E - ISSN: 2953-8203
P - ISSN: 2953-819X

YEREVAN STATE
UNIVERSITY

JOURNAL OF IRANIAN LINGUISTICS

Volume 2 - Issue 1 - 2025

JOURNAL OF IRANIAN LINGUISTICS

EDITOR-IN-CHIEF

Vardan Voskanian, Yerevan State University, Armenia

Volume 2 | issue 1

YEREVAN STATE UNIVERSITY PUBLISHING HOUSE

ASSOCCIATE EDITORS

Hakob Avchyan, Yerevan State University, Armenia

Artyom Tonoyan, Yerevan State University, Armenia

EDITORIAL BOARD

Chiara Barbati, University of Pisa, Italy

Desmond Durkin-Meisterernst, Freie Universität Berlin, Germany

Jila Ghomeshi, University of Manitoba, Canada

Geoffrey Haig, University of Bamberg, Germany

Arsalan Kahnemuyipour, University of Toronto Mississauga, Canada

Simin Karimi, University of Arizona, USA

Paola Orsatti, Sapienza University of Rome, Italy

Ludwig Paul, Hamburg University, Germany

Mohammad Rasekh-Mahand, Bu-Ali Sina University, Iran

Hassan Rezai Baghbidi, Osaka University, Japan

Pollet Samvelian, Université Sorbonne Nouvelle, France

Jaffer Sheyholislami, Carleton University, Canada

E - ISSN: 2953-8203
P - ISSN: 2953-819X

© YSU Publishing House, 2025
© Authors, 2025

JOURNAL OF IRANIAN LINGUISTICS
VOLUME 2 | ISSUE 1

CONTENT

VARDAN VOSKANIAN

Foreword **4-5**

NIMA ASEFI

Ewer, Garden and Gardening. An Edition of Berk. 25 and Revised Readings of Berlin 26, Berk. 11, and Berk. 122 Documents Belonging to the Pahlavi Archive of Hastijan **6-29**

JAKOB HALFMANN

On the Etymology of New Persian malang 'Intoxicated; Unorthodox Dervish' **30-35**

YOULI IOANNESYAN

The Progressive Tenses with the Verb "to have" as a Peculiar and Exclusive Feature of Western Iranian Dialects Spoken in Modern Iran **36-47**

MIDHAT SHAH

Persian Borrowings in Kashmiri: A Descriptive and Analytical Study **48-71**

ARTYOM TONOYAN

Caucasian Persian (Tati) Fragments in Armenian Script: A Study of the Lord's Prayer with Transliteration, Translation, and Comments **72-111**

Foreword

Volume 2 / Issue 1

"Language is the most massive and inclusive art we know, a mountainous and anonymous work of unconscious generations."
— Edward Sapir

This observation by Sapir serves as a fitting epigraph for the present volume of the *Journal of Iranian Linguistics*. The five articles gathered here reflect both the breadth of Iranian linguistics as a field and the depth of insight that emerges when scholars engage closely with its data—be it through archival manuscripts, dialectal features, etymological reconstruction, or the mechanisms of language contact.

The issue begins with ***Nima Asefi***'s edition of *Berk. 25* and revised readings of three additional documents from the Pahlavi Archive of Hastijan. Through careful philological and palaeographic analysis, Asefi offers new insights into administrative terminology, toponyms, and estate management practices in late Sasanian and early post-Sasanian Iran. His re-editions correct earlier readings and highlight the significance of the Hastijan corpus for understanding Middle Persian documentation.

Jakob Halfmann examines the etymology of the New Persian word *malang*, which is widely attested in Indian languages. After establishing its core semantic range—'intoxicated, drunk' and '(unorthodox) dervish'—he proposes a derivation from a previously unattested Bactrian loanword, ultimately linked to the Indo-Iranian root **mad-* 'to be/become intoxicated'. The article adds a new item to the growing list of Bactrian borrowings into New Persian and contributes to the documentation of the Bactrian lexicon by proposing a form not yet attested in existing sources.

Youli Ioannesyan investigates progressive tenses with the auxiliary verb "to have" in Persian and related Western Iranian dialects spoken in Iran. Drawing on a broad range of dialectal data, he argues that these constructions—absent from Dari and Tajik—are an indigenous and exclusive innovation of Iranian varieties within the borders of modern Iran. The article challenges earlier claims of foreign influence and offers a dialectologically grounded reassessment of these progressive forms.

Midhat Shah provides a comprehensive analysis of Persian borrowings in Kashmiri, with particular attention to the processes of nativization at phonological, morphological, and semantic levels. By situating these changes within both historical linguistics and contact linguistics frameworks, Shah

sheds light on the depth and complexity of Persian-Kashmiri interaction. Her study offers an important contribution to Indo-Iranian contact studies, addressing a longstanding scholarly gap.

Finally, ***Artyom Tonoyan***'s article brings attention to a rare and largely overlooked linguistic artifact: the Lord's Prayer translated into Caucasian Persian (Tati) and written in Armenian script. Through detailed transliteration, translation, and commentary, the article offers both a linguistic and historical analysis of the text and its context. It enriches the study of Caucasian Persian dialects and highlights the role of the Armenian script as a vehicle for preserving minority Iranian languages.

This issue would not have been possible without the dedication of many individuals. I would like to extend my sincere thanks to the editorial board for their guidance and to the reviewers for their careful assessments.

We hope that this volume will be of interest to all scholars of Iranian linguistics and will encourage further research and engagement with the many underexplored dimensions of this diverse field.

Vardan Voskanian
Editor-in-Chief
Journal of Iranian Linguistics

YEREVAN STATE UNIVERSITY PUBLISHING HOUSE

JIL 2/1 (2025): 6-29

Journal of Iranian Linguistics

Ewer, Garden and Gardening

An Edition of Berk. 25 and Revised Readings of Berlin 26, Berk. 11, and Berk. 122 Documents Belonging to the Pahlavi Archive of Hastijan[1]

Nima Asefi

University of Hamburg

doi.org/10.46991/jil/2025.01.01

Abstract: This article presents a new edition of the unpublished Middle Persian document Berk. 25 and revised readings of Berlin 26, Berk. 11, and Berk. 122, all from the Pahlavi Archive of Hastijan. The Archive, dating from late Sasanian to early post-Sasanian period, shed light on administrative practices, rationing systems, and terminology related to gardening and estate management in central Iran. The article offers updated transliterations, translations, and philological commentary, correcting prior readings by Gignoux and Weber. Particular attention is paid to lexical items such as *bāγbānīh* ("gardening") and geographic toponyms like *Kard-ābād-yazdbād*. Through detailed reanalysis, this study contributes to a more accurate understanding of the linguistic, economic, and administrative landscape reflected in the Pahlavi documents.

Keywords: Pahlavi Archive of Hastijan, Middle Persian, Cursive Pahlavi Script, Central Iran, Late Sasanian.

Nima Asefi
E-mail: nima.asefi@uni-hamburg.de
ORCID: https://orcid.org/0000-0002-6099-4516
Received: 14.01.2025
Revised: 10.06.2025
Accepted: 28.06.2025

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

© Nima Asefi, 2025

Introduction:

The publication of research findings on the documents that today should be referred to as the Pahlavi Archive of Hastijan[2], began in 1991 with an article

[1] All documents referenced with the abbreviation 'Berk.' are housed at the Bancroft Library, University of California, Berkeley. Documents cited with 'Berlin' are housed at Freie Universität Berlin and in a private collection in Berlin.

[2] In 2022, based on newly obtained evidence, I was able to identify the site where most likely the Pahlavi documents currently housed at the University of California, Berkeley, and the Freie Universität Berlin were first discovered. After I reported this finding to the Iranian Ministry of Cultural Heritage, Tourism, and Handicrafts, up

by Gignoux (1991). These studies were continued by Gignoux himself and Weber, resulting in significant advances in our understanding of the archive. Weber's involvement in the research, along with his valuable contributions, notably accelerated progress and greatly improved the precision of the interpretations. Despite their extensive efforts, a portion of the documents still requires rereading and retranslation. It is important to emphasize, however, that the possibility for younger researchers like myself to undertake such tasks today is entirely thanks to the remarkable groundwork and deep expertise of both Gignoux and Weber.

The study of the Pahlavi Archive of Hastijan—which includes rereading previously published documents, examining unpublished materials from the Berkeley collection, and analyzing newly discovered documents from the Hastijan cave—is essential for deepening our understanding of the final phases of the Middle Persian script and language, the history of the late Sasanian and early post-Sasanian periods, the historical geography of central Iran, etc. I began working on these documents approximately seven years ago, with my Ph.D. dissertation[3] devoted to their study. Over the years, by utilizing the writings of Philippe Gignoux and Dieter Weber, along with other supplementary resources, I have become familiar with the palaeographic features, linguistic traits, and historical context of these documents.[4] In addition to studying the unpublished documents, I have also revised several documents previously published by Gignoux and Weber. This article presents, for the first time, an edition of the unpublished document **Berk. 25**, along with revised editions of three previously published documents: **Berlin 26, Berk. 11, and Berk. 122**.

today, two archaeological excavations were carried out in the cave near the village of Hastijan, under the direction of Dr. Mohammad Reza Nemati. These efforts resulted in the discovery of a significant number of documents, clay bullae, and other valuable artifacts. I must emphasize that Berkley documents were formerly kept in a private collection and remained inaccessible to the academic community. Owing to the persistent efforts of Professor Azarpay, a private individual acquired the entire collection and generously donated it to the University of California, Berkeley, thereby enabling scholarly research on these valuable materials. For more information about the Pahlavi Archive of Hastijan, see Azarpay 2003, Asefi 2023a, 2024a: 2, and Asefi & Mirfakhraie 2024: 144-147. For more information about the result of the first season's excavation, see Nemati et al. (forthcoming).

[3] Under the supervision of Professor Mahshid Mirfakhraie.

[4] I have currently expanded my research to focus on Pahlavi documents from Hastijan, Fars, and Tabarestan as part of a project at the University of Hamburg under the supervision of Prof. Shervin Farridnejad.

I dedicate this article to Guitty Azarpay (1934–2024), Philippe Gignoux (1931–2023), and Dieter Weber, whose efforts, insight, and scholarly brilliance laid the very foundation for the study and advancement of research on these documents. Sadly, I never had the privilege of meeting them in person.

- **Edition of Berk. 25**

Catalog Number	Collection Name	Material	Dimensions (cm)	Lines	Verso	Bulla
25	The Bancroft Library	Textile	9.5 cm x 8.5 cm	9	Marked	At the bottom left

Transliteration	**Transcription**
1. wnd(?) Y PWN … [d'lyk] MN ZK Y [MN ml Y]	1. … ī pad… [dārīg] az ān ī [az mar ī]
2. yzd'np'nk' YLYDWNt [hlt wnd't 'wst'nd'l] ʿL bwn' ZNH BYRḤ	2. yazdānbānag zād [xrad-windād ōstāndār] ō bun ēn māh
3. 'rtwhšt' Y ŠNT XXX-III-III 'wcynyt Y dwtk' Y	3. Ardwahišt ī sāl 36 uzēnīd ī dūdag ī
4. 'wst'nd'l l'd cygwn PWN bwn' Y whwmn 'whrmzd	4. ōstāndār rāy čiyōn pad bun ī Wahman-Ohrmazd
5. Y PWN nmywl d'lyk npštn' W 'm'l ptš	5. ī pad Namēwar dārīg nibištan ud āmār pad-iš
6. krtn' 'p̄'yt' ḤS X sp̄wd ʿL whwmn 'whrmzd	6. kardan abāyēd may 10 sabōy ō Wahman-Ohrmazd
7. YḤBWNt W MN whwmn 'whrmzd PWN gwk'dy mwdr'n Y	7. dād ud az Wahman-Ohrmazd pad gugāy-muhrān ī
8. 'wstwb'l'n ptgl'b ptš YNSBWNtny	8. awetwārān padīrāy pad-iš stad
9. ck' dyn'p̄zwt' HTYMWNt	9. čak Dēnabzūd āwišt

Translation:

[wnd?/ Personal name], [storekeeper/warehouse-keeper] in [placename], from that which (belongs) to the bun[5], from the yazdānbānag zād[xradwindād ōstāndār] (governor) [account], this month Ardwahišt (2nd month) of the year 36[6], for (rāy line 4) the expense of the ōstāndār's family, as (it) must be written under the responsibility of Wahman- Ohrmazd who is a storekeeper/warehouse-keeper in Namēwar and making a reckoning over it, gave Wahman- Ohrmazd 10 ewers (sabōy) of wine, and took a receipt sealed by the trustees'/reliable witnesses' seals for it from Wahman- Ohrmazd. Dēnabzūd sealed the check.

Comments:

Line 1: The first word (Picture 1) appears to be a personal name, which functions as the subject of the verb *dād* in line 7. Although the individual's administrative title is unfortunately missing from this document, the presence of the particle *pad* following *Y* (= ī) suggests a well-attested and familiar syntactic structure. Typically, this structure consists of the personal name, immediately followed by *Y* (= ī), then the preposition *pad*, a geographic location, and finally the common administrative title *dārīg*. Accordingly, the most plausible reconstruction is: **(personal name) ī pad (geographic location) dārīg**. Compare this with the following examples:

[5] See Weber et al. 2008a, Macuch 2008; 2022, Ramble 2022, Asefi 2024b, and Zeini 2024.
[6] After the death of Yazdgird III.

Berk. 91, 1& 2[7] Dādēnwindād ī pad Yazdānābestān dārīg

Berk. 14, 1[8] Friyag ī pad Namēwar dārīg

Picture 1

Lines 2-3: The title *yazdānbānag*[9] is well-documented. Compare its spelling with the following evidence:

Berk. 84[10], 2 & 3

Berk. 85[11], 6

Tabarestan 21[12], 4

Although some parts of this section of the document are missing, substantial evidence supports the reading *zād xrad-windād ōstāndār*. Compare it with the following example:

Berk. 35, 1[13]

Line 4: For reading The name, Wahman-Ohrmazd, see Weber (2008b [2004]: 137).

Line 5: Namēwar is the most frequently encountered toponym in the Hastijan documents. It is currently known as *Nimwar*, a small town located near Delijan[14].

---

[7] Asefi 2023b: 9.
[8] Gignoux 2009: 124.
[9] Protected by divine beings.
[10] Asefi 2003b: 17.
[11] Gignoux (2010: 106) mistakenly read the word as suftag ōstāndār.
[12] Gignoux 2012: 72, Weber 2019b: 93.
[13] Asefi 2024a: 4.
[14] see Weber 2014 [2010]: 38.

Line 6: The use of the term *sabōy* is evidenced in three unpublished documents, **Berk. 7**, **Berk. 117**, and **Berk. 199**, as well as in **Berlin 17**, as follows[15]:

Berk. 7, 4-7

4) …ḤG G II **5)** W ḤS I sp̄wd ʿL plyʾtk' **6)** Y kʾlplmʾn YḤBWNt' ck' **7)** mʾhpylwc' ḤTYMWNt

4) … gandum grīw 2 **5)** ud may 1 sabōy ō Friyādag(?) **6)** ī kārframān dād čak **7)** Māhpērōz āwišt

… (he) gave Friyādag(?), the overseer, 2 grīw of wheat and 1 sabōy (ewer) of wine. Māhpērōz sealed the check.

Berk. 117, 1-5

1) hwtʾyk' ẒNḤ BYRḤ mtr' **2)** QDM ŠNT X IIII IIII ḤG **3)** G X O ḤS I sp̄wd **4)** ḤLYʾ X dwlk' TBNʾ I p 5) ʾsp̄st C glt' …

[15] Also see Weber 2023: 290.

1) xʷadāyīg ēn māh Mihr **2)** abar sāl 18 gandum **3)** grīw 10 may 1 sabōy **4)** sik 10 dōlag kāh 1 paymān **5)** aspast 100 gird …

Lord, this month Mihr (7th month) in the year 18: 10 grīw of wheat, 1 sabōy (ewer) of wine, 10 pails of vinegar, 1 paymān of straw, and 100 bundles of lucerne…

Berk. 199, 1-3

1) W ḤS Y LẒNḤ YWM … 2) I sp̄wd ʿL… 3) YḤBWNtn l'd …

1) ud may ī im rōz … 2) 1 sabōy ō … 3) dādan rāy …

For giving … 1 sabōy of today's wine …

Berlin 17, 2-3

Berlin 17, 2: ḤS III-II sp̄wd /may 5 sabōy; Wine 5 sabōy.
To support the connection for /sp̄/, see following evidence:

Berk. 8, 3[16]- sp̄ndrmt/spandarmad

Berk. 97, 15[17]- gwsp̄nd I/ gōspand 1

Line 8: Compare the word awestwārān with awestwārīh in **Berk. 42**.

Berk. 25, 8

Berk. 42, 3[18] pad awestwārīh ī

Line 9: Compare Dēnabzūd with following evidence and see Weber 2012 [2008]: 217):

Berk. 25

Berlin 8, 7[21]

Berk. 11, 20[20]

Berlin 5, 3[19]

[16] Gignoux 2013: 158.
[17] Weber 2015: 156.
[18] Gignoux 2010: 66.
[19] Weber et al. 2008a: 25.
[20] Gignoux 2010: 56, he could not read the name.
[21] Weber et al. 2008a: 41.

- For the reader's convenience and to distinguish the new readings from the earlier ones, I have bolded all the updated readings.

Catalog Number	Collection Name	Material	Dimensions (cm)	Lines	Verso	Bulla
26	Berlin – Freie Universität	Leather	12.1 × 10.7 cm	10	Marked	None

Adapted from: Weber et al. (2008a: Tafel XXVI)

Transliteration	**Transcription**
1. z'twy ī PWN CBW Y MT' **d'lyk**	1. Zādōy ī pad xīr ī deh **dārīg**
2. lwcyk Y **krtn'** Y **b'gp'nyh** Y PWN **b'g**	2. rōzīg ī **kardan** ī **bāγbānīh** ī pad **bāγ**
3. Y **krt'-'p̄'t-yẕdtp't** MN BYRḤ	3. ī **Kard-ābād-yazdbād** az māh
4.'rtwhšt Y ŠNT XXXX W YWM	4. Ardwahišt ī sāl 40 ud rōz
5. ddw' PWN 'twr' pr'c ʿD BYRḤ	5. Day pad Ādur frāz tā māh
6. 'mwrdt **y** nzdst' W YWM 'whrmẕd	6. Amurdād-**ē**[22] nazdist ud rōz Ohrmazd
7. PWN II BYRḤ W XX-II YWM	7. pad 2 māh ud 22 rōz
8. ḤG G IIII-IIII K II ʿL	8. gandum grīw 8 kabīz 2 ō
9. ... Y **b'gp'n (?)**	9. ...[23] ī **bāγbān(?)**
10. YḤBWNt	10. dād

Top of the document, vertically:

Transliteration	**Transcription**
z't'nplhw'	Zādānfarrox
ḤTYMWNt	āwišt

Translation:

Zādōy, who is the storekeeper/warehouse-keeper of the property of the village, for carrying out the gardening work in the garden of Kard-ābād-yazdbād, from the month Ardwahišt (2nd month) of the year 40[24] and the day Day pad Ādur (8th day) until the beginning of the month Amurdād (5th month) and the day Ohrmazd (1st day), for two months and 22 days, gave 8 grīw and 2 kabīz of wheat to …, the gardener.

[22] This is a genitive construction. For a complete explanation of this structure and the meaning and function of *nazdist*, see Asefi & Mirfakhraie 2024: 151, 158-159.

[23] It refers to a proper name, but I am uncertain about it. Weber tentatively considered this whole line as a proper name, reading it as *Dēn-Nēw-Rōdagbaxt*. The reading of *bāγbān* (gardener) appears plausible both orthographically and semantically, but a non-final reading should be considered.

[24] After the death of Yazdgird III.

Comments:

Berlin 26 was previously examined and published by Weber et al. (2008a: 109). His reading and translation are as follows[25]:

Transliteration	Transcription
1. ’šṭw(y)h Y pṭ' CBW Y MT’ d’ryk	1. Ašweh(?) ī pad xīr ī deh dārīg
2. lwcky Y kwṭk' Y b’y’sp̄’nyh Y pṭ' bg	2. rōzīg ī kōdak ī bāyāspānīh ī pad bay
3. y krṭ' ’c’t-…-p’t(?) MN BYRḤ	3. ī kard Āzād-…-pād(?) az māh
4. ’rṭwhšṭ Y s’l XXXX W YWM	4. Ardwahišt ī sāl 40 ud rōz
5. ddw' pt' ’ṭwr pr’c ʿD BYRḤ	5. Day pad Ādur frāz tā māh
6. ’mwrdt Y nzdst' W YWM ’whrmẕd	6. Amurdād ī nazdist ud rōz Ohrmazd
7. pṭ' II BYRḤ W XX-II YWM	7. pad 2 māh ud 22 rōz
8. ḤG *G* IIII-IIII *K* -# II ʿL	8. gandum grīw-8 kabīz-2 ō
9. dynnywlwṭkbhṭ'	9. Dēn-Nēw-Rōdagbaxt(?)
10. YḤBWNt'	10. dād

Senkrecht:

1. z’ṭ’nplhw'	1. Zādānfarrox
2. ḤṬYWMNṭ	2. āwišt

Translation

[1] Āštwēh(?), der Verwalter für die Angelegenheiten des Dorfes; [2]Wegen Ration(en) für das Kind aus der bāyaspān-Ehe, die vom Herrn [3] gestiftet (lit. gemacht), Āāzād-…-pād(?) vom Monat [4] Ardwahišt (2. Monat) des Jahres 40 (671/672 n. Chr.) und Tag [5] Day pad Ādur (8. Tag) bis hin zum nächsten Monat [6] Amurdād (5. Monat) und Tag Ohrmazd (1. Tag), [7] (also) für zwei Monate und 22 Tage [8] Weizen 8 grīw (und) 2 kabīz der (?) [9] Dēn-Nēw-Rōdagbaxt(?) [10] gegeben.

[25] I must note that in a subsequent article, Weber revised the name *Āštwēh* to *Zādōy*. Additionally, in his latest publications, Weber now favours the reading *gālīg* over *dārīg*. However, in an unpublished article and my dissertation, I have demonstrated that multiple Persian and Arabic sources confirm *dārīg* as the correct reading, rendering *gālīg* unacceptable- also see Asefi: 2023b, 5-6.

Lines 2 and 3:

As is evident from Weber's interpretation, he views the document as primarily pertaining to a marriage and the birth of a son. By emphasizing the meaning of *bāyāspānīh* (mission or embassy), Weber (Ibid.: 110) suggested that the document most likely falls within a legal context involving a *bāyāspān* marriage. He further noted that a son born from such a union would have been referred to as *kōdak ī bāyāspānīh*[26].

However, I would argue that this document is not related to legal matters but rather to agriculture. I believe that in the second line, what Weber read as *kōdak* is in fact the infinitive *kardan*. This interpretation is supported by the following evidence:

Berk. 189, 8
krtn'/ kardan

Berk. 103, 3
krtn'/ kardan

Furthermore, what Weber interpreted as *bāyāspānīh* is, in fact, *b'gp'nyh/bāγbānīh* (gardening).

The spelling of the term closely matches *bāγbānīh* and aligns semantically with the context found in the Hastijan documents. These documents provide substantial evidence of rations being delivered to individuals in exchange for performing specific tasks.

At the end of the same line, Weber interpreted the letter following the letter /b/ as a single letter /g/ and read the word as *bg/bay*. However, evidence suggests that the scribe may have used a continuous line to represent three consecutive prongs. For example, the particle *'y/ē* and the postposition

[26] Der Inhalt des Dokuments wird weiter wesentlich erhellt durch das Wort bāyāspānīh, mission, embassy' (so CPD 18) in der 2. Zeile. Eher wahrscheinlich ist aber, daß wir es hier mit dem rechtlichen Begriff der bāyāspān-Ehe zu tun haben, so daß es sich bei dem *kōdag ī bāyāspānīh* um einen Jungen handelt, der aus einer derartigen Ehe stammt.

l'd/rāy have often been written in this manner, as demonstrated in the following two examples:

Berk. 86, 10[27]
ē

Berlin 12, 5[28]
rāy

So I think what Weber read as *bay* should be *b'g/bāγ* (garden).

In the third line, after *ī*, Weber interpreted the first word as *kard*, treating it as an independent verb. However, it should be noted that the verb *kard* was used in the previous line, and its repetition is redundant in this context. It appears that what follows *ī* is, in fact, the name of the garden. Thus, *kard* is not an independent verb here but rather a component of the place name. I interpret the full name as *Kard-ābād-yazdbād* (the meaning is Yazdbād cultivated it).

The spelling of *Yazdbād* is comparable to *Yazd-duxt*, and *Yazd* as seen in the following evidence:

Berk. 98[29]**,** 4
yz̲dt' dwht'/
Yazd-duxt

Berlin 14[30] , 7
yz̲dt' dwht'/
Yazd-duxt

Tab. 1A, 3[31]
Tab 1B, 3[32]
yz̲dt'/Yazd

This form of place-naming is also attested in the Sasanian period. In the book *Tarikh-e Qom* (edited by Ansari, 2006), numerous garden and farm names from various districts are recorded. For example, Bāγ-e Marzbān[33] in

[27] Asefi 2023b: 12.
[28] Weber 2008: 57.
[29] Asefi 2023b: 2.
[30] Weber 2021: 48.
[31] Weber 2016a: 123.
[32] Weber 2016a: 124.
[33] Ansari 2006: 313.

Rudābadān and the farm of Yazdgirdābād[34] in Kāšān. However, I could not find the garden *Kard-ābād-yazdbād* in *Tarikh-e Qom*.

The book also includes a fascinating account of the creation of a garden at a location known as Fālizbānān: "One day, Webīl said in Persian: īδēn bālīz-ē shāhēd kardan, meaning here, a garden can be built. So they built the garden and named it accordingly." (Ibid: 242)

With these corrections, it becomes clear that the document concerns the payment of rations to a gardener who worked in a garden called *Kard-ābād-Yazdbād* for two months and 22 days. The total ration for this period amounts to 8 *grīw* and 2 *kabīz* of wheat.

Additionally, I propose a correction for the reading of a single word (without reassessing the entire text) from an ostracon previously read and published by Weber (2022b: 527). It appears that the word Weber read as *bayaspān* should be corrected to *bārbān* (gardener).

O. 104 + O. 108

Transliteration

1 d'ṭynk' YWM slwš
2 ḤS-II dwlk' 'L
3 (h)w(l)šyṭ' ZK(!) Y GBR' I
4 Y by'sp̄'n YḤBWN

Transcription

1 *Dādēnag rōz Srōš*
2 *may 2 dōlag ō*
3 *X^{w}aršēd ān ī mard 1*
4 *ī bayaspān dah*

Translation

[1] *Dādēnag* (on) day *Srōš* (17th day) [2] 2 pails of wine to [3] *X^{w}aršēd*, a man [4] of the envoy, should give!

Fig. 25. O. 104/O. 108, Brit. Mus. 131864 + 131868 Ed. Weber 1992, pp. 64–65.

Adapted from Weber (2022b).

[34] Ansari 2006: 355.

Catalog Number	Collection Name	Material	Dimensions (cm)	Lines	Verso	Bulla
11	The Bancroft Library	Textile	25 x 7.5 cm	21	Marked	At the bottom

Transliteration	**Transcription**
1. d'tynwnd't Y PWN	1. Dādēnwindād ī pad
2. yzd'n'p̄st'n bwnp'n MN	2. Yazdānābestān[35] bunbān az
3. Z̲K **ZYš** MN ml Y **yzd'np'nk**	3. ān **ī-š** az mar ī **Yazdānbānag**
4. 'wst'nd'l ʿL bwn' lwcyk	4. ōstāndār ō bun rōzīg
5. Y **ḤMR' IIII** Y MN ml	5. ī **xar 4** ī az mar
6. Y 'wst'nd'l W PWN **yzd'n'p̄st'n**	6. ī ōstāndār ud pad **Yazdānābestān**
7. **YKʿYMWNd** p'ncdhlwck	7. **estēnd** pānzdahrōzag
8. Y MN Z̲NH BYRḤ 'twr'	8. ī az ēn māh Ādur
9. QDM ŠNT XXX III III III	9. abar sāl 39
10. W YWM tyl pr'c	10. ud rōz Tīr frāz
11. KR' YWM ywdt ywdt	11. har rōz jud jud
12. ***k* II** W II glt' l'd	12. **kabīz 2** ud 2 gird rāy
13. ŠG G X-II 'sp̄st	13. jaw grīw 12 aspast
14. C-XX glt' TBN'	14. 120 gird kāh
15. II *p* ʿL gwšnsp̄	15. 2 paymān ō Gušnasp
16. Y 'wštlyk'n YHBWNt	16. ī uštarīgān[36] dād
17. W PWN gwk'dymwhlyh'	17. ud pad gugāymuhrīhā
18. Y 'wstwb'l'n ptgl'b̲	18. ī awestwārān padīrāy
19. ptš YNSB̲WNtny	19. pad-iš **stad**
20. ck' **dyn'p̄zwt**	20. čak **Dēnabzūd**
21. ḤTYMWNt	21. āwišt

Translation:

Dādēnwindād, the bunbān of Yazdānābestān, from what belongs to bun from the account of the governor protected by divine beings (=Yazdānbānag), for the ration of four donkeys, which belong to the governor and are in Yazdānābestan, For fifteen days, starting from the day Tīr onward, in this month Ādur (9th month) in the year 39[37], and for each day, separately (for each donkey), two *kabīz* and two bundles, (totally) twelve *kabīz* of barley and one hundred and twenty bundles of lucerne, two *paymān* of straw, gave Gušnasp, the camel-driver. And for that received a receipt sealed by trustees'/ witness's seals. Dēnabzūd sealed the check.

