Katarzyna Wąsala

Modern Spoken Persian in Contemporary Iranian Novels

V&R unipress

unipress

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

TRANSitions.
Transdisciplinary, Transmedial and Transnational Cultural Studies
Transdisziplinäre, transmediale und transnationale Studien zur Kultur

Volume / Band 4

Edited by / Herausgegeben von
Renata Dampc-Jarosz and / und Jadwiga Kita-Huber

The volumes of this series are peer-reviewed.
Die Bände dieser Reihe sind peer-reviewed.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Katarzyna Wąsala

Modern Spoken Persian in Contemporary Iranian Novels

An analysis of selected 21st century novels

With 28 figures

V&R unipress

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

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

The publication was funded by a grant from the Faculty of Philology under the Strategic Programme 'Excellence Initiative – Research University' at the Jagiellonian University in Kraków.

Cover image: 06/05/2019 Tehran, Iran, Famous night view of Tehran, Flow of traffic round Tohid Tunnel with Milad Tower and Alborz Mountains in Background, Tohid Tunnel one of longest urban tunnel in Middle East © Alexeiy, Adobe Stock.

Vandenhoeck & Ruprecht Verlage | www.vandenhoeck-ruprecht-verlage.com

ISSN 2751-8345
ISBN 978-3-7370-1638-4

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Contents

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Acknowledgements

The completion of the research project behind this book would not have been possible without the help, support and encouragement of wonderful and inspiring people I met on my way. First and foremost, I would like to express my gratitude to my supervisor, Prof. Kinga Paraskiewicz, for her assistance at every stage of the project and for her patience and understanding whenever it faced obstacles. I would also like to extend my sincere thanks to my assistant supervisor, Dr. Mateusz M. P. Kłagisz, for inspiring comments, invaluable advice and continuous support throughout this lengthy process, and to all the members of my PhD committee. I am also grateful for all the suggestions and comments from outside the Iranian studies field by my PhD thesis reviewers, Prof. Janusz Danecki and Prof. Tadeusz Szczerbowski.

A sincere and heartfelt thank you to Prof. Anna Krasnowolska, Dr. Renata Rusek-Kowalska, Dr. Karolina Rakowiecka-Asgari and the entire team of the Department of Iranian Studies for sharing their passion for the study of Persian language and literature, for all the fruitful discussions, insightful reflections and providing guidance whenever it was needed.

I would also like to thank Prof. Ireneusz Bobrowski and Prof. Rafał Górski for broadening my linguistic horizons during PhD seminars, and to all fellow PhD students with whom we would later engage into lively discussions over scientific matters—the discussions which more often than not were the beginnings of long-lasting friendships.

Last but not least, I wish to thank my family and friends for their patience and encouragement, especially in times of doubt. Without your support, it would not be possible to get to the place where I am today.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Chapter 1.
Introduction

Modern New Persian, a variety of Persian language used in present Iran,[1] is characterized among others by a substantial diversity between the language of writing and the language of speech. The gap between those two is so vast that foreign students of Persian need to be instructed on both varieties in order to be able to communicate effectively with native speakers. Differences lie not only within the realm of vocabulary and phraseology, but reach into the syntactic structures and morphology of the language. Intuitive understanding of the terms "spoken" and "written" suggests that the varieties are mutually exclusive and cannot trespass the borders of their prescribed medium, the reality, however, is much more complicated.[2] The characteristics of spoken variety can be found in dialogue parts of many fictional works and with the expansion of new technologies they are also being typed in all kinds of virtual spaces like social media, blogosphere etc. In general, an expansion of the spoken variety in the domains previously restricted to writing can be witnessed, while the written variety still holds its prestigious position of the language of literature, any kind of written academic discourse, official circumstances *et cetera*.

The history of application of spoken variety of language into fictional writing dates back as far as to the beginnings of modern Persian prose (1920s) and writers

1 Apart from Iran, Persian language is also spoken in Afghanistan (*dari*) and Tajikistan (*tajiki/toǧiki*). Those varieties are often referred to as dialects of Persian, but their official status remains unclear. The problem is not of a purely linguistic character and politics do play an important role here either. See for example Perry (1999), Beeman (2005).

2 In this work, I consequently apply the term Modern Spoken Persian (MSP) to this variety. It should be noted, though, that within the present circumstances regarding both structural differences and conditions of use, it might be time for a re-consideration of it and possible replacement with a term free of reference to the medium/usage (i. e., replacing "spoken" with a more relevant descriptive term). Further in the book, a discussion of the Persian term *mohāvere*, an original Persian name of this variety can be found; unfortunately as of yet, there is no widely accepted translation of this term used in the scholarship.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

such as Ǧamālzāde (1892–1997), Hedāyat (1903–1951), Čubak (1916–1998),[3] although we could actually trace the uses of spoken language in Persian literature even further back—in fact, the rise of the New Persian after the Muslim Conquest owes to the early poets who preferred to compose their poetry in a language spoken at the court, not written in the administrative records; a choice natural if the performative character of that poetry is taken into account,[4] but also seminal for the further development of the language after it started to be recorded in writing. Similarly, the so-called "spoken language"[5] was employed in some of the poetry of the constitutional period (early 20th century) by authors such as Āref Qazvini (1882–1934), who composed song lyrics (*tasnifāt*) as a reaction to political events, drawing from the everyday language and even slang.[6] There are "spoken" poems in *divans* of the well-known and respected contemporary poets using free verse, like Ahmad Šāmlu's *Man-o-to, deraxt-o-bārun* (from the collection *Āyda dar āyine* [Ayda in the mirror]) or Foruq Farroxzād's *Be Ali goft mādaraš ruzi* (from the collection *Tavallodi digar* [Another birth])[7]—but at the same time, using a word belonging to the "spoken" lexicon or a "spoken" sentence in translation of a foreign poem is seen as highly inappropriate.[8] Close

3 Mohammad-Ali Ǧamālzāde (1892–1997), Sādeq Hedāyat (1903–1951) and Sādeq Čubak (1916–1998) were all important intellectuals and prominent representatives of Persian literature, the founders of modern Persian prose. Ǧamālzāde (romanized also as Jamalzadeh) was even praised by critics of his time for an accurate representation of the current spoken language in his stories (Kamshad, Mozaffari 2016).

4 In Persian, the poetry is not 'written' but 'composed' or even literally 'spoken': instead of **še'r neveštan* 'to write poetry' either *še'r sorudan* 'to compose poetry' or *še'r goftan* 'to speak poetry' is used.

5 The quotation marks here indicate the popular rather than academic understanding of the phrase. However historically true, today the term "spoken" does not exactly match the characteristics of the variety in question (neither does the sometimes employed "colloquial," which may give a false impression that in Persian, the discussed difference resembles a difference between standard/formal and colloquial styles known from other languages, including for example Polish; yet the functionally-conditioned and much less optional opposition of the described varieties of Persian is not exactly of the same kind). The problems of terminology regarding written and spoken varieties of Persian will be addressed in the next chapter.

6 According to J. Matini, Āref "sometimes wrote melodious verses in a literary style, but, elsewhere he introduced slang that accorded well with the subject and mood; thus Malek-al-šoʿarāʾ Bahār described him as a "poet of the common people." His most important and impressive works are his taṣnīfāt […]. The taṣnīf had sunk to banality in wording and content, but he was able to impart a poetic quality to it" (Matini, Caton 1986: 391).

7 Ahmad Šāmlu (1925–2000) and Foruq Farroxzād (1935–1967) are two very prominent representatives of contemporary Persian poetry, whose poetry is valued both among the literary critics and readers throughout Iran and abroad.

8 This is an experience I had when translating poetry of Halina Poświatowska from Polish to Persian together with an Iranian poet, who would correct my "inappropriateness" right away. A similar experience was reported to me by a friend who translated a book of interviews with Krzysztof Kieślowski and did that into a spoken variety (they were transcribed conversations after all), but was then kindly asked by the editor to rewrite it into "proper Persian." (Both

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

relationship between the notion of correctness and literature is prevalent in the scholarship, whether conscious or not, which is well visible in the fact that the most reliable source of colloquial forms, that is, the descriptive grammar of Gilbert Lazard, presents the forms collected from… literature, which makes it sound paradoxical (see "Lazard's paradox" in chapter 5.1).

Generally speaking, the interesting question is not if the spoken variety can be used in writing, whether literary or of any other kind, because we know that it can. The question is therefore rather: how, under what circumstances and for what purpose it is done.

The problem of differences between spoken and written varieties of modern Persian has been studied from more than one angle, as will be presented in the third chapter, wherein several theoretical approaches to the subject are discussed. Still, in many grammatical discussions it is taken for granted, and either one or the other variety is dismissed to the side notes. Previously, there was a tendency to concentrate on the written, literary variety as the more proper and accurate rendering of the language. Nowadays, there appear textbooks which represent the opposite point of view and teach the spoken variety as the more useful and up-to-date one (for example Mace 2003, Rafiee 2001).

Some studies on this subject, regarding both classical and contemporary literature, have been published recently in Iran and abroad (see Orsatti 2015, Sayyedyazdi, Hakimi 2015; Dehqān, Hammāmči 2016), but those are mostly case studies, detailed works dealing with specific authors or texts. A general overview of the subject could shed more light on the ongoing interplay between spoken and written varieties of Modern New Persian language.

My personal interest in the application of spoken variety in literature and its functions was sparked by one of the novels analyzed here, namely *Kāfe Piyāno* by Farhād Ğa'fari, during a research stay and a state-funded scholarship at the Dehxodā Institute in Tehrān in 2011–2012. Being immersed into the spoken language of the streets of Iran's capital helped to trace its influences on the literary language of contemporary prose works, not only on the most intuitively distinguished levels of phonology (represented by spelling) and lexis, but also on the more inconspicuous levels such as verbal morphology or syntax. A preliminary attempt at the assessment of spoken forms and their functions in contemporary texts was conducted by me on the most renowned novel by Zoyā

books were later published: *Zendegiyam hame čiz-e man ast: mağmu'e-ye gofteguhā bā Kišlofski* [Kieślowski]. transl. by Anna Marcinowska, Tehrān: Qoqnoos (2017); Hālinā Pošfiātofskā [Halina Poświatowska], *Ādamak, matarsak va āšeq,* transl. by Ziyā Qāsemi, Kātāžinā Vānsālā [Katarzyna Wąsala], Tehrān: Sarzamin-e Ahoorāyi (2015). Some other remarks of similar character from coworking at the translations of Wisława Szymborska's poetry can be found in: Mārek Smužinski [Marek Smurzyński] (1997) Hezar nokte-ye bāriktar az mu [in:] *Negah-e nou* 34, Tehrān, pp. 177–183).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Pirzād, i.e., *Čerāqhā-rā man xāmuš mikonam* [I turn off the lights] and presented during an international conference *Power of Identity* at the University of Warsaw in 2013, where I showed how the spoken features serve as means of artistic creation (papers were later published as: M. Michalak and M. Rodziewicz (eds.) *In Quest of Identity. Studies of Persianate World*, Warszawa 2015). I have also discussed the problem of Persian diglossia during several conferences in Poland and abroad (*Asiam Explorare*, UAM, Poznań 2014; *Authenticity and Imitation in Translation and Culture*, SWPS, Warsaw 2015; *European Conference of Iranian Studies*, 9th edition, Freie Universität, Berlin 2019) and searched for its traces in translation on the example of *Mosāfer-e kučulu* [Little Traveller], the translation of Saint-Exupery's *Le Petit Prince* by Ahmad Šāmlu (Wąsala 2019). Those single, smaller studies can be said to have been the background research, setting ground for the project described in this book.

1.1 The aim of the research project

The main aim of this research project is to assess the presence of spoken variety of Modern New Persian in chosen contemporary literary writings published in Iran. Contrary to a tendency popular especially in Iran to focus on vocabulary and phraseology of colloquial language (see for example the introduction to Naǧafi's *Dictionary of Colloquial Persian*, Naǧafi 2000) or its phonological features (see Kalbāsi 2002), in this study I wish to put more stress on the morphological and syntactic features of Persian associated with the spoken variety, yet found in literature. The analysis is conducted with the following questions on mind:

1) in what parts of the text does "spoken language" appear (is it limited to dialogues, certain characters etc.),
2) on what circumstances does the presence of "spoken language" depend,
3) which characteristics of spoken variety are employed the most often and which do not make it into the printed word,
4) do the authors differ in their spelling or is there a unified orthography (and if there is, who created it)?[9]

The literature on Modern New Persian is not consistent in regard to the status of spoken and written varieties. There are terms which are used simultaneously in several, often conflicting meanings, there are also inconsistencies in translation between Persian and other languages. Therefore an additional, secondary aim to this work is to present the existing approaches, compare and

9 Forogh Hashabeiky (2005) in her comprehensive work on the orthography of Persian, unfortunately does not cover the spelling of forms not belonging to the written standard.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

contrast them in order to give a consistent satisfactory theoretical description of this phenomenon.

1.2 Source material

The literary material chosen for the analysis was taken from five contemporary novels published in the 21st century in Iran and enjoying relative popularity. The choice of books and authors was a deliberate one, meant to create a representation of different approaches to the use of Modern Spoken Persian within a literary text. From a wide range of possibilities, five books were chosen: 1) Dāryuš Mehrǧuyi (2010) *Dar xarābāt-e moqān* [In the tavern], Tehrān: Našr-e Qatre, 2) Mariam Ǧa'fari (2004) *Šahr-āšub* [Femme-fatale], Tehrān: Šādān, 3) Farhād Ǧa'fari (2007) *Kāfe Piyāno* [Café Piano], Tehrān: Našr-e Češme, 4) Zoyā Pirzād (2004) *Ādat mikonim* [We'll get used to it], Tehrān: Našr-e Markaz, and 5) Rezā Qāsemi (2001) *Hamnavāyi-ye šabāne-ye orkestr-e čubhā* [Nocturnal harmony of the orchestra of woods], Tehrān: Nilufar.

1.3 Organization of the book

The book is organized into six chapters.

The first, introductory chapter presents the general idea of the research project, some basic background and structural information on the book.

The second, theoretical chapter is devoted to the theoretical framework which has influenced the methods applied in this study. It begins with an overview of the variation in language and the problem of spoken versus written language in linguistics. In search of a best approach to the linguistic situation of contemporary Iran, various possibilities of approaching this kind of variation in language are discussed (linguistic variation and change, diglossia, functional linguistics, critical discourse analysis) which have influenced the chosen method.

The third chapter includes a brief background information on Persian language, its historical development and geographical varieties, the problems of its internal variation with focus on spoken and written varieties. It also gives an overview of the existing theoretical approaches to the spoken Modern New Persian in Iran and abroad and problems of terminological incoherence.

Chapter four gives an introduction to the analysis. It presents the details of chosen source material: authors' biographies, plot summaries and some stylistic remarks. It is also there where the process of the analysis is described step by step.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

The fifth chapter begins with a discussion of linguistic parameters of Modern Spoken Persian, which were unveiled during the analysis and are used to define and measure the MSP character of analyzed texts.

Chapter six is the final chapter, in which the conclusions of the analysis are presented and discussed in the light of theoretical concepts outlined in chapter two. It is also in this chapter that the possible areas of further research are suggested.

1.4 Notes on transcription and notational conventions

Transcription of Persian words in this book follows the scheme below:
آ – ā / ب – b / پ – p / ت – t / ث – s / ج – ğ / چ – č / ح – h / خ – x / د – d / ذ – z / ر – r / ز – z / ژ – ž / ص – s / ض – z / ط – t / ظ – z / ع – ‘ (not transcribed in initial position) / غ – q / ف – f / ق – q / ک – k / گ – g / ل – l / م – m / ن – n / و – v, u, ou / ه – h / ی – y, i, ey

Short vowels not represented in script: تَ – (t)a / تِ – (t)e / تُ – (t)o

The ezāfe is represented by *-e* after consonants and *-ye* after vowels and separated by hyphen; other morphological elements (plural suffix *-hā*, verbal prefixes *(ne)mi-/be-/na-*, parts of compounds etc.), unless otherwise stated, are not separated in the main body of the text.

Dates from the Persian solar calendar are calculated into their Georgian calendar equivalents for the purpose of clarity.[10] If necessary, the original Persian solar dates are indicated by “hš” (*heğrā-ye šamsi* ‘solar calendar’) and followed by their Georgian calendar equivalent in brackets, as in the example: 1397h.š. [= 2018].

Romanized Persian words and phrases appear in *italics* with a translation/definition in inverted commas (example: *goftāri* ‘spoken’). **Bold** is used for emphasis. Double inverted commas (quotation marks) are used for short quotes within the main body of text. Forms which belong to the spoken register, when compared to their written counterparts, appear in square brackets: [ğavunā] = *ğavānhā* ‘the youth.’

All the quotations from works in languages other than English, unless otherwise stated, appear in my translation. Also, for non-English titles of pub-

10 The solar calendar dates differ with Georgian calendar (GC) by 621+1 years. There is no difference in the length of the year, but Persian solar calendar (PSC) takes year 622 AD (Hiğra of prophet Muhammad) as its starting point and the year begins with spring equinox, which means that events from January, February and March will have their PSC date equivalent to GC-622 years, while those after March 21st become GC-621. Useful calendar converters can be found in the Internet, see for example: https://www.iranchamber.com/calendar/converter/iranian_calendar_converter.php.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

lications, my translation is included in square brackets after the original language versions.

In order to avoid repetition, the terms describing varieties of Persian language are often abbreviated as follows:

MStP	Modern Standard Persian
MSP	Modern Spoken Persian
MWP	Modern Written Persian

Glossing abbreviations generally follow the Leipzig Glossing Rules, with some exceptions resulting from specific features of Persian language:

Dom	direct object marker
Subj	subjunctive (either a verbal form or a verbal prefix indicating subjunctive form)
Indef	indefinite
Ep	enclitic pronoun
Temp	temporal
Cl	classifier
.pst	indicates that a past stem is used
.prs	indicates that a present stem is used
nomVerb	nominal part of the compound verb
ez	ezāfe

The examples presented in the analytical part of this book come from the five books which constitute the source material for the analysis. The examples are tagged by the book symbol and page number, while books, for convenience, are abbreviated as follows:

Dar xarābat-e moqān	DXM
Šahr-āšub	ŠĀ
Kāfe Piyāno	KP
Ādat mikonim	ĀM
Hamnavāyi-ye šabāne-ye orkestr-e čubhā	Hn

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Chapter 2.
Theoretical framework

The linguistic study of mutual relations between spoken and written variants of language within literary texts touches upon various areas of both linguistics, literary, and even sociocultural studies, therefore it requires a complex theoretical framework. Before any kind of analysis is begun, it is essential to establish the place that spoken and written varieties of Modern Standard Persian (MStP) occupy in modern Iran's linguistic situation. Therefore, constructing a theoretical and methodological framework will begin with the presentation of key terms and their definitions, later proceeding with Biber and Finegan's views on register and Ferguson's notion of diglossia, which are in author's opinion the most relevant frameworks suited for the task. Observations made by Biber and Finegan (1996), together with the theory of diglossia as presented for instance by Hudson (2002), can constitute a framework within which the two varieties of Persian being central to this analysis can be captured and defined. The second major concern of the present study is the relationship between spoken and written variants of certain structures of MStP. For its analysis, a term *linguistic variable,* borrowed from Labov's studies of linguistic change in progress (Labov 1994, 2001, 2010), will be employed. Even though the material chosen for this study is literary, texts are used merely as linguistic facts without any regard of their literary merit. Therefore methods applicable in literary studies (including literary stylistics and poetics) are not applicable, however some inspiration was gained from the methods of linguistic criticism (see, for instance, Fowler 1988, Labov 1972). The potential influence of spoken-written variation on text reading and interpretation is regarded here in terms of a side topic beyond the scope and purpose of the present research project.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

2.1 Variation in language. Basic definitions

It is a generally accepted view now that language is not homogenous, but an inherently variable system with different levels of variation. Despite this being stated, 20th century linguistics were concentrated on other aspects of language, and synchronic variation as a research problem was set aside. Ferdinand de Saussure with his *Course on General Linguistics* introduced the concept of language as a structured system, as well as the idea of a synchronic study of it (Saussure 1959).[11] Prague school structuralism and American structuralism continued on the path laid by Saussure, showing an interest in functions and behavioral aspects of language (see, for example, Jakobson 1960, Bloomfield 1933). Chomsky, on the other hand, was interested in the innate linguistic competence of an idealized speaker in a homogenous speech community, which gave way to the theory of transformational and generative grammar (Chomsky 1957). Their theories were largely based on language standards instead of a natural form of language as it is spoken. It was only in 1972 when Labov stated that "there is a growing realization that the basis of intersubjective knowledge in linguistics must be found in speech-language as it is used in everyday life" and mentioned ideological barriers to the study of language as such: the principle of studying structural systems of present and past in isolation, the idea that a sound change can be directly observed, while structural change is too fast to be done so (Labov 1972: xiii–xiv). Before that, while diachronic perspective on language dealt specifically with variation over time, synchronic studies tended to treat language as a uniform entity, usually on the basis of standard languages rather than the varied reality of actual use, or at least to consider the state of uniformity as desirable (Ferguson 1973: 38).

