-Eddine Mechehed.

value, they too may serve as evidence for the development of orthographic conventions. The Voyage of al-Warṭīlānī (1994), in which an 18th century Kabyle religious scholar recounts his travels, is of particular interest in that its original form presumably reflected Kabyle zaouias' practices; however, a study of the author's orthographic choices would have to be based on manuscripts rather than on later printed versions.

3.1.2 Customary law codes

At least since the Ottoman period, some Kabyle villages created manuscripts summarising their customary law, or *qanun*, a phenomenon recently surveyed by Gahlouz (2011). These were most often written down in Arabic, e.g. the *qanun* of the Maatka copied in 1859 (Millot 1922), or the pre-colonial *qanun* of the Ait Hichem dating to the 1830s (Millot 1928). Occasionally, however, they were written in Kabyle; the present survey examines the late 19th century *qanun* of the Ait Ali Ouharzoune scanned by Brugnatelli (2006) and partly translated (Anonymous 2009). Another reported example is that of the Ait Ouagennoun (Bernard and Millot 1933; Ould-Braham 1986). Unlike the zaouia manuscripts, these cases come from western ('Grande') Kabylie.

3.2 Colonial efforts

3.2.1 Language teaching materials

Given that Europeans and Americans are typically much more familiar with the Latin alphabet than with the Arabic one, it may appear paradoxical that textbooks and articles intended for their benefit made frequent use of the Arabic script. This fact, however, is important for the light it sheds on contemporary Kabyle practice. The earliest cases justify this practice on the basis of the speakers' own habits; thus Hodgson (1834, 38) writes somewhat paradoxically 'The Berber language, having no literature, has properly no alphabet of its own. It is written nevertheless by means of the Arabic characters, which are generally used for that purpose', while the dictionary of Brosselard (1844, II) – produced by a committee including one Kabyle, Sidi Ahmed ben el Hadj Ali of Bejaia – describes the Arabic orthography used as 'the system of transcription that appears most widely accepted', calling it 'the Berber alphabet' (author's translations, here and henceforth). Later works are more dismissive, describing Kabyle as an unwritten language (Delaporte 1836, 89; Ben Sedira 1887, III; Boulifa 1897, II), and presenting the writing of Kabyle in Arabic script as an occasional makeshift of the marabout class rather than a well-rooted practice. General Hanoteau's (1858) influen-

tial grammar of Kabyle did much to popularise this view: he states that '[t]hose who speak it borrow its characters from the Arabic alphabet, when they want to express their ideas in their language in writing, which they moreover do only very rarely, and always with a certain distaste. Kabyle, they say, is spoken, not written' (Hanoteau 1858, 1–2), and later added for good measure that 'All efforts made to find a book written in this language have been, and no doubt always will be, fruitless' (Hanoteau 1867, II).

Nevertheless, Hanoteau continued to include Arabic transcriptions in parts of his works, justifying them on the grounds that they make it possible for the learner to get a speaker of Berber to read the words written, and thus to correct the learner's pronunciation: 'This transcription, nonetheless, will not be lacking in utility for people who want to study the Kabyle language; it will allow them, in effect, to get the texts read for them by a Kabyle' (Hanoteau 1867, XII). The practicality of this solution is confirmed by other data. Rinn (1887) presents two Kabyle poems attacking Cheikh Aheddad, anonymously sent in Arabic script to a French general in 1872; these suggest that, at that period, even the small minority of Kabyle speakers who wanted to be seen as vocally pro-French were more accustomed to Arabic script.

The prolific output of René Basset would keep the practice of Arabic transcription going for some decades longer. In his *Manuel de langue kabyle* (1887), as throughout his extensive publications on Berber, he used Arabic characters alongside Latin transcription, on the grounds that (p. VIII) '[T]he natives, [w]hatever one may do, will not abandon the use of the Oriental characters', and that it allows problematic short vowels to be ignored in writing. In the meantime, however, this justification was becoming unnecessary: the demands of French administration were creating a minority of Kabyles literate in French.

For this new group, uncomfortably poised between the indigenous and French societies, the use of Arabic script could no longer be taken for granted. In a short grammar, Ben Khouas (1881, 5–6), presents a paragraph-long argument to conclude that '[s]ince [...] the Kabyles have much more frequent relations with the French than with the Arabs, especially when there is occasion to write something, it is preferable in every way, for writing Kabyle, to use French characters'. Textbooks by others in this situation at the same period, such as Ben Sedira (1887) and Boulifa (1897), dispense with Arabic characters entirely, or, like Boulifa (1913, 351), confine them to a single page by way of preparation for examinations. The same trend is observed in contemporary French linguists' work; Mouliéras (1892; 1893) uses Latin characters exclusively to transcribe the Kabyle stories dictated to him by informants he describes as illiterate, as does Gourliau (1893), while Luciani (1899, 33) publishes only the Latin transcription of poems he had actually received in Arabic transcription. In 20th century French-language works on Kabyle, whether by Frenchmen or by Algerians, Arabic characters are almost uniformly absent.

3.2.2 Missionary materials

Insofar as Christian missionaries' choice of orthography reflects their perception of existing usage, it casts light on Kabyle orthographic preferences in a similar way to language learning materials. However, sectarian factors also appear to have been relevant; Catholic sources consistently avoid Arabic script in this region, while Protestant ones show greater variability.

The earliest attempts to render parts of the Bible into Kabyle, from Arabic, took place at the behest of the American vice-consul Hodgson. The (Muslim) translator he commissioned, the marabout 'Si Hamet' of Bejaia, wrote his translation in the Arabic script. Parts of these (Anonymous 1833) were printed in a modified Arabic script and distributed in Algeria shortly after the French conquest, while another portion was later printed for purely linguistic uses (Newman 1847).

Serious efforts at Christian proselytisation in Algeria started a generation later, in 1868, when the newly appointed Catholic Archbishop of Algiers, Cardinal Lavigerie, founded the Pères Blancs; starting in 1873, they also established village schools in Kabylie (Dirèche 2007). From the start, their efforts, liturgical (Allemand-Lavigerie 1868; 1869) or lexicographic (Creusat 1873; Olivier 1878; Huyghe 1901), made exclusive use of Latin characters, again characterising Kabyle as an unwritten language.

Protestant evangelisation – initially based at Djemaa Saharidj near Tizi-Ouzou – largely followed the Catholic lead in this respect, justifying the choice on the grounds that French schools were in any case spreading literacy in the Latin script by then (Rutherfurd and Glenny 1900, 146; Cuendet, Hocart and Lamb 1893). Nevertheless, they hedged their bets with a supplementary printing of Mark and Luke in Arabic script (British and Foreign Bible Society 1904), specifically intended for students from traditional Islamic schools (Canton 1904, 16): 'Stirred by the energy of western life, the marabouts taught the Kabyle lads the Arabic character, and for these a transliteration of St Luke was in progress'. This would be the last missionary publication in Kabyle using Arabic script, as far as I have been able to find.

3.3 The early 20th century

3.3.1 Poetry

One of the earliest and most important printed works by a Kabyle author aimed partly at a Kabyle audience, Boulifa's (1904) anthology of Kabyle poetry, uses both Latin and Arabic characters. Since the author's pedagogical works intended for Frenchmen do not include Arabic script, this choice was most likely intended

to make the poems legible to a wider audience of literate Kabyles. Yet the expansion of French education in the region starting from 1880 (Dirèche 2007), along with the closing of zaouias, would leave the Arabic characters meaningless to later generations. Mouloud Feraoun (1960, 10–11) bears eloquent witness at once to Boulifa's success in appealing to a Kabyle audience and to the change in educational patterns that made this practice obsolete:

> They are incapable of reading the poems gathered by Boulifa in the Arabic transcription; if they tackle the French transcription, they can be sure of reading them well only when it comes to a poem that they already know [...] It is "the Book", the only book of young Kabyles. You find it in the villages, old and venerable, as it was published more than half a century ago [...] It is never complete; a number of its missing pages have been offered to a friend, to continue their clandestine career in others' hands.[1]

The use of Arabic script would still be continued, however, by those Kabyles who opted for a traditional or reformist (Ulama) education. Yacine (1987) presents an extensive corpus of reformist Islamic and nationalist poetry in Kabyle by Qasi Udifella, originally written down in vocalised Arabic letters by Hadj Boubekar in the 1940s in the Biban region in the extreme east of Kabylie. The choice of Arabic characters was natural, given their political alignment.

3.3.2 Popular music

In the 1950s and 1960s, Kabyle song titles (and sometimes lyrics) could often be found in Arabic script, as can readily be seen by examining the LPs of Cherif Kheddam, Slimane Azem, or El Hasnaoui. In fact, Cherif Kheddam, educated at the zaouia of Boudjelil in eastern Kabylie, often wrote his own lyrics in Arabic script (GEHIMAB 2006). This minor domain of Arabic-script Kabyle becomes less prominent in the 1970s (Idir's albums, to take an influential example, simply use Latin script), and later vanished almost completely.

1 Les poèmes recueillis par Boulifa, ils sont incapables de les lire dans la transcription arabe ; s'ils s'attaquent à la transcription française, ils ne sont sûrs de bien lire que lorsqu'il s'agit d'un poème qu'ils connaissent déjà [...] Il est "le Livre", l'unique livre des jeunes Kabyles. On le trouve dans les villages, vieilli et vénérable, tel qu'il fût édité, il y a plus d'un demi-siècle [...] Il n'est jamais complet : nombre de ses feuillets manquants ont été offerts à un ami pour poursuivre en d'autre mains leur carrière clandestine.

3.4 Official recognition

The post-independence Amazigh identity movement, reacting against the uniform Arab identity urged on them by the Algerian state, conspicuously avoided the use of Arabic script. The Neo-Tifinagh script proposed by the Académie Berbère quickly attained emblematic status for signs and mottos, while the Latin orthography later popularised by Mouloud Mammeri (1976) came to be generally adopted for longer works. However, political liberalisation following 1988 created arenas outside of Amazigh activism for the written use of Kabyle; already during the 1989–1990 election campaigns some parties created posters with slogans in Kabyle in Arabic script (not examined here). Soon afterwards, the Algerian state, facing intense pressure on multiple fronts, offered the Amazigh activists a limited compromise. Since 1995, the Algerian state has provided limited teaching of Tamazight (conceived as a single language) in selected schools; since 2002, Tamazight has been recognised as a national language, and in 2016 it was declared an official one as well.

3.4.1 Educational materials

The decision to teach Tamazight required the creation of textbooks and dictionaries acceptable to the state. As such, it brought previously unconcerned parties into the orthography debate. While the use of Arabic script was by and large strongly opposed by Amazigh activists, particularly in Kabylie, it was strongly favoured by significant segments of the state and of Algerian public opinion outside the region. The dictionary of Tagamount (1995), published just in time for this decision, reflects the debate: while primarily using Latin characters, it provides an Arabic transcription in the Kabyle-French/Arabic section, alongside the Arabic translations. This grudging concession to Arabic would not be followed up in subsequent dictionaries, however: while Kabyle dictionaries have since become a routine sight in Algerian bookstores, even those that give Arabic translations almost always avoid giving Arabic transcriptions.

Notwithstanding the widespread preference for Latin script (Sini 2004), Kabyle supporters of the Arabic script can still be found. A notable example is Salah Belaid, of the University of Tizi-Ouzou. He devotes nearly half a book (Belaid 1999) to the issue of which writing system to choose for Berber, coming down strongly in favour of Arabic. By way of examples, he gives a text transcribed into Tifinagh, Latin, and Arabic characters (159–166) and a glossary of neologisms drawing extensively on previous authors' work (215ff). Like most other writers on the subject of which script to choose for Kabyle, however, he does not appear

to have published any books in Kabyle, so there is little incentive for would-be readers to learn his orthographies.

Up to the present, the disagreement over script choice has been resolved by making official materials for the study of Berber available in three scripts: Latin, Arabic, and Tifinagh. In practice, Latin script overwhelmingly predominates in Kabylie, but official examples of the use of Arabic script can still be seen, e.g. Ministry of Education (2003), Hrouch et al. (2010). As will be seen below, the Arabic orthography used in these textbooks, described in detail by Gaci (2011), is rather eccentric, being chosen basically to maximise compatibility with the Latin transcription, and does not correspond to any of the practices seen elsewhere.

3.4.2 Religious materials

The post-1995 climate of relative openness to Tamazight affected not only the educational system, but also a domain whose control is almost equally important to the state: Islamic religious publications. The preference for Arabic orthography was, unsurprisingly, particularly strong in this domain – although even here, non-state-funded publications have used Latin script. A group led by the Kabyle religious scholar Si Mohand Tayeb has published first selections from the Qur'ān (2003) and later a full translation (2012), printed in Saudi Arabia with funding from Algeria's Ministry of Religious Affairs. In this work the Arabic script is used, with some additional characters. The decision to write in Arabic script attracted criticism from Kabyle activists, but was vigorously defended by the author (as quoted in Chachoua 2010): 'My principal long-term goal is to defend Berber against those who want to turn it away and distance it from Islam. [...] I have often, in many published articles and letters to various political authorities, called for the adoption of Arabic letters for writing Berber, and I would do the impossible to make this happen'. Rebahi (2009), in a booklet funded by the same ministry, gives two short religious poems in Kabyle, written in Arabic script, with translations.