Comments:

This document was previously examined and published by Gignoux (2010: 54-7).
In line 3, <ZYš> ī-š appears multiple times in the Pahlavi Archive of Hastijan and was previously always interpreted as 'YŠ/*kas* by Gignoux and Weber. However, Weber (2022a: 336), in his reading of **Berk. 24**, line 12, correctly

[35] This interpretation of the word is Arash Zeini's suggestion, which has been accepted by Weber. See Weber 2014 [2010]: 40.
[36] Gušnasp ī uštarīgān is Weber's suggestion Weber, 2022b: 533.
[37] After the death of Yazdgird III.

identifies it as *ī-š*. There is no doubt that this is the accurate interpretation, rather than *kas*.
Furthermore, in the same line, the word read by Gignoux as *xwadāyīg* should be revised to *yazdānbānag*.

Line 5: The first mistake in the reading of this line concerns the identification of the word used for the animal and its quantity. Gignoux interpreted the term as referring to a horse, with a quantity of one (asp 1). However, the word in question is the Aramaic heterogram ḤMR'/ xar and the correct number is four. This interpretation is confirmed by comparison with the following examples:

Berk. 75[38], 7
ḤMR' III-II: xar 5

Berk. 32[39]**,** 4
ḤMR' III: xar 3

Line 7: What Gignoux reads as *'ān nāmag'* should be corrected to YKʿYMWN-tny/ *estēnd*. In fact, this is the verb from the previous sentence, which indicates that these donkeys are in Yazdānābestān.

Line 12: The initial words read by Gignoux as *giyāh* should be corrected to k II/ *kabīz 2*. The key to reading and understanding this part of the document lies in the number of animals. In line 12, the daily ration for each animal is recorded by quantity and unit, but without specifying the type of ration. In line 13, the total ration, along with its name, is provided[40]. The main structure is: **jud jud** + **[Unit of measurement] [quantity for each. animal]+ rāy+ [Name of ration] [Unit of measurement] [total quantity].** In other words, lines 12 and 13 indicate that each animal received 2 kabīz and 2 bundles, daily. It is known that *jaw* (barley) is measured in *grīw* and *kabīz*, *aspast* (lucerne) in *gird* (bundles), and *kāh* (straw) in *paymān* or *dōl*. The following line states that a total of 12 *grīw* of *jaw* and 120 bundles of *aspast* have been given. Since four donkeys each received 2 *kabīz* per day, and the rations were for 15 days, each donkey should receive 30 *kabīz*, totaling 120 *kabīz*. Based on the *Tarikh-e Qom*, document named **Unknown II**, and **Berk. 219**[41], we know that in that region, each *grīw* was equivalent to 10 *kabīz*, so the provision of 12 *grīw* of barley recorded in the document exactly matches the 120 *kabīz* we calculated. Using the same calculation, the 120 bundles of *aspast* would also be distributed as recorded in lines 13 and 14. The total amount of straw was 2 *paymān*, which explains why straw is not mentioned in line 12, where the daily rations are listed.

---

[38] Unpublished.
[39] Gignoux 2010: 70, Weber 2010: 47.
[40] This numbering method is also observed in **Berk.100 and 119**. See Weber 2016b [2012]: 75.
[41] Gignoux 2009: 92.

Line 20: The name not read by Gignoux is Dēnabzūd. See the comment on line 9 of Berk. 25.

Catalog Number	Collection Name	Material	Dimensions (cm)	Lines	Verso	Bulla
122	The Bancroft Library	Leather	11 x 9 cm	10	Marked	None

Transliteration	**Transcription**
1. d'tynwnd't Y PWN yzd'n'p̄st'n	1. Dādēnwindād ī pad Yazdānābestān
2. bwnp'n MN Z̲K Y MN ml Y plhwtl	2. bunbān az ān ī az mar ī farroxtar
3. 'wst'nd'l ʿL BWN Z̲NH BYRḤ whwmn'	3. ōstāndār ō bun ēn māh Wahman
4. Y ŠNT XXXX-I lwcyk Y **'sm'nd't'**	4. ī sāl 41 rōzīg ī **Asmāndād**
5. **W ḤMR' III Y ʿL 'sm'nd't'**	5. **ud xar 4 ī ō Asmāndād**
6. **'p̄sp̄'lt' YKʿYMWNyt p'ncdh-lwck'**	6. **abespārd estēd pānzdah-rōzag**
7. **l'd MN YWM 'whrmz̲d pr'c ḤG *G* I**	7. **rāy az rōz Ohrmazd frāz gandum grīw 1**
8. ***k* III-II W ŠG *G* III ʿL 'sm'nd't'**	8. **kabīz 5 ud jaw grīw 3 ō Asmāndād**
9. YḤBWNt ck' **yz̲d'np[...]d'l(?)**	9. dād čak **Yazdānp...dār(?)**
10. ḤTYMWNt	10. āwišt

Transliteration:

Dādēnwindād, the storekeeper/ warehouse-keeper in Yazdānābestān, from the account of the fortunate governor to *bun*, in this month of Wahman (11th month), year 41[42], for the ration of *Asmāndād* and the four donkeys entrusted to *Asmāndād*, for fifteen days, starting from the day Ohrmazd onward, gave 1 *grīw* and 5 *kabīz* of wheat and 3 *grīw* of barley to *Asmāndād*. Yazdānp…dār sealed the *čak*.

Comments:

Line 4: Gignoux (2010, 92) has read this name as *Asmānzādān*. The final letter is written in a way that most likely represents a stroke. Compare it with /d't/ at the end of the name Dēdēnwindād at the beginning of the document.

Line 5: Gignoux did not read Lines 5 to the end of Line 8. These lines are being read for the first time in the present article.

Compare the spelling of *ḤMR'* with the following evidence:

Berk. 84[43], 4

Berk. 32[44], 4

[42] After the death of Yazdgird III.
[43] Asefi 2023b: 18.
[44] Gignoux 2010: 70, Weber 2010: 47.

Compare *abespārd estēd* in Line 6 with the following evidence:

Berk. 87[45]**,** 7

Compare *pānzdah-rōzag* with the same word in **Berk.11**:

pānzdah-rōzag
Berk. 11, 7

Line 7: Compare *az rōz Ohrmazd frāz* with the following evidence:

Berk. 122, 7

Berk. 33B, 3
ud rōz Ābān frāz

Unknown 2[46], 3
az rōz-iz Ohrmazd frāz

Line 9: some letters of the name is unclear. Although might initially appear to be ōstāndār, comparison with the evidence presented below indicates that this reading is unlikely. The word closely resembles Yazdānpādār, but two issues must be considered. First, the letters /'t/ are not visible in the available image. Second, the document is dated to the year 41, whereas all other known documents referring to Yazdānpādār were written between the years 34 and 40. In light of this, it is important to emphasize—as I noted in a lecture at INALCO[47]—that the evidence suggests there must have been

[45] Asefi 2023b: 18.
[46] Weber 2019a: 4.
[47] In April 2024, at the kind invitation of Prof. Samra Azarnouche.

documents mentioning Yazdānpādār in the year 41 which have not come down to us.

Yazdānpādār:

Berk. 78[48], 6

Berk. 33b, 8

Berk. 83, 10

Berk. 5, 7

Berk. 148, 10

ōstāndār:

Berk. 98, 5

Berk. 32, 8

Berk. 105, 4

Berk. 25[49], **4**

Berk. 43C, 4

BIBLIOGRAPHY

Ansari, M. R. (ed.) (2006/1385), *Tārīx-e Qom* [The History of Qom], Qom: Ketābxāne-ye Omoumī-ye Āyatollāh Marʿašī Naǰafī. In Persian.

Asefi, N. (2023a), "A New Middle Persian Document from Hastijan belonging to the Farroxzād Family", *Berkeley Working Papers in Middle Iranian Philology*, vol. 1, no. 3, pp. 1-14.

Asefi, N. (2023b), "Āzādmard in the Pahlavi Archive of Hastijan", *Sasanian Studies: Late Antique Iranian World / Sasanidische Studien: Spätantike iranische Welt*, vol 2, no. 1, pp. 1–26.

Asefi, N. (2024a), "Frāy in Seven Documents from the Pahlavi Archive of Hastijan", *DABIR*, vol. 10, no. 1, pp. 1-26.

Asefi, N. (2024b), *An analytical research on the parchment, leather, and textile manuscripts in Sasanian Middle Persian language from Qom* (PhD dissertation), Tehran: Institute for Humanities and Cultural Studies. In Persian.

[48] Unpublished.
[49] Unpublished.

Asefi, N., and Mirfakhraie, M. (2024), "Bāyegānī-ye asnād-e pahlavī-ye Hastiǰān" [Pahlavi Archive of Hastijan], *Journal of Persian language and Iranian Dialects*, vol. 8, no. 2. pp. 143- 163. In Persian.

Azarpay, G., 2003, "Bullae from the Pahlavi Archive at the University of California, Berkeley", in: *Ērān ud Anērān. Studies presented to Boris Ilich Marshak on the occasion of his 70th birthday (Transoxiana Webfestschrift Series 1)*, ed. M. Compareti, P. Raffetta, and G. Scarcia, Buenos Aires: Transoxiana. http://www.transoxiana.org/Eran/Articles/azarpay.html.

Gignoux, Ph. (1991), "Une nouvelle collection de documents en pehlevi cursif du début du septièmes siècle de notre ère", *Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres*, 135e année, N. 4, pp. 683-700. https://doi.org/10.3406/crai.1991.15035.

Gignoux, Ph. (2009), "Les documents économiques de Xwarēn", in: *Trésors d'Orient: Mélanges offerts à Rika Gyselen* (Studia Iranica 42), ed. Ph. Gignoux, Ch. Jullien, and F. Jullien, Paris: Association pour l'Avancement des Études Iraniennes, pp. 81–102.

Gignoux, Ph. (2010), "La collection de textes attribuables à Dādēn-Vindād dans l'archive pehlevie de Berkeley", in: *Sources for the History of Sasanian and Post Sasanian Iran* (Res Orientales 19), ed. R Gyselen, Bures-sur-Yvette: Groupe pour l'Étude de la Civilisation du Moyen-Orient, pp. 11-134.

Gignoux, Ph. (2012), "Une Archive post-Sassanide du Tabarestān (I)", in: *Objets et documents inscrits en pārsīg* (Res Orientales 21), ed. R. Gyselen, Bures-sur-Yvette: Groupe pour l'Étude de la Civilisation du Moyen-Orient, pp. 29-96.

Gignoux, Ph. (2013), "Les documents de Dādēn dans l'Archive de Berkeley/Berlin", in: *Commentationes Iranicae (Sbornik statej k 90-letiju V. A. Livšica)*, ed. P. Lur'e, and S. Tokhtas'ev, St. Petersburg: Nestor Istorija, pp. 157–165.

Nemati, M. R., Asefi N., and Sharahi I. (forthcoming), "New Discoveries from the First Excavation Season at Zel Cave in Hashtijan, Delijan, Iran."

Macuch, M. (2008), "Anhang VI: Zur juristischen Terminologie der Berliner Pahlavi Dokumente. in: *Berliner Pahlavi-Dokumente. Zeugnisse spätsassanidischer Brief- und Rechtskultur aus frühislamischer Zeit* (Iranica 15), ed. D. Weber, Wiesbaden: Harrassowitz, pp. 249–266.

Macuch, M. (2022), "Pahlavi legal documents from Tabarestān: The juristic contextof Tab. 12 and 26", in: *Sometimes Sasanian, always ēr* (Res Orientales 29), ed. R. Gyselen, Bures-sur-Yvette: Groupe pour l'Étude de la Civilisation du MoyenOrient, pp. 161–172.

Ramble, O. (2022), "Kerdīr's bun-xānag and funding foundations in Sasanian Iran." *Studia Iranica*, vol. 51, no. 2, pp. 163–203.

Weber, D., Krutzsch M., and Macuch M. (2008a), *Berliner Pahlavi-Dokumente*: *Zeugnisse spätsassanidischer Brief- und Rechtskultur aus frühislamischer Zeit* (Iranica 15), Wiesbaden: Harrassowitz.

Weber, D. (2008b [2004]), "Minuscula Pahlavica (Pahlavi Notes)", *Bulletin of the Asia Institute*, vol. 18, pp. 135-142.

Weber, D. (2012 [2008]), "New Arguments for Dating the Documents from the Pahlavi Archive." Bulletin of the Asia Institute, New Series, Vol. 22, pp. 215-222.

Weber, D. (2014 [2010]), "Villages and estates in the documents from the Pahlavi Archive: the geographical background", *Bulletin of the Asia Institute*, vol. 24, pp. 37-65.

Weber, D. (2015), "Spätsasanidische Preislisten im frühislamischen Iran", in: *Faszination Iran. Beiträgezur Religion, Geschichte und Kunst des alten Iran: Gedenkschrift für Klaus Schippmann* (Göttinger Orientforschungen, 3. Reihe: Iranica 13), ed. Sh. Farridnejad, R. Gyselen, and A. Joisten-Pruschke, Wiesbaden: Harrassowitz, pp. 233–243.

Weber, D. (2016a), "Court records of lawsuits in Tabarestān in the year 86/7 PYE (737 CE): a physiological examination", in: *Word and Symbols. Sasanian Objects and the Tabarestān Archive* (Res Orientales 24), éd. R. Gyselen, Leuven: Peeters Publishers, pp. 121-144.

Weber, D. (2016b [2012]), "Studies in some documents from the Pahlavi", *Bulletin of the Asia Institute*, vol. 26, pp. 61-95.

Weber, D. (2019a), "Two unknown documents from the Pahlavi Archive", *Sasanika Papyrological Studies*, no. 3. pp. 1-6.

Weber, D. (2019b), "Pahlavi legal documents from Tabarestān: Two claims and a re-evaluation of crop yields: a philological study of TAB. 21, 22, and 24", in: *Sasanian Persia and the Tabarestān Archive* (Res Orientales 27), éd. R. Gyselen, Bures-sur-Yvette: Groupe pour l'étude de la civilisation du Moyen-Orient, pp. 91- 116.

Weber, D. (2021 [2020-2021]), "Studies in some documents from the 'Pahlavi Archive' (3)": I. Breeding of poultry II. *Aspbād*, the 'caretaker', *Bulletin of the Asia Institute*, vol. 30, pp. 29–65.

Weber, D. (2022a), "Sasanian festivals in the documents from the Pahlavi Archive", *Sasanian Studies: Late Antique Iranian World / Sasanidische Studien: Spätantike iranische Welt*, vol. 1, no. 1, pp. 323- 345.

Weber, D. (2022b), "On the formal structure of checks in Pahlavi documents", in: *The Reward of The Righteous: Festschrift in Honor of Almut Hintze* (Iranica 30), ed. A. Cantera, M. Macuch, and N. Sims-Williams, Wiesbaden: Harrassowitz, pp. 515-548.

Weber, D. (2023a), "Cooking in 7th century Iran", *Sasanian Studies: Late Antique Iranian World / Sasanidische Studien: Spätantike iranische Welt*, vol. 2, no. 1, pp. 287- 313.

Zeini, A. (2024), "The covenant that binds: Ownership of life in late antique Zoroastrianism", in: *Deciphering the illegible: Festschrift in honour of Dieter Weber* (Iranica 33), ed. M. Macuch, and A. Zeini, Wiesbaden: Harrassowitz, pp. 461–480.

YEREVAN STATE UNIVERSITY PUBLISHING HOUSE

JIL 2/1 (2025): 30-35

Journal of Iranian Linguistics

On the Etymology of New Persian *malang* 'Intoxicated; Unorthodox Dervish'

Jakob Halfmann
University of Würzburg

doi.org/10.46991/jil/2025.01.02

Abstract: This brief article discusses the etymology of the New Persian word *malang*, which has been widely borrowed into Indian languages. First its semantic range is established as encompassing the two basic meanings 'intoxicated, drunk' and '(unorthodox) dervish', then a derivation from a Bactrian loanword into early New Persian is proposed, connecting it with the Indo-Iranian root **mad-* 'to be/become intoxicated'. The article thereby adds another lexical item to the list of Bactrian borrowings into New Persian and makes a contribution to the documentation of the Bactrian lexicon, as the presumable Bactrian source word is so far unattested in the known documents and inscriptions.

Keywords: Bactrian; etymology; loanwords; Persian

Jakob Halfmann
E-mail: jakob.halfmann@uni-wuerzburg.de
ORCID: https://orcid.org/0000-0003-3184-4022
Received: 29.11.2024
Revised: 27.06.2025
Accepted: 09.07.2025

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

© Jakob Halfmann, 2025

Introduction

The New Persian word *malang* seems to be relatively rarely used in present-day Iran and it is not listed in the standard etymological reference works like Horn (1893), Hübschmann (1895) or Ḥasan-Dūst (2014). It is, however, still well-known in Afghanistan, whence it has spread into many languages further east, especially within a Ṣūfī religious context.

In Steingass's (1892) dictionary, *malang* is defined as "A man barefoot and bare-headed, ecstasied; a religious enthusiast". Dehḫodā (1931) rejects the meaning 'naked from head to toe' (*sar-o-pā berahne*) and some others given by the *Farhang-i Rašīdī* as a "mistake" (*ġalaṭ*) and claims that 'drunk' is the only meaning of the word, citing also the collocation *mast-o-malang* with *mast* 'drunk'. The rejection of 'naked from head to toe' as a basic meaning of the word is likely correct, and also applies to the further misunderstanding as 'barefoot and bare-headed', that has found entrance into Steingass's (1892)

dictionary. This gloss probably entered the lexicographical tradition as an attribute to the meaning 'religious enthusiast' (vel. sim.), describing an unorthodox mendicant who walks about naked (a type of behavior that is historically well-attested for transgressive Ṣūfīs). In a similar way it appears in the *Maǧmaʿu l-Luġāt* of the *Qāmūs-i Kabīr-i Afġānistān* (Kabīr 2023), an online dictionary of the Persian of Afghanistan, which gives the following meanings:

> "A. mendicant (*faqīr*), drunk, tipsy (*sarxuš*), Sufi, inebriated (*maḫmūr*), (mystically) selfless (*bē-ḫud*), intoxicated (*mad-hūš*), (mystic) lover of God (*ʿāšiq-i Ḫudā*), gnostic (*ʿārif*), dervish, glad (*tar-damāġ*), *qalandar*
>
> B. persons without means (*bē-sar-u-sāmān*), barefoot and in need"

As in this latter source, the reference to (unorthodox) Ṣūfī mendicants is primary also in the borrowed forms in Pashto, Panjabi and Hindi/Urdu though additional meanings derived from this sense are also attested there:

- Pashto *maláng* – 'dervish, mendicant fakir; ascetic; careless person; in love'[1] (Aslanov 1966)
- Panjabi *malāg* – *n.* 'fakir, Muslim mendicant'; *adj.* 'carefree, indifferent to life' (Singh, Singh Gill & Joshi 1994)
- Hindi/Urdu *malang* – 1. 'a Muslim ascetic (esp. as in a state of trance)'; 2. 'a careless or inconsiderate person'; 3. 'a tall, robust man'[2] (McGregor 1993)

Papas's (2020) entry in the *Encyclopaedia of Islam* (3rd edn.) defines *malang* as "a word used in the Indian subcontinent and Afghanistan for a type of transgressive, if not antinomian, Ṣūfī" (Papas 2020: 117). He also offers a more precise description of the lifestyle of a *malang*:

> "*Malang*s often use hashish (smoked in a chillum pipe, drunk in a *bhāng* potion, or eaten in *pakoṛā* fritters), not to become passive but to strengthen their mystical experiences. They wear their hair long and wear women's dresses and heavy jewels (anklets, bangles, rings, and earrings). They dance and cry out with feminine voices, in order to transgress social or sexual norms and to make themselves the brides, even the prostitutes, of God. Often from poor backgrounds, malang dervishes, rather than entering into what they consider vain Ṣūfī discussions and theories, favour a spiritual life based on revelations during dances, drug hallucinations, and dreams." (Papas 2020: 117)

Papas (2020: 117) notes that "[t]he etymology of the word and the origin of the group are uncertain. Coming perhaps from Hindi or Persian and attested

---

[1] Transl. from Russ. "1) *derviš; niščenstvujuščij fakir*; 2) *asket*; 3) *bespečnyj čelovek*; 4) *vljubljonnyj*".

[2] The fourth sense 'a large white bird' is only attested lexicographically and may be erroneous or otherwise go back to a separate etymon.

since at least about 1008 [AH] / 1600 [CE], *malang* seems initially to have denoted certain members of a Ṣūfī group called *Madāriyya*; the scope of the term then broadened to individuals and communities comparable with respect to socio-religious practices."

From the point of view of semantic change, it is most likely that the meaning 'unorthodox Ṣūfī' developed from the more basic meaning 'drunk, intoxicated', especially since the religious practice of *malang*s prominently includes the consumption of intoxicating substances. This is also suggested by the collocation *mast-o-malang* mentioned by Dehḫodā. The origin of the term in Persian (rather than Hindi), on the other hand, seems assured based on the geographical distribution of the primary and secondary sense ('drunk' in Iran, but mostly 'dervish' and senses derived from it in India). The transmission of only the Ṣūfī sense into the Indian languages fits well into the historically attested sociolinguistic situation, in which Islamic concepts were transmitted to India via the medium of Persian.[3]

As for the date of first attestation, Papas's (2020: 117) dating to "at least about" 1600 CE probably relates only to the meaning 'dervish', since in the meaning of 'intoxicated' it already appears in much earlier works. A search in the *Ganjoor* corpus of classical Persian poetry gives a verse attributed by lexicographical works to the 10th – 11th century poet Labībī as its earliest attestation:[4]

zi ǧā ǧast čūn ātaš-ē bē-dirang dil az bāda-yi ʿišq mast-u-malang
He jumped from the place like a fire without delay,
the heart drunk and intoxicated from the wine of love.

Ever since Henning (1939), it has been acknowledged that the lexicon of New Persian contains a component of borrowings from Iranian languages once spoken further to the east, which may have entered the literary language in the early stages of its development in Ḫurāsān. After the discovery and edition of the Bactrian documents (Sims-Williams 2007 etc.), which resulted in a much better understanding of the Bactrian language, Lurje & Yakubovich (2017) were able to correct Henning's (1939) assumption that an unattested Sogdian dialect with the sound change **d > l* was the source of New Persian words that have this development. Instead, these words can more

[3] The alternative derivation of Hindi *malaṅg* in the sense 'a tall, robust man'/'stout-bodied man' from the Vedic hapax *mālánga-* 'epithet of a hornless goat', which is endorsed by Turner (1962-1966: T. 10089) based on Sharma (1959/1960; *non vidi*), is hardly convincing in itself and also does not account for the other senses of the word. The alleged Hindi variant ***mālaṅg* quoted there, which would need to be primary for the etymology to work phonologically, is not recorded by McGregor (1993), who also doubts the etymology ("?") and rather sees the source of the Hindi word in Panjabi, whence it may in fact have been transmitted. Much more likely is the assumption of a semantic development 'ascetic, fakir' > 'stout/robust man' in Hindi.

[4] https://ganjoor.net/labibi/abyat/sh74

straightforwardly be explained as borrowings from Bactrian, which shows **d* > *l* as one of its most characteristic features (cf. Kreidl 2021 for details and chronology).

I would like to argue that the word *malang* has to be added to the list of such borrowings from Bactrian. The shared Indo-Iranian root **mad-* 'to become intoxicated' (Cheung 2007: 253; Mayrhofer 1996: 299–300) is a likely source for a word with the basic meaning 'drunk, intoxicated' (cf. also the collocated New Persian word *mast* 'drunk' < **mad-ta-*) and *malang* could easily derive from this root if it is a Bactrian word with the typically Bactrian sound change **d* > *l*. It is also worth noting that a word from the same semantic sphere – Persian *mul* 'wine' < Bactrian μολο 'wine' – is already established as having been borrowed in this direction (Lurje & Yakubovich 2017: 324). Another point that speaks in favor of an origin in Bactrian, and against, e.g., a derivation from Pashto, which also shows **d* > *l*, is the suffix *-ang*. Nouns or adjectives ending in *-ng* (written ⟨-γγο⟩) are quite frequent in Bactrian, resulting in all cases where the *n* of a stem or primary suffix came into contact with the *k* of the frequent **-ka-* suffix after the syncope of the preceding vowel. Thus there is ⟨-αγγο⟩ < **-āna-ka-* (Sims-Williams 2007: 187), ⟨-ηγγο⟩ < **-aiṇa-ka-* (Sims-Williams 2007: 213), ⟨-γωγγο⟩ < **-gauṇa-ka-* (Sims-Williams 2007: 207), ⟨-λαγγο⟩ < **-dāna-ka-* (e.g. in βαγολαγγο 'temple, sanctuary'; Sims-Williams 2007: 200) etc.

A formation with the middle participle ending **-āna-ka-*, one of the sources of Bactrian ⟨-αγγο⟩ /-āng/, would be a very plausible candidate for the origin of *malang*, since intoxication is a typical middle voice activity, involving an action that affects the subject. The root **mad-* is attested with middle voice inflection, at least in some cases, both in Old Iranian and in Old Indo-Aryan. The short vowel in the second syllable of the Persian form can be explained with the Persian-internal sound change *āng* > *ang*, which is attested in some early New Persian texts (Lazard 1963: 182; Bernard 2025: 59–60, 65).[5] Based on the geographical origins of these texts, Lazard (1963: 182) attributes the change to the dialects of eastern Afghanistan. One of the works in question, the geographical treatise *Ḥudūd al-ʿālam*, was written in Guzgan, i.e., specifically on the territory of Bactria-Tokharistan. This makes it plausible that the sound change applied in the Persian dialects that were in contact with Bactrian and that the word **malāng* > *malang* was then transmitted from these dialects to other Persian-speaking regions.

A Bactrian word *μαλαγγο 'drunk' is not attested in the surviving Bactrian documents or inscriptions, but – if the etymology is correct – its former existence can be assumed based on the Persian word. Its attestation in Persian from the 11th century onwards is not too far removed from the latest surviving writings in Bactrian and allows for the assumption of transmission

[5] I owe this crucial observation to a comment from an anonymous reviewer.

into the Early New Persian of Bactria-Tokharistan and adjacent regions. The narrowing of its meaning to 'unorthodox dervish' may have come as late as 1600 CE in association with the Madāriyya order, a time when the Bactrian language was most likely long extinct.

BIBLIOGRAPHY

Aslanov, M. G. (1966), *Afgansko-Russkij Slovar' (Puštu)*, Moscow: Sovetskaja Ėnciklopedija.

Bernard, Ch. (2025), "An Early Judeo-Persian Rabbanite text: Vat. Pers. 61, its linguistic variety, its Arabic vocabulary, and the Targum Onqelos", *Journal of Jewish Languages*, vol. 13(1), pp. 22–76. https://doi.org/10.1163/22134638-bja10056.

Cheung, J. (2007), *Etymological Dictionary of the Iranian Verb* (Leiden Indo-European Etymological Dictionary Series 2). Leiden: Brill.

Dehḫodā, ʿA-A. (1931), *Loġat-nāme-ye Dehḫodā*. online version. https://dehkhoda.ut.ac.ir/fa/dictionary.

Ḥasan-Dūst, M. (2014), *Farhang-e Rīše-šenāḫtī-ye Zabān-e Fārsī*, vol. 1, Tehran: Farhangestān-e Zabān va Adab-e Fārsī.

Henning, W. B. (1939), "Sogdian loanwords in New Persian", *Bulletin of the School of Oriental and African Studies*, *University of London*, vol.10, no. 1, pp. 93–106.

Horn, P. (1893), *Grundriss der Neupersischen Etymologie* (Sammlung Indogermanischer Wörterbücher IV), Strassburg: Trübner.

Hübschmann, H. (1895), *Persische Studien*, Strassburg: Trübner.

Kabīr, Q. (2023), *Maǧmaʿu l-Luġāt. Qāmūs-i Kabīr-i Afġānistān*. http://www.afghan-dic.com/pages/dic.aspx?dict=dictCat1List1.

Kreidl, J. (2021), "Lambdacism and the Development of Old Iranian **t* in Pashto", *Iran and the Caucasus*, vol. 25, no. 2, pp. 175–193. https://doi.org/10.1163/1573384X-20210207.

Lazard, G. (1963), *La langue des plus anciens monuments de la prose persane* (Études linguistiques II), Paris: Librairie C. Klincksieck.

Lurje, P., and Yakubovich, I. (2017), "The myth of Sogdian lambdacism", in: *Zur lichten Heimat. Studien zu Manichäismus, Iranistik und Zentralasienkunde im Gedenken an Werner Sundermann* (Iranica 25), ed. Team "Turfanforschung", pp. 319–341, Wiesbaden: Harrassowitz.

Mayrhofer, M. (1996), *Etymologisches Wörterbuch des Altindoarischen*, vol. 2, Heidelberg: Universitätsverlag Winter.

McGregor, R. S. (1993), *Oxford Hindi-English Dictionary*, Oxford: Oxford University Press.

Papas, A. (2020), "Malang", in: *Encyclopaedia of Islam*, ed. K. Fleet, G. Krämer, D. Matringe, J. Nawas, and E. Rowson, Leiden: Brill.

Sharma, A. (1959/1960), *Beiträge zur vedischen Lexikographie. Neue Wörter in M. Bloomfields Vedic Concordance* (PHMA: Mitteilungen zur idg., vornehmlich indo-iranischen Wortkunde sowie zu holothetischen Sprachtheorie 5/6), München: In Kommission bei J. Kitzinger.