Synchronic variation of language is usually understood in two basic dimensions: according to user (dialectal variation) and according to situation of use (stylistic/register variation). There are certain key terms in use for the study of variation, those of crucial importance being: *variety*, *dialect*, *register* and *style*. Still, their working definitions tend to vary across subdisciplines of linguistics and even within them. The lack of clear-cut boundaries together with highly varied semantic range of those terms may lead to misinterpretations as well as purely academic discussions of whether certain phenomenon is a dialect according to one approach or a style according to another, both being equally justified. Hence I find it necessary to make an attempt on defining the above mentioned terms for the purpose of this project in order to avoid any kind of

11 Those concepts and principles were laid off by Ferdinand de Saussure during his lectures at the University of Geneva in the first decade of 20th century. The book was later compiled by his students and published posthumously in 1919 in French.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

possible misunderstanding. Since the approach of present research is mostly sociolinguistic, the focus is specifically on the definitions from this branch.

Variety is the most general term from the set. It is also neutral in the sense that it can refer to any kind of language, thus dialect, style, register are all language varieties.[12] A variety can be either very general (as it is with the spoken and written varieties) or very specific (i. e., a dialect of a certain social class within a certain geographical setting) (Trudgill 1992: 77). In some earlier works, this meaning of *variety* was shared with *dialect*, which according to Weinreich (1954: 389) was a term introduced to "designate the object of the description which is in fact a subdivision of the aggregate of systems which laymen call a single language" in structural linguistics. But he himself does not make use of the term *dialect* for the purposes of structural analysis, because the temporal and/or spatial attributes of *dialects* do not seem fit: they can be "adjacent or distant, contemporary or non-contemporary, prestigious or lowly" while "linguistic system in a strict structural view can only be identical or different" (Weinreich 1954: 389), and replaces it with *variety*.

As Ferguson wrote in his 1994 paper, a group of people which can be distinguished functionally in society in terms of their physical location, customs or any kind of interactional behavior is likely to develop markers of language structures and language use that will differentiate it from other groups.[13] Those markers will shape a dialect, which is a variation in people's speech according to where they come from or where they belong in a society (Ferguson 1994: 16–18). In other words, dialect can be defined as a variety of language according to the user (Moore 2004: 376) which is "regionally or socially distinctive and identified by a particular set of words and grammatical structures" (Crystal 2008: 142). It is possible that one of the dialects will become dominant as the official/standard form.[14] While in sociolinguistics we compare *dialect* with *style/register*, in the more general linguistic approach there exists some ambiguity on the level of *dialect* vs. *language*, and dialect is often used to describe less prestigious speech of the lower, uneducated classes, informal or rural speech. Haugen (1966: 922–

12 In Crystal's words, "a term used in sociolinguistics and stylistics to refer to any system of linguistic expression whose use is governed by situational variables" (Crystal 2008: 509). He notes, however, "for some sociolinguists, 'variety' is given a more restricted definition, as one kind of situational distinctive language—a specialized type of language used within a dialect, e. g., for occupational purposes."

13 The same assumption is made explicit in Crystal's (2008: 142) definition: "Any language with a reasonably large number of speakers will develop dialects, especially if there are geographical barriers separating groups of people from each other, or if there are divisions of social class."

14 According to Trudgill (1992: 24), however the term *dialect* is often used only for the non-standard or traditional dialects, standard varieties are "just as much dialects as any other dialects."

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

935) points out to two distinct dimensions of understanding those two terms: structural, where the defining characteristic is genetic relationship (if there are different speech-norms that descend from one earlier speech form, they will be called dialects, otherwise they are treated as separate languages). The other dimension is functional, depending on the use that speakers make of different linguistic codes. Here, language is a superposed norm in comparison with dialect and dialect can be said to be "a language that no one has taken the trouble to develop into what is often referred to as a 'standard language.'" For this understanding of *dialect*, however, he suggest application of the term *vernacular* instead (Haugen 1966: 927).

No less ambiguity and vagueness exist when it comes to *style* and *register*, which could even be labelled controversial if we consider all the existing views. They are both used on the same level as *dialect* (that is, a variety lower than the general notion of *language* and subordinate to it) and there are controversies regarding differentiation of dialect from *style* or from *register*, there is also no consensus as to the relationship of *style* and *register* themselves. It is often the case that what one author will call *style*, the other will refer to as *register*.

The trouble with *style* as a linguistic term is the existence of its more general meaning in the everyday language as "a manner of doing things" and, slightly more specific, as a set of characteristics of language use by certain individual (similar to *idiolect*).[15] According to Trudgill (1992: 72) it is a variety associated only with particular social situations which differs from other styles in terms of their formality, can be arranged on a continuum from very formal to highly informal or colloquial and "is in principle distinct from dialects and from registers: nonstandard dialect speakers can and do employ formal styles, and standard speakers can and do use informal styles."[16] In other words, it is a variety according to use (Moore 2004: 376).[17] Ferguson (1994: 28) finds the term so vague that in his opinion it is better to avoid it altogether. He points out that *style* is often used as an equivalent of *register*, or sometimes is restricted to the register variation in the aspect of formality (casual/formal), but it also appears in the

15 Compare: 'the manner of expression characteristic of a particular writer (hence of an orator), or of a literary group or period; a writer's mode of expression considered in regard to clearness, effectiveness, beauty, and the like' (OED 1989: 1008) and 'a distinctive manner of expression (as in writing or speech); a distinctive manner or custom of behaving or conducting oneself; also: a particular mode of living; a particular manner or technique by which something is done, created, or performed' (Merriam-Webster Online Dictionary; https://www.merriam-webster.com/dictionary/style, both definitions accessed 05/06/2018).

16 Changing from one style to another—or, better, moving along the continuum of styles—as the formality of a situation changes, or in order to change the formality of a situation, is known as style shifting (Trudgill 1992: 72).

17 But in reality the formal difference between dialect and style may as well be non-existent (Irvine 2001: 27–31).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

meaning of individual variation as opposed to shared conventions or variation within genres or registers. Crystal and Davy (1969: 61) use *style* as a cover term for functional varieties of language[18] (they do not use the term *register* at all and are very critical of it because of its generality). Yet there is no entry for *style* in the 6th edition of Crystal's *Dictionary of Linguistics and Phonetics* (2008), where it is linked to the entry for *stylistics*, reading as follows:

> a branch of linguistics which studies the features of situationally distinctive uses (varieties) of language, and tries to establish principles capable of accounting for the particular choices made by individual and social groups in their use of language. General stylistics deals with the whole range (or repertoire) of non-dialectal varieties encountered within a language; literary stylistics deals with the variations characteristic of literature as a genre and of the 'style' of individual authors. Applied stylistics is often used for the study of contextually distinctive varieties of language, especially with reference to the style of literary and non-literary texts (Crystal 2008: 460).

In Labov's writings, *style* is a function of attention paid to speech. This observation comes from sociolinguistic interview rather than spontaneous naturalistic speech. An important notion is that, according to Labov (1972: 208), this way "styles can be ranged along a single dimension" (from casual to formal), and attention paid to speech is supposed to underlie the patterns of style-shifting observable in interviews. Coupland, on the other hand, differentiates between *dialect style*, "stylistic variation in respect of variable features associated semiotically within 'social' or socioeconomic class differentiation and attribution within sociolinguistic communities," in which semiotic variants are considered that do not distinguish referential meaning, and *style* as "a way of speaking," which includes individual style choices of "what we choose to mean, to whom, when and where"—in other words, patterns of ideational[19] selection (Coupland: 189–190).[20]

The term *register* was introduced by some of the researchers, after their recognition of failing of *style*, as a term similar in its meaning, yet without the connotations with everyday language and layman's understanding. Perhaps the most recognizable works on register belong to Biber, who uses it as a general cover term for all situational varieties of language (as opposed to *dialects*, which are varieties associated with different groups of speakers (Biber 1994: 51). Ferguson is similar in his understanding of the term. For him, register variation correlates with different occasions of use (Ferguson 1994: 16). Trudgill restricts

18 And are, on the contrary, critical of the term *register* because of its generality.

19 The term *ideational* comes from Halliday's systemic-functional theory; it is one of the three metafunctions of language (the other being: interpersonal and textual) and denotes the ways in which language is used to represent our experience (Mann and Matthiessen 1991: 239).

20 For more discussion on the problems of *style* as a scientific term, see for example Ellis (1970) or Enkvist (1986).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

register to more specific understanding of a variety associated with particular topic, subject or activity (by which he means topics and activities like football, biochemistry or flower arranging; law and medicine are the examples of well-known technical registers)[21] (Trudgill 1992: 62–63). Crystal is closer to Biber, stating that *register* "refers to a variety of language defined according to its use in social situations" and that "in Hallidayan linguistics, the term is seen as specifically opposed to varieties of language defined according to the characteristics of the users (viz. their regional or class dialect), and is given a subclassification into field, mode and manner of discourse" (Crystal 2008: 409). To Beeman (2005), *registers* are specific clusters of speech variables controlled by the individuals within speech communities. He notes, however, that the term "can be deceptive, since they are rarely as unified as the term implies. In descriptive terms it is most often the case that speakers freely manipulate there variables to modify and shade their speech for specific social purposes" (Beeman 2005:1).

The most important difference between *register* and *style*, which can be found in all the above presented understandings, lies in the individually driven character of stylistic choices vs. functionally restricted choice of register. It is not the competence of an individual to choose the register freely (if we can speak of a choice of register at all, since in the light of the multi-dimensional analysis of registers, they are defined "from below" by the existence of certain co-occurring linguistic patterns), yet it is natural that choosing between formal and informal (or careful and casual, for that matter) speech is in the hands of the speaker. That is why I believe that the distinction between spoken and written language, even though existing on the extremely high level of generality, can be regarded as an example of registral variation—a functionally driven, non-optional variation applied by the speakers automatically, perhaps even unconsciously, and definitely not as a result of any decision process.

Problems of differentiation which arise here are in some part a consequence of the overlap of linguistic characteristics of written register and careful/formal style as well as spoken register and casual/informal style. Those problems will be addressed in later chapters on the linguistic material used for the analysis.

For the purpose of the present study, let us recapitulate the working definitions of the above discussed terms that will be used throughout this work:

1) *variety* – a general term which refers to any distinctive kind of language (dialect, register and style are examples of language varieties)

21 This understanding makes register roughly equivalent to *jargon*, which is pointed out by Trudgill himself: "Registers can identify speakers as being members of a particular peer group, and are for that reason often labelled jargon by outsiders who are not part of the group in question" (Trudgill 1992: 63).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

2) *dialect* – a variety of language, which is defined according to its users (either geographically or socially)
3) *register* – a variety of language defined according to its situational use and existing on different levels of generality
4) *style* – individual variation in language which can surface in different dialects and registers and is often connected with the level of formality

It is worth remembering that any instance of language can and will simultaneously exemplify all those categories (Ferguson 1994: 25).

2.2 Spoken and written language. Functional variation and the notion of register

One of the most widespread examples of variation is the one that exists between speech and writing. Just as other examples of internal linguistic variation, it was not given much attention before the 20th century. Historical studies were based on the written sources, and the state of language was assessed by what was recorded in script (Roberts and Street 1998: 117). With the idea of language as an abstract system represented by signs and introduction of terms *langue* and *parole*, commonly translated as "language" (the system) and "speech" (the realization of it) by Saussure, accompanied by the founding of disciplines such as sociolinguistics, the focus was put on the spoken language as primary, reducing the function of writing to the representation of the spoken form (Saussure 1959: 23–24) and thus causing a shift of interest of linguistics to the study of speech, this time with disregard to writing.[22] As long as the written language was regarded merely as a visual representation of the abstract system of language, secondary to the phonic representation in speech, there was no reason to discuss the written and the spoken as different codes.

The fact that we are dealing with two fundamentally different representations of one abstract system is stressed by Halliday (1985), who is closer to giving the two varieties equal status: "we achieve different goals by means of spoken and written language; but neither has any superior value over the other" (Halliday 1985: xv). At the same time he maintains the existence of a wide gap between them by saying that they are both forms of language with the same linguistic system underlying them, although they exploit different features of that system and their

22 Saussure blamed philology for confusing written text with the language itself, thus hindering the development of linguistics; at the same time he was also aware of the fact that the shift of attention to the phonological aspects of language led to excessive attention paid to phonological studies (Saussure 1996: 2a).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

power is gained in different ways (Halliday 1985: 100). The evolution of those two kinds of language took different directions because of the different functions they serve and as a result of a difference in medium. The prosodic features of spoken language are contrasted with the role of punctuation in writing, lexical sparsity and grammatical intricacy of the former with the lexical density and grammatical metaphor of the latter. Halliday puts emphasis on the fact that most of the highly valued texts and prestige functions of language have been taken over by writing, thus causing literate culture not to take spoken language seriously (Halliday 1985: 97). As Roberts and Street point out (1998: 117), the descriptions of differences between spoken and written language tend to be consistent in terms of the characteristics of both, which generally cover the features mentioned by Halliday and quoted above. A discussion on those two varieties leads inevitably to the discussion of orality and literacy,[23] which however are beyond the scope of this work. Instead, an important theoretical issue that arises here and has to be addressed is the question of nature of those two varieties of language. While in general use it is customary to speak of *spoken language* and *written language*, it is also obvious that none of them is actually a language itself, neither in the meaning of being independent of each other, nor in the meaning of constituting that ideal unified standard language by themselves. In the previous part of this chapter, I suggest that we can apply the term *register* to those two varieties as long as its high level of generality is highlighted. Their use is restricted functionally[24] and the choice is not optional for individual speaker, and they are not characterized by a certain group of speakers, as it would be in case they were dialects.

2.3 Analytical framework for the study of register

A comprehensive analytical framework for the study of register is given by Biber (1994), who notes that registers were previously often studied without adequate and proper theoretical framework: "in practice, most register studies have been atheoretical" (Biber 1994: 36). He surveys the frameworks existent-to-date within two groups: classificatory and descriptive. Classificatory framework in his

23 For an extensive study of the subject, see Ong (2002).

24 Theoretically speaking, we could say that *writing* and *speech* differ rather by medium (sound vs graphic signs), but the different character of situations in which they are applied (i. e. conversation vs academic writing) accounts for the functional understanding. Of course there are also borderline cases which are spoken in character yet written in terms of medium, especially nowadays in the era of electronic communication: e-mails, social media comments etc. (personal letters would serve as a non-digital example), blog entries, comic straps and so on. This returns again to the discussion of orality and literacy and the differences between oral and literary language (which include also lexical choices, syntax and stylistics); see Ong (2002).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

opinion dates back to at least to 1920s and Malinowski's (1923) distinction of magical and pragmatical language varieties, and they cover various typologies based on a range of different criteria: perceptions of reality, topics and text organizations and other linguistic and functional parameters. The multi-dimensional approach of Biber (1989) also falls into this category.[25] Classificatory approaches help distinguish different text types on the basis of specified characteristics. Descriptive frameworks are better developed, but not sufficiently explicit to use them for taxonomical purposes. Here the most important feature is the context of situation, and the most influential frameworks include that of Hymes (1974), designed to study the ethnography of communication and distinguishing sixteen components of speech situation,[26] Halliday's (1978) with three core features: field, tenor and mode, to define what is happening, who is taking part and what part is played by language, and Crystal and Davy's (1969)—here the situational varieties are covered by the term *style* and not *register*. In Crystal and Davy's framework there are eight dimensions falling into three categories: 1) characteristics of speaker, regional and social dialect and time, 2) discourse characteristics (medium and participants), 3) activity being done, relative social status of participants and idiosyncratic preferences of speakers.

The framework proposed by Biber (1994: 41–49) is based on presented approaches of Hymes, Halliday, and Crystal and Davy.[27] In his opinion, situational features, linguistic features, and functional and conventional associations between them are the typical three components of register studies and the proper framework should provide tools to analyze all three of them. One of his main goals is to specify parameters that will allow for explicit specification of differences between any pair of registers, so that it enables a proper comparison of

25 Multi-dimensional analysis is a methodological approach to register analysis, originally developed for the analysis of a full range of spoken and written registers within a language. It applies multivariate statistical techniques to the investigation of register variation within a prepared corpus. Two major steps in multi-dimensional analysis are: identification of salient patterns of linguistic co-occurrence, and comparing registers in the space defined by those patterns. Since individual features cannot reliably distinguish between registers, it is important to base the analysis on co-occurrence and alternation of those features instead. MD analysis have been applied to languages such as Somali, Korean, Taiwanese, Spanish and others (Biber 2004: 15).

26 Sixteen components of speech situation according to Hymes are: message form, message content, setting, scene, speaker (or sender), addressor, hearer (or receiver, or audience), addressee, purposes-outcomes, purposes-goals, key (tone or manner), channels, forms of speech, norms of interaction, norms of interpretation, and genres. The model of description which uses these components is called the SPEAKING model, where the acronym stands for: S(etting and scene), P(articipants), E(nds), A(cts sequence), K(ey), I(nstrumentalities), N(orms), and G(enre) (Hymes 1974: 56–62).

27 Further development of the ideas described above can be traced in a textbook of register analysis by Biber and Conrad (2009).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

registers. His eventual, long-term goal is creation of a situational framework with specified quantitative values for each parameter, so that the use of correlational statistic techniques would be possible. The list of parameters given is extensive, including both closed and open-ended parameters without any specified hierarchy. Unspecified parameters mark the level of generality—the more unspecified values, the higher is the degree of generality. Among the parameters there are: communicative characteristics of participants, relations between addressor and addressee, setting, mode/channel, relation of participants to the text, purpose and topic.

According to Biber, lexical choices do not typically mark registers by themselves, while grammatical choices sometimes do. In general, the indicators of register differences are core lexical and grammatical features that may come from different linguistic classes, among others: phonological features, tense and aspect markers, pronouns and pro-verbs, nominal forms, passive, dependent clauses, prepositional phrases, modals, reduced forms, discontinuous structures, distribution of given and new information, and speech act types. It is not the relative distribution of those features that distinguishes among registers, but the co-occurrence of those.[28]

A comprehensive linguistic analysis of a register requires consideration of a representative selection of these linguistic features. Such analyses are necessarily quantitative, because register distinctions are based on differences in the relative distribution of linguistic features, which in turn reflect differences in their communicative purposes and situations (Biber 1994: 35).