3.4.3 State media

A national radio channel using Tamazight (mainly but not exclusively Kabyle), Radio 2, has been maintained since independence. The question of script choice hardly arises for broadcasting, but since at least 2011 this channel has also had a webpage (Radio 2) at the national radio service's website. At no point has this page been translated into Kabyle or any other Tamazight variety. From 2011 to 2013, it was bilingual in Arabic and French, with Kabyle programme titles and headings

transcribed in Arabic script on the Arabic site and in Latin script on the French site. By 2014, this was no longer the case: since then the Radio 2 page has been available only in French, not in Arabic, and the programme titles and headings (still in Kabyle) are transcribed in Latin script (and later also Tifinagh). The only remaining trace of Arabic script in 2014 was the station motto at the top of the page, transcribed in Tifinagh, Arabic script, and Latin script – and by 2017 this had been replaced by the name of the station in Tifinagh script plus translations into French and Arabic. More recently, a Tamazight TV channel (TV 4) has been created; this station too makes frequent use of Arabic script to transcribe Tamazight titles or slogans, including Kabyle ones, but the texts in question remain very short.

In 2013, the state-owned primarily Arabic-language newspaper Ech Chaab, founded in 1962, started carrying columns in Tamazight written in Arabic script (Rubrique Radar 2013). Some, but not all, of these columns were in Kabyle, e.g. Lakhdari (2013).

3.4.4 Books on heritage

The most conspicuous contributor in this domain is the historian and ex-FFS deputy Mohamed Arezki Ferad. His history of his native Azeffoun (2003), primarily written in Arabic, includes many Kabyle poems and sayings written in Arabic script. In his study of Tipasa Berber proverbs (Ferad 2004), an attempt to document the oral traditions of a Berber dialect falling outside the Kabyle continuum proper, he acknowledges that the Latin script has taken the lead at present, but vigorously defends his choice of the Arabic script, arguing that 'in view of the richness of the Berber cultural stock recorded in Arabic characters (books, contracts…) and the cultural factors that have brought together the Berber and Arabic languages, the Arabic script is a candidate for filling the principal role in the advancement and development of Berber in the future' (63). The author has gone on to produce other books focused on Kabyle heritage, including an overview of the region (Ferad 2007) and a study of Kabyle Sufi poetry (Ferad 2011). More academically oriented Arabic-medium studies of Kabyle folklore also sometimes make use of Arabic script, such as the thesis of Ṭabarkān (2012).

4 Orthographic conventions and their contexts

The number of texts written in Kabyle in the Arabic script, while not enormous, is sufficiently large to raise the question of standardisation. How have different

authors dealt with the problem of representing phonemes with no direct Arabic equivalents? To what extent have they been influenced by one another's conventions, or more generally adopted similar conventions? When and in what milieux have standards, if any, emerged, and how long have they been maintained?

4.1 Challenges for orthography design

All of the phonemes of Classical Arabic as pronounced in Algeria are to be found in Kabyle except the stop *ḍ* (in eastern varieties *ḍ̣* is also missing). Since Arabic *ḍ* is usually pronounced *ḍ̣* in the region except in careful speech, this gives authors a choice between *ḍ* and *ḍ̣* to represent the phoneme *ḍ̣*. Kabyle, however, also has a number of consonant phonemes lacking in Classical Arabic, most conspicuously *g* and *ẓ* (shared with Algerian Arabic). It distinguishes affricates *č*, *ǧ* (mainly in geminates) from fricatives *š*, *ž* (conventionally transcribed *c*, *j*), while Arabic does not (depending on the underlying dialect, Arabic *ǧīm* is regionally realised either as *ǧ* or as *ž*). It also marginally distinguishes affricates *ts*, *dz* (the former is usually derived from underlying geminate *tt*, while the latter is very rare) from both *t*, *d* and *s*, *z*, while Arabic has no phonemic affricates. It has a broader distinction, usually predictable but contrastive in some contexts, between stops (*b*, *t*, *d*, *k*, *g*) and spirants (*b̲*, *t̲*, *d̲*, *k̲*, *ḡ*); of the spirants, only *t̲* and *d̲* have Arabic counterparts. The distinction between plain *l*, *r* and emphatic *l̲*, *r̲* is marginally contrastive both in Kabyle and in Algerian Arabic (for *l* it is arguably contrastive even for Classical Arabic), but is consistently not made in Arabic script; that between *š* and *ṣ̌* is even more marginal in Kabyle, and absent in Arabic. Kabyle also has a series of labiovelarised back consonants (k^w, g^w, x^w, $ɣ^w$), with no direct equivalent in Arabic, but often corresponding in loanwords to Arabic short *u*. Of these, note that the affricates *ts*, *dz* and the spirant series are not phonemic in Shilha and in Libya, simplifying the situation faced by would-be writers there.

The vowel inventory of Kabyle is rather smaller than that of Classical Arabic; whereas Arabic distinguishes short *a*, *i*, *u* from long *ā*, *ī*, *ū*, Kabyle contrasts one lax vowel, *ə* (conventionally transcribed *e*), with three tense ones, *a*, *i*, *u*. (In Shilha, the lax vowel no longer has any phonemic status.) Mapping the Kabyle system onto the Arabic one poses some difficulties. In Kabyle, the tense set are phonetically long in open syllables when stressed, and shorter elsewhere; this encourages a phonetic rather than phonemic transcription into Arabic characters. The lax vowel, however, is consistently short, so overtly representing length tends to create ambiguity with it. Moreover, word-finally, Arabic distinguishes *ā* (*alif mamdūdah*) from *à* (*alif maqṣūrah*, written with an undotted *yā'*) on purely morphological grounds not directly paralleled in Kabyle.

4.2 Observed solutions

Table 1 sums up the observed responses to the most important of the issues in 4.1, by source (for the numerals under 'Tense vowels', see 4.2.2). Note that eastern Kabyle varieties, used in several of these sources, do not have a phoneme *ḍ*. Only a limited selection of the Bible translation of 1904 was available.

Table 1: Orthographic choices by source.

	ḍ	g	ẓ	ts	č	ḅ	ḳ	ḡ	Cʷ	Tense vowels	ə
Zaouias:											
KA 21 (Gutova 2011)	-	ك	ʔ	ـَت	ش	ʔ	ك	ك	ʔ	2	ـَ
Poem in honour of Aheddad (GEHIMAB 2006, 3)	ض	ڤ	ز	ʔ	ʔ	ب		ڤ	ʔ	1	ـَ
Aheddad wordlist (GEHIMAB 2006. 5, Aïssani 2012)	ض	ڭ	ز	ʔ	ش	ب	ك	ڭ	ʔ	2	ـَ
Customary law:											
Ait Ali Ouharzoune (Brugnatelli 2006, Anonymous 2009)	ض	ڭ	ز	ʔ	ʔ	ب	ك	ڭ	-, ـُ	3	ـَ
Language teaching:											
Brosselard (1844)	-	ڭ	ز	ـَت	چ	ب	ك	ڭ	-	3	ـَ
Hanoteau (1867)	ض	ك	ز	ـَت	ش	ب	ك	ك	-	2	ـَ
Rinn (1887)	-	ڤ	ز	ʔ	ش	ب	ك	ڤ	ʔ	3	-
Basset (1887)	ض	ڭ	ز	ت	چ	ب	ك	ڭ	و, -	(various)	-
Bible translations:											
Anonymous (1833)	-	ڤ	ز	ش	چ	ب	ك	ڤ	ـُ	2	ـَ
BFBS (1904)		ڤ				ب				2	ـَ
Early 20th century:											
Boulifa (1904)	ض	ڤ	ز	ت	چ	ب	ك	ڤ	و, -	4	-
Udifella (Yacine 1987)	ظ، ض	ڤ	ز	ت	ʔ	پ، ب	ك	ڤ	ʔ	3	ـَ
Educational materials:											
Tagamount (1995)	-	ڤ	ز	ت	ʔ	ب	ك	ڤ	-	1	ـَ، ـْ
Belaid (1999) - text	ظ	ڤ	ز	ت	ʔ	ڤ	ك	ڤ	ʔ	4	-
Belaid (1999) – lexicon	ض، ظ	ʔ	ڒ	ـَت	ʔ	ب	ك	گ	ـُ	4	-
Hrouch et al. (2003)	ض	ڤ	ڒ	ت	ـّس	ب	ك	ڤ	-	5	آ
Ministry of Religion:											
Anonymous (2003), Tayeb (2012)	ظ	ڤ	ژ	ت	س ـت	ش ـت	پ	گ	چ	ـُ	3	ـَ
Folklore:											
Ferad (2003)	ض	ق	ز	ـَت	ش ـت	ب	ك	ق	ـُ	4	ـَ

Other minor points of variation have been omitted from the table. These rarely relate to consonants, although one source – the first Bible translation (Anonymous 1833) – makes the particularly eccentric choice to represent *ṯ* with a new character seemingly based on Greek *theta*: a *fāʾ* shape with a dot inside the loop (not in Unicode). However, earlier sources use the *maghribī* script, while modern ones use *naskhī*; this entails a difference in the dotting of *fāʾ* (under in *maghribī*, over in *naskhī*).

More commonly, differences relate to vocalisation. Initial geminates are often written with *al-*, particularly for coronal consonants. Some sources, such as the *qanun* of Ait Ali Ouharzoune, the poem in honour of Shaykh Aheddad, and the poetry of Udifella, frequently write word-final *-en* using *tanwīn (-an)*. Final *-a* is sometimes written with *à* or even with *tāʾ marbūṭah*, and final *-u* with *wāw+alif*. The use of *šaddah* to mark gemination is not always consistent.

The textbook of Hrouch et al. (2003) stands out in a number of respects. While other sources normally transcribe initial epenthetic schwa using an *alif*, this textbook systematically omits it. Likewise, most sources write initial *i* and *u* with *alif* (usually followed by *yāʾ* or *wāw*), but this textbook uses *yāʾ+hamza* and *wāw+hamza*. Almost all sources write clitics as part of the same orthographic word as their host, whereas this textbook separates them with a hyphen. In all these respects, and in others to be seen below, this textbook calques the Latin script orthography that was dominant by the time of its publication.

4.2.1 Consonants

A majority of the consonants are shared between Arabic and Berber; as a result, most consonants are represented identically in all texts. The exceptions, however, are informative.

The consonant *ḏ*, where present, is overwhelmingly written as *ḍād*, whose pronunciation (except in careful speech) is the same and which has a far higher frequency in Arabic than *ḏāʾ*. The only sources to use *ḏāʾ* consistently are the book of Belaid, a linguist teaching Arabic, and the Qur'an translation, whose authors had presumably been trained in recitation and hence in careful pronunciation of Classical Arabic phonemes; both would be expected to be unusually sensitive to the distinction in "proper" Arabic pronunciation between *ḍ* and *ḏ*.

In sources from the 19th century or earlier, *g* is normally represented by a *kāf*, with or without three extra dots. The choice to base the letter for *g* on *k* corresponds to traditions established in the region much earlier (cf. Ibn Khaldūn (1969, 32), van den Boogert (1997)). However, a three-dotted *qāf* is overwhelmingly predominant in 20th century sources, anticipated by the Bible translation of 1833 (but

not by Newman's (1847) transcription from the same source!) and by Rinn (1887). This reflects the influence of bilingualism in Algerian Arabic, where the reflex of classical *qāf* is often *g* rather than *q*, and corresponds to normal modern Algerian practice in Standard Arabic when transcribing place names.

No efforts are made to distinguish *ẓ* from *z* until the late 20th century – and even in those texts, actual usage is not entirely consistent. Hrouch et al.'s use of *zāy* with a circumflex may be a rare instance of Belaid's work influencing others, but may represent independent innovation; in either case, a possible source of inspiration at this late period is Sorani Kurdish, which uses ڕ. and ڵ. The Qur'ān translation's use of a *zāy* shape with three dots follows in the footsteps of Chafik (1990) and Baamrani (2003) for Moroccan Berber, and takes advantage of this character's presence in Unicode and in major digital fonts.

Very few sources seem to have felt any need to distinguish *ts* from its usual underlying source *tt*. Of those that do, Belaid (who adds a small *sīn* on top of the letter *tāʾ*) and the Qur'an translators (who use a cluster *t*+*s*) have already been observed to be particularly sensitive to minor details of pronunciation, while Anonymous (1833) may well have created the character (the letter shape of *sīn* plus the two dots of *tāʾ*) in response to the requests of the American consul who commissioned him.

The low textual frequency of *č* makes it difficult to generalise with confidence, but in texts predating the 1980s, it is normally represented as *šīn* except in works produced for foreigners.The latter often (but not always) use *jīm* with three dots, a letter borrowed directly from Ottoman Turkish. Among more recent works, Ferad and the Qur'an translators represent it as a cluster (*t*+*š*), again corresponding to common practice in Arabic transcriptions of place names in Algeria, while Hrouch et al. add a haček to *sīn*, probably calqued from the Latin orthography.

The treatment of spirants is further complicated by the existence of variation within Kabyle; not all dialects necessarily have *ḇ*, for example. In any case, we find that 19th century materials consistently write *bāʾ* in the positions where we might expect *ḇ*. Most 20th century sources do the same, but Udifella and the Qur'an translators both represent it as a *bāʾ* shape with three dots below – it is unclear whether the latter got it from the former, but in view of the religious content of Udifella's poetry it cannot be excluded. Belaid is strikingly inconsistent on this point (and the combination of *qāf* with three dots for *g* and *fāʾ* with three dots for *ḇ* that he uses in his sample text creates intolerable systematic ambiguity.)