Sims-Williams, N. (2007), *Bactrian Documents from Northern Afghanistan* (Corpus Inscriptionum Iranicarum Part II: Inscriptions of the Seleucid and Parthian Periods and of Eastern Iran and Central Asia 6.1.2), vol. II: Letters and Buddhist Texts. London: The Nour Foundation & Azimuth Editions.

Singh, G., Gill, M., and Joshi S. S. (1994), *Punjabi University Punjabi-English Dictionary*, Patiala: Punjabi University.

Steingass, F. J. (1892), *A Comprehensive Persian-English dictionary, including the Arabic words and phrases to be met with in Persian literature*, London: Routledge & K. Paul.

Turner, R. (1962), *A Comparative Dictionary of the Indo-Aryan Languages*, London: Oxford University Press.

YEREVAN STATE UNIVERSITY PUBLISHING HOUSE

JIL 2/1 (2025): 36-47

Journal of Iranian Linguistics

The Progressive Tenses with the Verb "to have" as a Peculiar and Exclusive Feature of Western Iranian Dialects Spoken in Modern Iran

Youli A. Ioannesyan

Institute of Oriental Manuscripts of the Russian Academy of Sceinces

doi.org/10.46991/jil/2025.01.03

Abstract: The article considers the verb forms with the auxiliary verb "to have" used in Persian and other western Iranian dialects of modern Iran to express a continuous action taking place at the moment of speech in the present or at a certain moment in the past. Drawing upon an extensive material collected from a large number of dialects, the author proves that the geographic area within which these forms are found is limited to the territory of modern Iran, which makes them a distinguishing and exclusive feature of western Iranian (Persian and non-Persian alike) dialects spoken in modern Iran. The author supports the idea that these verb forms are completely "native" to the linguistic area they are found in and, therefore, despite the view-point of some scholars, could not have originated under the influence of western or any other foreign languages.

Keywords: Persian language, Persian dialectology, Iranian dialectology

Youli A. Ioannesyan
E-mail: youli19@gmail.com
ORCID: https://orcid.org/0000-0003-2936-1128

Received: 13.02.2025
Revised: 19.06.2025
Accepted: 28.06.2025

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

© Youli A. Ioannesyan, 2025

Many Western Iranian languages and dialects have verb forms that express an action taking place at a given moment in time in the present or at a certain moment in time in the past. These forms have received various names in Iranian Studies literature, where they are often defined as two special tenses: "present continuous (concrete) tense" and "past continuous (concrete) tense"/ "concrete present" and "concrete past" or "the present and past progressive" (Rastorgueva & Èdel'man 1982: 537-538; cf. Thackston 2009: 204-205), or "periphrastic progressive" (Jeremiás 1993: 99-116). Since, in comparison with the present indicative and past continuous tenses, they are distinguished by a special emphasis on expressing an action in the

process of its implementation (an ongoing action), it would be proper, in my opinion, to also call these forms the "present progressive" and "past progressive" tenses,[1] the terms which we will apply to them in this article, while defining both of them, in the general sense, as "progressive action forms". It should be emphasized that these forms in Western Iranian languages and dialects are considered in Iranian Studies literature as "an innovation" (Rastorgueva & Èdel'man 1982: 538-539), which will be discussed below. This subject has been addressed by several authors in their works: V. A. Žukovskij (1888), A. Z. Rosenfeld (1948), L. S. Pejsikov (1960), M. Pistoso (1974), N. Nematollahi (2014, 2018), I. Dehghan (1972), etc., some of which will be considered in this article. It should also be noted that the present study is primarily dialectological and not typological, though it analyzes some non-dialectological aspects as well. Since a large number of examples collected from numerous published sources are used in this article, it is technically impossible to employ a uniform transcription for the examples from different Iranian languages and dialects. For this reason, examples are quoted in the original transcription of the sources they are cited from with non-substantial modifications, except for the transliteration of the Cyrillic script into Latin.

Progressive action forms are widespread in the Persian, Dari and Tajik languages, including their dialects. In Persian and Dari, they are especially characteristic of informal speech and are practically not found in the official style. In Tajik, these tenses are formed from the unchangeable participle of the notional verb and the personal form of the perfect or past perfect tense of the auxiliary verb *istodan* "to stand, to remain": *karda istodaam* and *karda istoda budam* – first-person singular of the present and past progressive tenses of the verb "to do", etc. (Rastorgueva & Èdel'man 1982: 538-539).[2]

In modern Dari, special complex forms are also used to express the same meaning, which also contain the unchangeable participle of the notional verb. But unlike the Tajik language, the auxiliary verb here is not *istodan*, but *raftan* "to go", which is inflected for person and number. The latter can appear in the present indicative or past tense forms, depending on whether the action it expresses refers to the present moment or to a certain moment in the past: *čerâ larzida mêri*? "why are you trembling?",[3] *tamâm-e šaw Šânâma xânda raftom*... "I read the Shah-nameh all night..." (Farhâdi 1955: 84). At the same time, in Dari, there is also a less common alternative

[1] Thus, unlike the past progressive tense, the past continuous tense in Persian can denote not only an ongoing action but also a repeated action.
[2] Transcription modified.
[3] Here and below examples from the Dari language are given according to the book by A. Farhâdi (1955) in a slightly modified transcription.

progressive action form with the compound verb *râyi* (lit. *râhi*) *budan* instead of the verb *raftan*, identical with the latter in meaning: *dida mêrên* = *dida râyi'sten*[4] "you see" (Farhâdi 1955: 125). Thus, it can be stated that, despite the fact that Dari and Tajiki use different auxiliary verbs, they are united by a common feature - the notional verb in both is employed in the form of a participle.

In the territory of Iran, in Persian and closely related to them Iranian dialects, progressive action forms for the present and future are also widely used. The difference between them and the above-described ones is that they are composed of the verb *dåštän* "to have" of the relevant tense (present indicative or past absolute) in agreement with person and number, and the personal form (not the participle) of the notional verb, cf. for example, in the Tehrani dialect: *män dåräm mixoräm* "I am eating", *to dåšši* (< *dåšti*) *mixundi* (< *mixåndi*) "you were reading" (Pejsikov 1960: 76, 77).[5]

One of the first European researchers to draw attention to these forms in Persian dialects was V. A. Žukovskij. In the 1880s, he wrote: "*In the Persian colloquial language, as in some modern Persian dialects, the verb* داشتن *[dåštän], in addition to its usual meaning "to have" ... has a special meaning, which in the literary language ... it does not have*" (Žukovskij 1888: 376).[6] This scholar identified two functions of these tenses: 1) the combination of the aorist of the verb *dåštän*, standing before the notional verb of the present indicative tense, "gives the latter the meaning of such a future tense that must occur immediately; therefore... resembles the present tense with *aller* in French..."; 2) the verb *dåštän* in the past tense "*before the Praeteritum... of another verb, gives the latter the meaning of such a past that has just happened - in French the present tense with venir de...*" (Žukovskij 1888: 376). V. A. Žukovskij gives examples from everyday speech: دارم مِیام [*dåräm miyåm*] "I'll be right there!", داشت رفت بازار [*dåšt räft båzår*] "he just now went to the market" and an example of folk poetry (satire):

اینها همه دُرُس شد حاکم طهران اُروس شد
کُلونل آمد از راه لَمسه پولِس ساز میزنه داره میرقصه.
(Žukovskij 1888: 376-377, also 1902: 250)

Inhå häme doros šod *håkem-e Tehrån orus šod*
Kolunel åmäd äz råh-e Lämse *pules såz mizäne dåre miräghse/miräkhse*

[4] *Râyi'sten* < *râyi (râhi)* + *(h)astên*
[5] The original transcription has been slightly modified here and below.
[6] All translations of the quoted material here and below are my own.

"All this came to order:
A Russian became the governor of Tehran,
The colonel came by way of Austria,
The police are playing music (sic),
Are about to dance."

While recognizing V. A. Žukovskij's undoubted merit in that he drew attention to these verb forms at the dawn of research in the field of Persian dialectology, it should be noted that their grammatical meaning was determined by him incorrectly, which was rightly pointed out, in particular, by A. Z. Rosenfeld, who wrote: "*... and the examples given by V. A. Žukovskij in the above-mentioned article should be translated not into the future, but into the present tense... [It is not that] 'I will come now, but I am coming.' The same is true for the following example from popular bazaar satire... 'is dancing (now)', i.e. 'the police are playing and dancing'*" (Rosenfeld 1948: 308).

A. Z. Rosenfeld's view was shared by L. S. Pejsikov. In his monograph on the Tehran dialect, he wrote: "*V. A. Žukovskij noted ... that combinations with داشتن [dåštän] have the meaning of the near future ... V. A. Žukovskij's conclusion is too 'broad', although the shade of the near future is sometimes there ... Thus, to the call 'come here' one can respond 'dåräm miyåm'... 'I'll be right there', but in another situation... 'dåräm miyåm dige!' [would mean] 'I'm coming!'...*" (Pejsikov 1960: 77). The example *dåre miraxse*..., L. S. Pejsikov rightly points out, is not interpreted by V. A. Žukovskij quite correctly: not "he is about to dance", but "(he) is dancing". The former author makes the following general conclusion: "*The feature of the verb dåštän noted by V. A. Žukovskij, which, in his opinion, corresponds to the French 'near' past tense of venir de with the infinitive of the conjugated verb, is not confirmed by the conversational examples I recorded, as well as numerous examples from modern literature. V. A. Žukovskij's example (داشت رفت بازار [dåšt räft båzår]) is not clear, since dåštän cannot stand before the past absolute/perfective form of an inflected verb at all. In this regard, it was also not possible to find a shade of meaning corresponding to the French Passé antérieur*" (Pejsikov 1960: 78). The general meaning of this verb form in the present tense is defined by L. S. Pejsikov as corresponding to the English present continuous (Pejsikov 1960: 76).
The verb forms described above have attracted the attention of Iranian linguists. Thus, the author of a thorough study on the Birjandi dialect (one of the southern Khorasan dialects of Persian) J̌. Reżāʾī not only denies them the legal right to exist in this dialect, but also believes that they originated in Persian "recently" and under the influence of European (!) languages, and explains their occasional occurrence in the speech of this dialect speakers by the desire to "conform to the norms of the modern Persian language" (Reżāʾī,

1998, 279).[7] This view of the situation is puzzling. Firstly, as N. Nematollahi rightly notes in an article on this topic, such a tense (or progressive construction with the verb "to have") does not exist in any of the foreign languages known to Iranians, including English, French, German and Russian. Rejecting the idea of these constructions being borrowed, she proceeds from the fact that their absence in the indicated languages excludes the possibility of their occurrence in Persian under external influence (Nematollahi 2014: 104, 110). Secondly, these grammatical forms began to be recorded by researchers in the late 19th – early 20th centuries not only in many Persian, i.e. dialects of the southwestern type, but also in northwestern Iranian dialects, and in the so-called "dialects of Central Iran", which share a number of common features with the northwestern ones. At the same time, the researchers who noted these forms at the turn of the 19th and 20th centuries collected their materials from among the predominantly illiterate and semi-literate population of Persia at that time, hardly subject to the influence of European languages. Thus, similar constructions are noted by V. A. Žukovskij, whose field research on various dialects dates back to the 80s of the 19th century, and by W. A. Ivanow and D. L. Lorimer in the dialects of the Zoroastrians of Yazd and Kerman (see below). The use of this tense in the Mazanderani language is extremely significant. The thesis about the appearance of these verb forms "under the influence of translations from European languages" looks even less convincing against the background of the early recording of their use in northwestern Iranian dialects and in the so-called "dialects of Central Iran" closely related to them. Since a detailed examination of examples is beyond the scope of this work, we will limit ourselves to only a few below.

In addition to the examples of Persian colloquial speech and folklore cited by V. A. Žukovskij above, in which the verb forms in question appear, this

---

[7] Original:

واگر اکنون کسی چنین فعلی به کار ببرد و مثلا بگوید «داشتم مرفتم» ... بیگمان زیر تأثیر زبان فارسی کنونی چنین کارواژه‌ای بر زبان خواهد راند، با آن‌که ... در خود زبان فارسی نیز چنین کاربردی تازگی دارد و یقیناً از طریق ترجمه از زبان‌های اروپایی به این زبان راه یافته است. (Reżāʾī, 1998, 279)

Transcription (mine):
vä ägär äknun käs-i čenin feʾl-i be kår bebäräd vä mäsälän beguyäd "dåštom maraftom" bigomān zir-e täʾsir-e zäbān-e fårsi-ye konuni čenin kårvåže-yi bär zäbān xåhäd rånd, bå ān ke … där xod-e zäbān-e fårsi niz čenin kårbord-i tāzegi dāräd vä yaqin äz tärigh-e tärjome äz zäbānhā-ye orupāyi be in zäbān rāh yāfte-äst.

English translation (mine):
"*If someone today uses such a verb [form], for example says: 'dāštom maraftom'…, then undoubtedly they do so under the influence of the modern Persian language, even though such a usage is also a novelty within Persian itself, and has undoubtedly entered the language through translations from European languages.*"

researcher in the same article notes a similar construction in the reworking of Hafiz's ghazal in the dialect of the village of Kafron near Isfahan. In it, the poet's words: فریاد رسی می آید [*faryādras-ī mīāyad*] "the helper is about to come" correspond to: *feŕyōd-rés-ī dōrū yūe* (Žukovskij 1888: 377). Similar forms are found in the dialect of the village of Sedeh, also located near Ishafan, cf.: *dåtom*[8]*-o šaame* (Rastorgueva & Moškalo 1997: 320) "I walked (at a certain point in time)" and in the dialects of the Jews of Isfahan, judging by the texts published by R. Abrahamian: *Tu fekr därbo, bišti ye ra'yet dáru älef činive* – "Il était dans ses penseès quand il vit un peysan moissoner son champe" ["he was lost in his thoughts when he saw a peasant collecting grass"] (Abrahamian 1936: 73, 104). A similar tense with a dialectal version of the verb "to have" - *důrtwun* is presented in the dialects of the Zoroastrians of Yazd and Kerman: *me 'důre 'rase* - corresponds to literary Persian: *man dåram miräsäm*, *ta' důri 'rasi - to dåri miräsi*, etc. cf. corresponding past tense forms: *me 'důrte 'rasůde*, *ta 'důrti 'rasůdi*, etc. (Ivanow 1935: 80-81).[9]

The Mazanderani language also has present progressive and past progressive tenses with the same auxiliary verb: *dɔrmɛ nɛvisɛme* "I am writing (now)", *dɔrni nɛvisɛni* "you are writing", *dɔšti nɛvéšti* "you were writing (at a given time)", *dɔštɛ nɛvéštɛ* "he was writing", etc. (Rastorgueva & Èdel'man 1982: 538-539).[10] It should be particularly emphasized that the above examples illustrate the use of these forms not in dialects of the southwestern type, but in northwestern ones (including the so-called "dialects of Central Iran").

J̌. Reżā'ī's assertion that modern speakers of southern Khorasani dialects (and perhaps even Khorasani dialects in general) in Iran use these constructions "exclusively under the influence of modern Persian" also seems controversial. Thus, describing the features of the Persian dialects of Khorasan at the dawn of the 20th century, W. A. Ivanow wrote: "*Common to colloquial speech all over Persia is the use of the verb dāshtan, in the Present tense for an action in progress: dårä muknä – he is now doing*" (Ivanow 1925: 253). Though, in the Khorasani texts themselves, published by Ivanow, the present author has not found these forms, but this fact should be explained by the peculiarity of the texts themselves - tales and quatrains, while the use of the constructions in question is more characteristic of texts of everyday content, not presented in W. A. Ivanow's Khorasani materials. However, the very statement of the fact by this outstanding researcher and expert on the dialects of Khorasan, although illustrated by him with a single example, is important evidence of

[8] dåtom < dåštom.
[9] The full paradigm of these forms in the dialects of the Zoroastrians is presented in Rastorgueva & Moškalo (1997: 320).
[10] Transcription slightly changed.

the use of progressive action forms in these dialects, at least already at the turn of the 19th and 20th centuries.

According to my observations, in the modern dialect of the city of Mashhad, the center of the Iranian province of Khorasan, forms such as *derän merän* "they are going/leaving now" are very common. They are also widely represented in stories of everyday content in this dialect, recorded in Arabic alphabet: دِره مِره *dere mere* "he is going", داشت پیاده مُشُد [*dåšt piyåde mošod*] "he was getting out of the car/bus", موُ که چِشمام داشت از حدقه مِزَد بیرون [*mu ke češmå-m dåšt az hadaghe mezad birun*] "my eyes [in surprise] bulged out of their sockets", داشتُم خِفه مُشُدُم [*dåštom xefe mošodom*] "I was suffocating [from the heat]".[11] An identical form is also registered in the Sistani Persian dialect: *malāng-ā dâra nō mxâra* "malang is also eating food[12]" (Ahangar 2010: 20).[13]

Bakhtiari is another example of the south-western dialects in Iran where the progressive action tense with *dåštän* (*dāsštan*) is recorded, cf. Bavadi dialect: *dāsštom film e-diδ-om* "I was watching a movie" (Taheri-Ardali et. al. 2025: 24).[14] It is significant that these forms, used in numerous southwestern and other Iranian dialects in the territory of modern Iran, are absent from dialects closely related to Persian, outside that country. They are not found even in those of them spoken in northwestern Afghanistan, such as Herati, which is a continuation of the single Khorasani area of Persian dialects. These constructions are not recorded in the Tati dialects of the Caucasus either. Thus, it can be concluded that the use of these tenses with the verb *dåštän* "to have" (in different phonetic variants) is an exclusive feature of the area of Persian and other Western Iranian dialects, that are spoken precisely in the territory of modern Iran, i.e. an isogloss that distinguishes these dialects from the broader spectrum of Iranian languages and dialects, including those belonging to the dialect continuum of closely related southwestern languages – Persian, Dari, and Tajiki.

These forms, as already noted, are considered in the relevant literature as "newly originated/of recent origin". Such a definition is justified if "recent origin" in this case implies that they occurred in the Persian language and its dialects later than other tenses used in them today, although it is currently impossible to determine how "late" they came into use in the language compared to the present indicative and past continuous tenses with

[11] The stories from which these examples are taken are available on the Internet site: http://www.bechemashad.blogfa.com/category/2/.
[12] Lit. bread.
[13] Transcription modified.
[14] Transcription modified.

the verbal prefix *mī-* (< *hamē*), which are not represented in the language of classical Persian literature either. In the latter, unprefixed verbal forms were employed to express the present action, which in modern language have acquired the function of the subjunctive mood. The fact that tenses with *dåštän* were not recorded in writing until the end of the 19th – beginning of the 20th centuries cannot serve as an argument in favor of their absence, if we take into account the extreme conservatism of the literary and official Persian language in the Middle Ages and in the Qajar era, which in no way reflected the tendencies of colloquial speech until the timid attempts to "democratize" Persian literature in the first half of the 20th century. On the other hand, the fact that these forms were recorded in the materials of European researchers as early as the 1880s (i.e., at the dawn of dialectological research) across a wide range of dialects suggests that they must have existed in the language prior to that time. Otherwise, they would not have been able to spread so quickly and take root in the speech of such wide strata of the population, speaking different social and territorial dialects, especially in the conditions of 19th century Persia with a predominantly illiterate population. In other words, it is quite logical to assume that at the time of their written recording, these forms could have already existed in the spoken language for more than one century. At the same time, it cannot be ruled out that their spread (but not their origination) could have been facilitated to a certain extent by the mutual influence of dialects and a certain common supra-dialectal conversational standard (*koine*), which is superimposed in each region on the local dialect and acquires local dialectal features. The influence of such a standard is especially noticeable in cities and regional centers. Perhaps this explains the fact that these forms are more characteristic of the dialects of cities (from Tehran to Kerman and from Shiraz to Mashhad) than of rural settlements, although they are also recorded in the latter. But, as noted above, it would not be correct to explain the appearance of any grammatical form or category in the speech of dialect speakers solely by their desire to "imitate" someone or some standard. The interaction between the supra-dialectal standard and local dialects is a much more complex and organic process, including mutual enrichment and the assimilation by territorial dialects of some common features inherent in the entire set of dialects of the area, features that, when they "fall on the soil of a separate dialect," become its integral part.

These forms or similar ones should have appeared in the language in any case, since they occupy an important niche in the verb system, i.e. a means of expressing an ongoing action at the moment of speech in the present or at a certain point in time in the past, and thus fill the obvious gap in the set of grammatical possibilities for expressing this type of action. The fact that without them this gap would have remained is confirmed not only by the

appearance of forms corresponding in meaning in Dari and Tajik (see above), but also by the occurrence of alternative ways of expressing such an action in some Iranian languages and dialects on the territory of Iran, for example, in Gilaki, Lari and Bandari, which reflects an objective general need for their existence.

In the Gilaki language, progressive action forms are constructed by combining the infinitive of the notional verb with the personal form of the insufficient verb "to be, to stay", outwardly very similar to the verb "to have": *nivištən dərəm* "I am writing", *nivištən dəri* "you are writing", *nivištən dərə* "he, she is writing", etc. At the same time, the past tense is represented by two variants: *nivištən dərə bum/nivištən dubum* "I was wroting", *nivištən dərə bim/nivištən dibim* "we were writing", *nivištən dərə bid/nivištən dibid* "you were writing", etc. ((Rastorgueva & Èdel′man 1982: 538).

In the Lari language/dialect, forms with the same meaning are derived from the infinitive by lengthening the final vowel *a*, the preverb *a-* and personal endings: *a-dedå-y* "he sees", *a-gotå-en ke* "they say that ...", etc.[15] M. S. Pelevin has recorded similar constructions formed from the past participle by means of the prefix *a-* and personal endings in the Bandari dialect of southern Iran: *me kār akerdum* "I am working now" (Pelevin 1998: 116).

Noting that attempts to discover the source of such constructions with the verb "to have" in Old and Middle Iranian, as well as in foreign languages, were unsuccessful, N. Nematollahi puts forward in the aforementioned article a hypothesis of their possible origin on the basis of the Persian language itself. And although this hypothesis is not without flaws and may raise questions (see below), the idea underlying it that these tenses were a consequence of the natural evolution of the lexical and grammatical means of the Persian language itself and the further development of their potential, represents a certain achievement. N. Nematollahi draws attention to the fact that a distinctive feature of these constructions, in addition to the use of a specific auxiliary verb "to have", is the use of personal forms of both the auxiliary and notional verb (both are inflected for person and number). This feature, this author believes, distinguishes them from the resultative-stative or perfective tenses, where only the auxiliary verb (or copula) get inflected,[16] and from constructions with an apocopated infinitive (Nematollahi 2014: 110).

[15] Cf. also Romaskevič (1945: 50).

[16] This, as has been noted earlier, is the difference between constructions with the verb "to have" in the Persian language and from progressive action forms similar in grammatical meaning in Dari and Tajiki.

N. Nematollahi puts forward a hypothesis about a possible connection between the progressive action forms and the so-called "Serial Verb Constructions" in which a sequence of two or more verbs is presented, functioning as a single predicate and describing a single event, specifically, combinations of the verb *bär-dåštän* with another, of which the first is something like a marker of the beginning of an action.[17] N. Nematollahi gives examples from the Persian language: *bär-dårim berävim…* "let us set off", lit. "take [and] go…", *in bud ke ånhå… bär-dåštänd telegråf kärdänd* "then they… set off", lit. "took [and] telegraphed", *be färängihå näguyid ke bär-dåränd… benevisänd*[18] "don't tell foreigners, or they (will) set off", lit., "take… write" etc. (Nematollahi 2014: 111-112). The above author suggests that in the process of grammaticalization of this serial verb construction with *bär-dåštän*, the latter could lose the preverb *bär-*, taking on the abbreviated form: *dåštän*. Its original semantics: an indication of an imminent action, was extended to other grammatical functions, more importantly denoting an ongoing action, which gradually became its main function (Nematollahi 2014: 112).

In the opinion of the present writer, at least two questions can be raised in connection with the described theory, which undoubtedly has the right to exist as a hypothesis, especially for lack of more convincing alternatives. Firstly, if the forms with *dåštän* go back to the serial verb construction with *bär-dåštän*, then how can one explain their use in those Iranian languages or dialects (for example, northwestern, see above), in which the verb *bär-dåštän* is absent or does not indicate an upcoming action? Secondly, among the Persian examples cited by N. Nematollahi, phrases with the verb *bär-dåštän* in the imperative and subjunctive moods are prevalent: *bär-dårim berävim…, bär-dåränd… benevisänd*. In light of the considered hypothesis, how can one explain the fact that the progressive action forms with *dåštän* are used exclusively in the indicative mood?

The main conclusions of the present study can be summarized as follows. The use of the analyzed tenses (constructions) with the verb *dåštän* "to have" in different dialectal and phonetic variants is an exceptional feature of the area of Persian and some other Western Iranian dialects, the distribution of which is limited to the territory of modern Iran. These temporal forms were a consequence of the natural evolution of the lexical and grammatical means of expression of the Persian language itself, and their origin cannot be explained by the "influence of foreign languages." As an integral and organic part of the Persian verb system, these constructions fill an important grammatical niche, i.e. a means of expressing an ongoing present or past

[17] Cf. similar combinations with the verb "take" in Russian: *ja vzjal i skazal…*, lit. "I took and said…"
[18] Transcription modified.

action. These forms are widespread within the entire geographical realm of Persian and other Western Iranian dialects of northern, central, southern and eastern Iran.

BIBLIOGRAPHY

Abrahamian, R. (1936), *Dialectes des Israélites de Hamadan et d'Ispahan et dialecte de Baba Tahir*, Paris: Librarie d'Amerique et d'Orient.

Ahangar, A. (2010), "A study of the verb system in the Sistani dialect of Persian", *Orientalia Suecana*, vol. 59, pp. 5-44.

Dehghān, I. (1972). "Dāshtan as an auxiliary in contemporary Persian", *Archiv Orientälni* (Praha), vol. 40, pp. 198–205.

Farhâdi, A.-Gh. (1955), *Le Persan Parlé en Afghanistan. Grammaire du Kâboli accompagnée d'un recueil de quatrains populaires de la Région de Kâbol*, Paris: Centre National de la Recherche Scientifique.

Ivanow, W. (1935), "The Gabri dialect spoken by the Zoroastrians of Persia", *Rivista Degli Studi Orientalni*, vol. 16, no. 1, pp. 31-97.

Ivanow, W. (1925), "Rustic Poetry in the Dialect of Khorasan", *Journal and Proceedings of the Asiatic Society of Bengal*, vol. 21, no. 3, pp. 233-313.

Jeremiás, É. (1993), "On the genesis of the periphrastic progressive in Iranian languages", in: *Medioiranica: Proceedings of the International Colloquium Organized by the Katholieke Universiteit Leuven, 21st–23rd May 1990,* ed. W. Skalmowski, and A. Van Tongerloo, Leuven: Peeters, pp. 99–116.

Nematollahi, N. (2014), "Development of the progressive construction in Modern Persian. In: *Proceedings of the 1st Conference on Central Eurasian Languages and Linguistics (ConCALL)*, ed. Özçelik, Ö., and Kent A., Bloomington: Center for the Languages of the Central Asian Region, pp. 102–114.

Nematollahi, N. (2018), "Have-progressive in Persian. A case of pattern replication?", *Diachronica*, vol. 35, no. 1, pp. 144-156. https://doi.org/10.1075/dia.17023.nem

Pejsikov, L. (1960), *Tegeranskij dialekt* [The Tehrani dialect]. Moscow: IMO. In Russian.

Pelevin, M. (1998), "Dialekt bandari Južnogo Irana" [The Bandari dialect from Southern Iran], *Strany i narody Vostoka*, vyp. XXX: Central'naja Azija. Vostočnyj Gindukuš. Pamjati Aleksandra Leonoviča Grjunberga [*Countries and peoples of the East*, vol. 30: Central Asia. Eastern Hindukush. In memory of Alexander Gruenberg], ed. I. M. Steblin-Kamensky, V. V. Kuščev, N. L. Lužeckaja, and L. Rzehak, Sankt-Peterburg: St. Petersburg for Oriental Studies, pp. 109-122. In Russian.

Pistoso, M. (1974), "L'ausiliare Dāštan in Neo-Persiano: Un Costrutto Linguistico Nord-Iranico?", *Oriente Moderno*, vol. 54, no. 4, pp. 157–162.

Rastorgueva, V., and Èdel'man, Dž. (1982), "Giljanskij, mazanderanskij (s dialektami šamerzadi i velatru)" [Gilaki, Mazandrani (including the Shamerzadi and Velatru dialects)], in: *Osnovy iranskogo jazykoznanija. Novoiranskie jazyki: Zapadnaja gruppa, prikaspijskie jazyki* [Basics of Iranian linguistics. New Iranian languages: the Western group, Caspian languages], ed. V. Abaev, M. Bogoljubov, and V. Rastorgueva, Moscow: Nauka, pp. 447-554. In Russian.

Rastorgueva, V, and Moškalo, V. (1997), "Dialekty Central'nogo Irana" [The Dialects of Central Iran], in: *Osnovy iranskogo jazykoznanija. Novoiranskie jazyki: Severo-zapadnaya gruppa. II* [Basics of Iranian linguistics. New Iranian languages: the Northwestern group. II], ed. V. Efimov, Moscow: Vostochnaya Literatura, pp. 195-329. In Russian.