In the multi-dimensional model of analysis, the computational tools are used to elicit features which co-occur in texts included in corpus. Rather than presupposing the existence of certain registers and finding proof in corpus, the researcher is presented with clusters of co-occurring linguistic parameters that he or she can then label as registers and analyze by certain dimensions. In case of a small, self-designed and non-digitized corpus, those steps will have to be reversed, yet once the linguistic characteristics of the register in question are known, such a change in procedure should not affect the results too much.

There remains a question of if and how to apply this kind of theoretical model of analysis to the works of fiction. One could say that fiction represents a kind of carefully created linguistic entity and will not reflect registral reality of the language in question, since the situational features are rather constant (narrative, monologue, dialogue). Other would perhaps argue that it is exactly the writer who

28 Varieties defined by shared sets of co-occurring linguistic features were identified as *speech styles* by Ervin-Tripp (1972) and Hymes (1974), while Halliday (1988) sees clusters of associated features with tendency to co-occur as a marker of register (Biber 1994: 35). *Style* is also a term applied by Labov (1972), who defines various styles on terms of the speaker's attention paid to speech.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

has both the knowledge and ability to apply a variety of styles and registers in his works, as long as he is not limited by other, outer constraints such as cultural-literary conventions of his language and culture. As Biber notes (1994: 132), "relevant situational context for a fictional text is the fictional world that the author creates in the text itself." The only constraint he finds for the investigation of fiction is the necessity of incorporating stylistic analysis as well (where stylistic choices are understood as those whose functions are associated more with aesthetic preferences than the real-world situational context of the register).

2.4 Diglossia: a special kind of hierarchized registral variation

The dichotomy of spoken and written varieties of language does not have to include any hierarchy. Despite of that fact, as Ferguson observes (1994: 18), it is also natural for the members of speech communities

> to hold certain evaluative attitudes toward variant forms. (…) Language change, that is, that aspect of conventionalization by which certain variants spread at the expense of others, is intimately connected with speakers' evaluative attitudes, and the analysis of language attitudes in relation to change is correspondingly important.

Diglossia is an example of a linguistic situation in which the evaluative attitudes of speech community and conventionalization lead to a hierarchized variation. The term "diglossia" was introduced into English by Ferguson in 1959[29] to describe a situation in which different varieties of the same language are used within a speech community under different, complementary conditions[30]:

> Diglossia is a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation (Ferguson 1959: 325–326).

29 The term itself was not a Ferguson's invention, in fact, and was already present in German and French scholarship before that time, but it was his article that put diglossia under worldwide attention. As Ferguson confessed himself later, that attention took a direction somewhat different from the one he hoped for, concentrating more on the question of whether a certain linguistic situation is, or is not, a case of diglossia, rather than on formulating a consistent typological approach (Ferguson 1991: 219).

30 It is meaningful that the preliminary version of that article was entitled "Classical or Colloquial – One Standard or Two," which marks relevance to the current situation of Persian, where the existence of two standards is recognized by scholars such as Sādeqi (1984), Rezāyi (2006), Kalbāsi (Eslāmzāde 2009), Mahmoodi-Bakhtiari (2018) and others.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

The two types of language in a diglossic situation are referred to as H[igh] and L[ow]. The L variety is used in regard to the vernacular (primary dialect), indicating its lower status, while H is used to represent the superposed variety. **The variation witnessed in diglossic situations is of a functional (registral) and not dialectal (societal) character** – switching between H and L correlates with situational context and occasion of use, regardless of language user's social position (Hudson 2002: 3, my emphasis).[31] It is contrasted with a more common situation of the so-called standard-with-dialects, where speakers use their local dialects for casual conversations at home, with family and friends, but switch to standard language for the more formal situations, public occasions and conversations with strangers or speakers of another dialect. The difference between standard-with-dialects and diglossia is supposed to lie in the existence of a structural gap between the written norm and spoken dialects of language (Wexler 1971: 336, Coulmas 1991: 126 as quoted in Hudson 2002: 11), or, as Ferguson himself puts it, "extensive differences between the grammatical structures of H and L" (Ferguson 1959: 333). There is, however, no exact or precise measure of "extensive," which led to some contrasting ideas on the importance of this criterion for distinguishing diglossia from other linguistic situations. It might be useful to refer to Joshua Fishman's notion that "without schooling the written/formal-spoken variety cannot be understood" (Fishman 1980: 4). H is a variety well-studied and well-described, with the existing norms of pronunciation, grammar and vocabulary which limit the internal variation. L, on the contrary, has no settled orthography, its pronunciation, grammar and vocabulary can vary widely (Ferguson 1959: 332).

Diglossia is also often contrasted with societal bilingualism—a situation in which a particular speech community uses more than one language. Societal bilingualism does not imply any status difference between the languages in use (Clyne 1997: 301). Although the original definition of diglossia by Ferguson was meant to refer to different varieties of one language, it was then broadened by Fishman, who rejected the need for diglossic codes to be structurally related (i. e., varieties of one language) and included also those situations, in which two different languages are employed, stating that:

31 It is of great importance to underline here that all the definitional criteria of diglossia mentioned in Ferguson's first formulation quoted here are true for the linguistic situation of Iran today. The written variety of Modern Persian fulfills all the requirements of the H variety (prestige, literary heritage, formal acquisition, no use in regular conversation, grammatical and lexical differences with L variety). However, the multiethnicity of Iran and the fact that other languages and dialects are also in use, often in bilingual or diglossic relations with standard Persian (for example, Azeri, Kurdish, Armenian as used by the minorities together with/complimentary to Modern Persian both in speech and in writing), tend to overshadow this observation.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

> diglossia exists not only in multilingual societies which officially recognize several 'languages' but, also, in societies which are multilingual in the sense that they employ separate dialects, registers or functionally differentiated language varieties of whatever kind (Fishman 1967: 30).

The relationship between H and L variety in diglossia can be seen as of a formal-informal character, therefore it is best studied from the theory of formal language use and social behavior perspective, as Hudson argues (Hudson 1994: 294). He compares two views on the notion of formality, a sociocultural (anthropological) and a discourse-functional one. In both of them, it is the functional characterization of linguistic varieties and not the social identity of the speaker that is important for categorization of them (Hudson 1994: 302).

Although Ferguson's definition describes diglossia as a relatively stable phenomenon, more recent studies show it somewhat differently (see Hudson 2002: 29–30) and, as Hudson points out, it is likely that instead of stability, the direction of change after the linguistic dichotomy is resolved should be taken into account as a marker of diglossia in comparison with other linguistic situations such as societal bilingualism. "In diglossic context (…) it is H that tends to be displaced by L through a process of structural convergence resulting in the emergence of a new standard more closely related to certain educated varieties of the vernacular" (Hudson 2002: 30), while in societal bilingualism it is the prestige variety that replaces the vernacular. Among the factors which act as motivators of the spread of vernacular varieties is its use in literature. The case of Tamil, brought by Britto (1991: 69–70), shows the introduction of L-variety into conversations between characters, which has resulted in it being seen appropriate for other monologic discourse (i. e., sermons, political speeches, lectures). Similarly, vernacular Greek spread from poetry through literary prose into critical writing, and from there into other written registers (Alexiou 1982: 159, as quoted in Hudson 2002: 31). The societal development leading to a decline of diglossia is connected with the spread of literacy and mass communication in the speech community. It can be also more generally connected with "the processes of modernization, urbanization, mercantilism and industrialization" (Hudson 2002: 32). Similar processes can be traced in regard to Modern Persian language, and there were studies undertaken in order to declare Modern Persian as either diglossic or monoglossic, or to define the extent of Persian diglossia. However, doubts arise as to whether labeling Persian diglossic does bring any new insights into the research of this language (see for example Spooner 2012, Rossi 2015). This problem will be discussed in the next chapter.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

2.5 Linguistic change in progress

If diglossia is not as stable as it was first thought to be, is it possible to observe the transitional stage between diglossia and a following step of development of linguistic situation? What tools could be used to account for such change and describe it?

The study of a linguistic change in progress is central to a model developed by Labov and his followers in the second part of the 20th century. Similarly to historical and comparative linguists, Labov is interested in the processes which lead to linguistic change, yet he supposes that this kind of change is structured and possible to be observed at the time in which it is operational instead of focusing on often scarce historical documents. The early formulations of a theoretical framework that would allow such a study of linguistic change are visible already in the writings of Labov's teacher and advisor, Uriel Weinreich, in his proposition of structural dialectology (Weinreich 1954).

Labovian model is often labelled a variationist paradigm. It is an empirical approach to the variability in language, which stems from the idea that this variability may be shown to be structured. In comparison with other branches of sociolinguistics, the focus of variationists is on understanding the variation and change in language rather than the nature of interaction between speakers and their behavior (Milroy and Milroy 1997: 45). Language change is perceived as a disturbance of a form/meaning relationship which results in a loss of cross-dialectal comprehension and mutual intelligibility (Labov 1994: 9).

Change in progress can be observed in two timeframes: in apparent time and real time. Observation in apparent time is the observation of how the variables are distributed across age levels in speech community. The existence of a correlation (or monotonic relationship) between age and the variable shows a probability of change in progress. The problem with such observation is that it might show patterns of age-grading (that is, a change that is repetitive in every generation) instead of a change in real time. That is why observations in apparent time should be complemented by real time studies, which rely on the observations of certain speech community in two distinct points in time (i. e., replicating a study after a certain period) in order to compare the results (Labov 1994: 45, 73).

Labov distinguishes between two types of linguistic change in regard to the manner in which they are triggered and by whom. *Change from above* is a type of change introduced typically by social classes from the peak of the socioeconomic hierarchy and is connected with the notion of higher prestige (often a borrowing from another speech community that is supposed to be more prestigious by the dominant social class). Such a change does not affect the vernacular patterns immediately, but is first reflected in careful speech, and it might even be inconsistent with the existing vernacular system. Therefore it is even likely not to be

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

integrated with the rest of the system and form a separate subsystem instead. This type of change, similarly to more conscious types of correction and style shifting,[32] is likely to show patterns of age-grading. *Change from below*, on the other hand, is a change that results from internal, linguistic factors, appears first in the vernacular and proceeds in the highly organized, phonetically controlled patterns (Labov 1994: 101). This kind of change, contrary to change from above, operates below the level of social awareness until it is close to completion. It might be introduced by any social class, however there is no record of the highest status classes acting as the innovators of that sort (Labov 1994: 78).

One of the core concerns of Labovian sociolinguistics and the study of linguistic change is the problem of transmission, i. e., how the change is transmitted between speakers. On the basis of his findings within American English, Labov formulates five principles of transmission of urban linguistic change, which are as follows:

1. Children begin their language development with the pattern transmitted to them by their female caretakers, and any further changes are built on or added to that pattern.
2. Linguistic variation is transmitted to children as stylistic differentiation on the formal/informal dimension, rather than as social stratification. Formal speech variants are associated by children with instruction and punishment, informal speech with intimacy and fun.
3. At some stage of socialization, dependent on class status, children learn that variants favoured in informal speech are associated with lower social status in the wider community.
4. Linguistic changes from below develop first in spontaneous speech at the most informal level. They are unconsciously associated with non-conformity to sociolinguistic norms, and advanced most by youth who resist conformity to adult institutional practices.
5. Linguistic changes are further promoted in the larger community by speakers who have earlier in life adopted symbols of nonconformity without taking other actions that lessen their socioeconomic mobility (Labov 2001: 437).

He links the transmission of change to the process of style shifting, saying that:

> Style shifting seems to be one of the keys to what we now see as the central problem of the theory of language change: the transmission problem. In the course of linguistic change, children learn to speak differently from their parents, and in the same direction that their parents learned to talk differently from their own parents (Labov 2001a: 85 in Eckert, Rickford 2001).

32 *Style shifting* is a concept parallel to *code-switching*. While *code-switching* is applied mostly in the situations of societal bilingualism or multilingualism, *style shifting* refers to the changes of code within one language, across a style continuum.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

This way, the concept of *style* (in the Labovian sense, that is, the function of attention paid to speech) becomes a central concern of the study of change in progress and comparison of styles (especially the opposite poles of the stylistic continuum: careful/formal vs casual/informal) can give us important insights into the direction that language change might be going to take.

It is important to note that however Labov's studies are preoccupied with sound change (see, for instance, Labov 1994: 9–112), it is possible to observe changes on other levels of grammar as well, even if the existence of morpho-syntactic variables is more difficult to study and describe than of their phonological counterparts (see below).

2.6 Linguistic variable

One of the key terms in studying the variation is a linguistic variable. The simplest definition of a linguistic variable is "two or more ways of saying the same thing" (Tagliamonte 2006: 70). According to an original definition by Labov, such a variable should be frequent, immune to conscious suppression to some extent, constitute an integral part of larger structures, be stratified and asymmetrically distributed over some order strata of society (Labov 1972: 8). A phonological variable is relatively straightforward—realizations of alternates differ by their phonological features, often within one word (see, for example, the analysis of sociolinguistic variable (s) in Bengali, Ferguson 1979). At the level of morphology and syntax, beyond which there is a question of possibility of semantic equivalence—the existence of morphological or syntactic variables means acceptation of a fact that two morphological or syntactic forms have the same underlying, referential meaning[33] (see Tagliamonte 2006: 71–72). A linguistic variable should also correlate with other (social or linguistic) phenomena. The study of linguistic variables is quantitative by nature. Their occurrence (or lack of it) in a particular, defined context gives important information on the progress of observed change in progress.

Biber and Finegan (1994: 6) suggest that the preoccupation with linguistic variable as a primary analytical construct restricts the scope of analysis to semantically neutral phonological variants and is one of the primary reasons for register studies being neglected in much of sociolinguistic research. Since registers serve different communicative purposes, they naturally vary in content as well as form rather than "saying the same thing," which is the basic definition.

33 In other words, variants need to serve similar discourse functions and strict semantic equivalence is not required (Sankoff and Thibault 1981: 208).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Therefore, variation across registers is more about different linguistic features than their semantically neutral variants.

> Neither sentences independent of their text nor texts independent of their social context will suffice; word lists will not suffice; intuitions will not suffice. Only a text suffices as an object of register analysis, and knowledge of the context in which it was created and the purpose for which it was intended is likewise crucial. There are no sentences in register analysis unless they form part of a text embedded in a social context. For spoken registers, the text is typically a transcript. For written registers, the text is given (Biber and Finegan 1994: 7).[34]

In the present work, the notion of linguistic variable has not been applied—at this point of the research, at which the features of the two registers (Modern Spoken Persian and Modern Written Persian) were elicited and described, it was too early to define clear cut privative opposites between them. Once it could be proved, though, that many of the differences between spoken and written registers of Persian are actually of a privative kind, which means that in most cases there is exactly one spoken equivalent for each written token (to which this opposition is relatable, of course, since not all of the units regardless of the level of language we consider do change with the change of register) and therefore it is possible to treat those equivalents as linguistic variables, a whole new set of research possibilities would be opened with the means developed by the variationists. An analysis of the occurrence of variables in certain contexts could support the hypothesis of spoken register gaining more power over written register over the course of time, and perhaps help present a study of a change in progress where non-phonological variables dominate.

2.7 Methods applied in present research

After an overview of the theories that influenced the methodological approach of present research, a method applied in the analysis of literary material studied for the project should be outlined. All the above approaches (stylistic, functional, diglossic, and variationist) help understand the linguistic situation of modern Iran better.

The relationship between spoken and written varieties of Modern Persian is of diglossic character to a great extent. However, as Spooner (2012) and Rossi (2015) underline, it is not clear whether our understanding of this relationship can

34 But if we look at the registers on the highest level of generality, will that still be true? Spoken and written registers (without further differentiation) may show certain variables which are semantically neutral and only medium-sensitive. Thus the **level of generality** should be the primary concern here if we want to look at those two varieties in terms of register.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

benefit from being given such a label. The majority of the studies on diglossia were concerned with its presence or absence (on which Ferguson himself expressed his disappointment in a 1996 paper *Diglossia revisited*, concluding almost forty years of research in this subject)—and it is no different in the study of Persian, where the question of Modern Persian being or not being diglossic has become a subject of some controversy, overshadowing the descriptions of structural impact this label would (or would not) have on the language.

Despite of its undeniable benefits for structured and systematic analysis of registers, accurate multi-dimensional register analysis could not be applied because of technical restrains, namely unavailability of a fully functional computational OCR tool for Persian language that would enable digitization of selected corpus and further proceeding of the digitized material. Although there exist some corpora of Persian language (Seraji 2015; Ghayoomi, Momtazi and Bijankhan 2010), the specified aim of present study to analyze structures present in certain literary works requests building a corpus on its own and cannot rely on the existing sources instead.

Contrary to the method of MD analysis, here the co-occurring linguistic features characterizing spoken register of Persian will be defined before the analysis of chosen language material. Such a change to the procedure is possible thanks to the existence of several descriptions of spoken Persian, even if those descriptions use a variety of label ranging from colloquial style of Persian to the dialect of Tehran or even Persian diglossia. A careful reading makes it possible to grasp those features of the described linguistic varieties which are repeatable in all of them, suggesting that they belong to a category superior to geographically limited dialect or user-restricted style. Characterization of the descriptions chosen to represent the register in question will be given in the next chapter.

Selecting the repetitive linguistic features which characterize spoken register of MSP on the level of phonology, morphology, syntax and vocabulary is thus the first step in the analysis. Of course one should bear in mind that not all the defining features will surface in printed text (this is especially true for the phonological features, in case of which there would be a competition between traditional/standard and "colloquial" spelling in words of lower frequency)[35]—indeed one of the questions behind the analysis to follow is: which elements do surface in a printed version of "spoken" text and which do not.

Once a set of characteristic linguistic features of spoken MSP is selected, a detailed manual analysis of the selected literary material will be conducted. However the absence of freely available optical character recognition for Persian will make it a time consuming and arduous task, it will at the same time help

35 Some of the problems acquainted here will be similar to those confronted by the corpus builders, described in Seraji 2015 or Ghayoomi, Momtazi and Bijankhan 2010.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

avoid certain difficulties resulting from the peculiarities of Persian language such as inconsistent orthography. The frequency with which selected features occur, altogether with their situational context (dialogue, monologue, narrative) and individual features of author's style, should give an overview of how spoken register is incorporated into what should be a written register overall.

A relatively small size of corpus does not suffice for an extensive statistical analysis similar to those conducted by variationists for linguistic change or within register studies as MD analysis, yet some simple statistical data will be drawn from the analyzed data-pool. If the hypotheses of this research project are correct, those simple statistics should visualize the scale of the phenomenon and enable some general remarks on the role of spoken register of Persian within literary texts. If present study yields satisfactory results, another step in ascertaining the role of spoken MSP in modern Persian speech community would naturally be the analysis of other realms of written register prone to spoken influences, such as press, radio and TV broadcasts, blog entries and the whole online sphere in a corpus-based research project.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Chapter 3.
State of the research

3.1 Defining Modern Persian

Any discussion on modern Persian should begin with definition of the kind of language being discussed. Easy and obvious as it may sound, it poses certain problems. Islamic Republic of Iran is a homoreligious[36] but multiethnic country with great linguistic diversity. While Persian is the only official language of the state, there are many different local dialects and varieties,[37] as well as of other Iranian (e. g. Kurdish), non-Iranian (e. g. Armenian) and also non-Indo-European (e. g. Azari) languages, also in use. Moreover, even the standard Persian of the state is not a monolithic entity, showing considerable variation within itself. This variation, once noted, was discussed by many scholars on different terms, also those that concern us here the most: on the speech—writing axis.[38] Above issues were underlined already 40 years ago by Gernot Windfuhr, who in the

36 According to the official census from 2011, 99,4% of the population (70 mln in 2006, 75 mln in 2011) declared themselves as Muslim. Christians, Jews and Zoroastrians are other officially recognized religious minorities, but according to the census, altogether they constitute 0,3% of the population (another 0,3% remains undeclared). The findings of 2016 census do not show data on religion.