Spirant *ḵ* and *ḡ* are almost never distinguished from stops *k* and *g*. Belaid makes sporadic attempts to do so, but the only source systematically distinguishing them is the Qur'an translation. The novelty of the latter's choice is clear from the signs chosen. The choice of *jīm* with three dots for *ḡ* can only reflect awareness of the Egyptian pronunciation of *jīm* as *g*, an awareness that would have

been unlikely in earlier periods, while that of *kāf* with a line on top for *ḵ* likewise reflects the assumption that an Algerian audience will not be familiar with the usage of this letter for *g*, an assumption which is certainly correct at present but which would have been much less well-founded in the 19th century, when some literate people would likely have been aware of Ottoman Turkish practice.

There is little consistency in the representation of labiovelarisation even within a single text, but insofar as it is represented it is usually represented using a *ḍamma*. Basset (1887) and Boulifa (1904) occasionally use a *wāw* instead.

4.2.2 Vowels

Due to the system incongruity discussed earlier, vowels are a much more conspicuous source of variation than consonants, despite being much fewer in number. The schwa is almost always written with a *fatḥa*, but for tense vowels there is considerable vacillation between short and long representations. A given word's transcription may differ even within the same manuscript, but five main distributions may be identified, falling along a continuum from preference for short vowels (greatest in early and religious sources) to preference for long ones (greatest in recent and secular ones):

1. All vowels are written short, without exception (except sometimes in Arabic loans, which optionally retain their original orthography). This is observed in the Aheddad poem, e.g. اِسُفْغِتْ <isufɣiṯ> *issufɣiṯ* 'he got him out', اَنْبِ <anbi> *ennḇi* 'the prophet' (an Arabic loan!), and in Tagamount's (1995) dictionary. In this system, Arabic vowels are used purely for their quality, and no effort is made to identify length with tenseness.
2. Most vowels are written short; in the rare exceptions (more or less common depending on the source), only one vowel per orthographic word is written long (excluding Arabic loans). This vowel is always tense, and normally in a penultimate or final open syllable. This is observed in KA21, e.g. اُرِمْنِعَرَ <urimniɛara> *ur imniɛ ara* 'it does not prevent', اَسَكُّذِيسْ <asakkuḏīs> *asekkuḏ-is* 'his seeing' (Gutova 2011, 56, 64); in the Aheddad wordlist, e.g. اُرَتْنُّدُّمَرَ <uratnuddumara> *ur tnuddum ara* 'do not get sleepy', اُرْتَّطِّسَارَ <urṯaṭṭisāra> *ur ṯeṭṭis ara* 'do not sleep'; in the Bible translations, e.g. إِنَّيَاسْ اَيَرْقَزْ <innayās ayargaz> *inna-yas ay argaz* 'he said to him: O man' (Anonymous 1833, Lk. 12:14), اَكَّقِي لَوَنَقَرَغْ <akkagī lawanaqaraɣ> *akkagi la-awen eqqareɣ* 'thus I say to you' (British and Foreign Bible Society (1904), Lk. 15:10, quoted in British and Foreign Bible Society (1965)); and in Hanoteau (1867), e.g. ذَكْجُفَارْ <ḏakjufār> *ḏeg (i)jufar* 'in the lower clothing' (p. 205), ذِمُلَبَنْ <ḏimulaban> *ḏ imulaben* 'are lizards' (p. 179). In most examples the long vowel seems to be stressed, and hence is in fact phonetically

longer; this orthography thus indicates a sensitivity to the length distinction that is phonemic in Classical Arabic but merely phonetic in Kabyle.

3. If a word contains any tense vowels, usually either one or two of them is written long (although, more or less frequently depending on the source, they may also all be written short). This is observed in Brosselard (1844), e.g. تْمَشْهُوثْ <ṯamašahūṯ> *ṯamašahuṯ* 'story' (p. 129), أُولَاغَرْ <ūlāɣar> *ulaɣer* 'there is no point' (p. 286); in the qanun of Ait Ali Ouharzoune, e.g. اَيَزِيضْ <ayazīḍ> *ayaẓiḍ* 'rooster', يَسْكِيدْبًا <yaskīdban> *yeskiddben* 'accusing of lying', اَزَالِيسْ <azālīs> *azal-is* 'its value'; in Rinn (1887), e.g. بابور <b'bwr> *babur* 'Babor (place name)' (p. 58), سعوذيو <sʕwḏyw> *s uʕuḍiw* 'with a horse' (p. 56); in Udifella's poems, e.g. اَكْيُوغَالْ <akyūɣāl> *a-k yuɣal* 'he will return you', التِّهُدِيثْ <attihudīṯ> *n teyhudiṯ* 'of malice', يَتْنَذِيَاسْ <yatnaḍiyās> *yettnaḍi-yas* 'he seeks for him'; in the Qur'an translations, e.g. لُمُورْ <lumur> *lumur* 'affairs', تِمِثَّلْ <ṯimiṯāl> *ṯimiṯal* 'example'. In the *qanun*, unusually, even schwas may optionally be written long in final syllables: يَرْوَالْ <yarwāl> *yerwel* 'he fled'. This practice too reflects phonetic reality to some extent, in that vowels before the main stress seem to have shorter duration.
4. All tense vowels are normally written long, except sometimes in antepenultimate or earlier positions. This is observed in Boulifa (1904), in Ferad (2003), e.g. أُوحْدِيقْ <ʾūḥḍīq> *uḥḍiq* 'polite' (p. 240), ثَازَالِيثْ <ṯāzālīṯ> *ṯaẓallit* 'prayer' (p. 257), أَذَمِيرُوحْ <ʾaḍamīrūḥ> *aḍ-am iruḥ* 'it will go for you' (p. 235), and in Belaid (1995), e.g. الطوموبيل \ الطوموبيل <ʾlṭ(w)mwbyl> *eṭṭumubil* 'car' (p. 166), ثمليليث أنتوثلايين <ṯmlylyṯ ʾntwṯl'yyn> *ṯimliliṯ n tuṯlayin* 'language contact' (p. 216). In that most tense vowels are written as long, this probably reflects interference from colloquial Algerian Arabic, which has the same vowel system as Kabyle, but with schwa as the normal reflex of Classical short vowels, and tense vowels for Classical long vowels.
5. All vowels are written using *matres lectionis*, even schwa (only in the textbook of Hrouch et al. (2003)). This practice can be explained neither in terms of Classical nor of Algerian Arabic; rather, like much else in this textbook, is obviously calqued from the Latin orthography which by that time was predominant for Kabyle. Examples: تيخسي <tixsi> '*ewe*', ئزآم <ʾizem> *izem* 'lion'.

Basset (1887) cannot readily be placed in any one category, since he quotes texts from several different sources.

In all but the last of these systems, schwa is written using a *fatḥah*, as is *a* in certain circumstances. This, to varying degrees, creates ambiguity between *e* and *a*.

5 Conclusions

Kabyle has repeatedly been written in Arabic script since the pre-colonial period, particularly in the east. This practice was a natural outgrowth of religious education, and as such was never suppressed entirely despite difficult conditions. However, despite a history of at least three centuries, only those conventions imported directly from Arabic are stable; where orthographic choices must be made, those choices vary substantially from one source to the next. A few long-term trends are observable, in particular the replacement of *kāf* with *qāf* as a basis for transcribing *g* and an increasing preference for matres lectionis over diacritics in representing vowels, but most of the variation is purely individual. A necessary precondition for this variation was the existence of significant divergences between Arabic and Kabyle phonology, notably greater than those observed for Shilha. While those divergences made the task of establishing a consistent orthography more difficult in principle, the fact that no one solution to these challenges has ever been consistently adopted for long primarily reflects historical rather than linguistic facts. At no point has any one writer's or school's Arabic-script work been sufficiently widely read to be imitated, and all but the most prominent of one generation's orthographic innovations have been forgotten by the next. This is to be expected given the repeated massive upheavals that have affected education in the region and the fact that teaching Kabyle has never been a primary goal of any of its educational systems, and is unlikely to change: while the belated introduction of Kabyle teaching after 1995 has provided a powerful new tool for standardisation, the general preference among today's Kabyle-language writers and teachers for Latin makes it practically certain that this standardisation will only affect the now-dominant, and already relatively stable, Latin orthography.

This chapter's conclusions are fairly solid insofar as they concern the print era; more data on political posters and album covers would be helpful, but appears unlikely to alter the overall picture. However, data on Kabyle manuscripts, particularly those dating to the pre-colonial period, remains frustratingly scanty. There is no reason to believe that all relevant manuscripts are known, and almost all of those that have been reported in the literature are still unpublished. Any conclusions concerning manuscript practices are thus necessarily somewhat tentative, and should be tested against newly discovered or newly published manuscripts as they become available.

References

Aïssani, Djamel (1998), 'Les écrits de langue berbère de la collection de manuscrits Oulahibib (Bédjaia)', in *Etudes et Documents Berbères*, 15–16: 81–99.

Aïssani, Djamel(2012), 'Le lexique manuscrit Arabe dialectal - Kabyle de la Zawiyya historique de Cheikh Aheddad', in Kamal Naït-Zerrad (ed.), *Linguistique de corpus: de la constitution à l'exploitation*, Paris: Inalco, 15–32.

Allemand-Lavigerie, Charles-Martial (ed.) (1868), *Catéchisme du diocèse d'Alger*. Algiers: Bastide.

Allemand-Lavigerie, Charles-Martial (ed.) (1869), *Aktab en tibratin d endjilen i idomniken tifestin ousoggas, s taqebailit / Livre des épitres et évangiles pour les dimanches et fêtes de l'année en langue kabyle*, Algiers: Bastide.

Anonymous (1833), *Extrait d'une traduction ms. en langue berbère de quelques parties de l'Ecriture Saine : contenant XII chapitres de S. Luc*. Londres: Société Biblique Britannique et Etrangère.

Anonymous (2003), *Sūrat al-Fātiḥah wa-l-'Ajzā' al-Thalāthah al-'Akhīrah Qad Samiʕa wa-Tabārak wa-ʕAmma wa-Tarjamat Maʕānīhā 'ilā l-Lughah al-' Amāzīghiyyah / Translation of the Meanings of Surah al-Faatihah and the last Three Juz' of the Quran in the Kabyle (Tamazight) Language*. Madinah: Mujammaʕ al-Malik Fahd li-Ṭibāʕat al-Maṣḥaf al-Sharīf.

Anonymous (2009), Le qânoûn des ait ali ouharzoune (en dialecte kabyle, transcrit en lettres arabes). http://aitaliouharzoune-bs.vip-blog.com/vip/article/4850323,-Le-qânoûn-des-ait-ali-ouharzoune(en-dialecte-kabyletranscrit-en-lettres-arabes)-.html (30 September, 2013).

Baamrani, Jouhadi Lhocine (2003), *Tarjamat maʕānī al-Qur'ān al-Karīm: Nūr ʕalā nūr / Tifawt f tifawt, bi-l-luġah al-'amāzīġiyyah*. Casablanca: Al-Najāḥ al-Jadīdah.

Basset, René (1887), *Manuel de langue kabyle (dialecte zouaoua) : grammaire, bibliographie, chrestomathie et lexique*. Paris: Maisonneuve & Ch. Leclerc.

Belaid, Salah (1999), *Fī al-Mas'alah al-'Amāzīghiyyah*. Algiers: Dār Hūmā.

Bernard, A., and Louis Millot (1933), 'Les qanouns kabyles dans l'ouvrage de Hanoteau et Letourneaux', in *Revue des études islamiques*, 7: 1–44.

Boogert, Nico van den (1997), *The Berber Literary Tradition of the Sous: with an edition and translation of "The Ocean of Tears" by Muḥammad Awzal (d. 1749)*, Leiden: Nederlands Instituut voor het Nabije Osten.

Boulifa, Si Ammar ben Saïd (1897), *Une première année de langue kabyle : dialecte zouaoua*. Algiers: Adolphe Jourdan.

Boulifa, Si Ammar ben Saïd (1904), *Recueil de poésies kabyles (texte zouaoua) traduites, annotés et précédées d'une étude sur la femme kabyle*. Algiers: Adolphe Jourdan. http://www.e-corpus.org/notices/104452/gallery/880441.

Boulifa, Si Ammar ben Saïd (1913), *Méthode de langue kabyle : cours de deuxième année. Etude linguistique et sociologique sur la Kabylie du Djurdjura. Texte zouaoua suivi d'un glossaire*. Algiers: Adolphe Jourdan.

Bouteflika, Abdelaziz (2002), 'Loi n° 02-03 du 27 Moharram 1423 correspondant au 10 avril 2002 portant révision constitutionnelle', in *Journal Officiel de la République Algérienne*, 41: 24.

Bouteflika, Abdelaziz (2016), Loi n° 16-01 du 26 Joumada El Oula 1437 correspondant au 6 mars 2016 portant révision constitutionnelle. *Journal Officiel de la République Algérienne*, 55: 14.

British and Foreign Bible Society (1904), *Evangile de St. Luc (caractères arabes)*. Algiers: Dépôt Biblique.

British and Foreign Bible Society (1965), *The Gospel in Many Tongues: Specimens of 875 languages in which the British and Foreign Bible Society has published or circulated some portion of the Bible*. London: British and Foreign Bible Society.

Brosselard, Ch. (1844), *Dictionnaire français-berbère (dialecte écrit et parlé par les Kabaïles de la division d'Alger)*. Paris: Imprimerie Royale.

Brugnatelli, Vermondo (ed.) (2006), 'Qanun At Ali u Herzun', Wikimedia. http://commons.wikimedia.org/wiki/File:Qanun_At_Ali_u_Herzun.jpg (8 August, 2018).