Reżā'ī, J̌. (1998 / 1377), Barrasī-ye gūyeš-e Bīrǰand. Dastūr va vāžehnāmeh: Bīrǰandī be Fārsī [Study of the dialect of Bīrǰand. Grammar and Bīrǰandī-Persian dictionary], Tehrān: Hīrmand. In Persian.

Romaskevič, A. (1945), "Lar i ego dialect" [Lar and its dialect], in: *Iranskiye Yaziki* [Iranian languages], vol. 1. Iranica 3, ed. V. Abaev, and I. Meščaninov, Moscow - Leningrad: Izdatel'stvo Akademii Nauk SSSR, pp. 31-86. In Russian.

Rosenfeld, A. (1948), "Vspomogatel'naja funkcija glagola *dāštan* v persidskom jazyke" [The auxiliary function of the verb *dāštan* in the Persian language], in: *Sovetskoe Vostokovedenie* [Soviet Oriental Studies], vol 5, Moscow-Leningrad: Izdatel'stvo Akademii Nauk SSSR, pp. 305-312. In Russian.

Taheri-Ardali, M., Borjian, H., and Anonby, E. (2025), "The Bavadi and their Bakhtiari Dialect", *Iranian Studies*, pp. 1-47. https://doi.org/10.1017/irn.2025.2.

Thackston, W. (2009), "An Introduction to Persian". Revised fourth edition. Bethesda, Maryland, 2009.

Žukovskij, V. (1888), "Osobennoe značenie glagola (داشتن) v persidskom razgovornom jazyke" [Special meaning of the verb *dāštan* in the conversational Persian language], *Zapiski Vostočnogo Otdelenija Imperatorskogo Russkogo Arxeologičeskogo Obščestva* [Proceedings of the Oriental Department of the Imperial Russian Archaeological Society], vol. 3. St. Petersburg: Tipografija Imperatorskoj Akademii Nauk, pp. 376-377. In Russian.

Žukovskij, V. (1902), *Obrazcy narodnogo tvorčestva. Pesni pevcov-muzykantov, pesni svadebnye, pesni kolybel'nye, zagadki, obrazcy raznogo soderžanija* [Samples of [Persian] folklore: songs of singers-musicians, wedding songs, lullabies, riddles, samples of various content], St. Petersburg: Tipografija I. Boraganskogo & Ko. In Russian.

YEREVAN STATE UNIVERSITY PUBLISHING HOUSE

JIL 2/1 (2025): 48-71

Journal of Iranian Linguistics

Persian Borrowings in Kashmiri: A Descriptive and Analytical Study

Midhat Shah
Louisiana State University
College of Humanities & Social Sciences

doi.org/10.46991/jil/2025.01.04

Abstract: This research examines lexical and structural borrowings from Persian into Kashmiri, with a particular emphasis on nativization techniques at the morphological, semantic, and phonological levels. Addressing a gap in Persian-Kashmiri contact linguistics, this research investigates the integration of Persian loanwords and the resultant structural innovations in Kashmiri. With a framework that blends contact and historical linguistics, the study investigates linguistic limitations that affect the borrowing process. Employing corpus analysis with diachronic and synchronic approaches, this study contributes to the understanding of language contact phenomena, particularly Persian-Kashmiri interactions. The results shed light on the sociolinguistic and historical influences that have shaped the vocabulary and structure of Kashmiri, providing insights into more general trends of language borrowing and language change in contact circumstances.

Keywords: Borrowings, Nativization, Contact-Induced Change, Persian, Kashmiri.

Midhat Shah
E-mail: mshah5@lsu.edu
ORCID: 0009-0007-8712-8160
Received: 25.09.2024
Revised: 02.05.2025
Accepted: 16.06.2025

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

© Midhat Shah, 2025

1. Introduction

In the field of historical linguistics and language evolution, the phenomenon of linguistic borrowing—a result of language contact—has long been the focus of academic study. In particular, the mechanisms of nativization and the restrictions governing such language transfers are examined, as well as the complex processes of lexical and structural borrowing from Persian into Kashmiri.

The degree and kind of linguistic borrowing are strongly influenced by the duration and intensity of language contact. Despite the extensive historical contact between Persian and Kashmiri, there has been a relative lack of dedicated scholarship examining Persian-Kashmiri language contact phenomena, particularly when compared to the substantial body of research on Indo-Persian linguistic interactions more broadly. This scholarly gap is particularly notable given Kashmir's unique position as a region of sustained Persian cultural and political influence. The present study addresses this lacuna by providing a comprehensive analysis of borrowing mechanisms and nativization processes specific to the Persian-Kashmiri contact situation.

Crystal (1997) posits a correlation between the geographic or social proximity of languages and the extent of lexical borrowing, which he identifies as the primary vehicle of language transfer and change. The historical context of the Persian-Kashmiri contact, rooted in the political and cultural ties between Kashmir and Iran, which were established in the 14th century, providing a rich terrain for investigating these linguistic phenomena. However, the integration of foreign linguistic elements is not unrestricted; it is constrained by the structural properties and rules of the recipient language. Winford (2003: 59) aptly notes that "borrowings are subject to different kinds of changes, as they vie for a place in the lexicon with the native items that may already occupy similar space." This observation underscores the complex interplay between borrowed and native linguistic elements.

The dynamic nature of borrowings, from their initial entry into a language to their eventual integration, is a crucial aspect of this study. Kay (1995: 72) elucidates this process: "Borrowings are especially open to modification, both on entering the language and with time… the flexibility of form and meaning of borrowings enables them to adapt easily to the structure of the host language and current trends and needs." This adaptability of borrowings forms a central theme in my analysis.

The current study expands significantly upon the foundational scholarship on Persian-Kashmiri contact, including Koul's linguistic studies (1986, 2005), Ashok K. Koul's lexical documentation (2008), and the comprehensive survey in Encyclopaedia Iranica (Koul 2012). By utilizing a larger variety of data, my research seeks to offer a more thorough understanding of the handling of lexical and structural borrowings. The methodology used includes thorough use of lexicographical resources to verify accuracy, as well as interaction with native speakers of Kashmiri and Persian.

This article is structured as follows: Section 1.1 presents the theoretical framework and socio- historical context underpinning this research. Section

2 offers a detailed analysis of lexical borrowings, with particular attention to the processes of nativization. Section 2.2 examines the semantic shifts observed in lexical borrowings. Section 3 explores structural borrowings and resultant linguistic innovations. Finally, Section 4 provides a comprehensive discussion of my findings and concluding remarks.

Through this study, I aim to contribute to the broader understanding of language contact phenomena, particularly in the context of Persian influence on Kashmiri. By elucidating the mechanisms of borrowing and nativization, I seek to shed light on the dynamic processes that shape language evolution in situations of prolonged cultural and linguistic contact.

1.1. Literature Review

Previous scholarship specifically addressing Persian–Kashmiri language contact has been limited but foundational. Koul (1986) examined mechanisms of stylistic diffusion from Persian into Kashmiri, identifying key patterns of linguistic influence. His comprehensive study (Koul 2005) offered a broader morphological and syntactic analysis of Kashmiri, while Ashok K. Koul (2008) provided the most systematic documentation of lexical borrowings to date. The authoritative *Encyclopaedia Iranica* entry (Koul 2012) presents a detailed overview of Persian elements across multiple linguistic domains. However, none of these works employ the integrated theoretical framework combining contact and historical linguistics adopted in the present study, nor do they offer a systematic analysis of nativization strategies across phonological, morphological, and structural levels. The present study builds upon these foundational contributions while expanding both the analytical scope and theoretical grounding. This segment offers a critical analysis of prior scholarship and recent research relevant to the processes of language borrowing and contact-induced change.

1.1.1. Contact-Induced Change

The field of contact linguistics has thoroughly investigated the linguistic changes that happen when speakers of distinct languages interact with prolonged interaction. These alterations have fueled language evolution throughout history, appearing in a variety of linguistic domains. Aikhenvald (2006:20) offers a nuanced perspective on the outcomes of such contact, positing that contact-induced changes can "alter the grammatical structure of the language or preserve it, depending on whether they restructure the

system or borrow a term." This observation underscores the complex and multifaceted nature of linguistic changes resulting from language contact.

Winford (2003) demonstrates that language maintenance scenarios—where minority languages interact with dominant linguistic groups—typically result in lexical and structural borrowing patterns driven by prestige factors, as evidenced in historical cases like Norman French influence on Middle English.

Bloomfield (1933) first introduced the idea of prestige in linguistic borrowing, proposing that borrowing usually entails the transfer of linguistic components from a more prestigious language to a less prestigious one. This theoretical framework offers a useful perspective for analyzing the historical infiltration of Persian lexical elements into the Kashmiri language. The elevated status of Persian as the language of governance in the 11th century created conducive conditions for its integration into the Kashmiri lexicon.

Building on these fundamental ideas, this research investigates the particular processes and results of Persian-Kashmiri language interaction. I hope to advance knowledge of contact induced language change by investigating the lexical and structural borrowing processes in this environment, especially when languages of different social and political standing are involved in language maintenance scenarios.

1.1.2. Linguistic Borrowing

The concept of linguistic borrowing, while ubiquitous in language contact studies, lacks a universally accepted definition. Thomason and Kaufman (1992:37) offer a comprehensive description, characterizing borrowing as "the incorporation of foreign features into a group's native language by speakers of that language." They further elucidate the terminological distinctions, designating languages that provide borrowed features as 'source languages' and those that incorporate these features as 'recipients' or 'borrowing languages'.

Trask 1996 adopts a broader perspective, conceptualizing borrowing as the transference of any type of linguistic element between languages due to contact. This definition encompasses a wide array of linguistic phenomena, from lexical to structural borrowings. The literature frequently employs terms such as 'donor language' or 'source language' to denote the origin of borrowed elements, while 'receiving language' refers to the linguistic system adopting these new forms. The borrowed linguistic material itself is commonly referred to as a 'loan' or 'borrowing'.

This idea is further developed by Aikhenvald (2006), who points out that a wide range of linguistic material is transferable. Phonemes, lexemes, prefixes, verbal categories, and even numerals may be included, highlighting how commonplace language borrowing is in contact situations. The degree and results of borrowing might differ greatly. Thomason and Kaufman (1992:74–76) offer a sophisticated borrowing scale that distinguishes between various borrowing stages and the characteristics that go along with them. This scale starts at "casual contact," which is defined as having little to no lexical borrowing; it goes up to "slightly more intense contact," which includes some lexical and minor structural borrowing; and it ends at "very strong contact," where substantial typological changes in the receiving language result from extensive structural borrowing. The language interaction between Persian and Kashmiri speakers can be categorized as "more intense," evidenced by substantial lexical borrowings accompanied by a limited number of structural borrowings.

1.1.3. Nativization

The process of nativization occurs when borrowed linguistic components are incorporated into the structure of the receiving language. The process of 'nativization', which is an integration mechanism, functions on various levels, varying from simple phonological modifications to intricate structural changes. The complexity of nativization tends to increase in tandem with the hierarchy of borrowability, with structural nativization presenting the most significant challenges.

As native speakers adopt and employ foreign lexical items, the process of borrowing and subsequent nativization unfolds. Winford 2003: 59 observes that this process "not only adds to the lexicon but also enriches its potential for further expansion. Phonological and morphological developments may also be triggered by foreign additions to the vocabulary." This finding emphasizes how borrowed words and phrases have the power to initiate more extensive grammatical shifts in the target language.

While the maintained language frequently experiences significant lexical borrowing, structural changes are usually restricted in language maintenance contexts. This tendency may be seen in the Persian-Kashmiri contact scenario, where the latter's nativization at different linguistic levels is a result of adopting Persian loans, but structural changes are still relatively small.

The subsequent sections of this study will explore the multifaceted nature of nativization, examining its manifestations across phonological, lexical, morphological, and structural domains in the context of Persian borrowings in Kashmiri. This comprehensive analysis aims to contribute

to our understanding of the intricate processes involved in linguistic borrowing and nativization in situations of prolonged language contact.

1.2. Socio-Historical Context

Historical and sociopolitical variables have a fundamental influence on the complex interactions between languages. It is a well-established phenomenon that dominant languages have an impact on minority speakers, and literature frequently acts as a vehicle for language dissemination. The 12th-century development of the Islamic Empire in India led to a notable infusion of Perso-Arabic vocabulary features into the languages of North India. As a reflection of the complicated linguistic environment of the time, the term "Perso-Arabic" here refers to both words of direct Persian origin and those of Arabic provenance that were conveyed through Persian.

Historical evidence consistently demonstrates that political and commercial interactions have been pivotal in fostering cultural and linguistic exchanges across societal strata. Muslim rule had an especially profound impact on Northern India prior to partition, permanently altering the linguistic fabric of the area. As stated by Bamzai (1962), Kashmiri culture was significantly influenced by Arab and Persian cultures, and these influences may still be evident today. The most notable result of this effect was the Kashmiri language and literature's Persian ancestry.

Kashmir and Iran established cultural and political ties in the 14th century, which facilitated the introduction of Persian lexical elements into Kashmiri linguistics. It is important to remember that although Persian borrowings make up the majority of borrowings, a substantial number of borrowed items trace their etymology to 9th-century Arabia, entering Kashmir via Persian. Examples include *nūr* 'light', '*ālam* 'universe', and *haqq* 'right', underscoring the complex trajectories of lexical borrowing. As documented extensively in the *Encyclopaedia Iranica* (Koul 2012), Persian elements permeated multiple domains of Kashmiri language and culture, with borrowings spanning religious, administrative, and literary vocabularies. Koul's (1986) analysis of style diffusion reveals that this influence operated not merely at the lexical level but also affected register and stylistic conventions.

Historically interconnected Arab and Iranian cultures, forged by political alliances, provide essential context for understanding these linguistic phenomena. Rezakhani (2011) reports that a synthesis of Arabic and Persian cultural and linguistic aspects was sparked by the Arab conquest of Iran in 651 CE, which came about as a result of the collapse of the

Sassanid Empire. The linguistic advancements that followed in Northern India were made possible by this historical event.

During the 11th century, when Muslims ruled a large portion of Northern India, there was a notable infusion of Perso-Arabic terminology into native languages. The long reign of this language left its mark on the lexicons of several Indian languages, to differing degrees. According to Safavi (2006), all Indian languages, such as Bengali, Punjabi, Marathi, Gujarati, Telugu, Hindi, and Kashmiri, possess a large number of Persian words and phrases without any native equivalents.

This historical synopsis offers an essential foundation for comprehending the sociolinguistic processes that enabled the widespread borrowing of Persian into the Kashmiri language. It emphasizes how crucial it is to take into account larger historical and cultural contexts when analyzing language occurrences, especially in areas with intricate histories of political upheaval and cross-cultural interaction.

1.3. Language Ideology and Diglossia

Ferguson (1959) introduced the idea of diglossia, which is the coexistence of high and low varieties of a language, each used in different formal and informal contexts. An essential framework for comprehending the processes of language contact and borrowing in multilingualcultures is provided by this sociolinguistic phenomenon.

According to Aikhenvald 2002:13, circumstances involving "*one-to-one language contact with diglossia may lead to a considerable influx of loanwords*" from the language that is socially or politically dominant. On the other hand, situations where there aren't any power differentials during language contact typically lead to *"mutual adjustment of structure"*. This theoretical framework provides insightful information about the type and degree of language borrowing in many contexts.

The diglossic link is especially important in the context of Persian-Kashmiri language interaction. Persian was the language of the state and was used primarily by the privileged strata of society. It was the principal language of court, government, and poetry. The lexical borrowing and phonological adaptation patterns were significantly impacted by this linguistic stratification.

One significant feature of this diglossic environment is the establishment of two distinct pronunciation patterns for borrowed words. Proficient in Persian and/or Urdu, the upper classes of society have mostly preserved the original pronunciation of borrowed phrases, especially in formal settings.

In contrast, the general populace has adopted nativized pronunciations, especially in informal settings. This phonological divergence not only reflects the sociolinguistic stratification of the speech community but also provides a fascinating case study of the processes of linguistic adaptation and nativization.

Dual pronunciation patterns present a rich field for sociolinguistic research, with the ability to clarify the intricate relationships between language ideology, social status, and linguistic practice. It emphasizes how diverse language contact outcomes may be, with borrowed pieces undergoing varying integration processes based on usage settings and sociocultural factors. Thus, the Persian-Kashmiri case illustrates how diglossic circumstances, influenced by social, political, and cultural forces, can result in complex patterns of language borrowing and adaptation. This viewpoint improves our comprehension of the intricate dynamics present in language contact scenarios, especially when there is a pronounced power imbalance between the interacting linguistic communities.

2. Lexical Borrowing: Processes and Nativization Strategies

Lexical borrowing is a common occurrence in cross-linguistic interaction that can take many different forms and occur in a variety of circumstances. This process frequently requires intricate nativization mechanisms in the target language, far from being a straightforward transference of lexical elements. Given its substantial borrowings from Persian in a variety of semantic areas, such as lexicons related to religion, administration, and law, the Kashmiri language offers an excellent case study for researching these nativization processes.

2.1. Nativization Strategies: Theoretical Framework and Mechanisms

The various approaches taken to nativize borrowed words highlight how complex language adaptation is in contact situations. These procedures help incorporate new terms more easily, but they may also cause pre-existing lexical paradigms to change in response to language influences from different cultures. As Haugen (1953:388) astutely observes, "every loan is part importation and part substitution," highlighting the dual nature of lexical borrowing as both an additive and transformative process.

2.1.1. Phonological Nativization

A key component of linguistic nativization is the phonological modification of borrowed lexemes. Usually, this technique entails replacing foreign

phonemes with the closest native equivalents. Hyman 1970 proposes a model of phonological borrowing based on phonemic comparison, positing that a language first finds the closest phonemes that encompass the phonetic quality of the foreign sound, and then the foreign sound is approximately phonemicized and subjected to the phonological constraints of that phoneme's members. Since they are accustomed to the phonotactic restrictions of their language, native speakers quickly modify borrowed words to fit these pre-established patterns. As part of this adaptation, the borrowed lexeme may be inserted, deleted, or substituted to align with the recipient's syllable structure and phonological rules.

In the context of Persian loanwords in Kashmiri, several distinctive nativization strategies are observable.

2.1.1.1. Vowels

There are a number of vowels that were nativized in Kashmiri, usually by an alternate vowel or vowel insertion. Below are some examples of how Persian vowels are nativized in Kashmiri.

2.1.1.1.1. Vowel alternations

The close-back rounded vowel *u* is replaced by the open-mid back rounded vowel *ɔ*: As indicated in the examples in (1) below:

(1) ***u* > *ɔ***

gulāb > *gɔlāb* 'rose',

ruxsār > *rɔkhsār* 'cheeks',

ruxsat > *rɔkhsat* 'leave/discharge'.

In the above examples, the velar fricative *x* is replaced by the velar stop *kh*. This change will be discussed in the consonant change below.

The open central unrounded vowel ā is replaced by the mid-central vowel *ə̄* (schwa but only when the following consonant is followed by the close front unrounded vowel i as indicated by examples in (2):

(2) ***ā* > *ə̄***

tājir > *tə̄jir* 'trader',

ǰāri > *ǰə̄ri* 'running/ongoing',

bāzi > *bə̄zi* 'play',

ɣāliban > *gə̄liban* 'probably'.

Comparing the borrowed word *tāb* 'fortitude/resistance', which does not change to *tə̄b*, its vowel remains the same. Also, in the word *gulāb* 'rose,' which changes to *gɔlāb* only the vowel in the first syllable changes. The open central unrounded vowel *ā* remains unchanged since the following consonant *b* is not followed by the vowel *i*. However, when the same noun *gulāb* is changed into the adjective *gulābi* 'rosy', its nativized form *gɔlāb* changes into *gɔlə̄bi,* the vowel *ā* changes to *ə̄,* because the final consonant is followed by the vowel *i*.

Kashmiri does not allow vowels *i, e, and u* in the initial position, so borrowings that start with these vowels are usually replaced. For example, when the close front unrounded vowel *i* is in the initial position, it is replaced by the voiced palatal semivowel *y,* as indicated by examples in (3):

(3) ***i* > *y***

insāf > *yinsālf* ~ *insāf* 'justice',

isteqbaal > *yistekbāll* ~ *istekbāl* 'welcome',

imkān > *yimkān* ~*imkān* 'possibility'.

Scholarly investigation is warranted due to the complex phenomenon of lexical borrowing between Kashmiri and Persian in the sociolinguistic context. The appearance of two distinct pronunciation patterns for borrowed lexemes is a prominent aspect of this linguistic situation. In phonological realization, this duality appears as a socially stratified dichotomy. The vast majority use nativized pronunciations, which have been phonologically adjusted to conform to Kashmiri phonotactic restrictions. Conversely, the upper classes of society, especially those who speak Persian and/or Urdu fluently, tend to retain pronunciations that are closer to the original Persian forms. This split in pronunciation reveals deeper sociolinguistic dynamics rather than being just a result of linguistic variety. The preservation of more "authentic" Persian pronunciations among the elite echelons of society can be interpreted through the lens of language ideology and prestige. Maintaining these pronunciations, particularly in formal settings, probably acts as a sociolinguistic marker, possibly indicating cultural sophistication, social standing, or education.

This tendency is consistent with more comprehensive theories of language ideology and sociolinguistic prestige. The elite pronunciation may have overt prestige in the speech community because of its ties to speakers of the historically dominant language. On the other hand, the nativized forms are more frequently linked to informal settings and the general public, while potentially carrying covert prestige as markers of local identity.

This scenario offers an intriguing case study on the relationship between linguistic ideology, social stratification, and language contact. It emphasizes

how intricately historical power relations, social standing, and phonological adaptation interact when there is a lot of lexical borrowing.

2.1.1.1.2. Vowel insertions

Instances of vowel insertions are also observed. For example, a vowel *ɨ* is added to break consonant clusters, as indicated by examples in (4):

(4) *darj* > *darɨj* 'record',
sabr > *sabɨr* 'patience',
garm > *garɨm* 'hot'.

2.1.1.2. Consonants

Consonants were mostly nativized by substituting them with native equivalents and deletion. Given are some of the examples of consonantal changes.

2.1.1.2.1. Consonantal Changes

The process of nativizing Persian loanwords into the Kashmiri language entails the methodical replacement of foreign stops and fricatives with their corresponding Kashmiri equivalents. This is what Hock 1991 refers to as a "system-based pattern of substitution." This procedure exhibits a dependable, consistent method for modifying foreign sounds while taking into account the phonological restrictions of the target language. These replacements adhere to observable patterns rather than being random, preserving Kashmiri's phonological integrity while adding new vocabulary. This methodical adaptation advances our knowledge of general principles in language borrowing and phonological change and offers insightful information about the phonological linkages between Persian and Kashmiri. For instance, Kashmiri adopts different stops and fricatives from Persian:

stops	ʔ →	ə
	q →	k
fricatives	x →	kh
	ɣ →	g
	f →	ph

The Persian glottal stop ʔ in the initial position changes to the mid-central vowel *ə*, as indicated by examples in (5):

(5) **ʔ > ə**

ʔitir > *ətir* 'scent/perfume',
ʔišq > *əšik* 'love',
ʔilm > *əlim* 'knowledge'.

Uvular stop *q* changes to velar stop *k*, as indicated by examples in (6):

(6) ***q* > *k***

qad > *kad* 'height',
qarār > *karār* 'rest',
yaqīn > *yakīn* 'belief'.

The voiceless velar fricative *x* is replaced by the velar stop *kh* as indicated by examples in (7):

(7) ***x* > *kh***

xarāb > *kharāb* 'bad',
xām > *khām* 'raw',
xoš > *khoš* 'happy',
xās > *khās* 'special'.

The voiced velar fricative *ɣ* changes to the voiced velar stop *g*, as indicated by examples in (8):

(8) ***ɣ* > *g***

ɣāliban > *gəliban* 'probably',
ɣaflat > *gaflat* 'heedlessness'

The labiodental fricative *f* changes to the voiced bilabial stop *ph* as indicated by examples in (9):

(9) ***f* > *ph***

fasād > *phasād* ~ *fasād* 'corruption',
fan > *phan* ~ *fan* 'abuse/curse',
fida > *phida* ~ *fida* 'sacrifice,'
farzi > *pharzi*~ *farzi* 'fake/false'

Koul and Kashi (1997: 295) observe a noteworthy phonological phenomenon in Kashmiri: the presence of a non-native sound, presumably [f], primarily in loanwords. This sound exhibits considerable variability in its realization,

alternating with the aspirated bilabial stop [pʰ] in many speakers' pronunciations. Furthermore, in formal or affected speech, it freely varies with [f]. This pattern of phonological adaptation and variation illustrates the complex interplay between native phonemic inventories and borrowed sounds. It emphasizes how phonological integration is dynamic in language contact scenarios, preserving the phonology of the source language while also adjusting it to the phonological system of the target language. This phenomenon offers significant fresh insights into the sociolinguistic variables affecting phoneme realization in borrowed lexemes as well as the processes of phonological nativization.

2.1.1.2.2. Consonantal Deletions

There are several borrowed words that undergo cluster reduction through deletions at various positions. For example, the dental unaspirated stop *t* is lost when it is preceded by alveolar fricative *s* or palatal *š*, as indicated by examples in (10):

(10) *dast* > *das* 'hand',
ābgošt > *ābgoš* 'meat cuisine',
šikast > *šikas* 'defeat'.

2.1.2. Morphological Nativization

The morphological nativization of Persian loanwords in Kashmiri presents a compelling case of linguistic integration, manifesting in both inflectional and derivational processes. This phenomenon demonstrates the adaptability of borrowed lexemes to the grammatical structure of the recipient language.

2.1.2.1. Inflectional Nativization

Inflectional nativization is particularly noteworthy, as Persian loanwords readily adopt Kashmiri inflectional morphology. Koul (1977) provides illuminating examples of this process. The loanword *batuk* 'duck' (from Persian *batax*) exemplifies this adaptation, undergoing pluralization (*batuk* M.Sg 'duck' > *batak* M.PL 'ducks'; *batič* F.Sg > *batiči* F.PL) and gender marking (*batuk* M, *batič* F) in accordance with Kashmiri morphological patterns. This pattern mirrors that of native Kashmiri words, as evidenced by the comparable inflection of *gagur* 'rat'. Moreover, Persian stems demonstrate compatibility with Kashmiri case marking systems. The dative case marker *-as* attaches to both borrowed and native stems (e.g., *batk-as*

'to the duck', *gagar-as* 'to the rat'), illustrating the seamless integration of loanwords into Kashmiri's case system.

Koul (1977) further elucidates this phenomenon, noting that the locative case marker *-an* readily affixes to Persian stems (*sahar* 'morning' > *sahran* 'in the morning'). Additionally, the feminine dative marker *-i* exhibit productivity in deriving adjectives from Persian nouns (*kitāb* 'book' > *kitābi* 'bookish'). This morphological integration aligns with broader patterns documented in Ashok K. Koul's (2008) comprehensive survey of Kashmiri lexical borrowings, though the present analysis extends beyond mere documentation to examine underlying nativization mechanisms.

This comprehensive morphological integration of Persian loanwords into Kashmiri's inflectional system underscores the language's capacity to assimilate foreign lexical items fully into its grammatical structure. It provides valuable insights into the mechanisms of linguistic borrowing and the adaptability of grammatical systems in language contact situations.

2.1.2.2. Derivational Nativization

There are not many examples of native Kashmiri prefixes attaching to a Persian stem because there are not many prefixes in Kashmiri itself, however, suffixes are common. The derivational suffix *-lad,* a nominalizer, has been extended from Kashmiri words like *phakhlad.* Below are some examples:

(11) Persian stem + Kashmiri suffix

Persian stem	***Gloss***	***Nativized form***	***Derived form***	***Gloss***
šikast	failure	*šikas*	*šikaslad*	someone who is destitute
sil	tuberculosis	*silɨ*	*silɨlad*	someone with tuberculosis
xasār	loss/damage	*xasāri*	*xasārilad*	something that is damaged

The morphological integration of Persian elements, while extensive, represents only one dimension of lexical adaptation. Equally significant are the semantic modifications that occur during the borrowing process. The phonological and morphological integration of Persian elements, while extensive, represents only one dimension of lexical adaptation. Equally significant are the semantic modifications that occur during the borrowing process.

2.2 Semantic Change in Lexical Borrowings

A plethora of words that were borrowed with no change in meaning, ranging from fruits like *gilās* 'cherry' and', *hendevane* 'watermelon' to body parts like *rag* 'blood vessel', *rān 'thigh',* and', *šooš* 'lung' to proper nouns like '*dilshad*', '*feroz', 'shams',* and *'aftāb'.* Besides phonological and morphological nativization, many borrowed words underwent semantic changes as well. In some cases, the meanings of borrowed words were extended to additional domains, like the Persian word *dur* which means 'pearl' or 'pearl earrings', extended to mean 'any kind of earring' in Kashmiri *duur*. Some words underwent specialization, i.e., the meaning that erstwhile covered a large sense becomes narrow. For example, the Persian word *bāng* means 'a loud cry', however, in Kashmiri, it is only used for 'the cry of a rooster' or 'the call to prayer of a *muezzin*' (person who gives the call to prayer at a mosque).
There have been numerous instances in Kashmiri poetry where Iranian lexical items have been used, mostly in the form of metaphors. Persian metaphors like *sarv,* literally the name of a cypress tree, is used to praise the height of someone; *durdān* 'pearl' for teeth, *sarv-e ravān* 'walking cypress' for praising gait, and *tīr-e misgān* 'arrows of eyelashes' for praising eyelashes. The last two are structurally new and different given the Kashmiri structure. There is a profusion of such Persian phrases in Kashmiri that were transported and used as frozen phrases. Due to structural constraints, they were not extended to native Kashmiri words, which will be discussed in the following section.