37 Windfuhr and Perry mention two groups of the varieties of Persian overall: Western (Iran) and Eastern (Afghanistan and Central Asia), with transitional varieties. The sociolinguistic registers of Persian of Iran are said to include Modern Standard Persian in its written norm, Colloquial Persian used in polite conversations and being a "normalized form of Colloquial Tehrani Persian […] which increasingly shows reflexes in the standard language" and a variety called *xodemuni*, a local variant that has not yet been normalized (Windfuhr and Perry 2009: 417).

38 One of the first scholars to notice and describe the internal variation of Persian was Aleksander Chodźko (1804–1891), an diplomat and orientalist of Polish origin, who mentioned them in his *Grammaire persane ou principes de l'iranien moderne*, published in Paris in 1852 (Calmard 1991: 503). See also Elwell-Sutton 1941, Hodge 1957, Lazard 1957, Jazayery 1970, Beeman 1976, 1986, Jeremiás 1984, Nağafi 2000, Kalbāsi 2002, Peisikov 2002, Perry 2003, Mękarska 2005 and others. Different approaches to the internal variation of Modern Persian will be discussed more broadly in the second part of this chapter.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

opening pages of his *Persian Grammar* states on the kind of Persian discussed in the book that:

> "**A definition is impossible**; the contemporary 'standard' dialect may be circumscribed as follows: it is the socially most prestigious dialect of Tehran which is *not yet*, but is fast becoming, the standard dialect of Iran. As such it is the geographically least marked dialect as opposed to the local dialect (Mundart) of Tehran. It has a written and a spoken, informal variant, here called contemporary colloquial, i. e. the style, which according to Hodge (1957: 366) is used between persons very well acquainted, although they may not be on familiar terms as regards politeness level" (Windfuhr1979: 7, my emphasis).[39]

3.2 Modern New Persian language. An overview

New Persian is an Iranian language belonging to the Western group of Indo-Iranian branch of the Indo-European language family. It has three main regional varieties known under their native names as *fārsi*—the official language of the Islamic Republic of Iran, *toǧiki*—the national language of Tajikistan, and *dari* —one of the official languages of the Islamic Republic of Afghanistan (beside *pashto*).[40] The variety discussed here is the Modern New Persian used in Iran. The number of speakers of this variety can be approximated to about 70 million (the population of Iran plus substantial diaspora in North America and Europe).[41] Historically speaking, it may be considered unique among the world's languages since it is the direct descendant of Old and Middle Persian—in other words, its development can be traced back to the second half of the 1st millennium BC and all the stages of it are known. The beginnings of the New Persian period in the history of Persian coincide with the fall of Sasanian dynasty and the Arabic

39 As we can see, Windfuhr (after Hodge) would classify the spoken variant of Persian as an informal, colloquial style, even though the distinction is made on the basis of medium (written – spoken) and not the level of formality (as it is for Hodge). Also interestingly, Hodge differentiates between deliberate informal and colloquial informal style. The one quoted above is colloquial informal, while deliberate informal is defined as used on ordinary occasions between persons not very well acquainted, or on what our culture would call 'formal' terms (Hodge 1957: 366); both definitions may equally apply to the spoken standard.

40 The status of those regional varieties is not clear and there is an ongoing discussion on them being regional varieties/dialects of one language or separate languages of common origin. Since the definition of 'language' is often influenced by non-linguistic, i. e. political, social and cultural, factors, different views can be equally justified (cf. Beeman 2005, Perry 1999 and 2007: 975, Mękarska 2005: 6).

41 Windfuhr gives an estimate of 60 million native speakers and 110 million total users of Persian in all the varieties (including users of Persian as a second language and speakers outside the three Persian-speaking countries); of which 35 million native/70 million overall speakers of Western Persian (i. e. Persian of Iran) (Windfuhr 2009: 418).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

conquest of Iran, which marked the end of the Middle Persian period. Although traditionally no fourth phase in the development of Persian is distinguished in its descriptions, it is not the New Persian of the Middle Ages, but a Modern New Persian that we deal with today—and in fact, it might be approaching another transitional period soon.

3.3 From old to modern: historical development of Persian

The development of Iranian languages is generally divided into three chronological stages: Old, Middle and New (Modern), and Persian is the only language for which all three stages are known (Skjærvø 2006: 344–377). It is thus the descendant of Old and Middle Persian.

Historically, those three stages would correspond to the periods of:
Old Persian – Achaemenid empire until Alexander the Great's invasion
(6th–4th/3rd centuries B.C.).
Middle Persian – until the fall of Sasanians and Arab invasion
(3rd century B.C.–7th/11th century A.D.)
New Persian – the Islamic period
(11th century A.D. onwards).

3.3.1 Old Persian

Old Persian is known to us mostly from the royal inscriptions of Achaemenid kings, namely Darius the Great (522–486 B.C.) and Xerxes (486–465 B.C.). It was a "highly inflectional language of the ancient Indo-European type" (Paul 2013), with both nominal and verbal inflection, written in cuneiform script. Although the only known remnants[42] of this language are inscriptions, the oldest one being Bisotun inscription from 521 B.C., there is a probability that it has already been spoken in the first half of 1st millennium B.C. What suggests this possibility is the fact that latest inscriptions include "mistakes" in word endings, that might in fact record a later stage of development ("post-Old Persian") (Skjærvø 2006: 344–377).

Typologically speaking, Old Persian represented a synthetic type of language, where the syntactic relations are represented by inflection. It had a complicated

42 This was the case until 2007 at least, when a clay tablet of administrative kind was found in the Persepolis Fortification Archive, casting shadow over the earlier presupposition that "practical" application of Old Persian was limited to prestige and display purposes, while almost no one was actually using the language in speech and writing in other circumstances (Stolper, Tavernier 2007: 6).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

system of nominal inflection: six cases (nominative, vocative, accusative, genitive-dative, instrumental-ablative, locative, three numbers (singular, dual and plural), and three genders (masculine, feminine and neuter). Adjectives were also declined. Thanks to the inflection, word order was not fixed. The verb, however, was already usually placed sentence-finally. As for the verbal system, Old Persian had three tenses (present, past, perfect), several moods (indicative, subjunctive, imperative, optative and injunctive). There were two voices: active and middle, plus a passive formation in *-iya-* (Skjærvø 2016).

3.3.2 Middle Persian

Invasion of Alexander the Great in 4th century B.C. serves as the opening date of Middle Persian period (cf. Skalmowski 1986: 143–147). However, the oldest examples of Middle Persian language belong to the reign of Sasanian dynasty in the south-western Persia, that is 3rd–7th century A.D. Sasanians' coming to power led to the replacement of the north-western Parthian language used in the court of Arsacid dynasty (3rd cent. B.C.–3rd cent. A.D.) by its close relative, the south-western Middle Persian (cf. Rastorguyeva 1966: 7–8). Similarly to the Old Persian period, also here the oldest remaining texts are royal inscriptions, this time of Ardashir I and his descendants (Shapur I and Narseh). The youngest texts, on the other hand, are religious compilations composed no later than 10th/11th century A.D. (cf. Rastorguyeva 1966: 8–11). The Middle Persian script in which they were recorded was of a very conservative character—once used, it did not allow for later adaptations. Again it was lexical and grammatical "mistakes" rather than a recorded phonological change that spoke of the evolution of the language (cf. Gignoux 1984: 17–18).

The movement from Old to Middle Persian was characterized by certain changes in structure from synthetic towards analytic type of language. Due to the changes in stress, final syllables of Old Persian were lost and so were the inflectional endings and gender distinctions. Thus to express the syntactic relations between words, adpositions were used, although the early Middle Persian still kept a two-case system, regularly represented in plural, and a marginally represented (for example in proper names) feminine gender. Personal pronouns had their enclitic forms, used as oblique case; moreover, they could have joined prepositions. The word order of a nominal phrase was not fixed, the head noun could have either preceded the modifier (that was the older construction) or come after, joined by a 'relative connector' *-ī* (which has then become a New Persian ezāfe; see Skjaervø 2009: 221). There was no definite or indefinite article; the numeral *ēw* 'one' could have been used to express specificity (see Skjaervø 2009: 203–212). As for verbal morphology, the categories comprised of person,

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

number, tense and mood (indicative, subjunctive, optative), plus stem formations. Verbal conjugation was based on two stems: past and present, with the present stem regularly formed from the past one (except for the irregular verbs where it was unpredictable); plus two sets of endings (present and imperfect). Present stem was used to form present and imperfect tense, while past stem was used for the past and perfect tenses. The compound verbs with a prenominal were frequent, and so were the compound verbs with a preverb, which could modify the original meaning of the verb (Skjaervø 2009: 214–219). The important notion about tense formation in Middle Persian is the fact that, unlike both Old and New Persian, it represented a split-ergtive type of language, that is, its past forms (preterite and perfect) of the transitive verbs were formed with an ergative construction, while for the instransitive, a nominative-accusative agreement was used (see Dabir-Moghaddam 2012, Skjaervø 2009: 219).

3.3.3 New Persian

Arabic invasion of the 7th century A.D., which brought with itself not only new religion, but also new language, weakened the position of Middle Persian as an official, written language. The Middle Persian script, often ambiguous and difficult to read, was replaced by Arabic script modified to the needs of Persian phonology:

> The process by which bilingual Persian intellectuals came to write Persian in Arabic characters was not through an accidental loss of rigor (…) but on the contrary a conscious striving for literary parity and comprehensibility on the part of secretaries, scholars and poets. Arabic was rapidly and selectively adapted to the needs of Persian by appropriation of the writing system and selected vocabulary. As in Luther's Germany, the drive by literati **to promote the vernacular to literary status** was actively patronized by ambitious independent princes, such as the Saffarids and Samanids in the east of Iran (Perry 2012: 77, my emphasis).

Unlike other countries invaded by Arabs during their Islamic conquest, Iran did not submit itself to the complete Arabization, neither in terms of language, nor in terms of culture and national identity. The first centuries after the conquest were not "two ages of silence"[43], but rather a transitional period, as Browne put it:

> [I]t is by no means correct to imply (…) that the two or three centuries immediately following the Muhammadan conquest of Persia were a blank page in the intellectual life of its people. It is, on the contrary, a period of immense and unique interest, of fusion

43 *Do qarn sokut* [Two centuries of silence] is a famous book by Abdolhossein Zarrinkub, published in Iran in 1957, that discusses the first centuries after the Islamic conquest.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

> between the old and the new, of transformation of forms and transmigration of ideas, but in no wise of stagnation or death (Browne 1908: 204).

It was in those times when the final stages of transition from Middle to New Persian were taking place. The transition itself leaves us again with many questions, some of which were addressed by Gilbert Lazard in *The Origins of Literary Persian* (1993) and other articles on the subject. The difficulties encountered by researchers lies in the scarcity of linguistic material dating back to the earliest periods of New Persian emergence (i. e., until 10th century). It is generally agreed that New Persian is a "more or less regular continuation of late Middle Persian as it was used in the Sasanian court, i. e. as a natural change of written language under the influence of developments in educated spoken language" (Utas 2009: 154). The base of this new language was formed by two main varieties of Persian used by the end of Sasanian reign, *pārsi* and *dari*, of which the former was in wide use also in writing. Lazard reconstructs the events as follows: Middle Persian (*pārsi*) was the only written language, reserved for official and literary use and fixed in its archaic form already before the beginning of the Sasanian period. *Dari* was used in everyday talk and originally differed slightly from the written language, but while the written remained static, the spoken tongue developed to such an extent that by the end of Sasanian period it was given a separate name. "*Pārsi* (Middle Persian literary) and *dari* were not, strictly speaking, two languages, nor even two separate dialects, but rather **two stylistic levels of the same language**, *pārsi* being the medium of administration, of religion and of written literature, and *dari* that of everyday oral communication, that is to say the common vernacular of Iran" (Lazard 1975: 599, my emphasis). Utas argues that such a hypothesis does not take the spoken language into account sufficiently and suggests that those two varieties might have been only a small part of a much more complicated dialect situation, while the process leading to the formation of New Persian might be seen as a process similar to koineization[44] (Utas 2009: 154–155). Rossi (2015: 215) mentions Middle Persian *Pārsīk* and *Darī* as an example of diglossia.

According to the notice attributed to Ibn Al-Moqaffa',[45] which describes the linguistic situation of Persia by the end of Sasanian period, there were in fact three languages in use: *pahlavi*, *pārsi* and *dari*. The first one was spoken in the region called Pahla (previously Media; the western part of Persia) and was close to Parthian. The second, *pārsi*, was the language of priests (*mobeds*), educated classes and scholars, i. e., the Middle Persian language, official language of the

44 The process of koineization is a type of dialect contact in which new variety of a language comes to existence as a result of contact between speakers of different (but mutually intelligible) varieties of that language (Kerswill 2000: 669).

45 Ibn al-Moqaffa' (ca. 721–757) *Kalīla wa Demna*.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Sasanian court and the Zoroastrian religion, in which all the administrative and religious texts were recorded. *Dari*, on the other hand, was the language **spoken** at the royal court (Ctesiphon) as well as the southern regions such as Xorāsān and Sistān, where it had gradually replaced Parthian; it had also included influences from eastern languages (Xwarazmian, Sogdian). It was called *pārsi-ye dari* and it spread to the east along with a muslim conquest of Sogdiana and today's Afghanistan—it was more acceptable than *arabi-ye dari*, the official administrative and religious language after the conquest (Lazard 1993, Lazard 2003).

Summarizing the above paragraphs, it could be said that the beginnings of modern standard New Persian (the written/ literary variant) date back to late Sasanian era, when people of the capital would speak *dari*, a language distinct from that of official writings (*pārsi*, i. e., written Middle Persian, to which the name *pahlavi*, originally referring to Parthian, was later attributed). Soldiers and administrative workers would then carry *dari* to Khorasan, where—accepting influence from other dialects—it would slowly replace them and become a spoken language of that region. Ruling dynasties, i. e., Saffarids or Samanids, accepted it as their official language of administration, which led to standardization of its orthography and continuous spread in the region. While there was a competition between Arabic, *pārsi* and *dari* in the western parts of Persia, the eastern provinces were not under such a big influence of the former two, which gave *dari* a chance to flourish and a body of literature to emerge (Lazard 1975: 597–602, Perry 2012: 70–93).

The first phase in the development of New Persian, which took place in 8th–12th century A.D., is called Early New Persian.[46] As for the differences between late Middle Persian and Early New Persian, they were not as remarkable as those between Old and Middle periods:

> The distinction between Middle and New Persian is due as much to convention, to extra-linguistic factors such as the historical break [i. e. the Islamization of Iran – KW], and to changes connected to this break such as the shift of script (from Pahlavi to Arabic), as to linguistic differences (Paul 2000/2013).

Morphologically, Early New Persian can be seen as the continuation of processes which took place between Old and Middle periods. The case distinctions between nouns were lost. The new plural suffix *-īhā* (based on the Middle Persian adverbial suffix) was established for inanimate nouns. The order of noun phrases was moving towards the head noun before the modifier, joined by the *ezāfe*, although there were still some adjectives which could have preceded the noun

46 This Early New Persian has later evolved into New Persian, again without a clear linguistic break. Paul (2000/2013) refers to a description by Lazard (1963), who mentions the decline of dialectal features in 13th century Persian, which would suggest reaching a more or less homogenous standard.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

they modified. Enclitic pronouns could have been attached to all parts of speech and serve any syntactic function. As for the verbal system, the two-stem tense formation is kept. The tenses are formed with the aid of auxiliary verbs and verbal particles rather than synthetically. The adverbs *(ha)mē* (NP *mi*) and *be* were added to the present forms to add a durative of progressive meaning or "emphasize its syntactic autonomy" (Paul 2000/2013), but were not grammaticalized until the 10th–13th century; they could have also been freely attached to all past tenses in Early New Persian. The ergative construction used with transitive verbs in Middle Persian was changed to the accusative. There appeared a past participle with suffix *-a* (later *-e*), on the basis of which a present perfect was formed. The Early New Persian suffix *-rā*, unlike its Middle Persian use, is already attested to denote a direct object, namely an animate one (compared with a New Persian, where it is used with definite direct objects regardless of their animacy. The Middle Persian subordinating conjunction *kū* 'that' has merged with *kē* 'who, which' and *ka* 'when' into a multifunctional particle *ki* (Modern Persian *ke*) (Maggi and Orsatti 2018: 45, Paul 2000/2013).

The above descriptions do not exhaust the characteristics of the historical phases of the development of Persian, but they present certain tendencies that can be seen throughout the subsequent periods. A useful review of the development of Persian throughout all three periods along with a survey of existing studies is given by Maggi and Orsatti (2018).

3.3.4 Modern Persian

Nowadays there still exists a popular belief that New Persian has not changed much since the times of Samanids (873–1005), poetry of Rudaki (ca. 857–941) and great epic of Ferdousi, the *Šāhnāme* (ca. 1010). And even though, as Paul (2002: 22) has it, "Any Iranian who has some knowledge of Persian literature knows that the Persian of Ferdousī (late 10th cent. A.D.) greatly differs from that of today, with respect to both grammar and vocabulary," he observes right away that:

> This is somewhat less obvious for Hāfez (14th cent.). Here it is rather the poet's highly complicated metaphorical style than his archaic grammar that makes his poems so difficult to understand. In the case of Sa'dī (13th cent.), some Iranians (especially the older ones) would certainly maintain that Sa'dī's is the "best" Persian prose that has ever been written, and in principle it should also be a model for the written Persian of today (Paul 2002: 22).

The 10th century texts are indeed intelligible for educated Iranians today, possibly with the aid of a glossary and some instruction, which serves as an evidence for

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

the point, even though it is more the liking of old literature that makes it accessible and not the language it is written in (Kalbāsi 2002: 36). Another factor which contributed to the view of New Persian as unchanging is the tendency to base descriptive grammars of the language on the classical literary texts (namely classical poetry), present in studies of this language at least until the second half of 19th century, and to some extent even today, while "not nearly enough has been done to study and describe the actual language as it occurs in texts, rather than to describe an ideal, standardized set of grammatical rules" (Windfuhr 1979: 165–169).[47] Only in 1852 does Chodźko begin the preface of his *Grammaire de la langue persane* with a quote from Mirza Ibrahim[48]: "I seek to teach the Persian of Persians and not the Persian only of books" (Chodźko 1883: VII). But still the traditional concept prevails, and in the beginnings of 21st century the concerns of some scholars are similar to those of Ibrahim and Chodźko 150 years earlier. "Contemporary Persian spoken in Iran, according to historical classifications and grammars, is the same language that was used in 11–15th centuries,"[49] writes Barbara Mękarska in her short article on the geographically and historically influenced varieties of Persian (Mękarska 2005: 8). It would be wrong, though, to think that no changes nor diachronic development of New Persian were noticed. As early as in 1895, Paul Horn identifies certain changes that occurred in literary language (Ger. *Neupersischen Schriftsprache*) (Horn 1895–1901: 16–17), thus suggesting further periodization within the "New" period of Persian. Phillot's *Higher Grammar of Persian* (1919) generally differentiates classical Persian uses from the modern Persian ones. Xānlari (1979) suggests the three periods: genesis and growth (10th–12th century), period of standard Persian (13th–19th century) and period of transformation and renaissance (19th century onwards) (Xānlari 1979, as quoted in Paul 2003: 24). According to Windfuhr, even if later in the 20th

47 This way, the descriptive grammars should rather be considered prescriptive.

48 Mirza Muhammad Ibrahim (1800–1857) was a native Persian teaching oriental languages at the East India Company College, who authored *A Grammar of the Persian Language* (1841) in which he seeks to present the language in its everyday use rather than the language of literature, since "I have had frequent reasons to think that the existing system of teaching might be improved, and that a more *idiomatic* and *living character* might be given to its Rules and Principles" (Ibrahim 1841: iii).