Brugnatelli, Vermondo (2011), 'Some grammatical features of Ancient Eastern Berber (the language of the Mudawwana)', in Luca Busetto (ed.), *Studies on Language and African Linguistics in Honour of Marcello Lamberti*, Milano: Qu.A.S.A.R, 29–40. http://unimib.academia.edu/VermondoBrugnatelli/Papers/1304580/Some_grammatical_features_of_Ancient_Eastern_Berber_the_language_of_the_Mudawwana_ (8 August, 2018).

Calassanti-Motylinski, Gustave Adolphe de (1905), Le Manuscrit arabo-berbère de Zouagha découvert par M. Rebillet : Notice sommaire et extraits. *Actes du XVIe Congrès des Orientalistes (Alger 1905)*, vol. II, Paris, 69–78.

Canton, William (1904), *A history of the British and Foreign Bible Society*, London: John Murray.

Chachoua, Kamel (2010), 'Radiographie de trois traductions du Coran en kabyle', in *Revue des mondes musulmans et de la Méditerranée*, 128: 231–245. doi:10.4000/remmm.6924 (1 October, 2013).

Chafik, Mohamed (1990), *Al-Muʕjam Al-ʕArabī Al-'Amāzīghī*. Rabat: Al-Maʕārif Al-Jadīdah.

Creusat, P. J.-B. (1873), *Essai de dictionnaire français-kabyle (zouaoua), précédé des éléments de cette langue*. Algiers: Adolphe Jourdan.

Cuendet, Eugène, T Hocart and H. G Lamb (1893), *Indjil n'Sidna-Aisa el-Masih´akken itsouaḱtheb s'Marqous*. London: Printed for the British and Foreign Bible Society.

Daumas, Eugène (1847), *La Grande Kabylie: Etudes historiques*. Paris: Hachette.

Delaporte, J.H. (1836), 'Vocabulaire berbère', in *Journal Asiatique*, 97–122.

Dirèche, Karima (2007), 'Les écoles catholiques dans la Kabylie du XIXe siècle', in *Cahiers de la Méditerranée*, 75: 17–27 (1 October, 2013).

Ferad, Mohamed Arezki (2003), *'Azaffūn: Tārīkh wa thaqāfah*. Tizi Ouzou: Dār al-Amal.

Ferad, Mohamed Arezki (2004), *Al-'Amāzīghiyyah: "ārā" ... wa-'amthāl (Tībāzah namūdhajan)*. Algiers: Dār Hūmā.

Ferad, Mohamed Arezki (2007), *"Iṭlālah ʕalā Manṭiqat al-Qabā"il*. Tizi Ouzou: Dār al-'Amal.

Ferad, Mohamed Arezki (2011), *Al-Ḥāj Muḥammad al-Saʕīd Tāzrūt: "Iṭlālah ʕalā al-Shiʕr al-Ṣūfī al-"Amāzīghī*. Tizi Ouzou: Dār al-'Amal.

Feraoun, Mouloud (1960), *Les poèmes de Si Mohand*. Paris: Editions de Minuit.

Gaci, Zohra (2011), Quel système d'écriture pour la langue berbère (le kabyle) ? Tizi Ouzou: Université Mouloud Mammeri de Tizi Ouzou Mémoire de Magister. http://www.ummto.dz/IMG/pdf/MEMOIRE_GACI_Zohra.pdf.

Gahlouz, Mustapha (2011), *Les qanouns kabyles : Anthropologie juridique du groupement social villageois de Kabylie*. Paris: L'Harmattan.

GEHIMAB (2006), *Les Manuscrits de Langue Berbère de la Kabylie*. Bejaia: GEHIMAB. http://www.gehimab.org/depliants/depliants/depliant21.pdf.

Gourliau, Ernest (1893), *La conversation française-kabyle, dialecte zouaoua, avec figuration en caractères latins*. Miliana: Légendre.

Gutova, Evgeniya (2011), *The Sanusi Creed in Kabyle Berber: Manuscript KA 21 from the Lmuhub Ulahbib Library*, Leiden: University of Leiden MA thesis.

Haddadou, Mohand Akli (2004), *L'alphabet berbère : Des écritures libyques aux transcriptions modernes*. Béjaïa: Azur Editions.
Hanoteau, Adolphe (1858), *Essai de Grammaire Kabyle*. Paris: Challamel.
Hanoteau, Adolphe (1867), *Poésies populaires de la Kabylie du Jurjura*. Algiers: Bastide.
Heggoy, Alf Andrew (1973), 'Education in French Algeria: An Essay on Cultural Conflict', in *Comparative Education Review*, 17(2): 180–197 (4 November, 2014).
Hodgson, William B. (1834), 'Grammatical Sketch and Specimens of the Berber Language: Preceded by Four Letters on Berber Etymologies', in *American Philosophical Society*, 4: 1–48.
Hrouch, Mouhoub, Ali Lounis, Remdane Achour and Kebir Boussaad (2010), *Aḏlis-iw n tmaziɣt: aseggas wis rebɛa n ulmuḏ amenzu*. Algiers: Office Nationale des Publications Scolaires.
Huyghe, G. (1901), *Qamus qbaili-rumi / Dictionnaire kabyle-français*. Paris: Imprimerie Nationale.
Ibn Ḥamādūsh, ʕAbd al-Razzāq Muḥammad (2001), *Kashf al-Rumūz fī Sharḥ al-ʕAqāqīr w-al-'Aʕshāb*, Beirut: Dār al-Kutub al-ʕIlmiyyah.
Ibn Khaldūn, Abū Zayd `Abd al-Raḥmān ibn Muḥammad (1969), *The Muqaddimah: An Introduction to History. (Abridged Edition)*. (Ed.) N. J. Dawood. (Trans.) Franz Rosenthal. Princeton: Princeton University Press.
Khouas, Ahmed Ben (1881), *Notions succinctes de grammaire kabyle, rédigées sous la direction du Sous-Préfet de Tizi-Ouzou et avec l'autorisation de M. le Préfet d'Alger*. Algiers: Adolphe Jourdan.
Lakhdari, Samira (2013), 'Ibn al-Faqīr' en Mulud Ferɛun es Teqbayliṯ. *Ech Chaab*. http://www.ech-chaab.net/ar/2012-06-09-14-44-43/%D8%A7%D9%84%D8%A3%D9%85%D8%A7%D8%B2%D9%8A%D8%BA%D9%8A%D8%A9/item/5108-%C2%AB%D8%A7%D8%A8%D9%86-%D8%A7%D9%84%D9%81%D9%82%D9%8A%D8%B1%C2%BB-%D8%A7%D9%86-%D9%85%D9%88%D9%84%D9%88%D8%AF-%D9%81%D8%B1%D8%B9%D9%88%D9%86-%D8%A7%D8%B3-%D8%AA%D9%82%D8%A8%D8%A7%D9%8A%D9%84%D9%8A%D8%AB.html.
Luciani, Jean Dominique (1893), 'El-H'aoudh : Manuscrit berbère de la Bibliothèque-Musée d'Alger', in *Revue Africaine*, 37: 151–180.
Luciani, Jean Dominique (1899), 'Chansons Kabyles de Smaïl Azikkiou', in *Revue Africaine*, 43: 17–33.
Mammeri, Mouloud (1976), *Tajerrumt n tmazight (tantala taqbayliṯ)*, Paris: Maspéro.
Mechehed, Djamel-Eddine (2007), 'L'organisation des notices de catalogage des manuscrits arabes ou berbères : cas de la collection Ulahbib, Kabylie', in *Les manuscrits berbères au Maghreb et dans les collections européennes*. Atelier Perrousseaux.
Millot, Louis (1922), 'Le qanoun de Maâtqa', in *Hespéris*, 2: 193–208.
Millot, Louis (1928), 'Le qanoun des Aït Hichem', in *Mémorial Henri Basset. Nouvelles études nord-africaines et orientales*, 153–167. (Publications de l'Institut Des Hautes-Études Marocaines XVIII). Paris: Paul Geuthner.
Ministry of Education (2003), Al-Wathīqah al-Murāfiqah li-Manāhij Māddat al-Lughah al-'Amāzīghiyyah bi-l-Lughah al-ʕArabiyyah. http://www.onefd.edu.dz/programmes/MOYEN/MOYEN/2AM/AmazigheA.htm.
Mouliéras, Auguste (1892), *Les fourberies de Si Djeh'a*, Paris: Ernest Leroux.
Mouliéras, Auguste (1893), *Légendes et contes merveilleux de la Grande Kabylie*. Paris: Ernest Leroux.
Newman, Francis William (1847), Third chapter of Matthew, in the Kabail (or Algiers) dialect of Berber ... translated by F.W. Newman. In James Richardson (ed.), *Third chapter of Matthew in the Ghadamsi dialect [by Ibn Mūsā ibn al-Qāsim], decyphered, punctuated*

and translated by F.W. Newman, ... - Third chapter of Matthew, in the Kabail (or Algiers) dialect of Berber ... translated by F.W. Newman - Familiar conversations in four languages: English and Arabic and Ghadamsee and Touarghee [by Ibn Mūsā]. - Vocabulary in Arabic and Ghadamsee [by Ibn Mūsā]. - Fragment of Grammar. The Ghadamsee and Touarghee languages. - ... Ben Mousa el-Kasem of Ghadames. London: T. E. Harrison.

Olivier, P. (1878), *Dictionnaire français-kabyle*. Le Puy: Freydier.

Ould-Braham, O. (1986), 'Un qanoun kabyle recueilli au XIXe siècle', in *Etudes et Documents Berbères*, 1: 68–77.

Radio 2. *Al-'Iḏāʕah Al-Jazā'iriyyah | Al-Qanāh Al-Ţāniyah*. http://www.radioalgerie.dz/ar/chaine2 (2 October, 2013).

Rebahi, Sadek (2009), Qaṣā'id ShiʕriyyahfīAl-Buʕd al-Rūḥī li-l-Turāth al-'Amāzīghī. In Ministry of Religious Affairs and Waqfs (ed.), *Al-Multaqā al-Rābiʕ li-l-Buʕd al-Rūḥī fī al-Turāth al-Waṭanī al-'Amāzīghī: Mukhtaṣarāt al-Mudākhalāt*, 53–54. Bejaia: Ministry of Religious Affairs and Waqfs.

Rinn, L. (1887), 'Deux chansons kabyles sur l'insurrection de 1871', in *Revue Africaine*, 31: 55–71.

Rubrique Radar (2013), *Louable initiative du journal arabophone : Le quotidien "Echaâb" lance une page en tamazight*. Liberté. http://www.liberte-algerie.com/radar/le-quotidien-echaab-lance-une-page-en-tamazight-louable-initiative-du-journal-arabophone-195254.

Rutherfurd, John, and Edward H. Glenny (1900), *The gospel in North Africa : in two parts*. London: Percy Lund, Humphries. http://archive.org/details/gospelinnorthafr00ruth.

Sedira, Belkassem Ben (1887), *Cours de langue kabyle*, Algiers: Adolphe Jourdan.

Sini, Chérif (2004), 'Quelle graphie pour écrire tamazight? Les raisons d'un choix', in *Nouvelles études berbères (II) – Le verbe et autres articles* (Berber Studies, 8), Colog*ne: Rüdiger* Köppe, 159–170

Ţabarkān, Salwā (2012), *Al-'Alghāz al-Shaʕbiyyah al-Qabā'iliyyah bi-minţaqatay Wāḍyah wa-ʕAzāzgah (Muqārabah niyāsiyyah)*. Tizi Ouzou: Université Mouloud Mammeri Masters' thesis. http://www.ummto.dz/IMG/pdf/MEMOIRE-3.pdf.

Tagamount, Azzedine (ed.) (1995), *Dictionnaire de Tamaziɣt : Edition trilingue Tamaziɣt-Français-Arabe / Munjid al-Lughah al-'Amāzīghiyyah, thulāthī al-lughāt: al-'Amāzīghiyyah, al-Faransiyyah, al-ʕArabiyyah*. Reghaia: Berti.

Ţayyib, Sī Ḥājj Muḥand Muḥand (2012), *Al-Qur'ān al-Karīm wa-Tarjamat Maʕānīhā "ilā l-Lughah al-"Amāzīghiyyah [al-Lahjah al-Qabā'iliyyah] / Leqṛan lɛaḍim d uteṛjem n lemɛani-is ɣer lluɣa n tmaziɣt (s teqbaylit)*. (Ed.) Riḍā Būshāmah and Muḥammad Ţāhir Tīqammūnīn. Madinah: MujammaʕalMalik Fahd li-Ţibāʕat al-Maṣḥaf al-Sharīf.

United Nations (1955), *Demographic Yearbook 1955*. New York: Statistical Office of the United Nations, Department of Economic and Social Affairs.

Al-Warţīlānī, Al-Ḥusayn ibn Muḥammad (1994), *Nuzhat al-'Anḍār fī Faḍl ʕIlm al-Tārīkh wal-'Akhbār*. (Ed.) Muḥammad Ibn Abī Shanab. (Al-Juġrāfiyā Al-'Islāmiyyah 192). Frankfurt am Main: Institut für Geschichte der Arabisch-Islamischen Wissenschaften.

Yacine, Tassadit (1987), *Poésie berbère et identité : Qasi Udifella, héraut des At Sidi Braham*. Paris: Editions de la Maison des Sciences de l'Homme.