3. Structural Borrowing and Innovations

The phonological, morphological, and syntactic systems of the receiving language can all be strongly impacted by the phenomena of structural borrowing in language contact situations. Although there has been significant phonological nativization of Persian lexical borrowings into Kashmiri, the process has not been consistent throughout the speech community. Alternative pronunciations continue to exist, especially among the upper classes of society, as demonstrated by earlier examples, most notably (3) and (9). The intricate sociolinguistic dynamics at work in the context of contact between Persian and Kashmiri are reflected in this linguistic stratification. In-depth analyses of the ways in which this linguistic interaction has shaped the morphological and syntactic structures of Kashmiri will be presented in the following sections, offering insights into the wider consequences of language contact on structural linguistic aspects.

3.1 Morphological Borrowings

One important component of structural borrowing in this language contact scenario is the integration of Persian derivational morphology into Kashmiri. The morphological system of Kashmiri has easily absorbed Persian derivational affixes, although inflectional morphemes usually do not transfer across languages. Interestingly, these borrowed affixes have proven to be remarkably productive, combining readily with native Kashmiri stems and going beyond their initial domain of Persian loanwords. The creation of new vocabulary forms has been made possible by this process, increasing the Kashmiri language's expressive potential in formerly unattainable ways. The phenomenon illustrates the dynamic nature of linguistic borrowing, where structural elements can be abstracted from their source and generalized within the recipient language's grammatical system. The following examples showcase the application of Persian derivational prefixes and suffixes to Kashmiri words, highlighting the creative linguistic innovations arising from this cross-linguistic influence.

1. Prefixation

(12) The prefix bi- 'without' is extended to native Kashmiri stems, mostly nouns and adjectives, generating words like:

Prefix (Persian) + stem (Kashmiri)

bi- + X

Kashmiri stem	***Gloss (stem)***	***Derived form***	***Gloss (derived form)***
šūb	etiquette	be-šūb	without etiquette, uncouth
ḍool	shape	be-dool	without shape
zuv	life	be-zuv	without life

The *bi-* prefix (meaning 'without') is extended to native stems but was originally borrowed from Persian into Kashmiri, as indicated by words like *biqarār* 'restless' and *biadab* 'mannerless'. However, it is important to note here that the *bi-* prefix is itself phonologically nativized in Kashmiri, changing the **c**lose-front unrounded vowel /i/ to the close-mid front unrounded vowel /e/. Many Kashmiri stems readily attach to the bi- prefix, which is the most productive. It is primarily an adjective-forming morpheme that attaches to nouns or adjectives.

2. Suffixation

Many suffixes are borrowed from Persian and are used with Kashmiri stems. For example, the suffix -*dār* (to hold, to bear) is very productive, and like other derivational affixes, it was also imported with Persian words like *māldār* 'wealthy', *tabardār* 'axe bearer', *dokāndār* 'shopkeeper', etc. The suffix *dār* is extended and is used as a nominalizer with native Kashmiri stems, mostly nouns and adjectives as illustrated in example (13):

(13) Kashmiri stem + Persian suffix:

[X] = (Kashmiri stem) + [-dār] → derived forms

	Kashmiri stems			***Derived forms***	***Gloss***
	gāš			**gāšdār**	something that bears light
[**X**] =	**gal**	+	**[-dār]**	**galdār**	someone who owns a grocery store
	lāg			**lāgdār**	something useful

3.3 Syntax

The syntactic borrowing from Persian into Kashmiri presents a compelling case of structural influence in language contact. Kashmiri typically follows a head-final structure in noun phrases, with determiners and modifiers preceding the noun and possessor pronouns in phrase-initial position. In Persian, on the other hand, possessor pronouns come last in phrases. Despite these basic structural distinctions, several Persian phrases that deviate greatly from Kashmiri's natural syntactic patterns have been included. This phenomenon challenges beliefs on the immutability of fundamental syntactic features by illuminating the possibility of deep structural borrowing under conditions of intense language contact. These Persian-style sentences have been successfully incorporated into Kashmiri, providing important new insights into the adaptability of syntactic systems and the formation of hybrid grammatical structures in multilingual settings. As an illustration,

(14) Kashmiri-

NP = [[DET (Possessor Pronoun)] + [Head Noun]]

K:	*mēn*	*kitāb*
	my.F.SG	book (N)
	'my book'	

K:	*m'ōn*	*ləḳdi*
	my.M.SG	boy (N)
	'my boy'	

(15) Persian-

NP = [[Head Noun-GEN] + [DET (Possessor Pronoun)]]

Pers:	*kitāb-e*	*man*
	book- EZ	my.SG
	'my book'	

Pers:	*pisar-e*	*man*
	son-EZ	my.SG
	'my son'	

In Kashmiri, the noun *kitāb* 'book' comes after the determiner *mēn* 'my'*;* however, in Persian, the noun *kitāb* comes before the determiner *man.* Similarly, when adjectives act as modifiers:

(16) Kashmiri-

NP = [[DET (Adjective)] + [Head Noun]]

K:	*jān*	*kitāb*
	good (Adj)	book(N)
	'good book'	

K:	*asl*	*ləḳdi*
	good (Adj)	boy(N)
	'good boy'	

(17) Persian-

NP = [[Head Noun-GEN] + [DET (Adjective)]]

Pers: *kitāb-e* *xūb*
book-EZ good (Adj)
'good book'

Pers: *pisar-e* *man*
boy-EZ my.SG
'my boy'

The syntactic divergence between Kashmiri and Persian is further exemplified in adjectival modification structures. In Kashmiri, adjectives precede the nouns they modify, as in *jān kitāb* 'good book'. Conversely, Persian employs a post-nominal adjectival placement, exemplified by *kitab-e xūb* 'good book'.

Persian syntax is characterized by the ezafe construction, which is denoted by the connecting element (-e). This form expresses possession and modifies the adjectival function, among other grammatical functions. Similar in function to a genitive case marker, the ezafe establishes a structural paradigm that is substantially different from the original syntax of Kashmiri.

One significant example of syntactic borrowing is the incorporation of Persian words using the ezafe construction into Kashmiri. The incorporation of Persian ezafe constructions into Kashmiri represents a form of lexicalized borrowing rather than productive syntactic innovation. These Persian-derived phrases appear to function primarily as fixed literary expressions, preserved in their original structural form without extending to broader syntactic productivity in Kashmiri. Even when comparable native Kashmiri structures already exist, they are nonetheless adopted, even though they are now mostly obsolete.

Examples (20-22) illustrate this phenomenon, juxtaposing borrowed Persian constructions with their native Kashmiri counterparts. The fact that these borrowed structures continue to exist while their native counterparts are becoming obsolete highlights the significant influence of Persian on Kashmiri grammar.

(18) Persian structural borrowing:

xūn *-e* *ǰigar*
blood(N) -EZ liver (N) [NP]

'blood of liver'

Native Kashmiri form:

ǰigar	*-uk*	*khūn*
liver(N)	[NP] -GEN	blood (N)

'blood of liver'

(19) Pers:

čašm	*-e*	*tar*
eye(N)	EZ	wet (Adj)[NP]

'wet eye'

K:

tar	*čašm*
wet (Adj)[NP]	eye(N)

'wet eye'

(20) Pers:

āb	*-e*	*hayāt*
water (N)	EZ	life [NP]

'water of life/life force'

K:

hayāt	*-uk*	*āb*
life (N) [NP]	EZ	water (N)

'water of life/life force'

The phenomenon of syntactic divergence between native Kashmiri structures and borrowed Persian phrases is noteworthy in this context of language contact. Even though the Kashmiri forms shown in (20)–(22) are grammatically correct, these native structures are now largely obsolete. On the other hand, Persian-derived structures are increasingly prevalent in formal writing, such as poetry and prose, as well as informal speech.

The key structural difference lies in the position of the head noun. In the borrowed Persian phrases, the head noun occupies the phrase-initial position, followed by its modifier. This arrangement contrasts sharply with the native Kashmiri syntax, where the head noun is phrase- final, preceded by its modifier.

These borrowed phrases frequently possess metaphorical meanings, as demonstrated by *xūn-e ǰigar* (literally 'blood of liver'), which is used to convey apprehension. This metaphorical use extends to sentential constructions, such as *me gav xūn-e ǰigar* 'I am anxious/worried'.

It is important to note that even though these Persian grammatical constructs have been extensively appropriated, their impact is limited to particular phrases. The structural pattern has not been expanded to become an efficient feature in Kashmiri syntax. Instead of a more extensive syntactic shift, this restricted borrowing points to a type of lexicalized syntactic borrowing.

It is noteworthy that a number of the lexical components of these phrases, such as *xūn* 'blood', *ǰigar* 'liver', *čašm* 'eye', *āb* 'water', and *hayāt* 'life', have also been independently borrowed and nativized in Kashmiri. This dual process reveals that while individual Persian lexemes undergo full nativization, ezafe constructions remain as frozen literary expressions, suggesting that structural borrowing in Persian-Kashmiri contact is lexically constrained rather than syntactically productive.

4. Discussion and Conclusion

The analysis of Persian-Kashmiri linguistic contact reveals a complex interplay of borrowing phenomena, demonstrating the multifaceted nature of lexical and structural adaptations. Empirical data indicate substantial phonological modifications in borrowed lexemes, encompassing processes of substitution, insertion, and deletion. The Kashmiri aspiration rule, exemplified in (11), where the dental unaspirated stop /t/ is elided in word-final position but retained medially, illustrates language-specific phonological constraints governing loanword integration.

A notable phonological tendency in Kashmiri involves the aspiration of word-final unaspirated consonants in loanwords, a phenomenon observed across Hindi-Urdu and Persian borrowings. In Kashmiri all consonants can occur in word-final position; however, Koul and Kashi (1997:298) point out that "as a general rule unaspirated word-final consonants are not tolerated. This seems to hold in many Hindi-Urdu, as well as Persian-borrowed words, as final deaspirates from these languages are aspirated in Kashmiri." This is evidenced by transformations such as /t/ > /th/ (e.g., *jannat* > *jannath* 'paradise') and /k/ > /kh/ (e.g., *paak* > *paakh*). As observed in section (2.1.1.2.1), the substitution of Persian-borrowed foreign stops and fricatives by perceived Kashmiri equivalents follows what Hock (1991) calls the system-based pattern of substitution common in borrowings. As illustrated in section (2.1.2), morphological nativization of Persian lexical borrowings at both inflectional and derivational levels is very productive and faces no major constraints.

Kashmiri remained resistant to structural borrowings of phrases but has freely borrowed derivational morphemes. A lot of borrowing results in inventions or the creation of new words that combine materials from the

source language and the receiving language. The borrowings from Persian also had an impact on Kashmiri morphology, particularly on the derivational morphology. Several derivational affixes were imported along with Persian words, and several of them were extended and used with native stems. As noted by Winford (2003: 62), "*Speakers do not have access to a list or a complete inventory of morphemes from which they can select. For instance, when Middle English speakers borrowed derivational morphemes like -tion, -able, etc., from French, it certainly wasn't because they isolated them in relevant French words and imported them independently of the stem to which they were attached. Nearly all the borrowed affixes were introduced with loanwords and were soon extended to native stems.*" Similarly, Persian morphemes used in Kashmiri were first imported with relevant Persian words and then used with native Kashmiri stems. It has been found that most nativization has been through analogical change: morphemes or stems have been changed to represent another item on the basis of analogy or apparent similarity. Talking of morphology, there is a general agreement among scholars that inflectional morphemes are notoriously difficult to borrow. According to Winford (2003:62), "… *morphological adaptation can be more difficult, especially if the recipient language has complex rules involving case, number, and gender.*" To illustrate, an example can be cited here: the pluralizing suffix in Persian is -hā, used as *bače* 'kid' > *bačehā* 'kids. The same pluralizing form is sometimes used in Kashmiri stems but is not very productive because Kashmiri masculine plurals are formed by palatalization and/or vowel change, e.g., *koṭ* 'boy' → *kəṭ* 'boys'. Feminine plurals are typically formed by adding the suffix *-i/-i* with a vowel and consonantal changes, e.g., *nər* 'arm'→ *nari* 'arms'. From the perspective of morphology, it is clear that derivational morphemes are more likely to be borrowed than inflectional ones.

Structural borrowing through syntax faces constraints due to the difference in the word order of the two languages. As illustrated in section (3.3), the Persian head noun is the initial phrase, and the modifier follows it. In Kashmiri, however, the head noun is phrase final, and the modifier comes before the noun.

As discussed above, borrowing due to contact encompasses not just the direct transfer of words but a range of other factors leading to innovations in the receiving language. From the discussion above it is evident that Persian had a profound effect on the Kashmiri language. Various Persian loans were used with little to no linguistic change, whereas others were nativized so much that they became indistinguishable from their Kashmiri counterparts. Kashmiri follows the same rules of borrowing and nativization in terms of phonology and morphology. However, it can be

said that due to the level of contact between the two languages and, more importantly, due to the grammatical particle *ezafe* found in the Persian language, structural borrowing faces constraints. Moreover, it is important to mention here that according to Musyken (1981), adjectives rank next to nouns on the borrowing scale. However, due to the ezafe construction in Persian, the borrowing of adjectives faces a challenge. Due to this, far fewer base adjectives have been borrowed from Persian than derived adjectives. The integration of borrowings led to changes at different levels, such as phonological adjustments in the pronunciation of the new items, morphological adjustments, and syntax. According to Winford (2010: 50), "the various types of integration demonstrate that so-called "borrowing" involves complex patterns of lexical change that create new lexical entries or modify existing ones in response to culture contact." It is observed here that the degree of structural and lexical borrowing does not always correlate. In cases of language maintenance, lexical borrowing serves as a vehicle for structural innovations in the receiving language, especially in derivational morphology. However, the recipient language remains highly resistant to foreign structural interference. This study's original contribution lies in its systematic documentation of nativization strategies across multiple linguistic levels, revealing that while Kashmiri readily adapts Persian morphological elements and syntactic borrowings remains constrained to lexicalized phrases. These findings build upon and extend the foundational documentation provided by Ashok K. Koul (2008) and the analytical frameworks established in earlier studies (Koul 1986, 2005, 2012), while offering new theoretical insights into the selectivity of borrowing processes and the constraints governing structural adaptation.

BIBLIOGRAPHY

Aikhenvald, A. Y. (2006), "Grammars in contact: A cross-linguistic perspective", *Grammars in Contact: A Cross-Linguistic Typology,* ed. A. Y. Aikhenvald, and R. M. W. Dixon, Oxford: Oxford University Press, pp. 1-66.

Bamzai, P. N. K. (1962), *A history of Kashmir: political, social, cultural: from the earliest times to the present day*, New Delhi: Metropolitan Book Company.

Bloomfield, L. (2015), *Language* (Routledge Revivals reprint of the 1935 edition), London: Routledge.

Crystal, D. (1997), *The Cambridge Encyclopedia of Language,* 2nd ed., Cambridge: Cambridge University Press.

Grierson, G. A. (1932). *A dictionary of the Kashmiri language*, Calcutta: Asiatic Society of Bengal.

Haugen, E. (1950), "The analysis of linguistic borrowing", *Language,* vol. 26, no. 2, pp. 210–231.

Hock, H. H. (1991), *Principles of Historical Linguistics*, Berlin: Mouton de Gruyter.

Hyman, L. M. (1970), "The role of borrowing in the justification of grammars", *Studies in African Linguistics*, vol. 1, no. 1, pp. 1-48.

Kay, G. (1995), "English loanwords in Japanese", *World Englishes*, vol. 14, no. 1, pp. 67–76.

Koul, O. N. (1977), *Linguistic Studies in Kashmiri.* New Delhi: Bahri Publications.

Koul, O. N. (1986), "Style Diffusion: Influence of Persian on Kashmiri", in: *Language, Style and Discourse*, ed. O. N. Koul, New Delhi: Bahri Publications.

Koul, Ashok (2008), *Lexical Borrowings in Kashmiri*, New Delhi: Indian Institute of Language Studies.

Koul, O. N. (2012), "Kashmir IV. Persian Elements in Kashmiri", *Encyclopaedia Iranica*, vol. XVI, fasc. 1, pp. 56–61.

Koul, O. N., and W. Kashi (2006), *Modern Kashmiri Grammar*, Hayettsville: Dunwoody Press.

Munshi, S. (2010), "Contact-induced language change in a trilingual context: The case of Burushaski in Srinagar", *Diachronica*, vol. 27, no. 1, pp. 32–72. https://doi.org/10.1075/dia.27.1.02mun.

Muysken, P. (1981), "Halfway between Quechua and Spanish: The case for relexification", in: *Historicity and Variation in Creole Studies*, ed. A. Highfield, and A. Valdman, Ann Arbor: Karoma Publishers, pp. 52–78.

Rezakhani, Kh. (2011), [Review of *Decline and Fall of the Sasanian Empire: The Sasanian–Parthian Confederacy and the Arab Conquest of Iran*, by P. Pourshariati], *Iranian Studies*, vol, 44, no. 3, pp. 415–419.

Thomason, S. G., and Kaufman, T. (1992), *Language Contact, Creolization, and Genetic Linguistics*, 1st pub. in 1988, Berkeley: University of California Press.

Thomason, S. G. (2001), *Language Contact: An Introduction*, Washington, DC: Georgetown University Press.

Trask, R. L. (1996), *Historical Linguistics*, London: Hodder Education Publishers.

Weinreich, U. (1953), *Languages in Contact: Findings and Problems*, New York: Linguistic Circle of New York.

Winford, D. (2003), *An Introduction to Contact Linguistics,* Malden, MA: Blackwell Publishing.

Winford, Donald (2010), "Contact and Borrowing", in: *The Handbook of Language Contact*, ed. R. Hickey, Chichester: Wiley-Blackwell, pp. 170–187.

YEREVAN STATE UNIVERSITY PUBLISHING HOUSE

JIL 2/1 (2025): 72-111

Journal of Iranian Linguistics

Caucasian Persian (Tati) Fragments in Armenian Script: A Study of the Lord's Prayer with Transliteration, Translation, and Comments[1]

Artyom Tonoyan

Yerevan State University

doi.org/10.46991/jil/2025.01.05

Abstract: The study of certain languages, particularly their history, has always greatly benefited from, and continues to rely on, the material of the Armenian language and its dialects, studied from both synchronic and diachronic perspectives, as well as the literature in Armenian script created in or translated into these languages. Within this context, the attention of scholars studying Western Iranian languages has primarily centered on two languages: Persian and Kurdish, the text corpora of both of which contain materials in Armenian script. To these two, Dimlī (Zazaki) can also be added, as separate words and expressions from this language appear in *Dersim* by Andranik (1900).

This article aims to expand the list by including another Iranian language: Caucasian Persian (Tati), text corpus of which contains several fragments in Armenian script, all published during the last quarter of the 19th century.

The article provides a general overview of Caucasian Persian fragments in Armenian script and examines the linguistic features of the *Lord's Prayer*, translated from Armenian into Caucasian Persian by Mkrtich Antonean-Dadayeanc' and published by Archbishop Mesrowb Smbateanc‘ in 1896.

Keywords: Caucasian Persian, Tati, Madrasa variety, fragments in Armenian script, Lord's Prayer, Archbishop Mesrop Smbateanc‘

Artyom Tonoyan
E-mail: artyom.tonoyan@ysu.am
ORCID: https://orcid.org/0000-0002-8740-730X

Received: 11.12.2024
Revised: 10.05.2025
Accepted: 28.06.2025

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

© Artyom Tonoyan, 2025

Introduction

[1] The main part of this article was written at the University of Hamburg during a research stay from July to September 2023, funded by the DAAD program *Research Stays for University Academics and Scientists*. The research was conducted at the invitation of Prof. Ludwig Paul, to whom I express my sincere gratitude. The project was titled *The Study of Armenian Traces on the Caucasian Persian (Tati) Dialect of Madrasa in Shamakhi: Comparative-Historical and Typological Aspects*. I am also grateful to Prof. Vardan Voskanian for his valuable comments on the text of prayer discussed in this article, as well as for kindly sharing his field notebooks and audio recordings from his research among the Armeno-Tats of Madrasa. Special thanks are due to my colleague and friend Dr. Hakob Avchyan for reading the final version of the article, offering insightful feedback, and assisting with proofreading.

Introduction

The study of certain languages, particularly their history, has always greatly benefited from, and continues to rely on, the material of the Armenian language and its dialects, studied from both synchronic and diachronic perspectives, as well as the literature in Armenian script created in or translated into these languages. The multilingual literature written in Armenian script, dating back to the 14th century and continuing thereafter (Kirakosyan 2020a: 60; Stepanyan 2023b: 31-43), encompass materials not only in the languages of peoples who lived alongside and interacted with Armenians in the Armenian Highland and its vicinity but also in the languages of other peoples with whom Armenians came into contact beyond the historical borders of their homeland. Therefore, among the languages, written in Armenian alphabet are not only Persian (Kirakosyan 2017; 2020a), Kurdish (Jndi 1962; Stepanyan 2007: 218-219; Yeranyan 2024), Georgian (Acharean 1941: 5; Stepanyan 2007: 223; Kirakosyan 2018: 47), Udi (Schulze 2005; 2016a; Maisak 2019), Greek, Syriac, (Stepanyan 2007: 223; Kirakosyan 2018: 47, Papazyan 1964: 216), Arabic (Papazyan 1964: 223), Turkish (Stepanyan 2023a; Stepanyan 2023b), Kipchak and Tatar (Stepanyan 2007; 2023a: 11-12; 2023b: 13-14), but also Polish (Grigoryan & Pisowicz 1964), Russian,[2] French,[3] etc.[4] Of these, the Kipchak materials preserved in Armenian script are the most extensive and have been the subject of continuous study for over 125 years by both Armenian (Alishan 1896; Stepanyan 2005; 2023a; 2023b;) and predominantly foreign researchers.

---

[2] See, e.g., Ms. 5954 preserved in the Matenadaran (*Catalogue of Manuscripts of the Mesrop Mashtots Matenadaran 2*, ed. Khachikyan & Mnatsakanyan 1970: 219–220), a manuscript copied in the last quarter of the 18th century and containing Russian folk songs written in the Armenian alphabet (Ms. 5954, fols. 53a–54b). I am grateful to Dr. Hasmik Kirakosyan for drawing my attention to this manuscript and for providing information on the relevant folios.

[3] See Ms. 5954, fol. 55b (*Catalogue of Manuscripts of the Mesrop Mashtots Matenadaran 2*, ed. Khachikyan & Mnatsakanyan 1970: 219–220).

[4] The Matenadaran preserves 425 non-Armenian manuscripts, all written in the Armenian script. These include 422 manuscripts in Armeno-Turkish (i.e., Turkish written in the Armenian script), 2 in Armeno-Persian (i.e., Persian written in the Armenian script), and 1 complete manuscript in Armeno-Kipchak, along with fragments preserved in 9 Armenian manuscripts. Armeno-Kurdish fragments are preserved in 6 Armenian manuscripts. Numerous religious and literary passages, as well as glossaries in the aforementioned languages, appear in Armenian-script form within Armenian manuscripts. Additionally, passages written in Arabic, Georgian, Greek, Russian, Polish, and French using the Armenian script are also found in Armenian manuscripts (data provided by Dr. Hasmik Kirakosyan, personal communication, June 2025).

Within this context, the attention of scholars studying Western Iranian languages has primarily centered on two languages: Persian (Papazyan 1964; Kirakosyan 2017, 2018, 2020a, 2020b, Kirakosian 2020; Rzepka 2006) and Kurdish (Jndi 1962; Stepanyan 2007: 218-219; Yeranyan 2024), the text corpora of both of which contain materials in Armenian script.

To these two, Dimlī (Zazaki) can also be added, as separate words and expressions from this language appear in *Dersim* published by Andranik in 1900.

Another Iranian language that deserves attention in this regard is Caucasian Persian (Tati), whose text corpus includes several fragments in Armenian script, all published during the last quarter of the 19th century. Among these, the *Lord's Prayer*, translated from Armenian into Caucasian Persian by Mkrtich Antonean-Dadayeanc' and published by Archbishop Mesrowb Smbateanc‘ in 1896, stands out as a notable example (Smbateanc‘ 1896). This article provides a detailed examination of this text and a general overview of other Caucasian Persian fragments in Armenian script.

On the other hand, this contributes a new specimen to the corpus of Persian translations of the *Lord's Prayer*, a subject extensively treated by Rzepka in a valuable article (Rzepka 2006) that gathers and examines ten versions of the prayer in Middle and New Persian, attested between the 9th and 20th centuries. Chronologically, two of these translations date to the 19th century (Rzepka 2006: 101-102) and thus correspond to the period in which the version discussed in this article was rendered. However, from the perspective of the study of the history of Caucasian Persian, and in comparison, with the translation by Smbateanc‘, even greater value lies in the translation (Rzepka 2006: 100) drawn from Papazyan's publication: a 15th-century translation preserved on fol. 145a of Ms. 7117 at the Matenadaran.

There is yet another reason to revisit this topic, which deserves mention here: it concerns the renewed source-critical assessment of the book by Smbateanc‘ (1896) in the context of studying the history of endangered minority languages spoken in the territory of Azerbaijan. Although some parts of the mentiond book, particularly the Udi version of the *Lord's Prayer* translated from Armenian, have drawn the attention of scholars specializing in Caucasian Albanian and its Udi descendant (Schulze 2005; 2016a; Maisak 2019), it has, to date, remained largely overlooked by researchers working on Iranian languages, including Caucasian Persian (Tati). In this context, the publication of the Caucasian Persian translation of the *Lord's Prayer* included in the mentioned book of Smbateanc‘ may prove significant in drawing scholarly attention to 19th-century samples of non-Persian Iranian languages written in the Armenian script.

And finally, this study also sets forth certain objectives within the context of the dialectology of Caucasian Persian (Tati). Specifically, although considerable work has already been carried out in the field of Caucasian Persian dialectology [5] and studies dedicated to individual dialects are currently available (see, e.g., Miller 1907; Grjunberg 1963; Hacıyev 1971; K'vač'aʒe 1988; Hüseynova 2002; Mammadov 2018; Suleymanov 2020a), there still remain understudied dialects, among which is the so-called Armeno-Tat dialect. This article represents an attempt to initiate a series of studies aimed at filling this gap, studies through which both recorded but unpublished texts and valuable previously printed texts that have escaped scholarly attention will be made accessible. One such text is the *Lord's Prayer* presented in this article.

Preliminaries on Armeno-Tats and their language

In academic literature, the term *Armeno-Tats* (Russ. *армяно-таты* [armjano-taty], Arm. *հայ–թաթեր* [hay–tʿatʿer]) has traditionally been used to refer to Armenians (see, e.g., Miller 1929: 13-17; Qalt'axč'yan 1970a; 1970b; 1970c; Chirikba 2015: 147) who lived in the territory of Azerbaijan during the 18th to 20th centuries.[6] Up until the outbreak of the Nagorno-Karabakh conflict, specifically, as late as 1988, Tati-speaking Armenians still lived in at least two villages: Madrasa in the Shamakhi district and Kilvar in the Shabran (formerly Davachi) district. Although by 1988 the area of settlement for Tati-speaking Armenians had significantly contracted, historically they

[5] A recent publication examines the history of scholarship on Caucasian Persian and its dialects, providing a critical assessment of the field. It presents a chronological and typological classification of the language data corpus and research produced between 1853 and 2024 (Tonoyan & Voskanian 2024).

[6] Among the terms used to refer to this group are *Tat-speaking Armenians* (Suleymanov 2023: 267), i.e., *թաթախոս հայեր* [tʿatʿaxos hayer] in Armenian (Qalt'axč'yan 1957: 86), and *татоязычные армяны* [tatojazyčnye armjany] in Russian (Akopjan 2006), as well as *Armenian Tats* (Schulze 2016b: 425–448), *Armenian-Tats* (Hakobian 2001) and *Armenian Persians*, i.e., *հայ-պարսիկներ* [hay–parsikner] (Qalt'axč'yan 1995) in Armenian. In addition to the aforementioned terms, there are also forms that emphasize local origin, such as *Mədrəsə Armenians* and *Kilvar Armenians* (Suleymanov 2020a: 28), as well as those that highlight religious affiliation, for instance *Tat-speaking Christians* (Schulze 2016b: 438). One also encounters the formulation *Tat-speaking Armenian Orthodox Christians* (Suleymanov 2023: 261), which encapsulates the group's linguistic, ethnic, and religious identities simultaneously. Regarding the attestations of this group's designation in 19th-century sources, Hakobian (2001: 125) notes four terms: *Armenian-Tats* (Arm. *հայ-թաթեր* [hay–tʿatʿer]), *Tat-Armenians* (Arm. *թաթ-հայեր* [tʿatʿ–hayer]), *Tat-speaking Armenians* (Arm. *թաթախոս հայեր* [tʿatʿaxos hayer]), and *Tat-Gregorians* (Arm. *թաթ-լուսավորչականներ* [tʿatʿ-lowsavorčʿakanner]. (2001: 125)

maintained a much broader presence across the region stretching from Upper Shirvan to Derbent.

For instance, in a work by Boris Miller, in a section devoted to the Armeno-Tats (Miller 1929: 13–17), more than fifteen settlements across the Upper Shirvan - Quba region are listed based on fieldwork and documentary sources (Miller 1929: 13–15). Traces of Armenian material culture, such as ruined churches, gravestones, and *khachkars* (Armenian cross-stones), have been found in these settlements, whose inhabitants, according to testimonies from neighboring villagers and written sources, were formerly Armeno-Tats (Miller 1929: 13–15). In addition to these fifteen settlements, Miller names over a dozen other villages that, while populated by Azerbaijanis in the 1920s, contained remains of Armenian material culture and were believed to have been formerly inhabited by Armenians. However, for these localities, Miller does not provide any evidence suggesting that those Armenians had spoken Tat (Miller 1929: 13–16).