49 This is a view Windfuhr opposes to more than 30 years earlier in his *Persian Grammar*, in which he gives a detailed presentation of studies published on Persian grammar throughout history. According to him, such an assumption ended with the recognition of structural grammar and focus on synchrony instead of diachronic studies: "This approach ended once and for all the assumption implicit in earlier studies that Persian was somehow a monolithic unchanging entity the purest example of which was Ferdowsi's *Šāhnāme* (early 11th century). (…) Once the stability of Modern Persian was recognized as a myth, the various periods of its development began to be recognized. **Most important was perhaps the recognition of the socio-linguistic layers of Persian, the difference between contemporary literary standard Persian and the layers of 'colloquials'"** (Windfuhr 1979:18).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

century the study of internal development of New Persian gave way to other concerns such as dialectology and studies of style in literature, certain periods could be distinguished: from the formative period of 7th–10th century and standardization of the language of poetry, through the period when prose was standardized (10th–12th century),[50] the classical period (until 15th century), the post-classical period of 15th–19th century and the contemporary period from 19th century onwards (Windfuhr 1979: 166).[51] For the purpose of clarity, the term "Modern Persian" as used throughout the book, refers to that latest stage of development.

3.4 The status of spoken and written Persian in contemporary writings in Iran and abroad

Scholarly opinions on the Modern Persian spoken and written in the Islamic Republic of Iran today are diverse. The differences lay not only in the perception of periods in the development of New Persian which were discussed above, but also in the approach to the issues of stylistic variation, standardization and normative rules. Since the main scope of this study is the variation between writing and speech, I will try to outline the present views on this aspect of modern Persian below.

3.4.1 Early writings on modern Persian (18th – early 20th century)

As it has already been stated, early writings on Persian were either of a diachronic-comparative character, or almost exclusively devoted to its classical, written form as witnessed in classical poetry. Many grammars were based on the variant of Persian used in India, and their chief purpose was to facilitate communication between English colonists and Indian higher educated classes:

50 The period of standardization of prose is called 'heroic' by Windfuhr in regard to *Šāhnāme*. The transition between this so-called 'heroic' and classical period is documented in Lazard's *La Langue des plus anciens monuments de la prose persane* (Lazard 1963).

51 Another Polish orientalist, Bogdan Składanek, in his *Introduction to Persian Grammar* dedicated mainly to Iranian Studies students, distinguishes two periods in the development of modern Persian: classical (until 15th century) and modern (from 15th century onwards), stating that his grammar of contemporary Persian deals with the latter, thus implying that contemporary Persian has not changed significantly for about five hundred years. He does mention the existence of the present day colloquial language in his book, but signals colloquial forms only as side comments to the literary ones (Składanek 1997: 21).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

> Whoever will study the Persian language according to my plan, will in less than a year be able to translate any letter from an Indian prince, and to converse with the natives of India, not only with fluency, but with elegance (Jones 1804: XVII).

For this use of Persian given by William Jones in 1771 in his *Grammar of the Persian Language*, indeed no focus on the actual Persian of Persians was needed. Similarly *The Persian Interpreter in Three Parts* by Edward Moises, printed in Newcastle in 1792, made no single mention of colloquial or spoken language and was devoted solely to the classical literature.[52] Around the same time, in a treatise on Persian and its regional varieties *Šaǧarat al-amānī*, a Khatri Persian-speaking scholar Mirza Muhammad Hassan Qatil specifically advises learners to avoid imitating the everyday speech of native speakers of Persian and devote to poetry and prose instead:

> Since the pillar of writing poetry and prose in Persian is the correctness of the language (*ṣiḥḥat-i zabān*) and the accuracy in following the native speakers (*durustī-yi tatabbu'-i ahl-i zabān*), the secretary and the poet must be aware of Persian peculiar vocabulary and conversation, and **the student of this discipline should not interfere with the current language (*rūzmarra*) of the native speakers (*ṣāḥibzabānān*)** and **make use of what he finds in their books**, considering himself an imitator, since there is a great difference between what is original (*aṣl*) and what is reported (*naql*). The imitator (*muqallid*) is held in consideration by the knower of the language (*zabāndān*) only through his hard work in poetry and prose (Qatīl 1872: 16; quoted after Pellò 2016: 209–210, my emphasis).[53]

It was not until the second half of the 19th century that the need to focus also on the actual state of language was recognized (cf. Mirza Ibrahim 1841, Chodźko 1852). This kind of approach was adopted by the authors of *The Vazir of Lankurān*, W. H. D. Haggard and G. Le Strange. Their book, which was published in 1882 in London by Trübner & Co., included not only a translation of a play by Fath Ali Axundzāde, but also "A text-book of Modern Colloquial Persian for the Use of European Travellers, Residents in Persia, and Students in India," as the second part of the title explains.

52 Perhaps the only notion of Persian as actually being spoken is a note on a quoted poem that was transcribed by W. Jones in *Asiatic Studies* where "he gives this poem a specimen of his plan of noting the Persic, and as coming nearest the true Isfahani pronunciation" (Moises 1792: 10).

53 *Šaǧarat al-amānī* is the first of three treatises Qatīl wrote on the various linguistic subjects, completed in ca. 1791 (Pellò 2016:207). Pellò comments also on a similarity between the above quoted observations by Qatīl and the fifty years later writings by Chodźko (1852), striving to describe the usual language of Persians instead of the artificial literary variety (Pellò 2016: 210).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

3.4.2 Early 20th century writings about Colloquial Persian; descriptive grammars

Some years after the writings of Chodźko and Mirza Ibrahim, the need of more focus on the language which was actually in use, was finally recognized and slowly, more books on the grammar of "Persian of Persians" started to appear: St. Clair Tisdall's *Modern Persian Conversation Grammar* (1902), Phillot's *Higher Grammar of Persian* (1919) and later *Colloquial Persian* by L.P. Elwell-Sutton (1941), which "is designed to cover the language of the common people and no more; polished speech can, indeed, only be acquired after a study both of Persian literature and of the Arabic language" (Elwell-Sutton 1941: vii). It is important to note here that those books were meant as textbooks for foreign use rather than as descriptive grammars. The authors, therefore, do not seek to explain and define their understanding of *colloquial* language more than in the quotations above. While it is clearly noticed that the language of people does differ from that of literature, theoretical attempts at the description of this notion are left to the study of dialectology or folklore (cf. Zhukovsky 1922 for instance). The publication of E. Denison Ross from 1931, entitled "Modern Persian Colloquial," is in fact a list of words collected by native Persian friend of the author as a supplement to the edition of Steingass's Persian Dictionary, without much further commentary.[54] As for the descriptive grammars of the second half of 20th century, the most influential ones are described in the analytical chapter of this book, since they were used as references for the analysis (see chapter 4, section 4.3.4, and chapter 5, section 5.1).

3.5 Contemporary theoretical approaches to the study of Modern Persian

Contemporary works on Modern Persian show more awareness of the internal variability of this language, which should receive more focus and attention from linguists. One of the examples of a great concern with the need of codification of a new standard of modern Persian is expressed by Mękarska, who states that "a unified colloquial standard of Persian does not exist" and descriptive grammars published both in Iran and abroad describe a literary language which no one speaks anymore and hardly anyone writes (Mękarska 2005: 9). Mękarska suggests that the formulation of new, modern standard[55] is necessary and a traditional

54 This marks a certain attitude towards colloquial Persian which associates it with vocabulary and phraseology along with a complete neglection of its grammatical features.

55 "The mass and depth of the changes to the language system demands codification of the modern language standard, which leads to the assertion that the history of the Persian

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

stratification as presented in Persian grammars from up until 1970s[56] is no longer valid, because other factors gain importance:

> firstly, there exists a definite yet troublesome difference between spoken and written language, secondly, classical literary language is now acquired only passively by reading classical literature, thirdly, the theoretical existence of a general literary language (Lazard's *persan classique simple* (1990: V) to be used in official, public written and spoken performances (which means that it should cover public speeches, Friday sermons, radio and TV broadcasts) is in fact limited only to writing; a speech will often start within this variety but soon the speaker is drawn to colloquial language, similarly news will be read out in literary language, but commented in colloquial. The same problem appears in education since there is a substantial difference between textbooks and oral lectures. (…) Modern Persian researchers are helpless, they come to the conclusion that there exists no literary standard of Persian today! (Mękarska 2005: 8).

Those radical views are not justified by what can be found in Iranian literature on the subject, though. There certainly are scholars who are aware of the situation and mention the existence of two standard varieties of contemporary Persian in their works. Although inconsistent terminologically to a great extent (we will return to this problem below), they are rather consistent in the opinion that a) there exist two standard varieties of contemporary Persian: written (*neveštāri*) and spoken (*goftāri*), and b) contemporary standard of Persian is based on the dialect of Tehran, which has spread throughout the country with help of mass media, education, urbanization and growing importance of the capital (Jahangiri 1980: 34, Sādeqi 1984: 19,[57] Rezāyi 2006: 25, Sedighi 2010: 12).

> Nowadays, apart from a written standard resulting from natural and historical processes, a spoken standard different from the written one has come to existence. **Standard spoken variety of modern Persian is the dialect of educated classes of Tehran.**

language should in fact be divided into four stages: Old-, Middle-, New- and Modern Persian." (Mękarska 2005: 8). She also considers application of a term *cultural dialect*, defined as 'language of the educated class, which either stands as a base for codification of a literary standard or clearly cooperates with it' (Polański 1993: 244), to the contemporary spoken variety, but comes to a conclusion that such a definition does not suit the linguistic situation of Iran.

56 She mentions four variants: 1) classical literary language understood by all traditionally educated speakers, 2) local literary language used in a large, historically and politically established region (Iran, Afghanistan, Tajikistan), 3) colloquial language (a modified variety of general literary language, influenced by local dialects of the most politically and culturally significant provinces) and 4) jargon (professional, social, court) (Mękarska 2005: 10).

57 According to Sādeqi, written standard is the language of educated class and intellectuals as well as the variety taught to foreigners. It has three distinct varieties (*gune*): scientific (*'elmi*), radio, tv and press (*rādio va televiziyun va ruznāmehā*), literary (*adabi*), which is then divided into poetry (*še'r*) and prose (*nasr*). The discussion on spoken standard is limited to the statement that nowadays it is the language used by people of Tehran. "Those who wish to speak standard Persian attempt to copy the dialect of Tehran" (Sādeqi 1988: 21).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

> **This variety, along with the written one, is used in official circumstances such as public speeches and media broadcasts.** (…) Today in our country, even in the most official situations, people use the spoken variety while speaking. Some of them, bearing in mind their profession and social status, might conform to the rules of written standard even while speaking, but that is an exception rather than a rule (Rezāyi 2006: 25–26, my emphasis).[58]

The existing theoretical approaches to the linguistic situation of Modern Persian and its internal variation (both in Iran and abroad) can roughly be grouped into four major streams, corresponding with trends in general linguistics.

3.5.1 Stylistic approach

In this approach, the focus is on formal and informal styles instead of spoken and written variants. It is the oldest and perhaps also the most influential paradigm in the study of variation within Modern Persian, both in Iranian and foreign scholarship. It encompasses writings of scholars such as Hodge (1957), Jazayery (1970), Afxami and Qāsemi (2003) to name a few. This is also the approach that can be traced in the background of grammars which do not deal with spoken Persian exclusively but do mention the issue (i. e. Lambton 1953, Lazard 1957).

Hodge's *Some Aspects of Persian Style* might be considered fundamental in the stylistic approach to variation in Modern Persian. It was "the first linguistic attempt at systematic analysis of stylistic variation in Modern Persian" (Hillmann 1981: 332), which has later inspired many researchers both in Iran and abroad.

Hodge based his observations on the standard spoken language of Tehran obtained from Iranian students and his observations in Iran and Afghanistan in 1955–1956. His use of printed sources was limited as they "are generally concerned with what is below defined as the formal style and are not sources for the language as a whole" (Hodge 1957: 356).[59]

58 This statement would go in line with Finegan and Biber's belief that with higher status speakers, "even their conversational norms will come to reflect more frequent use of literate forms" (Finegan and Biber 1994: 339)—although one could argue that in Iran, those who are most fluent in classical literary language are not necessarily those of higher social status (cf. Mękarska 2005).

59 Amongst those printed sources he mentions Elwell-Sutton's *Colloquial Persian* (1943) ("the only printed book which gives any extended presentation of informal material, even though the beginner cannot distinguish the informal from the formal"), Lambton's *Persian grammar* (1953) (which "has a brief and totally inadequate section on the informal, which she calls Colloquial Persian") and Levy's *The Persian language* (1951) (which "carefully defines the 'Persian language' to include only what we here call the formal") (Hodge 1957: 332).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Hodge describes two sets of phenomena under the term "style": the formal-informal and the familiar-polite-deferential. He notes the existence of a broad division between the formal and informal style, which is set mainly by morpheme alternation, while phonetic variation helps establish further distinctions (Hodge states specifically that his intention is to define styles on the basis of morpheme alternants and phonetic variation rather than the variety exhibited in the choice of words and in syntax) (Hodge 1957: 361). The most striking morpheme alternation is in his opinion the contrast of formal and informal verbal stems of the most common verbs as well as the present verbal endings (plus copula forms) (Hodge 1957: 362). As for the phonological differences, he mentions for instance the alternation before nasal consonants *m* and *n* of *ā* (formal) with *u* (informal) (Hodge 1957: 363). The distinction between formal and informal, however a broad one, is considered regular, Hodge mentions also the tendency of native speakers to use the formal speech to foreigners and to identify it with the standard, grammatically correct language. He recognizes two subvariants for formal (quotative and normal) and informal (deliberate and colloquial) styles, although those subvariants are meant to be the extremes of a continuum rather than sharply defined categories (Hodge 1957: 364).[60] Stylistic variation concerned with morpheme substitution rather than phonetic alternations is found in politeness levels, where there are four categories described: the familiar, polite, deferential, and royal (Hodge 1957: 366). The differences here lay mostly in the choice of verb forms (singular or plural, second of third person) between familiar and polite, and in the choice of verbs themselves as well as other words between polite and deferential levels (Hodge 1957: 367). However both formal-informal and familiar-polite-deferential variation is described as stylistic, an important distinction is made here that can be illustrated by a quote: "An utterance on a given politeness level may be either formal or informal. (…) [T]here is a coincidence of all politeness levels with both styles" (Hodge 1957: 368). It shows that those two phenomena do not belong to one category. The variation between formal and informal is of a more general kind (thus applicable to all politeness levels).

Another important work on the stylistic variation in Persian is Jazayery's *Observations on stylistic variation in Persian* (1967). Jazayery distinguishes internal variation in linguistic form that is "automatic within the linguistic structure" (Jazayery 1967: 447) from variation related to extralinguistic phenomena

60 Those style levels are established on the basis of form, not use, but there are also some general statements made about their use: the quotative for an utterance spoken carefully and deliberately (like repeating for emphasis or correction), the normal-formal for reading aloud, giving a speech, public broadcasting etc. as well as with foreigners, especially on the first encounter. The deliberate informal is for ordinary conversation between people who are not on very familiar terms, while the colloquial is for those very well acquainted (Hodge 1957: 366).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

such as dialectal variation (geographically based) or, within a dialect, variation determined by social context (level of politeness or formalness), which he names 'style.'[61] While the politeness-conditioned variation is treated only briefly, the rest of the paper is devoted to the differences between the 'formal' and 'informal' styles. The formal style is treated as basic here, while the informal is meant to be a deviation from that base (Jazayery 1967: 449). Again, similarly to Hodge, Jazayery brings attention to the phonological (or morphophonological[62]) variation as the most common and noticeable. This includes consonant cluster reduction in final syllables, dropping of certain syllables, assimilation within morphemes and words and the alternation between formal *ā* and informal *u* in certain phonological contexts as well as the differences within verbal system (different verbal endings, reduced present stems of some verbs, present perfect formation, variation in object marker and plural marker). Then there are the morphological and syntactic differences listed.[63] As for the vocabulary, here "the problem of formal versus informal (…) and that of politeness levels (..) is often closely tied to such phenomena as slang, or to **the relationship between speech and writing**" (Jazayery 1967: 455, my emphasis). At the end of his paper, Jazayery attempts at formulating some rules as to what makes an utterance informal: "the verb forms (stems and endings) must have the phonological shapes specified for the informal style. This is obligatory and constitutes the minimum requirement." (Jazayery 1967: 455).[64] The phonological forms of words are also important, but the lexical choices lose their importance. He makes an observation echoing that of Hodge (1957) as well, that "the stylistic differences based on politeness and those based on formalness operate independently of each other" (Jazayery 1967: 455). Differences between speech and writing should, in his opinion, be treated as a non-stylistic phenomenon, even though their relationship with the formal-informal distinction is quite close:

61 This distinction between dialect and style seems to be made more for the purpose of clarity, though, than theoretical necessity.

62 He uses the term 'morphophonemics' to refer to the 'phonological shapes of morphemes which have grammatical functions' (Jazayery 1967: 450).

63 Not all those observations seem to be valid, especially (5), where he states that "in informal style, the connective /-e/ (*ezāfe*) has almost completely disappeared" (Jazayery 1967: 452).

64 This observation in especially important in context of popular understanding of informal (or, as we would say, colloquial) variant of Persian by native Persians themselves. There is a tendency to associate it (in the *āmiyāne* understanding) language only with vocabulary and phraseology, with disregard to its grammatical (and even phonological) features. This attitude might be traced back to the works of Mohammad Ali Ǧamālzāde and Sādeq Hedāyat (first half of the 20th century) whose focus and interest in folk language would lay mostly on idiomatic expressions (pers. *estelāhāt*) and sayings that were collected and incorporated into their fiction, and is still present in today's scholarship as well as the general views.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

> **for the differences between speech and writing, these are best treated as something other than stylistic**, although the problem of the relationship between formal-informal distinction and the differences between speech and writing on the other hand has become more complicated in recent decades, as a result of certain cultural developments. The current situation can tentatively be phrased as follows: Writing *tends* to use the formal style (Jazayery 1970: 455–456, my emphasis).[65]

Abol-Hassan Naǧafi, one of the most notable scholars on colloquial language among Iranian linguists and the author of the *Dictionary of Colloquial Persian* [pers. *Farhang-e zabān-e āmiyāne*],[66] presents yet another view on the stratification of modern Persian and understanding of its colloquial variant. In the introduction to his dictionary, he differentiates three levels (*sath*) of Persian: literary (*adabi*), formal (*rasmi*) and colloquial (*āmiyāne*). Besides *āmiyāne* for 'colloquial,' he introduces the term 'daily language' (*zabān-e ruzmarre*; as an equivalent of *mohāvereyi*), those two being the components of spoken Persian language (*zabān-e goftāri*). He uses 'daily language' for the variety used by educated and partly educated classes in conversations with a familiar partner of equal status and also in some "free-style writing" (*zabān-e neveštāri-ye kamobiš āzad*), i.e., in dialogues between literary characters. "It is a language which family members use in their conversations, but will rather not use it when speaking to a stranger," he points out. Colloquial language in his understanding is not a sociopolitical but rather sociocultural term, which covers the expressions used in daily conversations (*kalamāt va tarkibāt-e mohāvere*) by partly educated classes who speak in a careless manner (*bi-qeyd-o-band soxan miguyand*) (Naǧafi 2000: 7).[67]

Another account of this kind of stylistic stratification is given by Afxami and Qāsemi, who distinguish five styles within Persian: deferential (*mohtaramāne*), formal (*rasmi*), semi-formal (*nime-rasmi*), colloquial *(mohāvereyi*) and slangy *(āmiyāne*) (Afxami and Qāsemi 2003: 5). Interestingly, their definition of style (Persian *sabk*) is based solely on the level of formality:

> People within a linguistic community, after assessing the level of formality of the situational context, the level of their familiarity with the addressee and social distance between them and the addressee, choose the proper form of the language known otherwise as "style" (*sabk-e zabāni*) (Afxami and Qāsemi 2003: 5).