Zeroual, Liamine (1995), 'Décret présidentiel n° 95-147 du 27 Dhou El Hidja 1415 correspondant au 27 mai 1995, portant création du Haut commissariat chargé de la réhabilitation de l'amazighité et de la promotion de la langue amazighe', in *Journal Officiel de la République Algérienne*, 34: 29.

Alessandro Gori

Beyond *ʿağamī* in Ethiopia: a short Note on an Arabic-Islamic Collection of Texts written in Ethiopian Script (*fidäl*)

Abstract: As in many other countries in Sub-Saharan Africa, Muslims in Ethiopia have produced a substantial amount of literature in their local languages, mostly using the Arabic script rather than the Ethiopic syllabary (*fidäl*), which has been connected with the Ethiopian Christian Church and state for centuries. Recently, after *fidäl* was adopted to write the language of the Harari Muslim people, a manuscript was written in which texts in Arabic and Old Harari were copied in Ethiopic script. I have analysed the manuscript in this paper, highlighting the strategies that the copyist followed in order to transcribe Arabic into *fidäl*, and have attempted to place the work within the general framework of Islamic-Christian relationships in Ethiopia.

1 *Fidäl*, Arabic and Islamic languages in Ethiopia: sketching a framework

The usage of the Ethiopian script (*fidäl*)[1] to write Arabic and other 'Islamic languages' of the Horn of Africa[2] is not completely unknown, but has so far been scarcely documented and analysed. While it is actually a relatively rare phenomenon, it does deserve some attention, as its study could be of some interest for linguists (particularly sociolinguists) and more generally for researchers investigating the social and cultural history of the Muslims of the Horn of Africa.

The data presented in this paper were compiled within the framework of and thanks to the research project 'Islam in the Horn of Africa: a Comparative Literary Approach' (IslHornAfr, European Research Council Advanced Grant 322849).

1 For a first introduction to Ethiopian script, see Frantsouzoff 2010.

2 I use the concept of 'Islamic languages' here in a very loose way as a generic term to refer to languages spoken (also) by Muslims and with no intention to touch on the issues discussed by Bausani 1967, 1978, 1981, Brenner and Last 1985, Zappa 2004, 2009 and Gori 2015.

https://doi.org/10.1515/9783110639063-012

Some marginal paratexts among Christian Ethiopian manuscripts in Gə'əz and Amharic have been found to include Arabic words, sentences and also short fully-fledged texts written in Ethiopian and/or in Arabic scripts.[3] Christian Ethiopian astronomical and calendar treatises (e.g. the *Baḥrä hassab*, 'The sea of calculation'),[4] contain many Arabic words and expressions. In the famous 16th-century Christian treatise *Anqaṣä amin* (the 'Door of the faith'), written in Gə'əz, Arabic-Islamic religious terms are quoted and transcribed using *fidäl* script.[5] Also, the 20th-century Amharic Christian *Sylloges* of *šayḫ Zäkaryas* contain several words and phrases taken from Arabic.[6] During my recent examination of MS EMML 6239, I identified a bilingual Arabic-Amharic version of *šayḫ* Zäkaryas' work written entirely in Ethiopian *fidäl*.[7] Arabic-Islamic words are also scattered throughout Christian Ethiopian magical literature, where they function as 'abracadabra' to enhance the mystical power of the texts.[8] The writing down of Islamic texts in Arabic or Amharic or their transliteration into one of these languages was usually carried out by or for the sake of Western scholars working in Ethiopia.[9]

3 Examples can be found among the manuscripts catalogued by the Hill Monastic Manuscript Library in Collegeville, MN, USA in the Ethiopian Manuscript Microfilm Library (EMML). See for example in MSS EMML 1056 fols 171b–174a (Gə'əz equivalent of Turkish and Arabic words); 1105 fol. 129b (name of Mary in Arabic); 1187, fol. 73b and 74a (names of the months in Arabic); 1168 2a (Arabic numerals); 1419 fol. 42b (Arabic names of the stars and of the months); 1373 fol. 1b and recto of the last rear guard leaf (exercising in Arabic); 1599, last rear guard leaf (Arabic numerals); 2011, fol. 72a (Arabic prayer in Ethiopian letters); 2054, fol. 247a (note in Arabic and Gə'əz); 2436, fol. 160a (prayers in Arabic written in Ethiopian letters); 2531, fol. 175b (explanation of Arabic and Hebrew words); 2548, fol. 4a (name of Mary in various different languages); 2630, fol. 165b (name of Mary); 3127, fol. 174b (pen trial); 4036, fol. 4b (theological formula Arabic-Amharic); 4836, fol. 78b (unreadable note); 4909, fol. 145b and fol. 149a (stamp of nəguś Mika'el of Wällo); and 4922 fol. 1b (a holy image printed in some Arab country).

4 Pietruschka/Amha Asfaw/Getatchew Haile 2003.

5 The *Anqaṣä amin* is a polemical work written by Ǝnbaqom a learned Muslim man who converted to Ethiopian Orthodox Christianity in the first half of the 16th century and subsequently wrote a defence of the Christian faith also using verses from the Qur'ān. The Ethiopic *Anqaṣä amin* text was edited (and translated into French) by Emeri van Donzel in 1969. For an analysis of the Arabic passages, see especially van Donzel 1969, 39–43.

6 On the *Sylloges of šayḫ Zäkaryas*, see Gori 2001 and Gori 2003. On the author, a learned Muslim man (c. 1845–1920), who converted to Christianity towards the end of the 19th century, see Gori 2014a.

7 I gave a presentation on this text at the *Colloque Manuscrits chrétiens et islamiques d'Ethiopie (XIIIe–XXe siècle)* in Paris on Dec. 12 2014 (see Gori 2014b).

8 A remarkable example of a magical prayer made up of passages of the Qur'ān is discussed by Marcel Griaule (see Griaule 1930, 109–110).

9 For example, see the texts published by Enrico Cerulli in 1926; and the collection of Qur'ānic

Moving from the Ethiopian Christian to the Ethiopian Islamic cultural landscape, the existence of Arabic-Islamic texts written in *fidäl* becomes less vague, at least during the modern era. A cultural and religious prejudice identifying Ethiopian script with Ethiopian Christianity may have hindered the usage of *fidäl* among Muslims at an early stage. However, since the mid-20th-century emergence of Islamic books printed in Amharic and Tigrinya, Ethiopian script has also begun to be used to write scattered words and phrases in Arabic. After the fall of the Socialist regime in 1991, a more or less religiously oriented Muslim book production began to blossom in Ethiopia, and *fidäl* became more and more widespread among Amharic- and Tigrinya-speaking Muslims who make use of it to transliterate passages from the Qur'an, Hadith and other fundamental texts of the Islamic tradition for books and booklets published in local languages. The subsequent birth and diffusion of a kind of Islamic calligraphic practice in Ethiopian script came to confirm a fairly generalised acceptance of *fidäl* among many Ethiopian Muslims.

An especially remarkable position in the general picture I am sketching here is occupied by Harari, the Semitic language of the city of Harar, spoken by the ethnic group of the Hararis. Traditionally written in Arabic script since at least the beginning of the 18th century,[10] Harari texts written in Ethiopian script have been found to have emerged at the end of the 19th century, after the walled city was conquered by Mənilək in 1887 and incorporated into the modern Ethiopian state.[11] The connection between the Ethiopian writing system and the Harari language became even tighter after the fall of the Imperial state (1974) and under the Socialist regime. Harari texts in Ethiopian alphabet were published abroad[12] and

passages compiled by *aläqa* Tayyä for Eugen Mittwoch in Berlin who published them in 1906 in Ethiopian script.

10 For a general description of Harari literature in Arabic script, see Banti 2005, 2010. The language of the texts written in Arabic script has been labelled 'Old Harari', because it shows some quite substantial differences from modern Harari as spoken and written in Harar nowadays.

11 See for example the texts written by the secretary of ras Mäkwännən for Casimir Mondon-Vidailhet and published by Carlo Conti Rossini in 1919 (see Wagner 2004, 355, note 5 with other examples).

12 See for example the *Suwār malasāyāčč ṭabā* ('Voice of the revolutionary youth'), a handwritten ideological journal produced by the Harari Students' Association in Egypt (*Harari ardāwigāčč ahadənnat mugād misrābe*). On a website entitled *Everything Harar* (http://www.everythingharar.com; last accessed 04/02/2018), which is run by Harari individuals from the Diaspora, I found two issues of the periodical (vol. 3 no. 2, January 1980 and vol. 3 no. 3 [originally 2], February 1980) which I was able to retrieve and download. The periodical was apparently produced in Egypt, but no information is available or traceable about its origin and periodicity. The content of the texts is extremely interesting from both a politico-cultural and from a purely linguistic point

in Ethiopia,[13] fostering the diffusion of *fidäl* among Hararis (first and foremost among the intellectual elite).

After the pro-Soviet regime collapsed in 1991, a relatively short period of incertitude followed: different options (Latin, Arabic, and Ethiopian scripts) for the writing of the Harari language were discussed, until eventually the choice officially fell on *fidäl*.[14] The appearance in 1992 of the first Harari-Amharic dictionary published entirely in *fidäl* paved the way to further development of the connection between the Harari language and the Ethiopian script.[15]

Fidäl is nowadays well established among Harari speakers and readers and has progressively become more deeply rooted in the writing and reading practices and in the literary production of Hararis. At the same time, Latin script has not completely been entirely abandoned and still seems to be quite widespread among Hararis living outside Ethiopia,[16] and it is extensively used on the web.[17]

What was probably also instrumental in this impressive success of the Ethiopian script among the Hararis was the diffusion of Amharic in the city of Harar and especially in the Harari diaspora. Numerically substantial and economically lively communities of Harari speakers can be found in Addis Ababa, Dire Dawa and many other cities of Ethiopia. These scattered groups of Hararis (in particular those living in the capital) are more and more exposed to Amharic, are practically bilingual and tend to use written Amharic for their daily business and in connection with official and bureaucratic issues. The combined effect of the shift

---

of view. In January 1986, the first issue of another magazine, *Šahan* ('The ray'), was published in *fidäl* in Rome by the *Ḥarargey tāriḫ-wā ādā dafdafti mugad* ('Society for the research on Harari history and customs'). This periodical was subsequently transferred to Toronto (see Wagner 2003a, 6–7 and Wagner 2004, 355, note 6).

13 One example is the magazine *Aner* ('Make it beautiful'), published by the Harari Cultural Association of Dire Dawa with an uncertain periodicity since 1988. According to Wagner 2003a, 1, the first two issues (Sept. 1988 and Oct. 1988) were written in Arabic script. On the internet portal *Everything Harar*, I found a later issue (probably issue no. 2, 1990) produced in Ethiopic script.

14 For the cultural (and political) background of this decision, see Gibb 1998, 256–7 and 266, note 18. The first issue of *Harari ṭabā*, the official newspaper of the Harari National League, was published in Hədar 1984 CE (= November–December 1991). This event might be considered a sort of official endorsement of the adoption of the Ethiopian script by the officials of the Harari regional state.

15 Abdurahman Mäḥamäd Qorram 1984 CE (= 1992).

16 It is estimated that one-third of the Harari people live outside Ethiopia. The most conspicuous Harari communities are now living in Australia, Canada and the USA (see Gibb 2002).

17 It must be noted that, since the Roman keyboard is surely the most widespread on computers, Amharic and Tigrinya written on the Internet often use a very simplified transliteration system based on the Latin script. This writing practice has not so far been studied by scholars of Ethiopian languages.

to *fidäl* to write Harari and the wider and more intense exposure to Amharic has reduced the knowledge and usage of Arabic script among the Harari people and the knowledge of Arabic itself.

As a matter of fact, while Arabic is known as a language of culture by a very restricted circle of learned men and is used as a spoken medium of communication by Harari emigrants returning from jobs in the Gulf, the usage of Arabic script to write Harari seems to be vanishing completely. Moreover, the capability to read and understand traditional Old Harari texts written in Arabic script is apparently declining dramatically. Manuscripts in Arabic script (both in Arabic and in Old Harari language) are not copied anymore but are substituted by computer-typed books and electronic documents. The recent production of *fidäl* reprints of the *Kitāb al-farā'iḍ* ('The book of obligations'), probably the most famous piece of Harari literature, proves the sorely felt need for editions of traditional texts targeting the growing number of readers who are unable to read Arabic script.[18] It is this general sociolinguistic framework which forms the context for the origin of the manuscript I am going to deal with in the rest of this chapter.

2 Arabic in *fidäl*: the case of a Harar manuscript

The manuscript I am about to discuss was photographed in Harar on 21 September 2003 by Dr Simone Tarsitani (Durham University)[19] during an ethno-musicological research mission he was conducting. The item is basically a collection of different texts written down by the late Mr Abdi Abubakar Sufiyan[20] with a blue ballpoint pen (some parts are in red) on a personal organiser (15cm x 20cm) for the year 1990 of the Ethiopian Calendar (i.e. year 1997–1998 of the Gregorian calendar; the Ethiopian year starts on 11 September). The manuscript pages lack proper numeration, therefore in the following I will locate the texts according to

18 For some general information about the *Kitāb al-farā'iḍ*, see Wagner 2005; the text has been published under the title *Kitabul fara'id* in Ethiopian script in Dire Dawa at *Khaläf Mattämiya bet* (i.e. printing press) s.d. The PDF of the text can be downloaded from. *Everything Harar* at http://www.everythingharar.com (last accessed 05/02/2018).