The villages of Mountainous Shirvan and Quba that had previously been inhabited by Armeno-Tats and Armenians were depopulated at the end of the 18th century—specifically after the 1796 military campaign to Eastern Transcaucasia led by Russian General Zubov. Following the withdrawal of his troops, the Armenian populations left these areas and resettled in various parts of the North Caucasus (Miller 1929: 13–15; Qalt'axč'yan 1957: 86; Akopjan 2006: 192–193), including Prikumsk, Kizlyar, Sarafan, Mozdok, Majar, Anapa, Tuapse, and Astrakhan (Miller 1929: 14).[7]

As a result, by 1988, only two villages populated by Armeno-Tats remained in Azerbaijan. However, amid the outbreak of the Nagorno-Karabakh conflict and the growing tension in Armenian–Azerbaijani relations, the residents of the villages of Madrasa and Kilvar were also deported between 1988 and 1990, along with Armenians living in other parts of Azerbaijan (Baku, Sumgait, Ganja, Shamakhi, Shamkhor, etc.).

According to Arsen Hakobyan, based on notes taken during fieldwork interviews with eyewitnesses of these events, the deportation of residents from both villages took place at the end of 1988. However, the inhabitants of Madrasa had attempted to initiate their relocation to Armenia as early as February 1988, immediately after the events in Sumgait (Akopjan 2006: 204; 206).

Most of the Madrasa Armenians moved to Russia, where they settled in Prokhladny, located in the Kabardino-Balkar Republic, as well as to Armenia, where they established the village of Dprevank in the Aragatsotn province

[7] For detailed information on the demographic distribution and deportations of Armenian-Tats from the early 18th to the late 19th centuries, see Hakobian 2001: 125-142.

(Akopjan 2006: 204–206). The Armenians of Kilvar, on the other hand, moved to Ukraine, particularly to Druzhkovka, and to the Russian Federation, where they primarily settled in the town of Budyonnovsk, located in the North Caucasus region (Akopjan 2006: 206).

Prior to the deportation that took place at the end of 1988, the sociolinguistic situation in the two villages of Madrasa and Kilvar had undergone significant changes over the course of at least the preceding hundred years.

In the case of Madrasa, during the second half of the 19th century and up until the early 20th century, the dominant and primary language spoken by the local population was Tati (Hakobian 2002: 76–78). Armenian served as the language of communication with neighboring Armenian villages, though not all inhabitants of Madrasa were proficient in Armenian. Those who did not know Armenian often attempted to communicate in Turkic, i.e. Azerbaijani (Hakobian 2002: 76–78).

Over time, the situation gradually shifted in favor of Armenian, due on the one hand to the establishment of an Armenian school in Madrasa in 1870–1872 (Qaltʿaxčʿyan 1970c: 83; Hakobian 2002: 76), and on the other hand to the dominant presence of Armenian in the surrounding linguistic environment (Hakobian 2002: 79).

It is a documented fact that, according to written sources from the last quarter of the 19th century, there were numerous individuals in Madrasa who did not know a single word of Armenian (Hakobian 2002: 76), and even into the 1930s and as late as the 1950s, there were elderly residents who could not speak Armenian at all (Hakobian 2002: 78; Qaltʿaxčʿyan 1957: 86).

It is therefore not surprising that, according to Miller, during the 1926 Soviet census, the inhabitants of Madrasa declared “Farsi” as their mother tongue and “Armenian” as their ethnicity (Miller 1929: 16).

Nevertheless, one of the most significant changes observed from the 1920s onward was the growing dominance of Armenian in Madrasa. From that period, Armenian gradually came to dominate, while Tati became confined to use as a home language, and even functioned as a secret language in other linguistic environments (Miller 1929: 16; Qaltʿaxčʿyan 1957: 87).

Based on his observations from the 1950s and 1960s, Qaltʿaxčʿyan already noted that Tati was no longer spoken in public spaces, gatherings, or community meetings in Madrasa. Its use had become confined to the household, and not even in all households, mainly among members of the older and middle generations (Qaltʿaxčʿyan 1970c: 83).

Essentially, over the span of approximately one hundred years, Caucasian Persian was almost entirely forgotten among the residents of

Madrasa, gradually being replaced by Armenian. Observations we conducted in Armenia during the 2010s revealed that only three to four individuals, aged between 80 and 90, were capable of translating sentences from Armenian into Caucasian Persian, and among them, only one could do so fluently. The rest remembered only a few words and expressions, typically those they had heard from their grandparents and, in some cases, their parents.

The status of the Armenian language was different among the residents of Kilvar. The Armenian school in the village, founded in 1875, operated intermittently until the first half of the 1910s (Hakobian 2002: 72–74), and during the Soviet period, likely continued until the mid-1950s (Qaltʿaxčʿyan 1970c: 83).

In contrast to Madrasa, bilingualism among the residents of Kilvar consisted of Caucasian Persian (Tati) as their mother tongue, and Azerbaijani as a second language or as the language used for communication with neighboring Azerbaijani-speaking villages (Hakobian 2002: 72–75).

Although several generations received instruction in Armenian thanks to the Armenian school, the language did not take root. According to Qaltʿaxčʿyan, by the mid-1950s the Armenian school was closed entirely at the request of the village population (Qaltʿaxčʿyan 1970c: 83). As he reports, the residents of Kilvar justified this decision by arguing that their children, upon graduating from the Armenian school, could not find employment in any institution, as all official correspondence in their region was conducted in Azerbaijani (Qaltʿaxčʿyan 1970c: 83).

According to Qaltʿaxčʿyan's observations from the 1960s, by that time the majority of Kilvar's children were no longer able to speak Armenian, although some of them could understand spoken Armenian to a certain extent Qaltʿaxčʿyan 1970c: 83). During those years, Armenian was known relatively well by members of the middle and older generations in Kilvar Qaltʿaxčʿyan 1970c: 83).

Consequently, due to the differing linguistic environments, the overall sociolinguistic situation in these two villages developed in substantially different ways. In Madrasa, the significant Armenian presence in Shamakhi and the existence of neighboring Armenian-populated villages ultimately led to the dominance of Armenian among its residents. In contrast, the weaker or altogether absent Armenian presence in the Shabran region diminished the role of Armenian among the people of Kilvar, contributing first to the consolidation of Azerbaijani as a second language, and eventually to its emergence as the dominant and primary language.

Remote fieldwork conducted by the author in the 2010s among Kilvar Armenians who had settled in the Russian Federation and Ukraine revealed that knowledge of Caucasian Persian within this community was in a

considerably better state. It was relatively easy to find individuals in their 60s and 70s who were able to translate everyday conversations and folklore texts from Azerbaijani and Russian into Caucasian Persian (Tati) without difficulty.

The endoglottonyms used by the Armenian-Tats of these two villages also differ. Among the residents of Madrasa, the terms *Parseren*, i.e. *Parsi* (Êfendyan 2004: 88–96) and *zuhunmu*, literally meaning "our language" (Schulze 2016b: 446), are attested as endoglottonyms. In contrast, the Kilvar Armenians consistently referred to the language as *Tati* in our interviews, and never used the terms *Parseren* or *Farsi*.

Although the languages of these two villages are often treated in scholarly literature as a single dialectal variety and collectively referred to as "Christian Tati", in reality they belong to distinct subgroups of Caucasian Persian, distinguished by phonological, grammatical, and lexical features. The variety spoken in Madrasa falls within the Shamakhi dialect group (Tonoyan 2015: 21–22; 192–193), whereas the Kilvar variety belongs to the Quba dialect group (Tonoyan 2015: 19–20; 191–192).

The Madrasa variety of Caucasian Persian (Tati): research overview and corpus

In the broader context of both Iranian dialectology and the study of Caucasian Persian dialects, the Madrasa variety remains one of the least studied. The only relatively extensive research devoted to it is the still-unpublished dissertation by Qaltʿaxčʿyan, defended in Yerevan in 1970 (Qaltʿaxčʿyan 1970a). This dissertation is particularly valuable for its fieldwork-based textual and lexical data, which will be discussed in more detail below, in the context of the corpus description.

In terms of published studies as well, there is little to present regarding this dialect variety. Only two articles written in Armenian by Qaltʿaxčʿyan are available. The first offers a general overview of the dialect's phonological, morphological, and lexical features (Qaltʿaxčʿyan 1957), while the second, written within the framework of contact linguistics, examines the influence of Armenian on the Caucasian Persian (Tati) dialect variety spoken in Madrasa[8] (Qaltʿaxčʿyan 1970b). Apart from these, no dedicated studies have

[8] A valuable study on the contact between Armenian and Caucasian Persian was carried out by Vardan Voskanian, who examined Armenian loanwords that appear not only in the Madrasa variety, but across all dialectal varieties of Caucasian Persian (Voskanian 2016: 240-248). In that article, Voskanian demonstrated that the contact between Armenian and Caucasian Persian occurred not only in specific areas inhabited by Tati-speaking Armenians, such as Madrasa, Kilvar, and formerly also Khachmasn, but throughout the broader region of historical Shirvan, that is, a

focused specifically on the Madrasa dialect.[9] Nevertheless, some works have drawn on material from Madrasa to a limited extent, and certain features of this dialectal variety have been examined in the context of comparative analyses (see, e.g., Hacıyev 1995: 115-121; Tonoyan 2018; Suleymanov 2020b).

The corpus of Madrasa Tati, comprising textual materials and glossaries (published and unpublished) as well as audio recordings, currently includes the following:

a single sentence written in Armenian script and published in *Mshak* in 1880 (Hakobian 2002: 77),

two short fairy tales (approx. 3 pages), published in 1894 by Lopatinskij (1894: 25–32),

the *Lord's Prayer* in Armenian script (7 lines), translated from Armenian and published by Smbateancʿ (1896: 547),

a short story about the construction of the Madrasa village church, written in the Armenian script (24 lines with interlinear translation into Armenian) and published by Smbateancʿ (1896: 548–549),

a short folklore text published by B. Miller in his collection of Tati materials (Miller 1945: 124-125),

15 texts (short folklore stories, fairy tales, and simple conversational texts with Armenian translations), included in the appendix of Qaltʿaxčʿyan's unpublished dissertation (Qaltʿaxčʿyan 1970a: 211–251).

a 22-page colloquial phrases, **simple interrogative, affirmative, and negative sentences** with accompanying Armenian translations, also found in the appendix of Qaltʿaxčʿyan's dissertation (Qaltʿaxčʿyan 1970a: 252–274),

a 75-page glossary, included in the same appendix (Qaltʿaxčʿyan 1970a: 275–350),

107 proverbs published by Hacıyev (1995: 115–121),

a 60-page wordlist (approx. 1,800 words and expressions) compiled in the village of Dprevank and preserved in prof. V. Voskanyan's unpublished field notebooks,

territory which, prior to its incorporation into the Russian Empire, roughly corresponded to the domains of the Baku, Shamakhi, and Quba khanates.

[9] Certain features of the Madrasa variety were also addressed by the author at ICIL-8. The material from this presentation has not yet been published as a standalone article; it is currently available only as a brief abstract in the abstract volume of the ICIL-8 conference (Authier & Tonoyan 2018: 67-68).

1 folklore text (3 pages), written in Yerevan and included in the appendix of A. Tonoyan's unpublished dissertation (Tonoyan 2015: 192–194),[10]

5 fairy tales (approx. 20 pages), written in Armenia (Dprevank, Charentsavan, and Etchmiadzin), preserved in A. Tonoyan's unpublished field notebooks,

3 hours of audio recordings, made by V. Voskanyan in 2001,

2 hours of audio recordings, made by A. Tonoyan in 2015.

Caucasian Persian (Tati) *Lord's Prayer* in Armenian Script (Smbateancʿ 1896)

The prayer under examination was published in 1896 in the book by Archbishop Mesrowb Smbateancʿ, who headed the Shamakhi diocese of the Mother See of Holy Etchmiadzin between 1887 and 1895. In his work *Description of the Saint Stephen Monastery in Saŧian and other monasteries, pilgrimage sites, villages and towns of Shamakhi diocese*, among other materials, he included a short story about the construction of the church in the village of Madrasa, written in the Armenian script (24 lines with interlinear translation into Armenian) (Smbateancʿ 1896: 548–549), as well as the *Lord's Prayer*, both in the Tat language (Smbateantsʿ 1896: 547). In the book (Smbateancʿ1896: 547), the record of the *Lord's Prayer* appears in Armenian script, written continuously in a single paragraph (see ill. 1).

In the section devoted to the village of Madrasa, Smbateancʿ begins with a description of the village's geographical location, historical background, population, and principal occupations (pp. 546–550). This contextual overview is followed by what may be considered a linguistic *field note*, in which he writes: *"In order to acquaint readers with the dialect of the people of Madrasa, which consists of Persian-Hebrew*[11] *words, I shall write the Lord's Prayer"* (Smbateantsʿ 1896: 547).

Following the Armenian-script version of the *Lord's Prayer*, Smbateancʿ adds the following sentence: *"I have also written a conversation, along with its Armenian translation, which is as follows"* (Smbateancʿ 1896: 547). The entire page 548 and the first half of page 549 are occupied by this conversation, which, as noted above, is a short story about the construction of the church in Madrasa village, written in the Armenian script with an interlinear Armenian translation (Smbateancʿ 1896: 548–549).

---

[10] This text forms part of an appendix in the dissertation, where the same folklore text is presented in 13 different varieties of Caucasian Persian (for details, see Tonoyan 2015: 175-193, also Tonoyan & Voskanian 2024: 204-205).

[11] *Ֆարս Հրէական բառեր* [Fars Hrêakan bařer] in Armenian original.

Smbateanc‘ concludes this valuable material on the Madrasa dialect with the following statement: *“The text presented here reflects the speech and heartfelt sentiments of the people of Madrasa. It was written, along with the Lord’s Prayer, by Mkrtich Antonean-Dadayeanc‘, a clerk-teacher from Madrasa, as a translation of the villagers’ feelings. After three years of preparation with us in Shamakhi, he was appointed as a clerk-teacher in the church.”*

«Փիար մունքի բիգեօկի, թեամուզ բաշը թունա նում թը.
բիով ջեանաթը բույու խիալթը. չի թահար բիգեօկի. բույու-
թահար բուտիւնեայ. նումնա իմունա հարրուզ թի բիմուն իմ-
ռուզ. բույութահար բէյիշ իմունա ղեարտ խուհունդամնա. բու-
յութահար մաբիար իմուն հարաբը խաթա. բույութահար ազատ-
քը աչար. քիանթունը ջաննաթ. բույութահար զիւյուրին. վա-
բայութահար շուքուր. ազըմ բէլա թաբա ախբը. ամին»

Ill. 1. *The Lord’s Prayer* in the book Smbateanc‘ (1896: 547).

A1. The *Lord’s Prayer* in Armenian script, as published in the original source (Smbateanc‘ 1896: 547, see ill. 1)

«Փիար մունքի բիգեօկի թեամուզ բաշը թունա նում թը. բիով ջեանաթը բույու խիալթը չի թահար բիգեօկի բույու-թահար բուտիւնեայ. նումնա իմունա հարրուզ թի բիմուն իմ-ռուզ. բույութահար բէյիշ իմունա ղեարտ խուհունդամնա. բու-յութահար մաբիար իմուն հարաբը խաթա. բույութահար ազատ-քը աչար. քիանթունը ջաննաթ. բույութահար զիւյուրին. վա-բայութահար շուքուր. ազըմ բէլա թաբա ախբը. ամին»

A2. Transliteration of original in Armenian script

P‘iar mownk‘i bigeôki t‘eamowz bašə t‘owna nowm t‘ə. biov ǰeanat‘ə bowyow xialt‘ə č‘i t‘ahar bigeôki bowyow-t‘ahar bowtiwneay. nowmna imowna harrowz t‘i bimown im-rowz bowyowt‘ahar bêyiš imowna qeart xowhowndamna. bow-yowt‘ahar mabiar imown harabə xat‘a. bowyowt‘ahar azat-k‘ə ač‘ar. k‘iant‘ownə ǰannat‘. bowyowt‘ahar ziwyowrin. va-bayowt‘ahar šowk‘owr. azəm bêla t‘aba axər. amin

A3. Selective modification of certain Armenian letters and their transliteration according to Iranological and Turkological systems, based on relevant phonetic correspondences.

Armenian original	*Armenological transliteration*[12]	*Iranological transliteration*[13]	*Türkological transliteration*[14]
Փ/փ	p‘	p	-
Թ	t‘	t	-
Ք	k‘	k	-
կ	k	k’	-
տ	t	t’	-
Չ	č‘	č	-
ղ	q[15]	q	-
ը	ə	ï	ı
ա	a	ä /ā[16]	-
ու	ow	u / ū	-
իւ	iw	yu / ü	ü
եա	ea	(y)ä / yā	-
եօ	eō	yo	ö

[12] The transliteration of Armenian is given according to the Hübschmann-Meillet system.

[13] The transliteration of Iranian material is based on the DMG system (version 1.0), with slight modifications made for consistency with the phonological features of the source text.

[14] In the case of the Turkological transliteration, only those phonemes are provided that occur in the words of Turkic origin found in the prayer text and whose pronunciation differs significantly from that of Armenian and/or Persian.

[15] Since the Armenian letter *[ղ]* appears only once in the prayer, specifically in the Arabic loanword *qarż*, it is preferable in this case to transliterate the Arabic *[ق]* as *[q]* rather than *[ḫ]*.

[16] Since the Armenian vowel [a] is phonetically distinct from both the short and long [a/ā] vowels of Persian, and given that the text of the *Lord's Prayer* contains an Armenian loanword with this vowel, the Armenian [a] is represented as *a* in transliteration. To avoid confusion and to preserve the phonetic contrast, the Persian short [a], which is traditionally transliterated as *a* (as opposed to the long [ā], marked with a diacritic), is here rendered as *ä*. This approach serves both to prevent ambiguity with Armenian *a* and to accurately reflect the phonemic distinctions among the various *a*-like vowels.

B1. The *Lord's Prayer*: transliteration of the Old Armenian original, Caucasian Persian translation in Iranological transliteration (DMG 1.0, with slight modifications), and its reconstructed pronunciation.

	Armenian original (translit.) /AO	*Caucasian Persian (Tati) translation / CPT*	*Reconstructed form and pronunciation / RFP*
1	Hayr mer or yerkins es,	Piär mūnki bigökī	Pi[y]ar[-]mūn ki[17] bi gökī
2	Sowrb ełicʿi anown kʿo:	tämuz bāšï tūnä nūm tï	Tamüz bāšï tūna nūm-tï
3	Ekescʿê arkʿayowtʿiwn kʿo:	biov ǰännätï	bī[y]ov ǰannat[-t]ï
4	Ełicʿin kamkʿ kʿo	būyu xīältï	būyu xī[y]āl-tï
5	Orpês yerkins ew yerkri:	čī ṭähär bigökī buyuṭähär budunya	čī ṭähär bi gökī böü ṭähär bu dünya[18]
6	Zhacʿ mer hanapazord	nūmnä imūnā härrūz	nū-m[ū]n-ā imūnā har rūz
7	Towr mez aysôr:	ti bimūn imrūz	ti bi[-]mūn imrūz

[17] In the source text, the form ⟨mūnkī⟩ is written as a single unit. It is separated here (mūn kī) using a space to reflect its morphological components. At the same time, *mūn-* is joined with *pīar* by means of a hyphen (*pīar-mūn*) to indicate its function as a personal pronominal clitic (cf. lit. Persian *pedaremān*, and also *pedaremun* "our father", where the shift *ā* > *ū* before *-n* reflects a regular sound change in colloquial Persian. Cf. also Turkic *atamız* "our father", where the possessive suffix *-ımız/-imiz* marks first person plural possession.

[18] Given that the locative preposition *ba* (< NP *ba* < MP *pad*) has become *bu*, providing evidence of vowel harmony (synharmonism), likely under the influence of local Turkic, it can be inferred that the traditional long final *-ā* in the word *dunyā* would have been pronounced short, i.e. *a*. On the other hand, in Azerbaijani the word is written and pronounced *dünya*, not *dünyə*. However, the Azerbaijani *a* is in any case not comparable in duration to the long *a* of Persian and cannot, by itself, serve as sufficient grounds for confidently asserting that a long *a* could have been pronounced in the syllable following *ü*. Unfortunately, the Armenian script in this case does not allow for the precise phonetic reconstruction, and this word is absent from the wordlist appendix of aQlt'axčʿyan's dissertation. Instead, for the notion of 'world, country,' an Armenian loan, i.e. *ašxarkʿ* is recorded (Qalt'axčʿyan 1970a: 277).

8	Ew tʿoł mez zpartis mer,	buyuṭähär behiš imūnā qärtxūhundāmnā	böü ṭahar behiš imūnā qart[]xūhund[-]ā[-]m[ū]nā
9	Orpês ew mekʿ tʿołowmkʿ meroc' partapanacʿ:		
10	Ew mi tanir zmez i pʿorjowtʿiwn	buyuṭähär mäbīār imūn hārābï xäṭā	böü ṭahar mabī[y]ār imūnhā[.]rā bï xaṭā
11	Ayl pʿrkea zmez i čʿarê:	buyuṭähär āzāt kï äčar	böü[.]ṭahar āzāt kï a[.]čâr
12	Zi kʿo ê arkʿayowtʿiwn	kīāntūnï ǰännät	kī[.]ān[-]tūnï ǰannat
13	Ew zôrowtʿiwn	buyuṭähär züürīn	böü[.]ṭahar züürīn
14	ew pʿarkʿ	väbuyuṭähär šukur äzïm	va[.]böü[.]ṭahar šukur azïm
15	yawiteans	bēla tā bï axïr	bela tā bï axïr
14	Amên	Āmin	Āmīn

B2. Final form

Piyar-mūn ki bi gökī
tamüz bāšï tüna nūm-tï
bīyov ǰannat-tï
būyu xīyāl-tï
čī ṭahar bi gökī böü ṭähär bu dünya
nū-mūn-ā imūnā har rūz
ti bi-mūn imrūz
böü ṭahar behiš imūnā qart-xūhundā-mūn-ā
böü ṭahar mabīyār imūnhā-rā bï xaṭā
böü ṭahar āzāt kï a čâr
kī āntünï ǰannat
böü ṭahar züürīn
va böü ṭahar šükür ʿaẓïm
bela tā bï axïr
āmīn

B3. An Interlinear morphological analysis

(1) *Piyar-mūn*[19] *ki* *bi* *gök-ī*
Father-PC.POSS.1PL REL LOC heaven-COP.2SG
"Our father, who are in heaven"

(2) *tamüz* *bāš-ï* *tü-n-a* *nūm-tï*
pure be$_1$-SBJV/IMP.PR.3SG you-<n>OBL[20] name-PC.POSS.2SG
"may your name be holy"

(3) *bī-y-ov* *ǰannat-tï*
ISP-<y>come$_1$-SBJV/IMP.PR.3SG paradise- PC.POSS.2SG
"may your kingdom come"

(4) *bū-y-u* *xīyāl-tï*
be$_2$-<y>-SBJV/IMP.PR.3SG will-PC.POSS.2SG
"may your will be carried out"

(5) *čī* *ṭahar*[21] *bi* *gök-ī*
what.INT kind.APP LOC heaven-COP.2SG
bö-ü *ṭahar* *bu* *dünya*
DIR-DIS kind.APP LOC earth
"just as in heaven, so also on earth"

(6) *nū-mūn-ā* *imūn-ā* *har* *rūz*
bread-PC.POSS.1PL-DOM we-OBL every day
ti *bi-mūn* *imrūz*
give$_1$-IMP.PR.2SG DIR-PC.POSS.1PL today
"give us today our daily bread"

(7) *bö-ü* *ṭahar* *be-hiš*
DIR-DIS kind.APP ISP-let$_2$-PR.2SG
imūn–ā *qart* *xūhund-ā-mūn-ā*
we-OBL debt owner-PL-PC.POSS.1PL- DOM
"also, forgive those who are the owners of our debts"

---

[19] In the Madrasa variety of Caucasian Persian, the possessive pronominal clitics are as follows: *-man* (1SG), *-tï(n)* (2SG), *-i* (3SG), *-mun* (1PL), *-šmun,* and *-šun (?)* (2PL), *šïn* (3PL) (Qalt'axč'yan 1970a: 115). These forms are nearly identical to those of the Upper Shirvan variety, where the corresponding forms are: *-man* (1SG), *-tï* (2SG), *-i* (3SG); and in the plural: *-mun* (1PL), *-šmun* (2PL), *-šun* (3PL) (Suleymanov 2020: 98).

[20] From a strictly grammatical standpoint, this can be classified as a dependent possessive pronoun (DPPR).

[21] The form *čī ṭahar* (or contracted *čīṭahar*), when used independently, conveys the meaning "how" or "in what manner" (cf. Pers. *četowr*). However, in finite manner clauses of the type *čī ṭahar ... böü ṭahar*, the construction functions as a correlative comparative structure and semantically corresponds to English *"just as … so also …"*. For a detailed discussion of the phonological changes and functional usage of this word of Arabic origin in Azerbaijani, see Pastor & Suleymanov 2023: 430-433.

(8)	*bö-ü*	*ṭahar*	*ma-bī-y-ār*
	DIR-DIS	kind.APP	NEG-ISP-<y>-bring$_1$-PR.2SG
	imūnhārā	bï	xaṭā
	we-Pl- DOM	DIR	temptation

"also, do not lead us into temptation"

(9)	*bö-ü*	*ṭahar*	*āzāt*	kï	*a*	*čâr*
	DIR-DIS	kind.APP	free	do$_1$- IMP.PR.2SG	SEP	evil

"in the same way, deliver us from evil"

(10)	*kī*	*āntü-n–ï*	*ǰannat*
	because	IPPR.2SG-<n>- COP.3SG	heaven

"for Yours is the kingdom"

(11)	*bö-ü*	*ṭahar*	*züürīn*
	DIR-DIS	kind.APP	power

"in the same way, the power"

(12)	*va*	*bö-ü*	*ṭahar*	*šükür*	*ʿaẓïm*[22]
	and	DIR-DIS	kind.APP	glory	immense

"and, also, the immense glory"

(13)	*bela*	*tā*	*bï*	*ʿaxïr*	*āmīn*
	in this manner	DIR	DIR	end	Amin

"In this way, until the end. Amin"

Comments

Overall, the *Lord's Prayer*, translated from Armenian into the Madrasa variety of Caucasian Persian, written in Armenian script, and published by Smbateancʿ, is quite concise in terms of length. Nevertheless, even in its brevity, it displays several noteworthy features. It is valuable not only for comparative analysis with other dialectal varieties of Caucasian Persian, but also for tracing internal developments within the Madrasa variety over several decades, particularly when compared with the data presented in Qaltʿaxčyan's dissertation.

The observations offered in this article concerning this version of the *Lord's Prayer* are organized into four categories: orthography, phonology, morphosyntax, and lexicon.

[22] Cf. Persian *ʿaẓīm* "great", "huge", "immense" "gigantic", "cosmic", "mighty" etc.

Orthography

The Caucasian Persian translation of the *Lord's Prayer* in Armenian script contains four instances of hiatus, which may be grouped into two categories:

a. *i* + *a*,[23] as in ***piar***, ***xīāl***, and ***mabīār*** (see A1, AO and CPT columns);
b. *i* + *o*,[24] as in ***biov*** (see A1, AO and CPT columns).

In the reconstructed forms presented in this article, i.e., ***pi[y]ar***, ***xī[y]āl***, ***mabī[y]ār***, ***bī[y]ov*** (see A1, RFP column), the presence of the glide *[y]* results from distinct phonological processes characteristic of both New Persian and Caucasian Persian. These processes include:

a. *-d-* > *-y-* in intervocalic position, as in ***pidar* > *piyar*** (cf. OIr. ***bauda-*** > ENP ***bōy*** > NP ***būy*** > CPers ***bū***);
b. epenthesis, i.e. insertion of the glide *[y]* between vowels, a highly productive phenomenon in New Persian and Caucasian Persian, as seen in ***mabī[y]ār***, ***bī[y]ov*** (see A1, RFP column).

The absence of the glide *[y]* in the relevant lexemes of the prayer, written in Armenian script, may be attributed to the conventions of Classical Armenian orthography, which, apart from a few loanword exceptions in Modern Eastern Armenian (MEA), such as ***kʿaos***[25] "chaos" and ***kooperacʿia***[26] "cooperation", permits the written representation of the glide *[y]* only in *a* + *vowel* and *o* + *vowel* sequences. In all other cases, including *i* + *a* and *i* + *o* sequences, although the glide *[y]* is present in pronunciation, it is typically omitted in writing (see Dum-Tragut 2009: 46-47).

Consequently, the translator of the prayer remained faithful to Classical Armenian orthography and deliberately omitted the glide *y* in the spelling of these four words, thereby artificially creating hiatuses of the *i* + *a* and *i* + *o* type.

This same principle appears to apply in Middle Armenian literature as well, where, with a single exception, the Persian word ***xīyār*** "cucumber" is spelled ***xiar***[27] rather than ***xiyar***[28] (Łazaryan 2009: 320). The same spelling is also

[23] Arm. *խ* + *ա:*
[24] Arm. *խ* + *ո:*
[25] Arm. *քաոս*
[26] Arm. *կոոպէրացիա*
[27] Arm. *խիար*
[28] Arm. *խիյար*

attested in several derived forms containing ***xīyār***, such as ***xiarak***,[29] ***xiaršampʿay***,[30] ***xiaršambay***,[31] ***xiaršampar***,[32] etc.