Afxami and Qāsemi state specifically that "language styles (*sabkha-ye zabāni*) are different from language varieties (*gunehā-ye zabāni*), which exist in every

65 In another article, however, published only four years earlier, he would write on spoken and written styles (Jazayery 1966).

66 It is worth noticing that Naǧafi uses Persian *āmiyāne* and not *mohāvereyi* to refer to colloquial language in a general manner.

67 This "careless manner" might be what Sami'i called "broken language" (*zabān-e šekaste*) and Rezāyi took as yet another way of referring to the spoken standard (Rezāyi 2006: 34, see below in the next section).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

language because of individual, social, geographical and historical factors" (2003: 6).[68] Also the fact that they apply *sabk* as the equivalent of *style* in the (socio)linguistic understanding, calls for attention, since however the English term 'style' is applied to spoken Persian in many of the scholarly writings, especially in the late 20th century (where it is largely described as an informal style), the Persian term *sabk* 'style' does not appear that often in any of the above quoted native studies of contemporary Persian in regard to the written, spoken and even colloquial variants. Most of the scholars would stick to *gune* ('variety, variant'), much less restricted term, while *lahǧe* ('dialect') seems to be used preferably on geographical grounds—as long as the dialect of Tehran is being referred to as *lahǧe-ye tehrāni,* the standard dialect of Persian becomes simply *fārsi-ye me'yār.*[69]

3.5.2 Functionalist approach

A functionalist view is based on the notion that difference between spoken and written language is the difference of medium, whereas language is essentially the same (cf. Halliday 1985). One of the important representatives of such approach is Iran Kalbāsi. What is important in Kalbāsi's point of view is that spoken (*goftāri*) variety of language is **not** the variety which is actually used in speech (*mohāvereyi* 'colloquial').

Irān Kalbāsi distinguishes colloquial standard from the dialect of Tehrān, and strongly opposes to such unification. She underlines the fact that colloquial variety of Persian should be free of any dialectal influence. While others (e.g., Jazayery 1967, Windfuhr 1979, Sādeqi 1984, Rezāyi 2003, Mękarska 2005) see the growing linguistic influence of Tehran as a cultural and political center of Iran, she specifically insists that becoming a standard bears no relation to being a capital or not (Eslāmzāde 2009: 10),[70] and that standard language is above all the language of education. She points out that "Persian language, apart from dialects such as Tehrani, Esfahani, Shirazi, Mashhadi, Kermani and so on, and two general varieties: written (literary) and spoken (colloquial), has also other, subsidiary varieties that depend on gender, age, occupation, social class and place of residence," thus suggesting further sociolinguistic variation, but also underlying that written and spoken varieties should be understood as superior to that. Among

68 Such understanding of *gune* would perhaps be closer to the notion of 'lect.' Generally, Afxami and Qāsemi in their study follow the approach of Hodge.

69 It is also worth noting that in consequence, the distinction between *sabk, gune* and *lahǧe* is rather vague.

70 Still the article entitled *Fārsi-ye goftāri va neveštāri* (Spoken and written Persian) deals with… educated Tehrani speech (Kalbāsi 2002).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

factors which contribute to the growing gap between speech and writing, she mentions the principle of least effort, which in her opinion is effective in most of the speech communities with similar results (Kalbāsi 2002: 50).

Kalbāsi is also very strict in differentiating spoken and written varieties of standard Persian (*fārsi-ye me'yār-e goftāri va neveštāri*) from colloquial Persian which she calls *mohāvereyi.* Unlike Rezāyi, who understands both "spoken" and "colloquial" as references to spoken standard,[71] in her interpretation, standard spoken and written varieties differ only by medium of use (people write *neveštāri* and read it out in *goftāri*, but do not speak *goftāri*), while the variety used in speech is *mohāvereyi*, distinguished from standard by means of phonology, vocabulary and grammatical structure and not included in standard language at all (ibid.). According to her, "contemporary Persian should be divided as follows:

1) standard written or literary language (*zabān-e neveštāri-ye me'yār* or *adabi*), in which we write,
2) non-standard written language (*zabān-e neveštāri-ye qeyr-e me'yār*), which is used for writing different dialects and accents (*lahǧe-hā va guyešhā*), i. e. writings of Sādeq Hedāyat, Mohammad Ali Ǧamālzāde and others,
3) standard spoken language or colloquial Persian (*zabān-e goftāri-ye me'yār yā fārsi-ye mohāvereyi*), which is a language of educated people, free of any dialectal or other language influence,
4) standard Tehran dialect, which is a dialect of people who were born and raised in Tehran (…),
5) non-standard Tehran dialect or Tehran vernacular (*lahǧe-ye Tehrāni-ye qeyr-e me'yār* or *lahǧe-ye āmiyāne-ye tehrāni*), which is a language of non-educated people of Tehran who were born in that city and who, depending on their profession and social class, might use their own peculiar words" (Kalbāsi 2003: 10).

3.5.3 Diglossic approach

The idea of Modern Persian being diglossic[72] was brought to attention by Hungarian scholar Éva Jeremiás in 1984.[73] Jeremiás studied differences between writing and speech in Modern Persian, coming to a conclusion that they do fulfill

71 'Spoken language (*zabān-e goftāri*) and colloquial language (*zabān-e mohāvereyi*) are one' (Rezāyi 2008: 27).

72 For a theoretical overview of diglossia, see previous chapter. Jeremiás notes at the very beginning of her paper that the differences she attributes to diglossia roughly match those understood as stylistic differences of language varieties used in formal and informal situations or the traditional distinction between classical (literary) and colloquial (spoken) language (Jeremiás 1984: 271).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

the necessary requirements of a diglossic situation as defined by Ferguson (1959). At first, she challenged Ferguson's own observation that Iran is a case of "standard-with-dialects" language situation, where one of the local dialects for some reasons gains more importance and starts being used and standardized above other dialects (becomes a superposed variety for those who did not use it their local variety before), while those dialects remain in use for informal situations among their speakers. According to Jeremiás, such statement would be true only if the Modern Standard Persian (its formal or otherwise written variant) was based on the local dialect of Tehran, which is obviously not the case if we consider the history and formation of New Persian and Modern Persian (Jeremiás 1984: 273–275, see also 1.2 in this chapter).

> Persian had a written (formal) variety since approximately the Safavid era (16th c.), which continued the traditions of the common old literary language, strictly speaking those of the classical prose texts, although it was very markedly different in grammar. It can be regarded as the immediate forerunner of today's Formal Standard Persian. (…) The spoken language, the means of everyday communication was very strikingly different from both this formal official style and the language of classical poetry. (…) This colloquial language is mainly based on the Tehran dialect, this is what tends to be standardized through the mass media today (Informal Modern Persian) (Jeremiás 1984: 275).

Jeremiás studied the ways in which linguistic situation of Iran is reflected in both Western and Iranian grammatical literature, describing Western writings as dealing mostly with prescriptive rules of the formal varieties, while Iranian writings are even more conservative and based largely on the body of classical and classicizing literature, with colloquialisms noted on the level of vocabulary and without any serious notice of differences in grammatical structure (Jeremiás 1984: 276–279). In her opinion, the stylistic approach to the differences between formal and informal variants of Persian in Western linguistic literature (which was also discussed in the above paragraphs) caused the changes which language has undergone to go largely unnoticed on the Western side, "because they [the Western grammarians – KW] refer the phenomena not fitting in the formal version of Modern Persian into the category 'stylistic variant' (archaic, classical, Arabic, colloquial, substandard, slang etc.)" (Jeremiás 1984: 276).[74] In Iran, the

73 Interestingly, Charles Ferguson, who is recognized as the father of diglossia (although it must be noted that he wasn't first to use this term), classified Iran's situation differently, as a "standard with dialects" situation (Ferguson 1959), which holds true until today. Perhaps a more adequate description would be: a diglossic standard with dialects?

74 By the end of her paper, a remark is made on the use of terms "variety" (formal/informal), "language" or "style" (colloquial/literary) without a detailed classification, that in Jeremiás's opinion is of no use since "our experience suggests that there are no clear-cut stylistic

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

preservation of classical norms is said to be due to the "special, uniquely intimate relation with classical literature" (Jeremiás 1984: 278). Then she dwells on the differences between formal and informal Modern Persian, examining their historical roots—she actually blames Hodge, Ferguson and Jazayery for taking into consideration only the easy-to-note linguistic facts such as phonetic and morphological alternations and the lexicon (Jeremiás 1984: 280)—so she herself claims to examine the much greater differences and presents compound progressive tense formed with *dāštan* (with thorough historical analysis), definite article *-e*, the different use of enclitic pronoun suffixes, preposition omission and changes in the word order.[75] All those examples are presented to have roots in historical forms of Persian instead of being a kind of colloquial innovation.[76]

The general conclusion Jeremiás draws from the analyzed data is that Persian is "a striking example of diglossia," a situation which dates back to the "establishment of a unified literary language" (Jeremiás 1984: 286). She notes also that analysis of code variation within a living language can help analyze the historical data and discover mechanisms of linguistic change. As long as five examples of linguistic discrepancies between H and L variants of language might not seem convincing enough to declare a language diglossic, the other circumstances do fulfill the necessary requirements (there are nine distinct features of diglossia and all of them are present: mutually exclusive function of the two varieties, a difference in prestige, highly esteemed literary heritage in H variety, L variety being the first to be acquired by children, H variety treated as standard and thus described in the grammar books and dictionaries as a default form of the language in question, stability of the diglossic situation, grammar, lexical and phonological differences between H and L varieties—all those criteria are met for spoken vs written Modern Standard Persian).

A completely opposite view is presented twenty years later by Perry (2003). Not denying the differences between colloquial and literary Modern Persian, Perry argues that that they do not justify the classification of Persian linguistic situation as diglossic by comparing it with the situation of classical Arabic vs colloquial Egyptian. By awarding certain score for each of the selected features

varieties. The various sections of the style-range cannot be clearly marked off, there are overlappings and mutual applications" (Jeremiás 1984: 287).

75 The differences in word order connected with the omission of certain grammatical morphemes and prepositions are to be bound with the difference of medium (thus not the formal-informal, but the written-spoken range) as they are often accompanied by changes in intonation, stress and pitch (which serve syntactic functions), means available only in speech (Jeremiás 1984: 285).

76 It is worth noting that similar observations were also made on Kabul Persian (for example Henderson 1975); they were also discussed by prof. Lutz Rzehak during his lecture "Dari – the Persian Language of Afghanistan" in the Department of Iranian Studies of Jagiellonian University on 5th October 2015.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(described as those which are diagnostic for diglossia) in both languages, Perry calculates the extent to which those languages are diglossic. The highest rank differences between high and low variants is given to diverse syntactic structures (differing between H and L or present in one of them only) and distinctive semantic features. Morphological differences come second, then there come morphophonological alternations and lexical pairs distributed exclusively between H or L. The least important factor are phonological alternations and cognate words that differ significantly in pronunciation between H and L. After presenting examples of those features and counting the scores, Perry comes to a conclusion that "on the basis of this mechanical quantification, Arabic would appear to be about three and a half times as diglossic as Persian" (Perry 2003: 24), with the greatest contrast in lexicon and morphology, attributed to the grammatical complexity of MSA[77] in terms of inflectional forms, compared to a significantly reduced number of inflectional forms in Modern Persian present from the Middle Persian period (Perry 2003: 21). Criticizing Jeremiás for her determination to prove Persian diglossic, Perry states the otherwise:

> However crude and debatable the methods, I think the evidence here is compelling that Persian is not nearly as dichotomous as Arabic in terms of a broad definition of diglossia. If, moreover, one accepts radical differences in morphology and a plethora of non-cognate lexical substitutes as salient diagnostic for diglossia, I believe I have shown that Arabic is, and Persian is not, diglossic in principle, i. e., in terms of the narrow definition. The real question to be answered is not "How diglossic is Persian?" but "How did Persian avoid diglossia?" (Perry 2003: 26)

As an answer to the posed question, he suggest the following: Firstly, unlike the purist movements that had lead to diglossia in Greek or Hindi of northern India (or modern Arabic to some extent), the purist movements in Persian were too moderate to have such an impact. Secondly, Persian throughout its history entered bilingual (not diglossic) relationships with Arabic at the time New Persian was evolving or Turkish after the Mongol conquest. Thirdly the fact that spoken Persian was used as the lingua franca of the Eastern Islamic world would enhance the tendency for it to be universal and avoid a diglossic breach. "The pressures for a fissure of this language of solidarity along diglossic lines were simply not there," he concludes (Perry 2003: 28).

However enticing the idea of measuring diglossia quantitatively appears, the comparison of two different language situations to prove one less diglossic than the other is worth more discussion. Is the fact that Arabic is more diglossic than Persian in terms of vocabulary and morphology enough to dismiss Persian diglossia at all—in favor of stylistic or sociolectal and dialectal variation? It might

77 Modern Standard Arabic.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

not be a "striking example of diglossia," as Jeremiás (1984: 286) would have it, but even if some of the differences between formal and informal standard (as Perry calls them) are actually simple transformational rules, they still are present, operative, and necessary for a speaker to shift between those standards freely. Without knowing the rules (or having learnt both varieties specifically), transformation is not possible and in case of a non-native speaker, mutual intelligibility might be at risk.

The issue is brought back to the attention in 2018 by Mahmoodi-Bakhtiari, who again declares Modern Persian diglossic, agreeing, however, with some of the remarks Perry made in his 2003 article (Mahmoodi-Bakhtiari 2018). First of all, he brings into question the inferior status of spoken Persian vs the formal/written standard, stating that:

> The drastic expansion of literacy, modernization, and growth of mass media in Iran during the recent decades has changed the status of spoken Persian as merely a dialect of Persian spoken by the inhabitants of Tehran. (…) Spoken Persian has gained a status of its own, to be studied as the major medium of speech among the Persian speakers of Iran (Mahmoodi-Bakhtiari 2018: 184).[78]

Then there comes a brief introduction into the study of diglossia, its definition and features, followed by what the author calls "three misunderstandings" (Mahmoodi-Bakhtiari 2018: 187) in regard of the low variety of language. First he mentions the lack of prestige which is supposedly characteristic of this variety, challenging this notion with a few examples from classical Persian poetry that supposedly use word forms belonging to the low variety. This argumentation does not seem valid though for at least two reasons—low and high varieties considered in diglossia both represent varieties of a contemporary standard, while the classical language of poetry is a historically distinct phase in the development of language, so there is no common ground for comparison; secondly, three out of four examples could rather be explained as a matter of meter (*licentia poetica*). The second misunderstanding is the expected lack of politeness within the low variety—yet we could use the argument of Hodge (1957) and Jazayery (1967), pointing out that politeness levels constitute an independent level of stylistic variation, so again we lack a common ground for such a comparison (and of course it is perfectly possible to formulate both polite and impolite utterances in low as well as high variety). The third misunderstanding is mixing up of diglossia and bilingualism. Those three misunderstandings (existent or not) might actually prevent laymen from accepting Persian as diglossic, but they should not be an obstacle for anybody more acquainted with the subject.

78 This statement is in accordance with the opinion expressed by Mękarska (2005: 8), quoted in the section 1.3.2 of this chapter.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Mahmoodi-Bakhtiari analyses the approach of Jeremiás (1984) and Perry (2003) and proceeds to show that "Persian is diglossic, although not to the extent that Arabic is" (Mahmoodi-Bakhtiari 2018: 190)—this way taking a stand somewhere in between the two previous outcomes. He notes that all nine features characterizing diglossia are present in Persian and it is only the lexicon that is not as dichotomous as in the other diglossic languages such as Arabic. His list of examples of diglossic features is quite extensive and they are organized into groups:

1) phonological alternations: vowel shifts and vowel harmony, vowel and consonant deletions, syllable deletions, non-syllabic cluster deletions, and insertion,
2) morphological alternations: free morphemes (personal enclitics, question words, some prepositions), bound morphemes (verbal endings), affixes, clitics, compounds (including a specific structure of noun/adjective + *bešo* denoting capability of reaching some state),
3) syntactic alternations: progressive constructions, adjective pluralization, definiteness marking, and especially differences in word order (with the biggest number of examples),
4) semantic alternations (24 examples of words which represent different meaning when used in spoken or written variety)[79] (Mahmoodi-Bakhtiari 2018: 191–210).

Having studied the examples, Mahmoodi-Bakhtiari concludes that "if we consider diglossia as a continuum, Persian may have a place on that continuum, although it may not necessarily be close to the classical diglossic languages" and replaces Perry's (2003: 26) question of "How did Persian avoid diglossia"? with "To what extent is Persian diglossic?" (Mahmoodi-Bakhtiari 2018: 210–211), thus leaving the discussion of Persian diglossia open—the extent of which is visible in a recently published *Oxford Handbook of Persian Linguistics* (2018), where (not far from each other) we can find the two statements quoted below:

> The answer to the question of whether or not Persian speech community is diglossic, depends on different definitions, criteria and interpretations linguists have had for the term since Ferguson's paper. On the basis of the extended meaning of diglossia, Persian speech community, like some other language communities around the globe, can be considered as diglossic, because its different standard and non-standard varieties or styles have different social functions, but in the framework of Ferguson's definition, Persian community is **not a good example for the term diglossia** (Modarresi 2018: 336, my emphasis).

79 Perhaps this group, rather than lexical sets of different words for the same item, could be used as a measure of Persian "semantic aspect" of diglossia.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

and:

> Persian demonstrates a **strong form of diglossia**, which here refers to situations where the language spoken by the people in a society differs considerably from the traditional written variant (Ferguson 1959). […] The recent explosion of blogs, microblogs (e.g. Twitter), forums and social networks has created a large amount of text written in the conversational or low-level variant of the language.
>
> The main differences found in conversational language involve phonological alternations, differences in word-formation rules, deviance from the verb-final word order found in high-level speech, and a large number of neologisms (newly coined words) and loans. In addition, code switching—occasions where a speaker alternates between two or more languages, or variants of the language, in the context of a single conversation—is often encountered in online social media text (Megerdoomian 2018: 472; my emphasis).

But perhaps the most important notion regarding Persian diglossia is made elsewhere (still in the same book), where the discussion of the 2500-year development of Persian it is concluded that:

> In actual fact, a distinction between a spoken variety—with a further stylistic differentiation between a formal, official, or educated spoken sub-variety, and an informal, familiar, or colloquial spoken sub-variety—and a literary or, for modern times, a standard variety should be taken into account not only for Contemporary New Persian, but, with the due differences, also for each period in the history of Persian (Maggi and Orsatti 2018: 28).

3.5.4 Sociolinguistic approach

Two different views can be distinguished within this group. The first one deals mainly with language planning, standardization, and language policies (Sādeqi 1984 and 2001, Modarresi 1990, Rezāyi 2003, Navidnia, Mehrani, Samar 2010). The second would be those studies which fall under the variationist paradigm founded by William Labov (Modarresi-Tehrani 1978, Jahangiri 1980), which focus on patterns of linguistic change mostly within spoken registers rather than the variation between speaking and writing.

Sādeqi (1984) in his paper *Zabān-e me'yār* [Standard language] explains the theoretical aspects of the process of standardization first: what constitutes a standard, how is it codified and defined, and what are its functions. As for the Persian language itself, Sādeqi clearly states that at the time being, there are two standard variants of Persian: the written and the spoken. Within the written variant three subvariants can be further distinguished: scientific (clear of emotional expressions, elaborate, elegant and precise), media (radio, television and press; less elaborate, with foreign borrowings) and literary (poetry and prose are

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

treated separately). He notes that similar subvariants exist in the spoken variant as well, without going into further details though (Sādeqi 1984: 21).