19 I am very grateful to Dr Tarsitani for providing me with the images of the manuscript and for authorising me to study them. I would also like to acknowledge his generosity in providing me with all the information in his possession about the origin of the manuscript.

20 Mr Abdi Abubakar Sufiyan was actively involved in the devotional practices at the sanctuary of Aw Basor in Harar and made many efforts to keep the tradition of the Mawlid recitation alive for the benefit of the new generation. The origin of the present manuscript in *fidäl* can be found in the context of Mr Abdi's cultural and religious activities (personal communication from Dr Tarsitani).

the dates shown on the daily agenda (every single page is devoted to a different day, except for the weekends: Saturdays and Sundays share the same page).

The texts are written almost continuously from the first day of the year (11 September) until 6 April. The rest of the organiser is blank, except for the pages of 4 September and 8 September (the pages in between these two dates have been torn out) and the page of the weekend 9–10 September (the last two days of the Ethiopian year). The manuscript contains texts in three languages written in Ethiopian script: Old Harari, Arabic and Oromo. Besides, English in Latin script and Arabic in Arabic script (8 September, *sūrat al-kahf*, Qur'an 18, verses 1–10) are also copied in the agenda.

The texts copied into the manuscript belong to what I have elsewhere referred to as the 'Harari *Mawlid* collection',[21] that is, the constellation of poetical and prose texts (mainly in Arabic, but interspersed with sustained wide sections in Old Harari) which is usually recited by Hararis to solemnise the feast of the birthday of the Prophet and on other important public and private occasions of a religious but sometimes also a secular nature. The collection appears in a wide diffusion of manuscripts (both relatively ancient and very recent) in Ethiopia.

While no comprehensive analysis of the manuscript tradition of the 'Harari *Mawlid* collection' has as yet been carried out, the following data are available:

> In the library of the Institute of Ethiopian Studies at the University of Addis Ababa, the following eight manuscripts contain the 'Harari *Mawlid* collection': 264:2, 273:3, 1855:2, 2662:2–3, 2663:2, 2664:2 (incomplete), 2665:2, 2666:2 (see Gori and others 2014 under the relevant entries). The Abdallah Sharif private museum in Harar[22] keeps 16 codices of the 'Harari *Mawlid* collection':[23] 1435 Abdulahi Collection 93 (not dated, but certainly 20th century); 1436 Abdulahi Collection 94 (not dated, but certainly 20th century); 1437 Abdulahi Collection 95 (not dated, but certainly post-1974); 1439 Abdulahi Collection 97 (not dated, possibly 18th century); 1440 Abdulahi Collection 98 (dated 1395/1975); 1441 Abdulahi Collection 99 (not dated, but certainly post-1972); 1442 Abdulahi Collection 100 (not dated, possibly 19th century); 1443 Abdulahi Collection 101 (not dated, possibly 18th century); 1527 Abdulahi Collection 185 (dated 29 May 1918); 1528 Abdulahi Collection 186 (dated 1142/1729–30); 1529 Abdulahi Collection 187 (dated 9 July 1875); 1569 Abdulahi Collection 227 (not dated, possibly 20th century); 1576 Abdulahi Collection 234 (not dated, possibly 19th century); 1690

21 See Gori 2010 for a first description of the structure of the collection.

22 www.unesco.org/new/en/culture/themes/dynamic-content-single-view/news/opening_of_the_permanent_exhibition_of_the_sherif_harar_city (last accessed, 06/02/2018).

23 References are given according to the current provisional numbering used for PDF files realised in May 2011 by Jeremy Brown within the framework of the Ethiopic Manuscript Imaging Project directed by Professor Steve Delamarter, to whom I am very grateful for allowing me to study the items.

Abdulahi Collection 348 (not dated, but certainly before 1810); 1691 Abdulahi Collection 349 (not dated, but 20th century); 1692 Abdulahi Collection 350 (not dated, but 20th century).

The 'Harari *Mawlid* collection' has also been printed several times both abroad and in Ethiopia.

In Egypt, the work has been printed twice: in 1350/1931 (127 pages) and in 1366/1947 (100 pages), under the title *Mawlid šaraf al-ʿālamīn* ('The birth of the honour of the universe'). Two editions published in Ethiopia (around 1992–1993 and in 2000) under the title *Kitāb ʿunwān al-šarīf bi-al-mawlid al-šarīf* ('The book of the noble title on the noble birth') are actually photomechanical reproductions of two manuscripts. The 1992 edition (160 pages) contains a manuscript dated 1412/1992–1993 and written by an anonymous copyist. The 2000 edition (171 written pages) reproduces a manuscript dated 26 *ramaḍān* 1421/22 December 2000 and written by the famous copyist Ibrāhīm Muḥammad Wazīr. A third Ethiopian printed edition was published in Dire Dawa at an unknown date with the title *Mawlid šaraf al-ʿālamīn* ('The birth of the honour of the universe') at the expenses of Mahdī *ḥāǧǧ* 'Abdallāh. It seems to be a reprint of one of the above-mentioned Egyptian books. Finally, in *šaʿbān* 1426/September 2005, a computer-typed text (175 pages) was circulated in Addis Ababa under the title *Kitāb mawlid šaraf al-ʿālamīn* ('The book of the birth of the honour of the universe') by the Harari scholar Abū Bakr Ṯābit (Sabit).

This conspicuous number of testimonies proves the high esteem that the text collection enjoys among Hararis both inside and outside the city of Harar; however, despite this, research on the origins and the first diffusion of this text collection is still scanty and insufficient.[24]

One of the most remarkable features of the 'Harari *Mawlid* collection' is that while it is structured according to a relatively stable general framework, the verses and prose sections which actually make up the textual constellation and the sequence in which they are put together show a high degree of variation from manuscript to manuscript and from book to book.

Among the testimonies of the 'Harari *Mawlid* collection' I am aware of I was not able to identify a direct model for the manuscript I present here.

Moreover, for the specific purposes of this chapter, I will only focus here on a number of main Arabic texts which I managed to identify, leaving to another occasion a full description of the unknown or unidentified Arabic texts and the analysis of the Oromo and Harari sections.

24 In addition, the very famous and somewhat controversial Harari scholar 'Abdallāh al-Harārī (d. 2008) realised an abridged version of it (*muḫtaṣar*) and published it at an unknown date in Beirut at the *Dār al-mašārīʿ* (the publishing house of the so-called Ahbash movement).

When lacking a well-established title, poetic texts are identified by their incipit; for cross-reference, I will use the September 2005 edition of the 'Harari *Mawlid* collection'.

Here is a list of the texts I managed to identify in the *fidäl*-Arabic part of the manuscript:

- 6–8 October: poem *Ṭuf bi-ḥānī* ('Walk around my tavern');[25]
- 8–9 October: prayer for the Prophet and *šayḫ* Abādir; incipit: *Ṣalātun wa-taslīmun wa-azkā taḥiyyatin ʿalā Aḥmad al-muḫtār ṭaha al-mafāḫir* ('Peace, greeting and best salutations to Aḥmad, the chosen, the possessor of the glorious deeds'); *tawassaltu bi-šayḫ al-Abādir* ('I ask the intercession of *šayḫ* Abādir');[26]
- 10–14 October: poem *Ḥādī al-ʿīs adrik kaʾs al-hanā'* ('O Camel Driver, get the chalice of happiness');[27]
- 15 October, first part of the page: poem *Ḫuḏ yamīnan* ('Take a right hand') attributed to the renowned Yemenite mystic master 'Abdallāh Ibn 'Alawī al-Ḥaddād (1634–1720);
- 21 October (second part of the page)–27 October: poem on different epithets of the Prophet along a model structure: – [Muhammad who is] *ṣāḥib al-*X [possessor of something, in an *iḍāfa* construction], e.g. 1st line: *Ṣāḥib al-muʿǧizāt al-qāti'āt* ('The possessor of the irrefutable miracles'; 2nd line: *Ṣāḥib al-barāhīn al-sāṭiʿa* 'Possessor of the clear proofs');[28]
- 27–28 October: *Asmāʾ allāh al-ḥusnā* ('The beautiful names of God');
- 6–7 November: hymn in praise of the Prophet; incipit *Ya waǧha al-dīni wa-al-karam* (O face of the religion and of the generosity);

25 *Mawlid* edn 2005, 12–13; attributed to 'Abd al-Qādir al-Ǧīlānī. The text of the poem as preserved in the '*Mawlid* collection in Harar' is different from that usually featured in *'Abd al-Qādir's Dīwān* (see e.g. al-Ǧīlānī s.d: 157–164).

26 *Mawlid* edn 2005, 124–5; šayḫ Abādir is the patron saint of Harar and the most revered holy man in the city (for a general introduction to his personage, see Wagner 2003b). A *taḫmīs* of the text is published in Wagner 1975, 47–53.

27 *Mawlid* edn 2005, 84–93. Tašṭīr (possibly produced by an unknown 'Abdallāh b. Ǧa'far al-'Alawī al-Yamanī) based on the poem *Inna lamʿa al-barq min ḥayfi minā* ('As the gleam of the lightning feared the desires ') composed by the Sudanese author Muḥammad b. Ṭāhir al-Maǧdūb (1842–1929), a member of the Burhāniyya brotherhood. The text copied into the manuscript under discussion extends to the verse: *Šafīʿu al-ḫalqi fī ḥayratihim* ('Intercessor of the creatures in their bewilderment').

28 I am here transcribing the phrase as it should be according to the rules of Classical Arabic: the transliteration system of Arabic used in the text is specifically analysed in part 3 of this chapter.

- 7–13 November: hymn in honour of the Prophet; incipit: *Al-ḥamd lillāh al-qadīm/ğalla rabbī 'an al-miṯāl* ('Praise to God the Eternal, my Lord is beyond any similarity');
- 14–21 November: another section of the above-mentioned poem *Ḥādī al-'īs adrik ka's al-hanā'* ('O Camel Driver, get the chalice of happiness'); from the verse *man arāda al-naǧwa* ('the one who wants the upland') to the end *sayfuhu munṣalitan mā wahana* ('His sword is drawn and he is not weak');
- 21–22 November: the *qaṣīda Bānat Su'ād* ('Su'ād is gone', also called *al-Burda* 'The Mantel', like al-Buṣīrī's poem) by the famous Ka'b b. Zuhayr (first a fierce opponent of Muhammad, then one his Companions) with anonymous *taḫmīs Hānat Su'ād* (Su'ād arrived', up to the verse *Hayfā' muqlatin 'aǧzā' mudbiratin lā yuštakā qiṣarun minhā wa-lā ṭūl* ('Thin appears when she comes and well-proportioned when she turns away, not too short nor too tall , no complaints about her height');[29]
- 22 November–3 December: poem in honour of Abū Bakr (with *taḫmīs*); incipit: *Aya sā'ilī 'an madḥi man ğalla fī al-ḏikr* ('O the one who demands about the one who is honoured in the *ḏikr*');[30]
- 4–5 December: poem in praise of 'Abd al-Qādir al-Ǧīlānī; incipit: *'Alā al-awliyā' alqaytu sirrī wa-burhānī* ('On the saints I have thrown my secret and my proof');[31]
- 6–7 December: poem in praise of 'Abd al-Qādir al-Ǧīlānī; incipit *Yā quṭb al-wuǧūd bi-kum ṣafā wuǧūdī* ('O pole of the existence through you my existence is purified'); written in two columns;[32]
- 8–9 December: invocation to God; each verse starts with one different letter of the Arabic alphabet in our manuscript reproduced in a very coarse hand; incipit: *Ilāhun wāḥidun munfaridun bi-ḏātih* ('Only one God, alone in His essence');[33]
- 10–13 December: poem in praise of the Prophet; incipit *Ṣalātun salāmun ka-miski al-ḫitām* ('Prayer and greeting like the musk of the end'; attributed to 'Abdallāh Ibn 'Alawī al-Ḥaddād);[34]
- 13–15 December: poem in praise of the Prophet known from the Arabic-Harari refrain: *Nūr salām nūr salām* ('Light and greeting, light and greeting');

29 The bibliography about Ka'b and his *Bānat Su'ād* is extremely wide: see at least the article Sells and Sells 1990.

30 *Mawlid* edn 2005, 70–78.

31 *Mawlid* edn 2005, 100–101.

32 *Mawlid* edn 2005, 94–95.

33 *Mawlid* edn 2005, 98–99.

34 *Mawlid* edn 2005, 79-82.

Arabic incipit: *Šahidnā bi-annā Allāha zakkā Muḥammada* ('We testify that God attested the truth of Muḥammad');[35]

- 15–18 December: poem in praise of the Prophet known from the Arabic-Harari refrain: *Āw ayo salām āw ayo salām* ('Father, mother salute! Father, mother, salute!); Arabic incipit*: Ǧamālun badā min fawq al-ʿizzi dalla muḥtāǧun fa-qāla hazzī* ('A beauty appeared from above the high level a needy pointed out and said "Swing!"');[36]
- 18–19 December: mystical poem; incipit: *Bi-ḥaqqi allāh riǧāl allāh* ('For God's sake, o men of God');[37]
- 20–22 December: invocation; incipit: *Habū lī min faḍlikum habū lī* ('Grant me from your grace, grant me!');
- 24–29 December: hymn in praise of the Prophet; opening formula: *Šay' lillāh rasūl allāh yā sayyidī ḥabīb allāh al-madad rasūl allāh* ('O dear to God, messenger of God! O my lord, beloved of God! Help, messenger of God!'); incipit: *Abtadi'u bi-smi llāhi madḥa ḫayri ḫalqi llāh* ('I commence in the name of God the praise of the best of God's creatures');[38]
- 4 September: prayer for the Prophet and *šayḫ* Abādir (the same as in 8–9 October); incipit: *Ṣalātun wa-taslīmun wa-azkā taḥyatin ʿalā Aḥmad al-muḫtār ṭaha al-mafāḫir* ('Peace, greeting and best salutations to Aḥmad, the chosen, the possessor of the glorious deeds'); *tawassaltu bi-šayḫ al-Abādir* ('I ask the intercession of *šayḫ* Abādir').