This orthographic principle may likewise be operative in another Persian translation of the *Lord's Prayer* written in Armenian script and preserved in Matenadaran (Ms. 7117), dated to the early 15th century, where the word ***pʿiar*** ("father") appears in the same form. V. Papazyan (1964: 219), however, emended it as ***pidar***,[33] possibly overlooking the dialectological context, i.e., the possibility that the prayer could have been composed in Caucasian Persian (Tati), in which case the form ***piyar*** should not have been replaced with ***pidar*** (see Appendix)[34].

The same orthographic pattern is also found in Ms. 8492 (1717–1721, Shamakhi–Tiflis) and Ms. 3044 (1780, Ganja), both preserved in the Matenadaran and both containing Persian translations of the Gospels written in Armenian script.[35]

Moreover, although the Madrasa variety of Caucasian Persian allows hiatus, particularly of the ***ou*** and ***öü*** types, as seen in forms such as ***gouz*** "whip", ***nour*** "lake", ***ǰouri*** "broom", ***xound*** "lord", ***dagröündan*** "to change, to dress",

---

[29] Arm. *խիարակ*

[30] Arm. *խիարշամպայ*

[31] Arm. *խիարշամբայ*

[32] Arm. *խիարշամպար*

[33] Pers. پدر

[34] In April 2016, during a private conversation, Prof. Vardan Voskanian informed me that the so-called "Persian" translation of the *Lord's Prayer*, written in the Armenian alphabet and preserved in Matenadaran Ms. 7117, may represent one of the earliest extant texts in Caucasian Persian (Tati). According to Voskanian, several of its phonological and grammatical features suggest that it should be identified with Caucasian Persian (Tati). Probably unfamiliar with the Caucasian Persian (Tati) language, Pʿapʿazyan interpreted all deviations from standard Persian in the text as the result of misreadings of a presumed original in Arabic script or, more frequently, as deviations and irregularities caused by dialectal influence and a disregard for normative literary forms (Pʿapʿazyan 1964: 218–219). Both the translated fragment recorded by M. Smbateancʿ in Madrasa and the growing body of research on various dialects of Caucasian Persian support the conclusion that Voskanian's observation was well-founded.

[35] Although H. Kirakosyan, who conducted a detailed study of the two aforementioned manuscripts (see Kirakosyan 2020a), notes in one of his articles that the semivowel *-y-* is inserted before the Persian plural suffix *-ān* and following the vowels *-ā*, *-i*, and *-u*, the examples she provides do not include any cases of *-y-* insertion after the vowels *-i* or *-u* (see Kirakosyan 2018: 108–109). On the contrary, the manuscripts contain several instances where the semivowel *-y-* is absent even after the vowels *-i* and *-u*, as in *biapi* (Kirakosyan 2020: 17, cf. NP *bīyābī* "may you find"), *isayian* (Kirakosyan 2020: 21, cf. NP *īsāyīyān* "Christians"), *siay* (Kirakosyan 2020: 37, cf. NP *sīyāh* or *siyah* "black"), *xial* (Kirakosyan 2020: 39, cf. NP *xīyāl* "thought"), *bïgowit/bïgowid* (Kirakosyan 2020: 187; 218–219, cf. NP *begūyēd* "say!"), and *mi gowim* (Kirakosyan 2020: 223, cf. NP *mē-gūyēm* "we say / we will say").

höüz "cistern", ***vartöür*** < Arm. ***vardavar***[36]), the word for "father" is nevertheless consistently attested with the medial glide *[y],* i.e. ***piyar***, in both the unpublished dissertation by Qaltʿaxčʿyan (1970a) and in my 2014 field recordings.[37]

The use of -*y*- in the form ***bêyiš*** (see A1 and A2), instead of -*h*-, is likely to be influenced by Classical Armenian orthographic conventions. The verb is ***hištan*** (inf.) "to leave, to permit," which in this context appears in the imperative form (2SG), i.e. ***be+hiš***. It is possible that the translator, following the rules of Classical Armenian orthography and pronunciation, where initial *y*- is pronounced *[h]*, rendered ***hiš***- as ***yiš***-. This may also have been reinforced by analogy with Armenian words such as ***yišel*** "to remember", ***yišecʿnel*** "to remind", and ***yišatak*** "memorial", which, although spelled with initial *y*-, are pronounced as *[h]* (i.e., ***hišel, hišecʿnel, hišatak***). Thus, influenced by both phonetic coincidence and graphic analogy, the translator may have represented the imperative form of ***hištan*** as ***bêyiš***.

Particularly noteworthy in the prayer is the spelling of the Persian word ***dunyā*** ("world") as ***tiwnēay***.[38] First, according to Acharyan's law, attested in several Armenian dialects, including that of Shamakhi, the word-initial voiced dental *d*- is devoiced and palatalized, becoming *t*. As a result, the word takes the form *tiwnea*. Then, following traditional Armenian orthographic conventions, which require the insertion of the semivowel -*y* after final -*a* and -*o*, the word receives a final -*y*, producing the written form ***tiwnēay*** (see A1 and A2).

The word ***āzāt*** "free" (see B3(9) and C1(21)) also appears to be a result of Armenian orthographic influence (cf. Arm. ***azat***, itself an early Iranian loan), as one would have expected azad with final -*d* instead (cf. NP ***āzād***). The translator likely followed Armenian orthographic conventions and rendered the word in the spelling commonly used in Armenian, disregarding the need to represent its actual pronunciation. This is further supported by the forms ***āzād*** "free" and ***āzād sātan*** "to free" mentioned in Qaltʿaxčʿyan's dissertation (1970a: 275), as well as by several instances in the audio recordings made by Vardan Voskanian, in which the word is pronounced as *āzād*.

[36] It is a centuries-old Armenian festival where participants drench each other with water, a custom still observed today.
[37] Cf. also UpSh., Madr., ArDagh., Sur. *piyar*.
[38] Arm. *տիւնեայ* (see A1)

Phonology

Among the phonological features, the following may be noted:

1. Reduction of the rounded vowel *u* in interconsonant position, likely due to strong stress on the first syllable. This occurs in the case of the 1PL DPPR *mūn*, when followed by the DOM *ā*, as in ***nū-mnā < *nū-mūnā*** ("our bread" + DOM), ***qeart-xuhunda-mnā < *qart-xuhundamūnā*** ("our debtors" + DOM) (see A1, A2, B1/CPT and RFP columns, B3(6),(7)).
2. The phonological form of ***qarż*** (< Arabic) attested in the prayer is particularly noteworthy. It appears in the compound ***qart-xuhund***, meaning "creditor", "debtor", where *qart* means "debt" and ***xuhund*** means "owner." The final voiceless stop *-t* in *qart* appears irregular, given that in other dialects of Caucasian Persian (including UpSh and ArDagh) the word is attested with a final *-d*, i.e., *qard* (Suleymanov 2020: 232; Soltanov & Soltanov 2013: 102; Grjunberg 1963: 141). In Qaltʿaxčʿyan's dissertation (1970a), two forms are recorded: ***gart*** "debt" (p. 286) and ***qārd-xound*** "creditor, debtor" (p. 285), the latter preserving the original *[q-]*, displaying *[-d]* in final position (as in other dialects of Caucasian Persian), and showing a shift from short *[a]* to long *[ā]*.

   This word is of particular interest for two reasons. On the one hand, the dialects of Caucasian Persian preserve an archaic pronunciation of Arabic *[ż]*,[39] in contrast to Azerbaijani and Modern Persian, where Arabic *[ż]* is realized as *[-z]* in final position.[40] On the other hand, the pronunciation *-t* in the *-rt* consonant cluster, as found in the prayer in Armenian script, may reflect dialectal influence from Armenian, particularly from the Shamakhi region, where the voiced stop *[d]* is regularly devoiced to *[t]* after *[r-]* in ***-rd*** clusters. Interestingly, the form *γart* attested in Ms. 7117 (see Appendix) indicates that the ***-rd > -rt*** shift had already been operative in borrowed words by at least the 15th century. Although the vowel length in the 15th-century form ***γart*** is difficult to determine, the spelling ***qart*** found in the Tati translation of the *Lord's Prayer* published by Smbatʿeancʿ clearly suggests that the vowel was pronounced short.
3. The compound *qart-xuhund* also contains a phonological sequence, i.e. *uhu*, within its second component ***xuhund*** (< ***xudāwand*** (?)), meaning "lord" or "master," which is atypical for the Madrasa variety. Unlike other Tati dialects, which retain the glottal consonant *h* in intervocalic position,

[39] Arabic [ض].
[40] Cf. NP *qarż* (Steingass 1892: 964), Az. *qərz* (ADİL 2006(3): 122).

particularly in environments such as *u(C)u*,[41] the Madrasa dialect exhibits loss of this consonant due to intervocalic lenition.[42] Notably, in mid-20th-century field data recorded by Qaltʿaxčʿyan, the word meaning "lord", attested in the *Lord's Prayer* as *xuhund*, appears in two alternative phonological forms, both of which reflect the loss of *[h]* through intervocalic lenition. In one case, the forms ***xound*** ("lord, director"; Qaltʿaxčʿyan 1970a: 306) and ***qardxound*** ("creditor"; Qaltʿaxčʿyan 1970a: 285) are documented; in the other, the form *xuund* with the meaning "lord" is attested (Qaltʿaxčʿyan 1970a: 235, 264). These attestations suggest that the change ***–uhu– > uu*** is a relatively recent development in the Madrasa variety of Caucasian Persian, likely emerging or becoming stabilized between approximately 1900 and the 1950s.

4. Another noteworthy point is that the word ***gök***, used in the prayer to denote the concept of "heaven," appears in Qaltʿaxčʿyan's dissertation with the pronunciation ***gög*** (Qaltʿaxčʿyan 1970a: 288). This form may either reflect a case of word-final ***-k > -g*** voicing that occurred during the aforementioned period (1900–1950s), or it may suggest that the original form was ***gög***,[43] which, due to dialect-specific devoicing, was rendered as ***gök*** in the prayer text (see B1, B2 and B3).
5. Another change observable in comparison with Qaltʿaxčʿyan's dissertation is the labialization ***-i > -u*** in word-final position (Auslaut) in the 1PL personal pronoun *imūn*, which occurs in the prayer only in its oblique form, i.e. ***imūnā*** (see B3(6)),[44] functioning essentially as a dependent possessive pronoun. In the *Lord's Prayer* published by Smbatʿeancʿ, this form is attested as ***imūnā***[45] whereas in Qaltʿaxčʿyan's data it appears as ***umūnā*** (Qaltʿaxčʿyan 1970a: 340).[46]

---

[41] Cf. UpSh., ArDagh., Af., Sur. *zuhun* "language" (Suleymanov 2020: 110, 142; Soltanov & Soltanov 2013: 159; Grjunberg 1963: 141, 159; Mammadova 2018: 56), UpSh., ArDagh., Af. *duhun* "mouth" (Suleymanov 2020: 323; Soltanov & Soltanov 2013: 62; Grjunberg 1963: 26), as well as Sur. *dahan* "mouth" (Mammadova 2018: 334).

[42] Cf. *zuun* "language", "tongue" and *duun* "mouth" (Qaltʿaxčʿyan 1970a: 296, 292).

[43] Cf. SW Osm. *gök* (before vowels *gög* …), Tkm. *gö:k* (*gö:g* …). According to Clauson (1972: 708), these forms, along with Az. *köy*, suggest that the original form may have been *gö:g*.

[44] Cf. UpSh. *imun, umun* "we" (Suleymanov 2020: 97), and *imuna* (Suleymanov 2020: 334), and *umuna* "our" (Suleymanov 2020: 166).

[45] See B3(6).

[46] Notably, both in the field notebooks of Vardan Voskanian and in my own recordings from the mid-2010s, only the forms *umun* "we" and *umunā* "our" are attested; there are no instances of the forms *imun* or *imunā* in use.

Morphosyntax

The prayer exhibits several interesting morphosyntactic features, some of which are relatively straightforward and easy to explain, while others are more complex and less transparent.

1. Among the more accessible features is, for instance, the addition of the nominal plural suffix *-hā* to personal pronouns in the plural (i.e. i*mūnhā*),[47] a grammatical innovation also found in modern Persian and Azerbaijani. Examples include Pers. ***mā-hā*** "we," ***šomā-hā*** "you" (pl.), and Az. ***bizlər*** "we," ***sizlər*** "you" (pl.).
2. An interesting case is the word ***züürīn***, which appears in the prayer with the meaning "power."[48] This lexeme appears to present an unusual case of deadjectivalisation, assuming that the suffix ***-īn*** is etymologically identical to the Persian adjectival suffix ***-īn*** (< Middle Persian ***-ēn***) as in ***āhanīn*** "made of iron", ***čūbīn*** "wooden", ***pašmīn*** "woolen," etc.[49]
3. Particular attention must be paid to instances where past verbal stems are used in subjunctive and imperative moods, most notably the form ***būyu*** meaning "may it be" or "let it be" (see B3(4)), and ***behiš***, with the sense of "forgive" or "let [one] be forgiven" (see B3(7)). What is especially noteworthy, however, is the coexistence of two different verbal stems for the same mood and tense, namely, the non-past subjunctive, in the case of the verb ***biran*** (< NP ***būdan***) meaning "to be, to become." In one case, we find the form ***bāš***, which, while constituting a grammatical innovation in New Persian, has since become regularized and widely used in modern standard Persian. In the Madrasa variety, however, its usage remains marginal, although it does appear in the prayer. In the other case, the form ***būyu*** is employed—again with the meaning "may it be" or "let it be"—which is atypical not only for modern Persian but also for Caucasian Persian and all of its other dialects, with the exception of Madrasa. At the same time, this pattern recalls the coexistence in Early New Persian of innovative forms based on ***bāš-*** alongside the older ***bovāδ*** (SBJV, pr., 3SG) form, suggesting that ***būyu*** may reflect a preserved grammatical archaism in the Madrasa dialect of Caucasian Persian.

---

[47] See *imūnhā-rā* in B3(8).

[48] See B3(11).

[49] Even more notable is the case of denominalisation observed in the Madrasa variety, where *zūr*, which originally meant "power", has been attested in Qalt'axč'yan's records as having acquired the adjectival meaning "strong", "powerful" (see Qalt'axč'yan 1970a: 273). It should be noted that, in addition to the form *zūr*, the word list appendix of Qalt'axč'yan's dissertation also includes the form *zuur*, though with its expected meaning of "power" (Qalt'axč'yan 1970a: 296).

An analysis of the corpus collected by Qalt'axč'yan further reveals a noteworthy distributional contrast between the two dialects (Madrasa and Kilvar) with respect to the expression of subjunctive.

In Madrasa, the dominant form is ***bū-***, attested in 7 cases, with ***bāš-*** appearing only once. In contrast, in Kilvar, the corpus records 3 instances of ***bāš-***, and one case of the highly archaic ***bav-*** form (see the examples below).

Ex. Madrasa **(1)**[50]

Imruz šaangum xəntril mistanəm qonaγmun ***buyind****.*

Kilvar (**1**)

Imbruz šaangum xahiš misaxtänum ***bašind*** imuna qonaqmu.

Transl. "Tonight I ask you to be our guest" (Qalt'axč'yan 1970a: 255).

Ex. Madrasa **(2)**

P'oy gušt čak düšü, kerakur t'əmin ***buyu****.*

Kilvar **(2)**

P'ey gušt xub düšüki, xurek läzätin ***bavu****.*

Transl. "Let the meat cook well, so that the food is delicious" (Qalt'axč'yan 1970a: 257).

Ex. Madrasa **(3)**

Änt'ün ***bašə*** *u bar, k'ämini k'i bə čat'inet'i kurabiränbu.*

Kilvar **(3)**

T'üna ***bašu*** *u mähsul, k'i bä čät'iniyuz večira bu.*

Transl. "Let the harvest that was gathered with difficulty be yours" (Qalt'axč'yan 1970a: 269).

Ex. Madrasa **(4)**

Ägä girk ***nabuyi****, män vamuxt'än namdunum.*

Kilvar **(4)**

Ägä därsk'it'ab ***nabašu****, mä na mutunum amuxt'ä.*

Transl. "If there is no textbook, I cannot study" (Qalt'axč'yan 1970a: 273)

Ex. Madrasa **(5)**

Adä, bä män či k'ənəm, e, bärdand, bärdand, dä, ǰunt'an xaš ***buyə****, saγutün buyə, yeloon az dava saγ – salamat' vagardənd...*

[50] In these examples, it has been deemed appropriate to refrain from making any changes and to retain the original transcription and orthography (including the separation or concatenation of words) as used by Qalt'axč'yan.

Transl. "Oh boy, what can I do? Well, if they took it, they took it… May you live, may there be health, and may our children return safely from war" (Qaltʿaxčʿyan 1970a: 226).

Ex. Madrasa **(6)**

Härči mxastän ***buyu*** *män in kilärä berimän zän mevagnəm.*

Transl. "Let whatever they want be—I'm marrying this girl" (Qaltʿaxčʿyan 1970a: 244).

Ex. Madrasa **(7)**

sayutün ***buy***

Transl. "Let there be health" (Qaltʿaxčʿyan 1970a: 331).

As for the second instance of the use of a past stem in the imperative mood, i.e. ***behiš***, which appears in the prayer with the meaning "forgive" or "let go", this form is unusual not only for other dialects of Caucasian Persian but even for the Madrasa variety itself. The appearance of the past stem of ***heštan/hištan*** in the imperative may be, on the one hand, the result of a typographical error, possibly caused by the printer's unfamiliarity with the language and a misreading of the handwritten Armenian letter [և] as [շ]. After all, it is the ***hil-*** present stem of ***heštan*** that is recorded in the 15th-century Armenian-script version of the *Lord's Prayer* (see Appendix), following the verbal prefix.

Moreover, even assuming there is no typographical error and that *hiš-* is indeed the intended form, the simplification of the ***št*** cluster to ***š*** also appears problematic. In Caucasian Persian, including the Madrasa variety, clusters containing *-t* tend to preserve this segment more robustly in processes of reduction (cf. ***sātan*** "to do" < ***saxtan***). In the lexical appendix to Qaltʿaxčʿyan's dissertation (p. 310), the verb ***heštan*** is glossed as "to let go", with a comparison to Persian ***heštan*** also offered. The verb additionally appears in two compound forms: ***darun heštan*** "to let in" (p. 291) and ***rāv heštan*** "to release, to forgive" (p. 339).

Unfortunately, the form in question does not occur in the imperative in the textual corpus of Qaltʿaxčʿyan's dissertation, which prevents further verification. It is attested only once in the non-past subjunctive (irrealis modality), and in that instance the *[l]* of the ***hil-*** present stem is retained, which again supports the possibility of a typographical error. See the example below:[51]

[51] In this and the following examples as well, Qaltʿaxčʿyan's transcription has been preserved.

Ex. Madr. (1) *In zän mït'k' mistän, axi in t'ämbäl märdä bəxunä taynä čitär* ***beylə****, axi az näxardän imürə.*

Transl. "This woman is thinking about how to leave this weak man alone at home, he'll die of hunger!" (Qaltʿaxčʿyan 1970a: 222).

This stem, ***hil***, is used in the imperative in the form ***bil*** **(<** ***behil*****)** in several Tati dialects in so-called *let*-constructions, but it is not attested with such a function in the Madrasa variety. There, instead of ***bil***, the form ***poy*** (present stem of ***poyistan*** "to stand") is used. See the examples below:

Ex. Madr. (1) *A Horomsim, moy-piyärt'ənа, qoşiyatəna näärä k'ənə,* ***p'oy*** *vinənd či čaka ärüs birey.*

Transl. "Hey Horomsim, call your father and mother, and the neighbors too—let them see what a fine bride you've become" (Qaltʿaxčʿyan 1970a: 222).

Ex. Madr. (2) ***P'oy*** *gušt čak düšü, kerakur t'əmin buyu.*

Transl. "Let the meat cook well, so that the food is delicious" (Qaltʿaxčʿyan 1970a: 257).

Ex. Af. (3) *Darz tār-ü,* ***bil*** *xišk bašū.*

Transl. "The sheaves are wet, let them dry " (Afurja variety, Grjunberg 1963: 160).

Ex. Gand. (4) *… ou biyar, pur san,* ***bil*** *düšü.*

Transl. "Bring water, make fire, let it boil " (Afurja variety, Grjunberg 1963: 195).

4. Due to direct translation from Armenian, the position of the adjectival modifier within the possessive attributive structure has also been altered (see B3(4). In Caucasian Persian, Azerbaijani, as well as in modern colloquial Eastern Armenian and several Armenian dialects, the canonical structure of the possessive attributive construction is as follows:

   DPPR (= ADJ MOD) + (n. (= head) + PC)

For example:

Caucasian Persian	*mana kitab-man*
Azerbaijani	*mənim kitab-ım*

MEA dialects *im girk-s*
"my book"

In the Caucasian Persian translation of the prayer that follows the Classical Armenian model, however, we encounter a possessive attributive structure in which the adjectival modifier follows the head noun, producing the following order in the Classical Armenian column of B1:

DOM + n. (= head) + DPPR (= ADJ MOD)

Strikingly, the same word order, where the head precedes the adjectival modifier, is preserved in the Caucasian Persian translation as well, thereby deviating from the typologically expected structure of the language:

(n. (= head) + PC + DOM) + DPPR (= ADJ MOD) → *(nū+mūn+ā + imūnā)*

Glossary

The vocabulary found in the prayer is quite interesting both in terms of loanwords and in the use of certain word forms and expressions that are native Tati, yet specific only to the Madrasa variety, thus distinguishing it from other dialects of Caucasian Persian.

1. Two loanwords stand out for their uniqueness. The first is the Turkic loanword ***gök***, meaning "sky," which appears not to occur in other dialects (see C1(3) below). The second is ***čâr***, meaning "evil," which likewise does not occur in other dialectal varieties and appears in the Madrasa dialect (including the translation of the prayer) as a loanword of Armenian origin (see B3(9), also C1(23) below).
2. In total, only two native Turkic words are used in the prayer, i.e. ***gök*** and ***bela*** "in this manner" (see B3(13) and C1(29) below). Two additional words, ***ṭahar*** (see B3(5), C1(13)), and ***şükür*** (see B3(12), C1(27)), although of Arabic origin,[52] appear, based on their phonological form, not to have entered the dialect directly via Arabic or even Persian, but rather through Azerbaijani or a Transcaucasian Turkic variety. In these two cases, the trajectory of borrowing can thus be reconstructed approximately as follows: Arabic > New Persian > Azerbaijani or a Transcaucasian Turkic variety > Madrasa variety of Caucasian Persian.

[52] For *ṭahar* see also Pastor & Suleymanov 2023.

3. The compound form ***böü ṭahar***, as used in the prayer, performs two distinct functions. In the first instance, it appears within the construction ***čī ṭahar ... böü ṭahar***, corresponding to the English correlative structure *"just as ... so also"* (see B3(5)), and is thus equivalent to the *"... so also"* segment in English. In all subsequent occurrences, i.e. B3 (7), (8), (9), (11), 12) ***böü ṭahar*** functions as a discourse continuity marker (DCM) with the meaning *"also"* or *"in the same way"*, and corresponds semantically to the Modern Persian form ***hamčenīn***.
4. A particularly interesting lexical example is the imperative form ***āzāt kï*** "free (me), save (me)" (see B3 (9)), in which the second, verbal component derives from the present stem of the verb ***kardan***, i.e. ***kï-*** or ***kïn-***. This example is noteworthy precisely because it is not attested in other Tati dialects, due to a semantic shift whereby the verb acquired a pejorative or cursive meaning, roughly equivalent to the Latin *coire*. This semantic development subsequently led to the avoidance of the form and its eventual removal from normative usage. In its place, the original (etymonic) meaning was transferred to another verb, *saxtan*, which in this language came to fulfill the same functions as *kardan* does in Persian. Interestingly, the derivative *satan* (from *saxtan*) has gained wide usage in Madrasa as well and, following the pattern observed in other Tati dialects, has come to perform the same functional role as *kardan*, including in compound verb constructions. However, a number of forms derived from *kardan* do occur in Qaltʿaxčʿyan's corpus, albeit only in the present tense, and all such instances should be considered archaisms. These forms are as follows.

Ägä zu mebrey, nuna hazər ***k'ənəm***.

Transl. "If you're leaving early, I'll prepare the meal" (Qaltʿaxčʿyan 1970a: 273).

Ägä umun mxastänink' gändümä çak' k'ura k'ınink', umun t'aba şaangum be k'ar ***k'ınink'***.

Transl. "If we want to gather the wheat properly, we must work until evening" (Qaltʿaxčʿyan 1970a: 272).

*U ama bə k'inəmän, namak yazmiş****k'ənə***.

Transl. "He came to me to write a letter" (Qaltʿaxčʿyan 1970a: 257).

Ye ruz gözät ***k'ən*** *invini.*

Transl. "Watch for a day, observe, and you will see" (Qaltʿaxčʿyan 1970a: 242).

Adä, bä män či ***k'ənəm****, e, bärdand, bärdand, dä, ǰunt'an xaš buyə, saγutün buyə, yeloon az dava saγ – salamat' vagardənd...*

Transl. "Oh boy, what can I do? Well, if they took it, they took it… May you live, may there be health, and may our children return safely from war" (Qalt'axč'yan 1970a: 226).

Essentially, in terms of the semantic development of *kardan*, the Madrasa dialect has followed the same trajectory as the rest of the Tati dialect continuum—with the important distinction that it has retained certain archaic forms built on its present stem.

C1. Word list from the Caucasian Persian translation of the *Lord's Prayer*

1	***piyar*** (n.)[53]	"father"	< Ir. origin	cf. MP *pidar*,[54] NP *pidar, padar* [55]
2	***ki*** (rel. pn.)	"that", "which", "who"	< Ir. origin	cf. MP *kē*, NP *ki* , also Az. *ki*[56]
3	***gök*** (n.)[57]	"heaven", "sky"	< Turk. origin	cf. OT *kö:k (g-)* "sky", "sky-coloured", "sky", "blue", "blue-grey"[58], Az. *göy* "sky", "blue", "green"[59]

[53] Cf. UpSh., Madr., ArDagh., Sur. *piyar*. For the purpose of comparison with the dialects of Caucasian Persian, including later forms recorded in Madrasa, the dialectal examples given here are drawn mainly from the following sources: Suleymanov (2020) for the Upper Shirvan variety (Lahij and adjacent territories); Qalt'axč'yan (1970a) for the Madrasa variety; Mammadova (2018) for the Surakhani variety; and Soltanov & Soltanov (2013) for the Arüsküsh and Daghqushchu variety.

[54] The Middle Persian words used are taken from the Pahlavi dictionary compiled by MacKenzie (1966).

[55] The new Persian words are given according to the dictionary of Steingass (1892).

[56] An Iranian loan, attested across multiple Turkic languages, including Turkish and Azerbaijani of Oghuz group, as well as some non-Turkic languages spoken in the territory of present-day Turkey (for some details see, e.g. Haig 2001: 200-202).

[57] Cf. UpSh. *havo*, *osimon* (Suleymanov 2020: 170, 329) Madr. *gög* (Qalt'axč'yan 1970a: 288), *yergink* (Qalt'axč'yan 1970a: 313), ArDagh. *āsiman, hovva, 'arš,* Sur. *āsmān* (Bérésine 1853: 21), *zavār* (Mammadova 2018: 30; 185), *hovo* (Mammadova 2018: 265), Xalt., Gand. (Grjunberg 1963: 183; 209).

[58] The Old Turkic word forms are taken from Clauson's dictionary (1972: 708-709). Old Turkic refers to those Turkic varieties spoken approximately between the 5th and the early 12th centuries (Róna-Tas & Berta 2011: VII).

[59] The Azerbaijani examples are drawn from the four-volume *Explanatory Dictionary of the Azerbaijani Language* (ADİL 2006).

4	***tamüz*** (adj.)[60]	"pure", "clean", "holy"	< Arab. origin	cf. NP *tamīz* Az. *təmiz* Arm. *tamiz, tamuz*[61] (dial.)
5	***biran*** (v.), pr. stem *bāš-*, pst. stem *bū-*	"to be", "to become"	< Ir. origin	cf. MP *būdan* "to be", "to become", NP *būdan* "to be", "to become", "to exist"
6	***tüna*** (pn.)[62]	"your"	< Ir. origin[63]	cf. NP *tō rā* [64]
7	***nūm*** (n.)[65]	"name"	< Ir. origin	cf. MP *nām*, NP *nām*
8	***āmaran*** (v.) pr. stem o-/*ā-*/*a-*, [66] pst. stem *āmār-*[67]	"to come"	< Ir. origin	cf. MP *āmadan*, NP *āmadan*
9	***ǰannat*** (n.)	"heaven"	< Arab. origin	cf. NP *ǰannat* Az. *cənnət* Arm. ǰannatʿ, ǰanatʿ, (dial.) ǰennetʿ

[60] Cf. ArDagh. *pāk* "clean" (Soltanov & Soltanov 2013: 119), UpSh. *tamiz* "pure" (Suleymanov 2020: 142), Madr. *tamiz* "clean", *tamiz sātan* "to clean", *surp* "holy" (Qaltʿaxčʿyan 1970a: 297-298, 338), Sur. *gülamöra, tamiz* "clean" (Mammadova 2018: 337, 346).

[61] The Armenian dialectal forms used here for comparison are drawn from the *Dialectological Dictionary of the Armenian Language* (HLBB: 2001–2012).

[62] Cf. the form *türa* in all other varieties of Caucasian Persian.