Similar aspects (and similar opinions) on language standardization appear in Rezāyi (2006). First part of his paper is devoted to the problems of defining a language standard and differentiating it from other language forms that do not necessarily have clear-cut boundaries. Rezāyi points out that some of the important problems in defining standard language are the extra-linguistic elements which enter the discussion as well as the interference of the other terms such as official language (*zabān-e rasmi*),[80] common language (*zabān-e moštarak*), national language (*zabān-e melli*) or written language (*zabān-e nevaštāri*), so that some definitions are more accurate for the above mentioned than for the standard itself. Moreover, the term *standard* is of a relative character—some language variants are more and some are less standard, given the inherent lack of uniformity that characterizes language in use (compared to abstract linguistic norms) (Rezāyi 2003: 21–22).

It is important to note that not all linguists dealing with standardization do acknowledge the existence of two standards. Some do not differentiate between the language of writing and that of speech (perhaps reducing it to the level of styles within standard, although the authors do not specify it anywhere in their article)[81]:

> "By standard Farsi we mean that variety which is used in print, and which is normally taught in schools. **It is also the variety spoken by educated people and used in news broadcasts** and other similar situations. Apart from the preceding internal and external factors that increase standardization, speaking the standard variety is associated with a kind of prestige, so the speakers of other dialects try to adopt it, and modify their dialect according to the standard one" (Navidnia, Mehrani, Samar 2010: 23, my emphasis).

3.5.5 Anthropological approach

Another set of works on variation in Persian are those which combine linguistic variation with cultural codes of behaviour and deference. This is the approach adopted and developed by William O. Beeman in his writings, namely *Status, Strategy and Style in Iranian Interaction, The Hows and Whys of Persian Style*

80 In Persian, the term *rasmi* (official, formal) can be applied both to the formal style of language and to its official, state-recognized form.

81 It might be worth noticing that quoted observation is not referenced in literature, so it seems to be a personal account of the authors. In fact, this article's references are limited to English language scholarship on sociolinguistics without any sources on Persian, so perhaps all the observations should be considered personal opinion of the authors.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(1977), and *Language, status and power in Iran* (1986) or Reza Assadi's *Deference: Persian Style* (1980).

Beeman himself notes that his approach is both pragmatic and, necessarily, ethnographic, in the sense that not only forms and their variation matter, "but also the 'cultural logic' by which language behavior in general is interpreted" (Beeman 1977: 271).

To Beeman and Assadi, phonological variation in Persian can be closely linked to the system of *ta'ārof*, the complex code of courtesy and politeness present in any Iranian social interaction. To the laymen, system of *ta'ārof* might seem shallow and superficial, restricted to certain polite expressions exchanged during social interactions with Iranians. Assadi claims that it is "a generic term which denotes a myriad of verbal and non-verbal deferential behaviors in Persian" and points out that Hodge (1957) and Beeman (1973) "classify *ta'arof* as a **stylistic aspect of the Persian (Farsi) language** inherent in its speakers" (Assadi 1980: 221, my emphasis).[82] Beeman sets it at the core of any interaction between Iranians, a mean of constant negotiation of one's social status and maintenance of social relationships on all levels, symmetrical and asymmetrical alike. Together with another important concept, *birun/zāher* 'outside' and *andarun/bāten* 'inside,' they form the dimensions controlling social interactions: social ethic (asymmetrical – symmetrical), individual ethic (restricted expression – free expression) and arena of activity (outside – inside) (Beeman 1977: 277). Beeman makes an attempt at matching those dimensions with the variation of Persian sound system.

> In standard Persian there is a great deal of variations withing the sound system of speech as it is differentially produced in varying sociocultural contexts. Some of this variation might properly be considered morphophonemic in nature; however, several purely phonetic shifts do occur in the language. The most common of these shifts is the reduction of sounds in some contexts which occur in others (Beeman 1977: 271).

The three principal sound reductions considered in his study are (1) the reduction of /h/ and /'/ in all except phrase-initial position, i. e., *dā'i* 'maternal uncle' [dāi], *sag-hā* 'dogs' [sagā] (2) deletion of /r/ in the final and post-consonantal position, i. e. *sag-rā didam* 'I saw the dog' [sago didam], *digar* 'other' [dige], (3) the reduction of final consonantal clusters, i. e., *saxt* 'difficult' [sax], *dust* 'friend' [dus], etc.[83] According to Beeman's observations, those three di-

82 Although both Hodge (1957) and Beeman (1973, 1977) reflect on the influence *ta'ārof* has on the linguistic forms, none of them specifically stated that it is a stylistic aspect of language, but rather a social phenomenon having direct influence on the language. Neither does the Assadi's observation that Iranians may continue to use *ta'ārof* also in foreign language by simply translating the expressions (Assadi 1980: 222) support the stylistic interpretation of it.

83 A more extensive list of examples is given in Beeman 1977 or Beeman 1986.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

mensions show direct relation with the sound variation in language, namely sound reductions increase in symmetrical relations with free expression and 'inside' activities. It is important to note that the presented phonetical realizations belong to one of the *stylistic poles* of Persian standard, as Beeman calls it (that is, the colloquial as compared to formal):

> In the discussion that follows, I have attempted to move away from the tendency to describe stylistic variation in terms of discrete "levels" and toward description in terms of scaled polar opposite tendencies (Beeman 1986: 97).

Hodge (1957, discussed above in the present chapter) considered variation within the sound system style-sensitive, while the politeness level and other factors determining social interaction (i. e., the three dimensions described by Beeman) would surface in wording: choice of verbs (not verbal forms) and employing certain polite expressions. According to Hodge, politeness can be expressed in both styles. Beeman's observations complicate this statement as the linguistic expression of politeness is not unidimensional to him and is in fact a matter of negotiating one's social position through language. The choice of style and application of certain sound changes is affected by the social ethic, individual ethic and arena of activity[84]:

> I now wish to suggest that there is a **direct correspondence between the dimensions** that I have identified as determining the contextualization in Iranian interaction **and the dimensions of stylistic variation within the sound system of the language.** […] I will attempt to show that the speakers' perceptions of socio-cultural context correspond to their actualization of the sound system in language (Beeman 1986: 108, my emphasis).

3.6 Terminological plurality

The problem of terminological plurality in regard of Modern Standard Persian in its spoken and written varieties can easily be noticed even in the short review of theoretical approaches to this linguistic phenomenon presented above. The two varieties are termed differently according to the paradigm in use by certain researcher (thus sometimes they become styles, sometimes the H and L varieties, sometimes formal and informal or literary and colloquial standards), and sometimes they vary even within a single paper.

84 *Social ethic* is "the perception of symmetry and asymmetry between oneself and others," *expressional ethic* "concerns a person's feeling about his *ability* to operate within different contexts the degree of "freedom" he feels in personal expression in interaction," and *arena of activity* relates to how the immediate situation influences the behavior as it is defined as *birun* (outer) or *andarun* (inner) (Beeman 1986: 107).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Peisikov (1960/2002) points at the multiplicity of terms being applied to the dialect of Tehran (widespread in Iran) and lack of recognition of the divergence between this dialect and spoken language (*zabān-e goftāri*) from each other in the academic writings.[85] He lists a number of terms: *zabān-e fārsi-ye āmiyāne*, *zabān-e goftāri-ye moštarak*, *zabān-e avām*, *zabān-e goftāri-āmiyāne*, *goftār-e āmiyāne* (attributed to F. E. Korsh, V. A. Zhukovsky, A. A. Romaskewicz, Yu. N. Marr, E. E. Bertels, B. V. Miller, L. I. Zhirkov and others), French *langue populaire* [āme], *langue parlé* [goftāri], *langue persane vulgaire* [fārsi-ye āmiyāne] (from A. Christensen, H. Massé), English: colloquial Persian [fārsi-ye goftāri], vulgar Persian language [zabān-e fārsi-ye āmiyāne], conversational language of the uneducated classes [zabān-e goftegu-ye tabaqāt-e bisavād], modern Persian colloquial [fārsi-ye goftāri-ye emruz] (from D. C. Phillott, E. Denison-Ross, W. St. Clair Tisdall), *goftār-e rāyeğ-e fārsi* (R. Golonov), *zabān-e bāzāri, bāzāri* (V. Ivanov) and finally *zabān-e avāmāne*, *zabān-e avām* by M. A. Ğamālzāde, S. Nafisi, Y. Rahmati and others (Peisikov 1960/2002: 16–18). This multiplicity, in his opinion, stems from lack of clear-cut boundaries between the dialect of Tehran and spoken Persian language (*zabān-e fārsi-ye goftāri*).

It seems, however, that the distinction between Tehrani dialect and modern spoken standard might just be the tip of the iceberg. Deciding which of the original Persian terms does relate to the modern spoken standard is just as troublesome. The choice between *fārsi-ye goftāri*, *fārsi-ye āmiyāne* and *fārsi-ye mohāvereyi* is not obvious and the boundaries fuzzy. Is *goftāri* (lit. 'spoken') simply referring to the sound representation of *zabān-e me'yār* ('standard language'), as Irān Kalbāsi suggests (Kalbāsi 2003: 10) or is it rather used synonymously with *mohāvereyi* to refer to the conversational variety used by educated classes, as Rezāyi (2003) puts it? Adding *āmiyāne* to the equation does not clarify much, and there is *qeyr-e rasmi* ('informal,' often used to replace 'colloquial') in line that can also be applied to approximately the same variety. Of course, it is possible to split down *goftāri* into all the above and try to single out the differences, but I would like to focus on the opposite and try to find a common denominator for the standard varieties used in speech.

From all the Persian terms mentioned above, three are used the most in regard to the discussed phenomena. Those are: *fārsi-ye goftāri*, *fārsi-ye mohāvereyi* and *fārsi-ye āmiyāne. Goftāri* 'spoken' is naturally opposed to *neveštāri* 'written,' and their clear opposition would make them good designated of the two standard registers used in Iran nowadays. However, as I tried to underline, registral variation is of a functional character; writing and speech do not refer to function, they rather define the medium. *Āmiyāne* is a term roughly equivalent to 'collo-

85 Sādeqi (1984), Rezāyi (2003), Windfuhr (1979) do intentionally merge the two, stating that the dialect of Tehran forms the modern spoken standard (as was quoted in the present chapter).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

quial': it is characterized by a familiar use, lack of formality, lesser prestige. With those characteristics, it fits the category of style and can be understood as a level lower than informal.

The most troubleful term, then, is *mohāvereyi*. Contrary to *āmiyāne*, it does not necessarily involve lower prestige and lack of formality. Although one does not associate *mohāvereyi* with formal or deferential utterances, it is perfectly possible to ask someone politely to come by saying *šomā tašrif biyārin* (with a contracted verbal stem and adequate verbal ending) instead of *šomā tašrif biyāvarid* (without the contractions) 'please come (lit. bring your honor).' It is the variety in which the changes between writing and speech surface: asked about the variety of language in which 3rd person singular present verbal ending is *-e* and verbal stems are contracted, a Persian speaker is likely to label it *mohāvereyi*[86] (neither *goftāri* nor *āmiyāne*). I believe this is a term that, in common understanding, refers to the MSP (spoken) standard of Modern Persian. It is no longer a local Tehrani dialect, since it is used and intelligible throughout the country. It is not a style, for its application is conditioned by function, and it is not exactly a *goftāri* variety, since that term refers to the medium and does not imply any changes in structure by itself.

<table>
<tr><th></th><th>Regional variety/ ethnolect</th><th>Dialect</th><th>Register</th><th>Style</th></tr>
<tr><td rowspan="5">Persian</td><td rowspan="4">Persian (Iran)</td><td rowspan="2">standard (zabān-e me'yār)</td><td>spoken (MSP) (mohāvereyi)</td><td>deferential</td></tr>
<tr><td>written (MWP)</td><td rowspan="2">formal (rasmi)</td></tr>
<tr><td rowspan="2">local dialects (lahǧehā-ye mahalli): Tehrani, Esfahani, Shirazi, Hamadani etc.</td><td rowspan="2">less generalized registers</td></tr>
<tr><td>informal (qeyr-e rasmi)</td></tr>
<tr><td>colloquial (āmiyāne)</td></tr>
<tr><td>Dari (Afghanistan)</td><td colspan="3">(own set of varieties on all levels)</td></tr>
<tr><td>Toǧiki (Tajikistan)</td><td colspan="3">(own set of varieties on all levels)</td></tr>
</table>

Table 1. Stratification of the varieties of Persian (source: own research).

The table above is an attempt to summarize the discussion of the varieties of Persian. It shows the discussed variation on all the three levels of dialect, register and style. It should be noted that there is no direct horizontal correspondence

86 This observation was made during my research stay in Tehran in 2012. Asked for a book including *āmiyāne* language, booksellers would search for a book with colloquial vocabulary and idiomatic expressions; a book including the grammatical forms deviating from written (or, so to say, proper and correct) standard were not described as *āmiyāne* but as *mohāvereyi*.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

between the columns, that is, "Persian – Persian of Iran – standard dialect – spoken register" can exemplify any of styles and not only deferential or formal; the same is valid for written register. I have decided not to label the regional variety/ethnolect as dialect for the purpose of clarity. I have not included the Persian term *guyeš* 'speech,' which is often used for the local dialects of lesser importance (but sometimes it seems to be used interchangeably with *lahǧe* 'dialect,' or even be higher in status). As for Afghan and Tajik Persian, they have their own varieties of all the discussed sorts (see for example Spooner 2012, Henderson 1975, Rzehak 2001), which however are beyond the scope of the present study.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Chapter 4.
Analysis of source material

4.1 General remarks

As has already been stated, the main aim of this research project was to assess the presence of spoken variety of Modern New Persian in contemporary literary writings published in Iran. In the present era of online communication, social media and mobile phones, where the economy of expression takes precedence over the strict adherence to the grammatical rules, a change in literary language towards a closer representation of speech can be expected, especially in a linguistic community with a structural gap as wide as it is between Modern Written Persian (=MWP) and its spoken counterpart Modern Spoken Persian (=MSP), where the common imaginary ideal literary language is still modeled on the language of classical poetry which took its final shape about 16th century (Paul 2002: 33). Since literature is such a powerful tool of shaping people's perception of language, my assumption is that it could 1) be conservative in regard to language in order to preserve its presumed ideal state, 2) accept only those elements which are already grounded well enough to be considered a familiarized part of the spoken standard, 3) by allowing those elements and reproducing them, slowly introduce them to the written standard. With those assumptions in mind, I believe the analysis of a chosen corpus of contemporary prose can be useful in assessing which elements of the MSP are considered acceptable enough to enter literary language on one hand, and representative enough to give a clear idea to the reader that certain passage belongs to the spoken standard, on the other.

The analysis was aimed to be both qualitative and quantitative, the former eliciting features of Modern Spoken Persian register (later abbreviated as MSP) within the corpus and the latter counting their frequencies. The results shed some light on the research questions such as:

1) Are there any patterns (possibly genre- or author-sensitive) of application of MSP features in literary texts?

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

2) Are MSP features limited to dialogue parts or can they appear outside direct speech too?
3) Which grammatical category of features is mostly represented?
4) Are there significant differences in spelling of the MSP features between authors?

The project draws upon linguistic data from five popular 21st century novels (about 10 000 characters / 7–10 pages of each for the qualitative, and between 30–45 000 characters / 30 pages of each for the quantitative analysis, which makes an overall of 150 pages and 184 000 characters). Certain restrictions were posed so that such a small sample could fulfill the needs of the study. Since the aim of this research is to trace spoken forms and their influences on the language of literary texts recently, books published earlier than the 21st century are not included as a source material (which does not imply that they cannot be brought up as a reference). I have excluded experimental works either, as in those pushing the boundaries of literary language might go much further than in popular literature. Also, it was important to find works which reach wide audience, because it is especially those that have the power of shaping people's understanding (both conscious and unconscious) of literary language. I have restricted my choice to novels, excluding poetry and short prose forms such as short stories (Persian *dāstān* or *dāstān-e kutāh*, a genre that is very popular in Iran) for the purpose of getting longer coherent passages of text and decided to use middle parts of chosen books rather than the openings or endings as those could be a matter of more conscious linguistic creation due to their specific role in the storyline. The authors of chosen books come from different regional, ethnic and professional background and represent both female and male gender. While it is of course far too small a sample to assess the relevance of those factors to the linguistic characteristics of books and draw general conclusions, yet some features and peculiarities of the analyzed books might in fact bear a connection to the author's background.

Certain doubts might arise when the literature is chosen as a primary source material in a research project connected with the spoken register of a language, which is naturally associated with oral communication and characterized among others by qualities not represented in writing or print such as pitch, intonation. The analysis conducted as a part of the present research project is not designed to describe the spoken register as it is and on the whole, but to see how it surfaces in print, whether and how can it influence the literary language. Novels, especially those which are or at least pretend to be realistic, create a representation of the surrounding world, of which language naturally is a part. Consciously or unconsciously, the writers convey their imaginary picture of the spoken register by reproducing it in dialogues between characters or in their monologues and

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

thoughts. At the same time, again either consciously or not, they are influenced by this spoken register in their thinking and writing, which may impact their writing also outside the parts of text modeled to represent speech or thoughts. Even if the speech of characters is a result of careful and meticulous stylization, it still is a representation of author's imaginary picture of the spoken language and therefore a valuable source of data. Obviously, not all the features of spoken register will appear in the literature despite their presence in the actual speech. The prestige of literary language is guarded by editors and publishers and not all forms will be allowed to stay in the text after the process of edition and proofreading. My working assumption would be that if a form endures the process and is accepted—it is a proof that such a form is already familiarized, grounded and accepted in language. It is those familiarized and grounded elements of spoken register and their possible influence on the norms governing the written standard that is of interest to the present research project.

4.2 Description of the chosen source material

In this section, the selected books and their authors are presented in more detail. For each book, there is a biographical sketch of the author with information on age, place of birth, education and professional background, followed by a plot summary, editorial and printing details such as number of pages, number of lines per page and characters per line. At the end, a short description of the general style of the book and its linguistic properties is given.

A few remarks need to be made here. First, not all the biographical information on the authors was readily available in reliable printed sources, therefore it was necessary to resort to the online resources, sometimes even those of inferior provenance such as blogs, press reprints, interviews, or meeting recordings. The footnotes give the additional comments and information on the sources whenever necessary.

The editorial and print details were important to calculate the relevant number of pages undergoing the qualitative analysis. For the results to be representative, it was substantial that a comparable amount of text undergoes the analysis. Therefore not a number of pages, but a number of characters with spaces was calculated. I also included additional information on page layout (blank pages, extra spaces, chapter organization).

The observations on the style of the book and its linguistic properties are mostly my personal observations made during the initial stages of the analysis (the first reading). Consequently, they are of an *a priori*, subjective character, and they may (or may not) find proof in the final results of the analysis.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

4.2.1 Dāryuš Mehrǧuyi, *Dar xarābāt-e moqān* [In the tavern]

Tehrān: Našr-e Qatre, first published in 2011/2012 [1389h.š.] (6 reprints, 304 pages)

Dāryuš Mehrǧuyi (Pers. داریوش مهرجویی) is a film director, screenwriter, producer and writer.[87] He was born on 8th December 1939 in Tehrān. Being one of the most famous Iranian film directors, he is also considered the founder of the Iranian New Wave movement [Pers. Mouǧ-e nou-e sinamā-ye Irān] in cinema in the 1970s. He is also a member of the Iranian Academy of Arts [Pers. Farhangestān-e Honar-e Ǧomhuri-ye Eslāmi-ye Irān].