3 *Fidäl* for Arabic: an easy adaptation?

The relative proximity of the Arabic phonetic system to that of many of the Semitic languages of Ethiopia and Eritrea could generate the idea that the usage of *fidäl* to write Arabic might be a relatively simple adaptation process. I believe that the following few observations on the way Arabic has been transcribed in the manuscript under analysis only partially confirm this idea.

First of all, the linguistic interference represented in the manuscript is that between Arabic and Harari. To briefly highlight the main differences between the

35 *Mawlid* edn 2005, 62.
36 *Mawlid* edn 2005, 57–58.
37 *Mawlid* edn 2005, 52–54.
38 *Mawlid* edn 2005, 121–24.

phonologies of these two languages, Wolf Leslau's contribution on the Arabic loanwords in Harari is useful as a reliable reference:[39]

Harari, on the one hand, possesses four phonemes which are absent in Arabic: a voiceless palato-alveolar affricate /č/; an 'emphatic' (ejective, glottalised) voiceless palato-alveolar affricate /č̣/; a palatal nasal /ñ/; and a voiced velar stop /g/. Arabic, on the other hand, has seven phonemes which do not exist in Harari: a voiceless (inter)dental fricative /t̲/; a voiced (inter)dental fricative /d̲/; an 'emphatic' (pharyngealised) voiced alveolar fricative /ẓ/; an 'emphatic' (pharyngealised) voiceless alveolar fricative /ṣ/; an 'emphatic' (pharyngealised) voiced dental stop /ḍ/; a voiced velar fricative /ġ̣/; and a voiced pharyngeal fricative /ʿ/.

Moreover, there are four further important differences 1) the so-called 'emphatic' consonants are pronounced as ejectives (glottalised) in Harari, while they are basically pharyngealised in Arabic; 2) in Harari the voiceless velar fricative [ḫ] appears only as an intervocalic variant of the voiceless velar stop /k/ as a consequence of spirantisation and thus has no phonemic status. In Arabic loanwords, /ḫ/ passes to /k/ but is sometimes preserved; 3) Harari has a voiceless pharyngeal fricative /ḥ/, but no voiceless glottal fricative /h/. Arabic /h/ is thus rendered as /ḥ/ in Harari, but in some cases the original /ḥ/ is preserved; and, finally, 4) the vocalic system of Harari is still an object of discussion among linguists, but latest research points to the phonemic value of vowel quantity.[40] It thus seems that in terms of both vowel quantity and vowel quality the phonology of Harari can easily accommodate Arabic words.

In the end, Leslau's analysis clearly demonstrates that the passage of a relatively substantial amount of Arabic loanwords (about 300) into Harari vocabulary did not take place according to a stable correspondence among phonemes of the two languages. The inconsistent phonetic adaptation is apparent when considering the way the Arabic phonemes (in particular those unknown to Harari) are rendered in Harari.[41]

It can be surmised that the picture is complicated by the fact that the Arabic loanwords' mode of entry into Harari is twofold: 1) oral, through one of the forms of spoken Arabic to which the speakers of Harari are exposed;[42] and 2) written,

39 See Leslau 1956, and Leslau 1957, which contains a comprehensive description of the differences between the Arabic phonological system and that of other Semitic languages of the Horn of Africa.

40 For a detailed discussion on this, see Garad-Wagner 1998, 157–168, and for a more general description, see Wagner 1997, 487–488.

41 For the details, see in particular the recapitulative table in Leslau 1956, 21.

42 Leslau (1956, 22–23), indicates a number of Arabic dialects as possible sources for at least

through a learned milieu, especially that of Arabic teachers and Islamic scholars who are well acquainted with the Classical Arabic of the Qur'an and the theological tradition.[43]

In the case of the manuscript I am presenting here, the issue of the Arabic dialects remaining behind or underneath the words is not relevant. Nevertheless, the idea that an oral dimension of the texts has influenced their transcription cannot be excluded from these considerations.

The main features of the transcription system used in the manuscript for depicting Arabic words in Ethiopian script can be briefly sketched as follows:

- voiceless (inter)dental fricative /t̲/: written as ሰ <s> for voiceless alveolar fricative /s/; e.g. ሱቡቱ for *ṯubūtu* ('His immutability');
- voiced (inter)dental fricative /d̲/: written as ዘ <z> for voiced alveolar fricative /z/; e.g. ቢዘቲሒ for *bi-ḏatihi* ('By His essence');
- 'emphatic' (pharyngealised) voiced alveolar fricative /ẓ/: written as ዘ <z> for voiced alveolar fricative /z/; e.g. ዛሐሩ for *ẓahara* ('It appeared');
- 'emphatic' (pharyngealised) voiceless alveolar fricative /ṣ/: written as ሰ <s> for voiceless alveolar fricative /s/; e.g. ሲዶቁን *ṣidqun* ('His truthfulness');
- 'emphatic' (pharyngealised) voiced dental stop /ḍ/: written as ዳ <d> for (non-emphatic) voiced dental stop /d/; e.g. ዲያኡን for *ḍiyā'un* ('Lights');
- voiced velar fricative /ġ/: written as ኀ < ḫ > for voiceless velar fricative /ḫ/, not only in intervocalic position but also in initial position; e.g. ኀኒዩን for *ġaniyyun* ('Rich');
- voiced pharyngeal fricative /ʿ/: written as ዐ <'> for voiceless glottal stop /'/; e.g. ዐሊዩን for *ʿaliyyun* ('Sublime');
- voiceless velar fricative /ḫ/: written as ኀ <ḫ> for voiceless velar fricative /ḫ/, not only in intervocalic position but also in initial position; e.g. ኀረቅቱ for *ḫaraqtu* ('I tore');
- voiceless pharyngeal fricative /ḥ/ and voiceless glottal fricative /h/: mostly both written as ሐ <ḥ> for voiceless pharyngeal fricative /ḥ/; sometimes the Arabic voiceless glottal fricative /h/ is 'correctly' written as ሀ <h> with the

some of the Arabic loanwords in Harari, without being able to localise all of them in one single dialect. His conclusions are somewhat ambiguous. After remarking that 'one is inclined to think that not one single Arabic dialect is to be taken as the source of Harari, but various Arabic dialects', he says a few lines later: 'As for the historically possible source of the Arabic loanwords in Harari, a South Arabic dialect seems to be more likely to be considered'.

43 Leslau (1956) did not at all consider the possibility that at least some of the Arabic words he discusses could have entered Harari vocabulary from a learned, written source and not from a spoken variant of the language. An assessment of the loanwords taking this possibility into account might provide some fresh insights into the topic.

Ethiopian letter for the voiceless glottal fricative /h/; however the letter ሀ <h> is sometimes also used for rendering the voiceless pharyngeal fricative /ḥ/ and vice versa; the Ethiopian letter ሐ <ḥ> sometimes also represents /h/; e.g. ሐሊሙ for *ḥalīm(u)*('Patient') and ሀዚ for *hazzī* ('Swing!'); but also ኢሀሳን for *iḥsān* ('Doing good deeds') and ዛሐሩ for *ẓahara* ('He appeared').

The quality of the Arabic vowels /a/, /i/, /u/ is always reproduced accordingly in the text in *fidäl*. As for quantity, only /ā/ is sometimes rendered with the fourth order of the Ethiopian syllabary (<Ca>), which in the orthography of modern Harari also represents the long *a* (short /a/ written with the first order of the syllabary (<Cä>), e.g. ኸረቅቱ for *ẖaraqtu* ('I tore') and ዋሒዱን for *wāḥidun* ('Unique'); however the fourth order long /a/ can also be used for an Arabic short /a/, e.g. ዛሐሩ for *ẓahara* ('He appeared').[44] As for Arabic /i/ and /ī/, both are always written with the third order of the Ethiopian syllabary (<Ci>), as well as /u/ and /ū/ being reproduced with the second order of the *fidäl* (<Cu>).

Diphthongs /aw/ and /ay/ are more or less consistently rendered in *fidäl* (e.g. ለው for *law* ['If']), but many times the writing in *fidäl* seems to point to an assimilated pronunciation of the /a/ to the following /w/ or /y/ e.g. የውም *yowm* for *yawm* ('Day'), and ኸይሩ *ẖeyr(u)* for *ẖayr(u)* ('Benefit').

It is worth noting that in many instances the word division does not correspond to the one used in Arabic; thus forms like ቢዛቲሒ for *bi-ḏatihi* ('By His essence') can be found together with ፈቀልቢለ የውመ (*falqalbil yowma*) for *fa-qalbī al-yawma* ('Then my heart today').

To sum up the collected data, the rendering of the Arabic texts in *fidäl* can be considered as being only partially accurate. Failure to represent many Arabic phonemes makes the transcription to a great extent imprecise: it is true that the Ethiopian syllabary lacks specific signs for some of the Arabic phonemes and for distinguishing the vowel quantity of /i/ and /u/, but no creative effort has been made by the copyist to fill the gaps in the writing system he was using and create a one-to-one sign set of transliteration.

The written rendering in *fidäl* of many of the Arabic sounds which are absent in Harari follows the same pattern as the phonetic adaptation of the Arabic loanwords into Harari. The word division in Ethiopian script is quite inconsistent and often only oddly corresponds to the Arabic original. The diphthongs are written in a way that hints at an assimilation of the short Arabic /a/ to the following semi consonant.

44 See Wagner 2004, 358–359 for a discussion of the treatment of the vowels in the *fidäl* orthography of Harari.

These elements possibly indicate that the writing process of the manuscript was carried out while hearing someone from Harar reciting the texts or while the copyist himself was reading them aloud. As a matter of fact, the writing down of the Arabic texts does not adhere to any written model: the copyist does not even try to rewrite in *fidäl* a written Arabic *Vorlage*, but mostly transcribes the texts the way they would sound, if they were read by a Harari speaker. The copyist does not seem to have any intention to produce a well-thought-out system of transliteration of Arabic with the Ethiopian syllabary, he just aims at providing the faithful who are unfamiliar with Arabic script with a tool to access the texts and recite them on different religious occasions.[45]

To my knowledge this is so far the only known example of a substantial amount of complex Arabic texts written in Ethiopian syllabary. The general cultural trends in the Islamic communities in Ethiopia will tell us whether it will remain only an individual, isolated effort or the first step of a linguistic and literary development among the Ethiopian Muslims.

References

Abdurahman Mäḥamäd Qorram (1984 Ethiopian Calendar = 1992), *Čuqti kətab*. *Yä'adärəññaamarəññа mäzgäbä qalat*, Addis Ababa.

Banti, Giorgio (2005), 'Harari literature', in Siegbert Uhlig (ed), *Encyclopaedia Aethiopica, vol II (D–Ha),* Wiesbaden: Otto Harrasowitz Verlag, 1028–1030.

Banti, Giorgio (2010), 'The literature of Harar until the end of the 19th century', in Alessandro Gori and Biancamaria Scarcia Amoretti (eds), *Islam in Etiopia*: *bilanci e prospettive*. *Atti del Convegno Internazionale. Università degli Studi di Napoli Federico II, 18–19 giugno 2008* (Civiltà del Mediterraneo, 16–17), Roma: Edizioni di Storia e Letteratura, 149–181.

Bausani, Alessandro (1967), 'Per una lettura comparata delle lingue islamiche', in *Atti del Terzo Congresso di Studi Arabi e Islamici, Ravello, 1–6 settembre 1966*, Napoli: Istituto Universitario Orientale, 145–56.

Bausani, Alessandro (1978), 'Incontri linguistici in Asia: esperienze di un orientalista', in *Lingue a contatto nel mondo antico*. *Atti del Convegno della Società Italiana di Glottologia, Napoli, 12 e 13 maggio 1978* (Orientamenti Linguistici, 7), Pisa: Giardini, 11–33.

45 The usage of somewhat impressionistic transcriptions of Arabic religious texts based on a phonetically very imprecise oral reproduction of Arabic to meet the needs of the common faithful who do not read Arabic is very common in all non-Arabic speaking Islamic countries. See for example the dozens of different booklets for the performance of the ritual prayers in use in Turkey (commonly referred to as *namaz hocası*) which contain Arabic texts written in Latin script according to current Turkish pronunciation of Arabic (e.g. *Elhamdulillâhillezî ce'alel-mâe tahûren ve ce'alel-İslâme nûra*).

Bausani, Alessandro (1981), 'Le lingue islamiche: interazioni e acculturazioni', in Alessandro Bausani and Biancamaria Scarcia Amoretti (eds), *Il mondo islamico tra interazione e acculturazione*, Roma: Istituto di Studi Islamici, 3–19.

Brenner, Louis, and Murray Last (1985), 'The Role of Language in West African Islam' in *Africa*: *Journal of the International African Institute,* 55: 432–446.

Cerulli, Enrico (1926), 'Canti amarici dei musulmani d'Abissinia', in *Accademia Nazionale dei Lincei. Rendiconti* serie VI, II: 433–447 (reprinted in Cerulli, Enrico. 1971, *L'islam di ieri e di oggi*, Roma: Istituto per l'Oriente, 245–256).