[63] *tü* + DOM, i.e. *na (<rā)*.

[64] Compare purely in terms of etymological properties, but under no circumstances in terms of semantic or functional equivalence, particularly in structures such as *tüna nūm tū* "your name".

[65] Cf. ArDagh., UpSh., Gand. *num* (Soltanov & Soltanov 2013: 161; Suleymanov 2020: 206; Grjunberg 1963: 191), Madr. *nïm, num* (Qaltʿaxčʿyan 1970a: 320, 228, 238), Sur. *nām* (Mammadova 2018: 340).

[66] In the prayer, the present stem (i.e. *o-/ā-/a-*) of the verb *āmaran* "to come", is attested in the non-past subjunctive (/imperative) form *biyov*, which is translated as "may he/she come" or "let him/her come" (cf. Az. *gəlsin* "let him/her come"). Taking into account the usual shift *-ā(h)* > *-ov* in Auslaut observed in the Madrasa dialect (e.g. *gīyov* < *gīyāh* "grass"), one may hypothesize the existence of a parallel form *biyā(y)* in other Armeno-Tat varieties, a form also attested in later field material from Kilvar, such as *biye* "let him/her come", "may he/she come" (Qaltʿaxčʿyan 1970a: 251).

[67] The infinitive also appears in the forms *imaran* and *omaran* in certain compound verbs, such as *vārīš imaran* "to rain" and *tur omaran* (< EA *dur gal*) "to be pleasing" or "to appeal to" (Qaltʿaxčʿyan 1970a: 336, 338).

10	***dāran*** (v.), pr. stem. *ti(n)-*, [68] pst. stem *dār-*	"to give"	< Ir. origin	cf. MP *dādan* NP *dādan*
11	***xīyāl*** (n.)	"thought", "will"	< Arab. origin	cf. NP *xīyāl* Az. *xəyal, xiyal* "dream" Arm. *xial, xayal, xïyal* (dial.)
12	***či*** (pn.)	"what"	< Ir. origin	cf. MP *čē* NP *či*
13	***ṭahar*** (n.)	"kind"	< Arab. origin	cf. NP *ṭour* Az. *təhər* Arm. *t'avur* (dial.)
14	***bi/bö/bï*** (prep.)	"in", "at" (loc.), and "to", "into" (dir.)	< Ir. origin	cf. MP *pad* NP *ba(h)/ be(h)*
15	***ü*** (pn., in böü)	"that"	< Ir. origin	cf. MP *ān* NP *ān* CP *ūn*
16	***hištan*** (v.), pr. stem. *hil-*, pst. stem *hiš(t)-*	"to let", "to leave", "to permit"	< Ir. origin	cf. MP *hištan* cf. NP *hištan*
17	***imūnā*** (pn.)[69]	"our", "us"	< Ir. origin	cf. NP *mā rā* only for the meaning of "us"
18	***qartxuhund*** (n.)	"debtor", "creditor"	*qart* < Arab. *xuhund* < Ir.	cf. NP *qarż* "debt" cf. NP *xudāwand* "lord", "master", "possessor"
19	***(ā)vārdan*** (v.)[70], pr. stem. *ār(n)-*, pst. stem *āvārd-*[71]	"to bring"	< Ir. origin	cf. MP *āwurdan* NP *āwardan*
20	***xaṭā*** (n.)	"fault", "error", "miss"	< Arab. origin	cf. NP *xaṭā*

[68] See B3(6).
[69] *imūn+ā* (< NP *rā* < MP *rāy* < OP *rādī/rādiy*). For the context, see B3(6).
[70] See also Qalt'axč'yan 1970a: 277.
[71] Cf. UpSh. *vordān* (Suleymanov 2020: 377), Sur. *āvārdan* (Mammadova 2018: 157).

21	***āzāt kardan*** (v.), pr. stem *kĭ-*(?), pst. st. *kard-*(?) [72]	"to free", "to liberate"	< Ir. origin	cf. NP *āzād kardan*
22	***a*** (prep.)[73]	"from"	< Ir. origin	cf. MP *az* (< OIr. **hača-*) NP *az* CP *a* (e.g., *a kojā* "from where?")
23	***čâr*** (n.)	"evil"	< Arm. origin	cf. Arm. *čʿar (չար)*
24	***āntü(n)*** (pn.)[74]	"yours"	< Ir. origin	cf. NP *az ān-i tō*
25	***züürīn*** (n.)	"power"	< Ir. origin	cf. MP *zōr* "power" and *zōrīg* "powerful" NP *zūr*, *zor*, *zorī* "power"
26	***va*** (cnj.)[75]	"and"	< Arab. origin	cf. NP *va*
27	***šükür*** (n.)	"glory", "praise"	< Arab. origin	cf. NP *šukr* Az. *šükür*
28	***ʿaẓïm*** (adj.)	"great", "huge", "immense"	< Arab. origin	cf. NP ʿaẓīm Az. əzim
29	***bela*** (adv.)	"thus", "in this way"	< Turk. origin	cf. Az. *belə*
30	***tā bï*** (prep.)[76]	"till", "up to", "untill"	< Ir. origin	cf. NP *tā be* Az. *n.(+dat.) + qədər*

[72] Cf. āzād sātan "to free", "to liberate" in Qaltʿaxčʿyan's dissertation (1970a: 275).
[73] For the context, see B3(9).
[74] This independent possessive pronoun has also been attested in the later period in Madrasa as *antün* (Qaltʿaxčʿyan 1970a: 104). Cf. also UpSh. *anti* (Suleymanov 2020: 98).
[75] This conjunction has extremely limited usage across all varieties of Caucasian Persian, including Madrasa, and its occurrence in the prayer is not entirely typical for the dialect.
[76] Although composed of two separate prepositions—*tā* ('until, to') and *bï* (1. 'to, towards', 2. 'in, at', cf. MP *pad*, and NP *ba*)—this fixed combination functions as a single complex preposition meaning 'till, up to' in a temporal sense. Despite their individual meanings, the sequence *tā bï / tābï* forms a grammatical unit in usage, closely corresponding to the equivalent construction in standard Persian. Qaltʿaxčʿyan (1970a: 168) also defines it as a single preposition, without discussing its origin or etymology. Similarly, Suleymanov (2020: 223) holds that, synchronically, this combination should be regarded as a single preposition.

31	***axïr*** (n., intj.)	"end"	< Arab. origin	cf. NP *axar* Az. *axır* Arm. *axr* (dial.)
32	***āmīn*** (intj.)	"Amen", "so be it", "truly"	< Hebr. origin	cf. NP *āmīn* Az. *amin* arm. *amin*

In lieu of a conclusion

The Madrasa variety of Caucasian Persian is now an extinct dialect. Following the Kilit dialect of the Tatic subgroup within the Northwestern Iranian (NWI) branch, once spoken in Nakhijevan, this is the second Iranian dialect of the Transcaucasian region to have become extinct. In this case, the extinct variety belongs to the SWI branch, and despite its highly archaic character and remarkably interesting grammatical features, it too has fallen out of use.

Regrettably, the linguistic material documented in the Madrasa variety remains unpublished. The *Lord's Prayer*, which forms the focus of this study, is one of the few surviving fragments printed in the Armenian script and has received little scholarly attention. It is now time for the linguistic data from the Madrasa dialect to be systematically analyzed and published, so that this variety becomes accessible to scholars working in Iranian dialectology and historical linguistics.

From the perspective of Iranian dialectological research, particularly the study of lesser-known minor dialects and varieties within the SWI group, and given the need to publish still-unreleased materials on sub-varieties of Caucasian Persian, the translation and publication of Qalt'axč'yan's dissertation would be of considerable value.

More broadly, studying Caucasian Persian, and specifically the Madrasa variety, within a comparative-historical framework could significantly deepen our understanding of the historical development of Persian. It may, for example, help address questions such as the grammatical problem raised by Carina Jahani (2008: 171). Lesser-studied minor languages and dialects of the SWI group, such as Caucasian Persian, may well serve as a *kelīd-e moškelgošā*, a "key to unlocking complex problems".

Appendix. The 15th century "Persian" translation of *Lord's prayer*, written in Armenian script (Ms. 7117, fol. 145a, Matenadaran), published by Hakob Papazyan (1964: 219)

Original (as published by Papazyan)	**Papazyan's Reconstruction (in Persian script)**
Փի[դ]արի ման քի բար ասմանի։	پدر مان که بر آسمانی
Փաքաստ ու նամի թու։	پاک است و نام تو
Բիայաթ բեհէշտի թու։	بیاید بهشت تو
Հաստ ու քամի թու, համչին քի բար ասմանի ու դարյալամի։	هست و کام تو همچنین که بر آسمانی و در عالمی
Նանի մա մուդամ բըդէ բմա ըռoզ-բառoզ։	نان ما مدام بده بما روز بروز
Բէհէլ դարտի մարա, քի մա բէհիլամ դարտի դարտդարանի մարա։	بهل قرض مارا که ما بهلیم قرض قرضئاران مارا
Մադէ մարա բըդադա[ու] բըբալա, բըռահան ազդադայ ու ազբալա։	مده مارا بقضا و ببلا برهان از قضا و از بلا
Անի թուսթ բէհէշդ։	آن تست بهشت
Ազ թուսթ մoճիզատ։	از تست معجزات
Աբաթըլ յաբաթ։	ابدالابد
Ամին։	آمین

	Transliteration of the original (Hübschmann-Meillet system)	**Rzepka's transliteration of the original 2006: 100), based on Papazyan's publication**
1	P'i[d]ari man k'i bar asmani:	P'idari man k'i bar asmani.
2	P'ak'ast ow nami t'ow:	P'akast u nami t'u.
3	Biayat' behêšti t'ow:	Biayat' behêšti t'u.
4	Hast ow k'ami t'ow, hamč'in k'i bar asmani ow daryalami:	Hast u k'ami t'u, hamč'in k'i bar asmani u dar yalami.
5	Nani ma mowdam bədê bma ərôz-barôz:	Nani ma mudam bədê bəma ə rôz-barôz.
6	Bêhêl łarti mara, k'i ma bêhilam łarti łartdarani mara	Bêhêl γarti mara, k'i ma bêhilam γarti γartdarani mara.
7	Madê mara bəłada[ow] bəbala,	Madê mara bəγada[u] bəbala,
8	bərahan azładay ow azbala:	bərahan azγaday u azbala.
9	Ani t'owst' bêhêšd:	Ani t'ust' bêhêšd.
10	Az t'owst' môčizat:	az t'owst' môğizat.
11	Abat'əl yabat':	Abat'əl yabat'.
12	Amin:	Amin.

Abbreviations (glosses)

-	morpheme separator
<…>	epenthesis, the insertion of a non-etymological segment
[…]	the reconstruction of a segment
1	present stem
2	past stem
1	first person
2	second person
3	third person
ADJ	adjectival
APP	approximative
COP	copula
DCM	discourse continuity marker
DIR	directional preposition
DIS	distal
DOM	differential object marking
DPPR	dependent possesive pronoun
IMP	imperative
INT	intterogative
IPPR	independent possesive pronoun
ISP	imperative-subjunctive prefix
LOC	locative preposition
MOD	modifier
OBL	oblique
PC	pronominal clitic
PL	plural
POSS	possesive
REL	relativiser
SBJV	subjunctive
SBJV/IMP	subjunctive/imperative
SEP	separative preposition
SG	singular

Abbreviations (general and glottonyms)

adj.	adjective
adv.	adverb
Af.	Afurǰa (Quba) variety of Caucasian Persian (Tati)
AO	Armenian original
Arab.	Arabic
ArDagh.	Arüsküsh-Daghqushchu (Central) variety of Caucasian Persian (Tati)
Arm.	Armenian
Az.	Azerbaijani (Transcaucasian Turkish)
cnj.	conjunction
CP	Colloquial Persian

CPers	Contemporary Persian
CPT	Caucasian Persian translation
dat.	dative
dial.	dialectal
EA	Eastern Armenian
ENP	Early New Persian
ex.	example
Gand.	Gandov (Shabran) variety of Caucasian Persian (Tati)
Hebr.	Hebrew
inf.	infinitive
intj.	interjection
Ir.	Iranian
Madr.	Madrasa (Shamakhi) variety of Caucasian Persian (Tati)
MEA	Modern Eastern Armenian
MP	Middle Persian
n.	noun
NP	New Persian
NWI	Northwest Iranian
OIr.	Old Iranian
OP	Old Persian
Osm.	Osmanli, i.e. Ottoman Turkish
OT	Old Turkic
Pers.	Persian
pn.	pronoun
pr.	present
prep.	preposition
pst.	past
rel.	relative
RFP	reconstructed form and pronunciation
Sur.	Surakhani variety of Caucasian Persian (Tati)
SW	south-western [language group]
SWI	Southwest Iranian
Tkm.	Turkmen language
Turk.	Turkic
UpSh.	Upper Shirvan variety of Caucasian Persian (Tati)
v.	verb
Xalt.	Xaltan (Quba) variety of Caucasian Persian (Tati)

BIBLIOGRAPHY

Acharean, H. (1941), “Ałvanicʿ girə” [The alphabet of Caucasian Albanian], *Tełekagir SSŘM GA haykakan filiali* [Bulletin of the Armenian branch of the Academy of Sciences of the SSSR], no. 3-4 (8-9), pp. 3-11. In Armenian.

ADİL (2006), *Azərbaycan dilinin izahlı lüğəti* [Explanatory dictionary of the Azerbaijani language], 4 vols., ed. Ə. Orucov, Bakı: Şərq. In Azerbaijani.

Akopjan, A. (2006), "Tatojazyčnye armjane: 15 vekov ot "armjanstva" k "armjanstvu"" [Tat-speaking Armenians: 15 centuries from "Armenianhood" to "Armenianhood"], in: *Južnyj Kavkaz: Territorii. Istorii. Ljudi* [South Caucasus: Spaces. Histories. People], ed. E. Gerasimova, and N. Ležava, Heinrich Böll Stiftung, Tbilisi: Diogene, pp. 190-210. In Russian.

Alishan, Ł. (1896), *Kamenic': Taregirk' hayoc' Lehastani ew Rowmenioy, hawastč'eay yawelowacovk'* [Kamenits. Annals of the Armenians of Poland and Romania, with authentic documents], Venice: San Lazzaro. In Armenian.

[Eric'yan] Andranik (1900), *Tersim. Čanaparhordowt'iwn ew tełagrowt'iwn* [Dersim: Travel and Topography], Tiflis: Mn. Martiroseanc' Press.

Authier, G., and Tonoyan, A. (2018). "Defining features of the Madrasa dialect of Armeno-Tat", in: *8th International Conference on Iranian Linguistics (ICIL-8). Program and Abstracts*, Yerevan: YSU Publishing House, pp. 67-68.

Bérésine, É. (1853), *Recherches sur les dialectes persans*, Kazan: Imprimerie de l'Université de Casan.

Catalogue of Manuscripts of the Mesrop Mashtots Matenadaran, vol. 2, ed. L. Khachikyan and A. Mnatsakanyan, Yerevan: Publishing house of the Academy of Sciences of the Armenian SSR, 1970.

Chirikba, V. (2015), "Between Christianity and Islam: Heathen Heritage in the Caucasus", in: *Studies on Iran and The Caucasus*, ed. U. Bläsing et al., Leiden: Brill, pp. 145-191.

Clauson, S. G. (1972), *An Etymological Dictionary of Pre-Thirteenth Century Turkish*, Oxford: Clarendon Press.

Dum-Tragut, J. (2009), *Armenian: Modern Eastern Armenian*, Amsterdam: John Benjamins.

Jndi, H. (1962), "Mesropyan grerə ew k'urd žołovrdi kulturan" [Armenian alphabet and the culture of Kurds], *HSSŘ GA Tełekagir hasarakakan gitut'yunneri* [Bulletin of the Academy of Sciences of the Armenian SSR], no. 5, pp. 67-74. In Armenian.

Êfendyan, P. (2004), "Madrasac'ineri erklezvut'yan harc'i šurǰ" [On the problem of bilingualism among Madrasa residents], *Lraber hasarakakan gitut'yunneri* [Herald of Social Sciences], issue 1, pp. 88-96. In Armenian.

Jahani, C. (2008), "Expressions of future in Classical and Modern New Persian", in: *Aspects of Iranian Linguistics*, ed. S. Karimi, V. Samiian, and D. Stilo, Newcastle: Cambridge Scholars Publishing, pp. 155–176.

Łazaryan, R., and Avetisyan, H. (2009), *Mǐǰin hayereni bařaran* [Dictionary of Middle Armenian], 2nd ed., Yerevan: YSU Publishing House. In Armenian.

Qalt'axč'yan, S. (1957), "Hamařotaki Madrasayi hay-t'at'ereni masin" [Briefly on the Armeno-Tat of Madrasa], *Haykakan SSR Gitut'yunneri akademiayi*

tełekagir hasarakakan gitut'yunneri [Bulletin of the Academy of Sciences of the Armenian SSR – Social Sciences], issue 4, pp. 85-96. In Armenian.

Qalt'axč'yan, S. (1970a), *Madrasac'ineri t'at'eren lezun* [The Tati language of the residents of Madrasa], Candidate of Sciences dissertation defended at the Academy of Sciences of the Armenian SSR Sector of Oriental Studies, Yerevan. In Armenian.

Qalt'axč'yan, S. (1970b), "Hayereni tarrerə hay-t'at'akan barbařum" [On the elements of Armenian language in the Armeno-Tat dialect], *Lraber hasarakakan gitowt'yownneri* [Herald of Social Sciences], issue 3, pp. 72-77. In Armenian.

Qalt'axč'yan, S. (1970c), "Erklezvut'yan harc'i šurǰ" [On the problem of bilingualism], *Sovetakan mankavarž* [Soviet Pedagogue], issue 5, pp. 82-83. In Armenian.

Grigoryan, V., and Pisowicz, A. (1964), "Hayatař leheren vaveragrerə" [Polish documents in Armenian script], *Banber Matenadarani* [The Bulletin of Matenadaran], vol. 7, pp. 225-236. In Armenian.

Grjunberg, A. (1963), *Jazyk severoazerbajdžanskix tatov* [The language of the Tats of Northern Azerbaijan], Leningrad: Izdatel'stvo Akademii Nauk SSSR. In Russian.

Hacıyev, M. (1971), *Tat dilinin Qonaqkənd ləhcəsi* [The Qonaqkənd dialect of Tati], Baku: Elm. In Azerbaijani.

Hacıyev, M. (1995), *Toy-düyün xəzinəmizdən: Azərbaycan tatlarının tarixi-etnografik oçerki* [From our wedding treasure-house: a historical and etnographic sketch of Azerbaijani Tats], Baku: Tat Cultural Centre "Azeri". In Azerbaijani.

Haig, G. (2001), "Linguistic diffusion in present-day East Anatolia: from top to bottom", in: *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, ed. A. Y. Aikhenvald and R. M. W. Dixon, Oxford: Oxford University Press, pp. 195–224.

Hakobian, A. (2001), "Hay-t'at'eri tełabašxowacowt'ean ew tełahanowt'iwnneri patmowt'iwnic' (XVIII dari skizb - XIX dari verǰ)" [From the history of the demographic distribution and deportations of Armenian-Tats (early 18th - late 19th cc.)], *Haigazian Armenological Review*, vol. 21, pp. 125-142. In Armenian.

Hakobian, A. (2002), "T'at'axos hayeri lezvakan mšakuyt'i šurǰ (18-rd d. - 20-rd d. skizb)" [On the linguistic culture of Tatspeaking Armenians during 18th-early 20th cc.], *Echmiatsin*, no. 9, pp. 70-82. In Armenian.

HLBB (2001-2012), *Hayoc' lezvi barbařayin bařaran* [Dialectological dictionary of the Armenian language], 9 vols., ed. A. Sargsyan, L. Hovhannisyan, N. Sargsyan, R. Toxmaxyan, and R. Urutyan, Yerevan: Gitutyun. In Armenian.

Hüseynova, G. (2002), *Lahıc tatlarının dili* [The language of the Tats of Lahıc], Baku: Nurlan. In Azerbaijani.

K'vač'aʒe, M. (1988), *Masalebi gomborel laiǯta met'q'velebis šesc'avlisatvis: šesavali, t'ekst'ebi, leksik'oni* [Materials for the study of the dialect of the Gombori Laij: Introduction, texts, vocabulary], Tbilisi: tbilisis universit'et'is gamomcemloba. In Georgian.

Kirakosian, H. (2020), "The Orthographic Rules of the Eighteenth Century Armeno-Persian Gospels of the Matenadaran (Ms. 8492, Ms. 3044)", *Iranian Studies*, vol. 53, no. 1-2, pp. 295–330.

Kirakosyan, H. (2017), *Catalogue of the Persian manuscript fragments of Matenadaran. Bilingual Armenian-English edition*, Yerevan: Nairi.

Kirakosyan, H. (2018), "Maštocʿyan Matenadarani hayatař parskeren erku jeřagir Avetaranneri hišatakaranneri masin (MM. 3044, MM. 8492)" [On the colophones of two Armeno-Persian manuscripts of the Matenadaran (Ms. 3044, Ms. 8492)], *Echmiatsin*, no. 5, pp. 47-61. In Armenian.

Kirakosyan, H. (2020a), *Matenadarani hayatař parskeren Avetarani jeřagrerə. kʿnnakan usumnasirutʿyun* [The Manuscripts of Armeno-Persian Gospel of the Matenadaran. Critical study]. Yerevan: Tigran Mets. In Armenian.

Kirakosyan, H. (2020b), "Ganjakyan hayatař parskeren jeřagir avetaranə vičabanakan tʿargmanutʿyunneri hamatekʿstum" [The Armeno-Persian Handwritten Gospel from Gandzak in the Context of Interreligious Polemic Translations], *Banber Matenadarani*, vol. 29, pp. 281-298. In Armenian.

Lopatinskij, L. (1894), "Armjano-tatskie teksty" [Armeno-Tati texts], *Sbornik materialov dlja opisanija mestnostej i plemen Kavkaza* [A collection of materials for describing the places and tribes of the Caucasus], vol. 20, section 2/2, pp. 25-32. In Russian.

MacKenzie, D. N. (1986), *A Concise Pahlavi Dictionary*, 2nd ed., London: Oxford University Press.

Maisak, T. (2019), "Perevody «Otče naš» v istorii udinskogo jazyka" [Translations of the Lord's Prayer in the Udi language], *Rodnoj jazyk* [Mother tongue], no. 1, pp. 114-145. In Russian.

Mammadova, N. (2018), *Eléments de description et documentation du tat de l'Apshéron, langue iranienne d'Azerbaïdjan* [thèse de doctorat non publiée]. Paris: INALCO (Institut national des langues et civilisations orientales).

Miller, B. (1929), "Taty, ix rasselenie i govory (materialy i voprosy)" [The Tats, their settlement and dialects (materials and issues)], *Obščestva obsledovanija i izučenija Azerbajdžana* [Society of investigation and study of Azerbaijan], vol. 8, issue 7, pp. 3-33. In Russian.

Miller, B. (1945), "Tatskie teksty (Materialy po govoram tatov Sovetskogo Azerbajdžana)" [Tati texts: Materials on the dialects of the Tats of Soviet Azerbaijan], in: *Iranskie jazyki* [Iranian languages], vol. 1 (Iranica 3), ed. V. Abaev, and I. Meščaninov, Moscow - Leningrad: Izdatel'stvo Akademii Nauk SSSR, pp. 107-126. In Russian.

Miller, V. (1907), *Tatskie ètjudy II. Opyt grammatiki tatskogo jazyka* [Essays on the Tati language II: Towards a Grammar of Tati], (Trudy po

vostokovedeniju 26 [Works on Oriental Studies 26]), Moscow: Tipografija Vjač. Al. Gatcuk. In Russian.

Papazyan, H. (1964), "Mesropatař aylalezu grakanutʿyan masin" [On the foreign literature in Armenian characters], *Banber Matenadarani* [The Bulletin of Matenadaran], vol. 7, pp. 209-224. In Armenian.

Pastor, D., and Suleymanov, M. (2023), "Sort of grammaticalized: morphology of new approximative suffixes in Persian and Azeri", *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, vol. 173, no. 2, pp. 411-440.

Róna-Tas, A., and Berta, Á. (2011), *West Old Turkic. Turkic Loanwords in Hungarian*, Wiesbaden: Harrassowitz Verlag.

Rzepka, M. (2006), "Ey pedar-e mā. ˝Ojcze nasz˝ w języku perskim", in: *In the Orient where the Gracious Light... Satura orientalis in honorem Andrzej Pisowicz*, ed. A. Krasnowolska, K. Maciuszak & B. Mękarska, Kraków: Księgarnia Akademicka, pp. 97-106. In Polish.

Schulze, W. (2005), "Towards a history of Udi", *International Journal of Diachronic Linguistics*, vol. 1, pp. 55-91.

Schulze, W. (2016a), "Textual resources for Udi", in: *Kaukasiologie heute: Eine Festschrift zum 70. Geburtstag von Heinz Fähnrich*, ed. N. Reineck, and U. Rieger, Greiz: König, pp. 361-381.

Schulze, W. (2016b), "Armenian Tats", in: *A Handbook of the Minorities of Armenia. A Sociocultural and Sociolinguistic Survey*, ed. I. Schulze, and W. Schulze, Hamburg: Verlag Dr. Kovač, pp. 425-448.

Smbateancʿ, M. (1896), *Nkaragir Surb Stepʿannosi vanacʿ Sałiani ew miws vanōrēicʿ ew uxtatełeacʿ, ews ew kʿałakʿacʿn giwłōrēicʿ, orkʿ i Šamaxwoy tʿemi* [Description of the Saint Stephen Monastery in Sałian and other monasteries, pilgrimage sites, villages and towns of Shamakhi diocese], Tpʿxis: Tparan Movsēs Vardaneancʿi. In Armenian.

Soltanov, A., and Soltanov, M. (2013), *Tati-türki, türki-tati lüğət (Ərüsküş-Dağ Quşçi ləhceyi)* [Tati-Azerbaijani, Azerbaijani-Tati dictionary (The dialect of Ərisküş - Dağ Quşçu)], 2nd edition, Baku: Qanun.

Steingass, F. (1892), *A comprehensive Persian-English dictionary*, London: W.H. Allen & Company.

Stepanyan, H. (2007), "Hayatař tʿurkʿeren, tʿatʿareren, łpʿčʿałeren, kʿrderen ew hunaren grakanutʿyan masin" [On the literature in Turkish, Tatar, Kipchak, Kurdish and Greek languages, written in the Armenian alphabet], *Merjavor ew Mijin Arewelkʿi erkrner ew žołovurdner* [The Countries and Peoples of the Near and Middle East], vol. 26, pp. 218-224. In Armenian.

Stepanyan, H. (2023a), *Hayatař tʿurkʿeren grakanutʿyunə* [The Turkish literature in Armenian alphabet], Yerevan: Hovsepyan LLC. For English edition see Stepanyan 2023b. In Armenian.

Stepanyan, H. (2023b), *Armeno-Turkish Literature*, Yerevan: Hovsepyan LLC.

Suleymanov, M. (2020a), *A grammar of Şirvan Tat* (Beiträge zur Iranistik, Band 46), Wiesbaden: Dr. Ludwig Reichert Verlag.

Suleymanov, M. (2020b), "Oblique marking and adpositional constructions in Tat: A mosaic of dialectal convergence and divergence" (= Chapter 13), *Advances in Iranian Linguistics* (Current Issues in Linguistic Theory, vol. 351), ed. R. Larson, S. Moradi, and V. Samiian, Amsterdam-Philadelphia: John Benjamins, pp. 275-300.

Suleymanov, M. (2023), "Tat, an endangered language of Azerbaijan, and its speakers: a historical and sociolinguistic overview", in: *Linguistic Diversity in Azerbaijan: Present State and Future Challenges*, ed. J. Garibova, and E. Ragagnin, Ankara: Grafiker, pp. 255-272.

Tonoyan, A. (2015), *Kovkasyan parskeren. patmahamematakan usumnasirutʿyun* [Caucasian Persian: a comparative-historical research]. PhD dissertation defended at the Institute of Linguistics of the National Academy of Sciences of Armenia. Yerevan. In Armenian.

Tonoyan, A. (2018), "Kovkasyan parskerenum bun apařni žamanakajewi kazmutʿyan oroš ařanjnahatkutʿyunneri šurǰ" [On certain peculiarities in the formation of the absolute future tense in Caucasian Persian], *Arewelagitutʿyan harcʿer* [Journal of Oriental Studies], vol. 14, pp. 366-380. In Armenian.

Tonoyan, A., and Voskanian, V. (2024), "Caucasian Persian (Tati): History of Study, Current State and Perspectives", in: Mohammad Afshin-Vafaie & Pezhman Firoozbakhsh (eds), *Studia Persica 23 (in Memory of Dr. Mahmoud Afshar Yazdi)*. Tehran: Dr. Mahmoud Afshar Endowment Foundation, pp. 183-216.

Voskanian, V. (2016), "Oroš haykabanutʿyunner kovkasyan parskerenum" [Some Armenisms in Caucasian Persian language], *Arewelagitutʿyan harcʿer* [Problems of Oriental Studies], vol. 12, pp. 240-248. In Armenian.

Yeranyan, G. (2024), *Êǰer hay-kʿrdakan patmutʿyunicʿ. Mas A. Kʿrdaxos hayer, hayatař kʿrderen tpagir grkʿer* [Episodes from Armenian-Kurdish history: Vol. 1: Kurdish-speaking Armenians, printed books in Kurdish language with Armenian script], Yerevan: Matenadaran. In Armenian.