Conti Rossini, Carlo (1919), 'Testi in lingua harari', in *Rivista degli Studi Orientali,* 8: 401–425.

Donzel, Emeri Johannes van (ed. and tr.) (1969), *'Ĕnbaqom: Anqasa Amin (La porte de la foi), apologie éthiopienne du Christianisme contre lilslam à partir du Coran.* Leiden: E. J. Brill.

Frantsouzoff, Serguei A., (2010), 'Script, Ethiopic', in Siegbert Uhlig and Alessandro Bausi (eds), *Encyclopaedia Aethiopica,* vol. IV (O–X), Wiesbaden: Otto Harrassowitz Verlag, 580–585.

Garad, Abdurahman, and Ewald Wagner (1998), *Harari-Studien*. *Texte mit Übersetzung und grammatischen Skizzen und Glossar*, Wiesbaden: Otto Harrassowitz Verlag.

Gibb, Camilla (1998), 'Religious Identification in Transnational Contexts: Being and Becoming Muslim in Ethiopia and Canada', in *Diaspora: A Journal of Transnational Studies,* 7: 247–269.

Gibb, Camilla (2002), 'Deterritorialized People in Hyperspace: Creating and Debating Harari Identity over the Internet', in *Anthropologica,* 44: 55–67.

al-Ǧīlānī, ʿAbd al-Qādir, (s.d.), *Dīwān ʻAbd al-Qādir al-Ǧīlānī*, edited by Yūsuf Zaydān, Bayrūt: Dār al-Ǧīl.

Gori, Alessandro (2001), *La 'Silloge' di šah Zakkāryās*. *Testo arabo originale, introduzione traduzione e note di Alessandro Gori* (Accademia Nazionale dei Lincei, Atti, Classe di Scienze morali, storiche e filologiche, Memorie. Scienze morali, storiche, filologiche, serie 9, 13, 4: 453–583), Roma: Accademia Nazionale dei Lincei.

Gori, Alessandro (2003), *La 'Silloge' di šah Zakkāryās sull'Islam – Versione amarica*. *Testo amarico originale, introduzione, traduzione e note di Alessandro Gori* (Accademia Nazionale dei Lincei, Atti, Classe di Scienze morali, storiche e filologiche, Memorie. Scienze morali, storiche, filologiche, serie 9, 16, 2: 57–319), Roma: Accademia Nazionale dei Lincei.

Gori, Alessandro (2010), 'Texts in the *Mawlid* Collection in Harar: some first critical observations', in *African Study Monographs,* Supplement 41: 51–62.

Gori, Alessandro (2014a), 'Zäkaryas', in Alessandro Bausi and Siegbert Uhlig (eds), *Encyclopaedia Aethiopica,* vol. V (Y–Z), Wiesbaden: Otto Harrassowitz Verlag, 116–117.

Gori, Alessandro (2014b), Ağamī *in the opposite way. An Arabic manuscript text written in Ethiopian script (EMML 6239)*. Unpublished presentation delivered at the *Colloque Manuscrits chrétiens et islamiques d'Ethiopie (XIIIe-XXe siècle)* 12–13 December 2014, Bibliothèque nationale de France, Paris.

Gori, Alessandro (2015), 'Languages and literatures of the Muslims of the Horn of Africa: some first general reflections', in Paolo Nicelli (ed.), *L'Africa*, *l'Oriente mediterraneo e l'Europa*: *tradizioni e culture a confront* (Africana Ambrosiana, 1), Milano, Roma: Biblioteca Ambrosiana – Bulzoni Editore, 119–126.

Gori, Alessandro, Anne Regourd, Jeremy Brown, Steve Delamarter, and Demeke Berhane (2014), *A Handlist of the Manuscripts in the Institute of Ethiopian Studies*, *Volume Two*: *The Arabic*

Materials of the Ethiopian Islamic Tradition, (= Ethiopic manuscripts, texts and studies, 20). Eugene, Or.: Pickwick Publications.

Griaule, Marcel (1930), *Le livre de recettes d'un dabtara abyssin*, (= Travaux et mémoires de l'institut d'ethnologie, XII). Paris: Institut d'ethnologie.

Leslau, Wolf (1956), 'Arabic Loanwords in Harari', in *Studi Orientalistici in Onore di Giorgio Levi Della Vida*, Vol. II, Roma: Istituto per l'Oriente. 14-35 (reprinted in Leslau, Wolf (1990), *Arabic Loanwords in Ethiopian Semitic*, Wiesbaden: Otto Harrassowitz Verlag, 119–140).

Leslau, Wolf (1957), 'The Phonetic Treatment of the Arabic Loanwords in Ethiopic', in *Word,* 13: 100–23 (reprinted in Leslau, Wolf (1990), *Arabic Loanwords in Ethiopian Semitic*, Wiesbaden: Otto Harrassowitz Verlag, 213–236).

Mawlid (2005), *Kitāb mawlid šaraf al-ʿālamīn*. Edited by Abū Bakr Ṯābit (Sabit), Addis Ababa: Dire Printing S.C.

Mittwoch, Eugen (1906), 'Exzerpte aus dem Koran in amharischer Sprache', in *Mitteilungen des Seminars für Orientalischen Sprachen zu Berlin, 9*: 111–147.

Pietruschka, Ute, Amha Asfaw, and Getatchew Haile (2003), 'Baḥrä ḥassab', in Siegbert Uhlig (ed.), *Encyclopaedia Aethiopica,* vol. I (A–C), Wiesbaden, Otto Harrassowitz Verlag, 445–446.

Sells, Michael Anthony - Sells, M. J. (1990), ''Bānat Suʿād': Translation and Introduction', in *Journal of Arabic Literature,* 21/2:140–154.

Wagner, Ewald (1997), 'Harari', in Robert Hetzron (ed.), *The Semitic Languages*, Abingdon, New York: Routledge, 486–508.

Wagner, Ewald (2003a), *Harar: Annotierte Bibliographie zum Schrifttum über die Stadt und den Islam in Südostäthiopien* (Aethiopistische Forschungen, 61), Wiesbaden: Otto Harrassowitz Verlag.

Wagner, Ewald (2003b), 'Abādir ʿUmar ar-Riḍā', in Siegbert Uhlig (ed.), *Encyclopaedia Aethiopica,* vol. I (A–C), Wiesbaden: Otto Harrassowitz Verlag, 4–5.

Wagner, Ewald (2004), 'Die Verwendung der äthiopischen Schrift für das Harari', in Verena Böll, Denis Nosnitsin, Thomas Rave, Wolbert Smidt, and Evgenia Sokolinskaia (eds), *Studia Aethiopica: In Honour of Siegbert Uhlig on Occasion of His 65th Birthday*, Wiesbaden: Otto Harrassowitz Verlag, 355–360.

Wagner, Ewald (2005), 'Farāʾiḍ, Kitāb al-', in Siegbert Uhlig (ed.), *Encyclopaedia Aethiopica*, vol II (D–Ha), Wiesbaden: Otto Harrassowitz Verlag, 491–492.

Zappa, Francesco (2004), *L'islamizzazione della lingua bambara in Mali. Tra pubblicistica scritta e letteratura orale* (= Supplemento n. 2 della Rivista degli Studi Orientali, vol. LXXVII), Pisa/Roma: Istituti Editoriali e Poligrafici Internazionali.

Zappa, Francesco (2009), 'Écrire l'Islam en Bambara', in *Archives de sciences sociales des religions,* 147: 167–186.

List of Contributors

Dmitry Bondarev is a researcher and head of West Africa projects at the Centre for the Study of Manuscript Cultures, University of Hamburg. His research interests cover African linguistics, literacy studies, the history of writing in Ajami, and the palaeography and codicology of West African manuscripts.

Nuria de Castilla is Professor of History and Codicology of the Book in the Islamicate World at the Ecole Pratique des Hautes Etudes (EPHE, Paris), where she has been employed since 2016. She is currently the scientific co-ordinator of the ERC project 'Saadian Intellectual Cultural Life' (2016–2021), which focuses on the study of the Arabic Collection kept at the Monastery of San Lorenzo de El Escorial, and is also head of the project 'Corana. Production and Transmission of the Qur'an in the Western Islamic World'. Her main research topics are the history of the book, the history of the library, Islamo-Christian relations in the Early Modern Period, Aljamiado literature, and the Qur'an in the Islamic West.

Nikolay Dobronravin is a professor at the Department of World Politics in the School of International Relations, St Petersburg State University, Russia. He graduated from the same university in 1983 (in African Studies) and earned his doctorate (Dr. habil.) on West African literary traditions in Arabic script in 2000 (a short version of it was published in Russian in 1999). Besides writing articles in Russian, English and Portuguese, he has published Hausa historical songs (2004, in Russian, based on the study of Ajami manuscripts) and authored a chapter on literacy among Muslims in nineteenth-century Trinidad and Brazil in Slavery, Islam and Diaspora (2009). His current research in this field focuses on the history of Ajami in Sudanic Africa and in the diaspora.

Alessandro Gori is Associate Professor of Arabic Language and Literature at the Department of Cross-Cultural and Regional Studies at the University of Copenhagen, Denmark. His main field of specialisation is the Islamic manuscript tradition in the Horn of Africa, a subject on which he has published extensively.

Branka Ivušić is a comparative linguist who has been working on language contact and the use of Arabic script in the early modern Balkans and adjacent areas for some time. In her PhD project at the Centre for the Study of Manuscript Cultures, she is exploring an early modern multilingual manuscript in Arabic script. She works as a lecturer in German linguistics and language history at the University of Leipzig and at the University of Halle.

Paola Orsatti is Associate Professor of Persian Language and Literature at La Sapienza University, Rome, where she formerly also taught the history of the Persian language. Her research focuses chiefly on the history of Persian, Persian classical literature, the history of Persian studies in Italy and Europe, the history of Oriental manuscript collections, and the palaeography and codicology of Islamic manuscripts. Besides authoring a number of scholarly articles, she has published *Il fondo Borgia della Biblioteca Vaticana e gli studi orientali a Roma tra Sette e Ottocento* (Citta del Vaticano, 1996), *Appunti per una storia della lingua neopersiana. Parte I: parte generale – fonologia – la più antica documentazione* (Rome, 2007), *Corso di lingua persiana* (with Daniela Meneghini, Milan, 2012) and edited *The Persian Language in History* together with Mauro Maggi (Wiesbaden, 2011).

Paolo La Spisa is Associate Professor of Arabic Language and Literature at the University of Florence, Italy. After conducting field research in Syria and Lebanon a number of times, he was appointed a Marie Curie Fellow at the Université catholique de Louvain (Louvain-la-Neuve, Belgium), where he worked from 2008 to 2010. His main research interests are Christian Arabic literature and Muslim–Christian relations in pre-modern times, which involves the analysis and description of the Middle/Mixed Arabic linguistic features found in Arabic manuscripts. His main publication is a critical edition of the theological treatises by Sulaymān Ibn Ḥasan al-Ġal-n, the Palestinian Bishop of Gaza (10th–11th cent.).

Jan Schmidt was a lecturer in Turkish studies at Leiden University, the Netherlands, until his retirement in December 2016. He has published a host of articles and books on Ottoman-Turkish history and literature. He also compiled a catalogue of the Turkish manuscripts kept in Dutch public institutions, particularly the Leiden University Library, which appeared in four volumes between 2000 and 2012.

Florian Sobieroj catalogues Arabic manuscripts on the KOHD project (cataloguing and digitising Middle Eastern manuscripts), which is run under the auspices of the Gottingen Academy of Sciences and Humanities. He began work on the project in 1995 and has published three catalogues in the Union Catalogue of Oriental Manuscripts in German Collections (VOHD) series, including descriptions of Arabic manuscripts as well as some Persian and Turkish manuscripts kept at the Bavarian State Library in Munich and some major collections in Thuringia. Besides cataloguing Islamic manuscripts, he takes an interest in Sufism, focusing on its manifestations in Arabic and Persian literature. He has written a number of articles and books on manuscript studies and Sufism, including an editio princeps of one of the works by al-Qushayrī.

Lameen Souag is a researcher at LACITO (CNRS) near Paris and focuses on the linguistic history of northern Africa with particular emphasis on language contact. He has undertaken descriptive fieldwork on several minority languages of the Sahara, notably Korandje in south-western Algeria and Siwi Berber in western Egypt. His first book, Berber and Arabic in Siwa (Egypt): A Study in Linguistic Contact, was published in 2013.

Jan van der Putten is Professor for Austronesian Languages in the Department of South-east Asian Studies (Asien-Afrika-Institut) at the University of Hamburg and is affiliated with the Centre for the Study of Manuscript Cultures. In his research, he explores Malay texts and their cultural contexts in different periods and a variety of media.

Esther-Miriam Wagner is Director of Research at the Woolf Institute, Cambridge. Her own research focuses on the Cairo Genizah, Judaeo-Arabic, sociolinguistics, the Arabic Prize Papers and Yiddish. Miriam is editor-in-chief of al-Masāq. Journal of the Medieval Mediterranean, and an Affiliated Researcher at the University of Cambridge. She is also a Research Associate at the Taylor-Schechter Genizah Research Unit, University of Cambridge.

Index of Persons

https://doi.org/10.1515/9783110639063-013

Index of Polities, Ethnic Groups, Kingdoms, States, Towns, and Religious Denominations

Index of Languages and Scripts

Keywords