After completing primary and secondary education in Iran, in 1959 he left for the United States, where he received his college education at the University of California, Los Angeles. Initially a student at the Department of Cinema, he later decided to change his major to philosophy and gained his degree there. Upon graduation, he founded an English-language literary magazine *Pars Review* in order to familiarize western readers with contemporary Persian literature. In 1965 he moved back to Tehrān and started working as a journalist and screenwriter. He taught English and literature in Tehrān's Center for Foreign Language Studies [Pers. Markaz-e Pažuheši-ye Zabān-e Xāreǧi] and gave lectures on literature and cinema at the Center for Audiovisual Studies Tehrān [Pers. Markaz-e Motāle'āt-e Sam'i va Basari] at the University of Tehrān.

His filmography as a director consists of 23 films of which perhaps the best known is *Gāv* [The cow], his second feature film from 1969, that gained him both national and international recognition. Interestingly, though himself a screenwriter, he chose to base the movie on a story by a renowned writer and a friend of his, Qolāmhoseyn Sā'edi. As Hamid Dabashi has it, Sā'edi was Mehrǧui's literary alter ego (Dabashi 2007: 110). He also points out that throughout his career, Mehrǧui's most successful films have been adaptations of works of fiction that were not his own (Dabashi 2007: 109). Mehrǧuyi has received numerous national and international awards for his films.

As for the literary career, Mehrǧuyi's debut novel *Be xāter-e yek film-e bolande la'nati* [Because of a stupid long movie] was published by Našr-e Qatre in 2010. The book enjoyed some popularity and had five reprints. *Dar xarābāt-e moqān* [In the tavern], his second novel, was published subsequently in 2012. It was followed by *Safar be sarzamin-e fereštegān* [A journey to the land of angels] in 2017.

87 The biographical information on Dāryuš Mehrǧuyi comes from two main sources: Dabashi, Hamid (2007) Chapter IV: Dariush Mehrjui, The Cow [in:] *Masters & Masterpieces of Iranian Cinema*, Washington, DC: Mage Publishers, pp. 107–134; and Wakeman, John (1988) *World Film Directors.* Vol. 2. The H. W. Wilson Company, pp. 663–669.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Dar xarābāt-e moqān (abbreviated as DXM) is a story in which topics such as religion, philosophy, politics and migration all come together and plots merge into a story worthy of an action movie. Mohammad, a 22-year-old devout Muslim from a traditional merchant family, goes to the United States to study business administration at the University of Pennsylvania. There he meets Matilda, a Catholic girl from an Italian family. Their relationship, beginning with philosophical and religious discussions, leads finally to a marriage despite cultural and religious differences. Initial difficulties in finding a job end with Mohammad accepting a job offer in a casino run by some Italian friends of his father-in-law, by whom he is referenced for the position. The family lives a relatively peaceful life and stays away from politics until the events of September 11th, 2001, which reshape the reality of Muslim migrants in the US harshly. Mohammad loses his job and is advised by friends to deny his religion. Even the thought of such a concept puts him, an obedient Muslim strongly attached to his religious practices, in absolute turmoil, but faced with lack of other options, he finally seeks advice from a local Muslim cleric (the same with whom he consulted about the job in casino), who issues a fatwa allowing Mohammad to proceed with the denial. In the course of events it turns out that Mohammad has developed a rare ability to see and detect suspicious criminal activity. Mohammad begins to cooperate with FBI, but despite the will to cooperate and lack of ties to any terroristic activity, he is in fact held in captivity in Guantanamo prison with no contact with his family of whatsoever. With the aid of an Iranian lawyer he manages to contact Italian friends from the casino. They organize his escape and Mohammad flies with his family to the Bahama islands, where with newly gained mental powers he earns some money in the stock market, buys a hotel and starts his life anew.

The title of the book is a reference to a poem by Hāfez, a well-known 14th century Persian poet, which reads in its opening line:

در خرابات مغان نور خدا می بینم وین عجب بین که چه نوری ز کجا می بینم
Dar xarābāt-e moqān nur-e Xodā mibinam vin aǧab bin ke če nuri ze koǧā mibinam[88]

that translates roughly to: "I see the God's divine light in the tavern. See the miracle of such a light in such kind of a place!" That line serves as a motto to the story and points at the protagonist's difficult life choices, often contradictory to his religious beliefs, still not making him give up his faith.

The book was first published in Tehrān in 2011/2012 by Našr-e Qatre and has had six reprints so far. For the present analysis, the first edition was used. It consists of 304 pages of approximately 24 lines per page (65 characters per line).

88 *Divān-e Hāfez* (1384h.š.), Tehrān: Ǧeyhun, p. 379.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Most of the pages are fully covered with text. There are no specified chapters, sections are divided by three asterisks. The dialogue lines introduced by a line break do not appear often. In the detailed analysis, pages 50–56 were included which approximates 10514 characters (163 full lines), while the quantitative analysis included pages 50–80.

Dar xarābāt-e moqān is a realistic first-person retrospective narrative, declared so by the narrator himself in the opening lines. He considers writing about past events a kind of autotherapy: "be neveštan-e dāstān-e badbaxtihām mipardāzam (…) čun eyn-e ravānkāvi mofti mimune" [I turn to writing about my misfortunes (…) because it is like psychoanalysis, only free-of-charge, DXM 7] and uses direct address "az šomā če penhān" [as you know well, DXM 7] to make an initial bond with the reader. The time is specified to be soon after the terrorist attacks of 9/11, in the aftermath of which the protagonist loses his job—a misery that triggers the whole retrospective. The first-person narrative is employed consequently throughout the story and most of it, due to the retrospective character, is told in the past tense. The interesting linguistic feature of this book is a consistent use of the *mohāvere*, that is, the spoken register of Persian. The language is neat and careful and the MSP forms consistent and consequently applied. Even though using MSP to construct a whole narrative is not common in contemporary Persian fiction, the author manages to avoid an impression of artificiality.

4.2.2 Maryam Ǧa'fari, *Šahr-āšub* [Femme fatale]

Tehrān: Šādān, first published in 2004/2005 [1383hš] (3 reprints, 595 pages)

Maryam Ǧa'fari (Pers. مریم جعفری) is an Iranian writer and the author of more than twenty novels.[89] She was born in Tehrān in 1975. Having great interest in literature, she started writing at the age of fourteen, strongly influenced by books such as *Gone with the Wind* by Margaret Mitchell or Danielle Steele's novels. Her literary debut, a collection of three short stories entitled *Sāye-ye qamgin-e ešq* [Sad Shadow of Love], was published in 1995 despite discouragement from her teachers, a decision she admits to regret until now. The book did not meet with the approval of critics, prompting her to stop writing and commit to her family instead. The birth of two children and the early diagnosed terminal illness of her daughter (whom she finally lost in 2005), have taken their toll on her and reshaped her worldview. After some years, encouraged by her husband, she got

89 The biographical information on Maryam Ǧa'fari is provided by her publisher on their official website (shadan.ir) in the section *Āšenāyi bā ahl-e qalam* [Get to know our writers; retrieved 29.04.2021].

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

back to writing and published a novel called *Ešq-e Šivā* [The Love of Shiva]. Contrary to her previous experience, this book was received well by the readership.

Maryam Ğa'fari has worked with a number of publishers. Currently, she is associated with a publishing house Šādān located in Tehrān.

The complete list of her publications consists of the following titles: (i) *Sāye-ye qamgin-e ešq* [Sad shadow of love], (ii) *Ešq-e Šivā* [The love of Shiva], (iii) *Tanhā bā to* [Only with you], (iv) *Afsus-e yek negāh* [Alas a look], (v) *Zanğir-e ešq* [Chain of love], (vi) *Farāmuš nakon* [Don't forget], (vii) *Be entezārat xāham māND* [I will wait for you], (viii) *Hasrat-e didār-e to* [Longing to meet you], (ix) *Sālhā-ye bikasi* [Solitary years, 2 volumes], (x) *Tak-savār-e ešq* [Single rider of love], (xi) *Ro'yā-ye man* [My dream], (xii) *Taqdir-e šum* [Dark fate], (xiii) *Šabhā-ye qorbat* [Nights of foreignness], (xiv) *Āšenāyi az harir* [Acquaintance delicate as silk], (xv) *Bi-qarār* [Restless], (xvi) *Tarāne-ye nime-šab* [Midnight song], (xvii) *Mastāne-ye ešq* [Drunk with love], (xviii) *Golhā-ye šab-bu* [Gillyflowers], (xix) *Bibāzgašt* [No return], (xx) *Kuh-e šišeyi* [Glass mountain], (xxi) *Šahr-āšub* [Troublemaker], (xxii) *Zamzame-ye bād* [The murmur of the wind], (xxiii) *Ārām ğān* [Peaceful soul].

Šahr-āšub (abbreviated as ŠĀ) is a fictionalized biography of Foruq Farroxzād (1935–1967), a prominent Persian poetess. In the preface to the book, Bahman Rahimi writes that:

> *Šahr-āšub* is different [from other works about Farroxzād's life]. It is a novel which begins in the early teenage days of a girl called Foruq and instead of telling her story only to those who already know it, it aims at targeting a larger audience, who might have known her just by some recollections of others, and by letting them into the details of her everyday life, make them follow the works of that girl who one day became Foruq Farroxzād (Rahimi 2003: 6).

The beginning of the story is set in Foruq's secondary education days (after finishing nine grades of school, she moved to a secondary vocational school [*honarestān*] for girls to learn painting and sewing). At first, the rare visits of her father, colonel Farroxzād, recently married to his second wife, are described, when the children are given orders regardless of their or their mother's will. The reader enters a patriarchal, traditional family in which little or no voice is given to women and children. After the marriage of her older sister Purān, Foruq finds herself attracted to her 16 years older cousin Parviz Šāpur, a press satirist and civil servant, whom she eventually marries despite the objections of her father, only to find that a housewife's role does not satisfy her ambitions. Even though her husband belongs to a different generation and represents an approach radically different to that of Foruq's father in terms of family relations, still she deems the marriage unhappy. The events known from Farroxzād's biography follow: birth

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

of the first and only child, beginnings of her literary career and all the controversies around her poetry, separation from her husband and eventual divorce. Travels abroad, film career, acquaintance and relationship with Ebrāhim Golestān, a filmmaker and the owner of a film studio Golestan Films, are described, allegedly from the point of view of Foruq herself, which enables the reader to sympathize with the poet. The book ends with Foruq's dead in a car accident and her funeral at the Zahir-od-Doule[90] cemetery in Tehrān.

The book was first published in Tehrān in 2004/2005 by Entešārāt-e Šādān and has had three reprints so far. For the present analysis, the second edition (Tehrān: Šādān, 2006/2007) was used. It consists of 595 pages of approximately 28 lines per page (76 characters per line), but lines are often broken because of the frequent use of dialogues and many pages are left half-blank. In the detailed analysis, pages 50–58 were included which approximates 10419 characters (138 full lines), while the quantitative analysis included pages 50–80.

Šahr-āšub is a typical third-person realistic narrative. Since it is a fictionalized biography, it is strongly grounded in reality, which however is subject to author's interpretation and modelling. Dialogues are prevalent and they often consist of very short exchanges:

"Foruq (…) porsid:
Xodet lāzemeš nadāri?

Sohrāb bā mohabbat goft:
Engār to bištar lāzemeš dāri."

"Foruq (…) asked:
– You don't need it yourself?
– You seem to need it much more – Sohrāb replied kindly." [ŠĀ 24]

The language of dialogues has the characteristics of MSP register, while the narrative parts tend towards the written standard with little or no exceptions. This is well visible in the detailed analysis, where on the list of nonstandard forms the line numbers are often all odd or all even and the forms appear every other line (the dialogues are either introduced according to the Persian convention by the verb *goftan* 'to speak,' *porsidan* 'to ask,' *ǧavāb dādan* 'to answer,' etc., or by a line break and a pause, or quite often by both, therefore the dialogue parts are usually separated by a one-line introductory verb as in the example above). On the whole, the style is simple and close to the every-day casual language. In comparison with the other analyzed books, the language of *Šahr-āšub* is not as careful and coherent. The MSP and MWP forms are used inconsistently at times

90 The Zahir-od-Dowle cemetery (Persian: ظهیرالدوله, romanized also as Zahir-od-dowleh) is located in Darband, in the northern part of Tehrān. It is a resting place of many artists (musicians, writers and poets).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(especially as far as the number of verb is considered) and there is an overall feeling of certain linguistic clumsiness. Within the dialogues, the reader can sense a frequent use of colloquial phrases and sayings, for instance: "Kodum guri budi?" ['where on Earth have you been?' ŠĀ 30], sometimes even to the point of artificiality.

4.2.3 Farhād Ǧa'fari, *Kāfe Piyāno* [Café Piano]

Tehrān: Našr-e Češme, first published in 2007/2008 [1386hš] (60 reprints, 266 pages)

Farhād Ǧa'fari (Pers. فرهاد جعفری) is an Iranian journalist, writer and political activist.[91] He was born on 19th September 1965 in Šusef (South Xorāsān province). He completed his primary and secondary education in Šusef and Birǧand, then left for Mašhad to study. He holds a bachelor's degree in judicial law from the Free Islamic University [Pers. Dānešgāh-e Āzād-e Eslāmi], Mašhad branch.

The beginning of Ǧa'fari's press career was his collaboration with *Tus* [Tus; a city near Mašhad] newspaper from 1990 to 1996, in 1993–94 as a head of the Art and Literature group and as a chief editor in 1994–96. After that, he was a honorary collaborator of many websites and newspapers. In the years 1999–2002 he worked as the editor of a socio-cultural weekly *Xāvarān* [The East] and the owner and chief editor of a socio-political weekly *Yek-haftom* [One-seventh], which was then closed by the government.

As for political activity, Ǧa'fari run for parliament [Islamic Consultative Assembly, Pers. Mağles-e Šourā-ye Eslāmi] in the 5th parliamentary elections in 1996. He got substantial support and was in fact elected, but then suddenly withdrew (actually, was forced by the government to resign after being elected). After that experience, he resumed his journalistic activity and writing in favor of the reformists until the publication of his debut novel *Kāfe Piyāno* in 2008. The novel became a bestseller and was reprinted 30 times in the first three years after publication.

A sudden decline in the popularity of *Kāfe Piyāno* and its author was noted in 2009 after his unexpected support of conservative Mahmud Ahmadinežād in the 10th presidential elections.[92] Ǧa'fari had stated his reasons for backing Ahmadi-

91 It is difficult to find a reliable source of information on this author. The biographical details found above are based on the personal communication with Farhād Ǧa'fari which took place in April 2021. Ǧa'fari has kindly provided his resume, corrected inaccuracies and filled missing information that I could gather from scattered press articles available online.

92 Ahmadinežād, a conservative former mayor of Tehrān, was then running for a second term of presidency. His first presidency was associated with violations of human rights, Holocaust denial and hostility towards Israel and the United States. He was also criticized for economic policies. His election for the second term was considered fraudulent by many and sparked the

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

nežād on his personal blog *Goftamgoft,* still he was widely criticized for this move both by literary circles and by his readership. Several purchased copies of *Kāfe Piyāno* were even returned to the publisher as a sign of protest.[93] Still, the novel remains one of the bestselling books of last two decades, with the number of reeditions reaching sixty according to the publisher's website.[94]

Apart from the traditional journalism, Ǧa'fari was also active in the online sphere, publishing articles both on his personal blog and social media. His blog *Goftamgoft* was eventually closed by the government, together with his deprivation of writing and publishing content on newspapers, social media, websites and blogs from 2013 until present.[95]

In June 2018 Ǧa'fari had been arrested and held in Vakilābād Prison of Mašhad because of his journalistic activity in the field of politics. After facing trial and being convicted to 13 years of prison and 150 lashes for his writings, he fled Iran and now lives in exile in Turkey.

Kāfe Piyāno (abbreviated as KP) is a story of a journalist and writer who, not having succeeded in his previous career path, decides to run a café in order to collect his wife's dowry and divorce her. Since many of the book characters are real (including the narrator, his daughter Gol-Gisu, his wife Pari-Simā and his best friend Homāyun) and many of the events have actually occurred in real life, it could be considered quasi-autobiographical. Rather than following a linear plot, the author presents us with short sketches revolving around people he meets in the café, each chapter devoted to one character or event from author's life that he finds important or amusing. The chapters usually cover a short period of time, i. e. a day, and the whole story happens in between the author's separation from his wife and finishing collecting the dowry. In the course of a story, the author shares his opinions on various aspects of life, from the importance of proper clothing to the relations between him and his father on one hand and him and his daughter on the other. The slow pace with which the story develops and the often

Iranian Green Movement, a massive wave of protests that took place throughout the country in 2009–2010.

93 According to Ǧa'fari himself, about 200–300 copies were returned and at that time about 150 000 had already been sold (personal communication with the author on 28–29 April 2021).

94 Official website of Našr-e Češme at www.cheshmeh.ir/Book/663/15800/%DA%A9%D8%A7%D9%81%D9%87%20%D9%BE%DB%8C%D8%A7%D9%86%D9%88 [retrieved 29. 04. 2021].

95 It is a fate shared by his later attempts at online activity; the latest website at jafari.ws, accessible in December 2020, is also no longer available. For now, Ǧa'fari has an active official Instagram profile and surprisingly, his previous Facebook page is also accessible at: https://www.facebook.com/%D9%81%D8%B1%D9%87%D8%A7%D8%AF-%D8%AC%D8%B9%D9%81%D8%B1%DB%8C-farhad-jafari-196810230499456/ [accessed on 29. 04. 2021, though the latest publication on that site is dated 04. 09. 2015].

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

humorous and ironic reflections and insights mark the distinctiveness of this novel in comparison to others.

The book was first published in Tehrān in 2007/2008 by Našr-e Češme and has had 60 reprints so far. For the present analysis, the 35th edition (Tehrān: Češme, 2012/2013) was used. It consists of 266 pages of approximately 22 lines per page (70 characters per line). Some pages are left half-blank, usually at the beginnings and ends of chapters, and some (but not all) the dialogues are signaled by line breaks. In the detailed analysis, pages 50–57 were included which approximates 10286 characters (148 full lines), while the quantitative analysis included pages 50–80.

Kāfe Piyāno is a first-person narrative with frequent use of direct address (second person plural) to create the feeling of dialogue or familiarizing with the readers: "I said all this just to tell you" [in hame harf zadam barāye inke be-tān beguyam, KP 9], "if you saw him" [agar u-rā bebinid, KP 53]. The style of the book resembles that of a casual conversation and the language can be considered colloquial, although grammatically speaking, it stays closer to MWP in the narrative and MSP forms are limited to the dialogues, direct and indirect quotes and chapter titles. The conversational style is achieved by a frequent use of phrases such as *mixāham beguyam* [I want to say], *fekr mikonam* [I think] etc., demonstratives and enclitic personal pronouns and intricate sentences with a sense of carefully designed sloppiness. Unlike *Šahr-āšub*, though, here the sloppiness is more a matter of careful writer's design than a result of his lack of competence.

4.2.4 Zoyā Pirzād, *Ādat mikonim* [We'll get used to it]

Tehrān: Našr-e Markaz, first published in 2004/2005 [1383hš] (70 reprints, 266 pages)

Zoyā Pirzād (Pers. زویا پیرزاد) is a translator, novelist and short story writer. She was born in Ābādān in 1952 of Armenian mother and Iranian father of Russian origin.[96] She spent her school years in Ābādān, then moved to Tehrān. She began her literary career as a translator of Lewis Carrol's *Alice in Wonderland* [*Ālis dar sarzamin-e aǧāyeb*, published by Našr-e Markaz in 1997] and a collection of haiku *Āvā-ye ǧahidan-e quk* [The sound of jumping frog]. Married and a mother of two boys, she now lives in Yerevan, Armenia.[97] Her books were translated to

96 Biographical information on Zoyā Pirzād based on Šarifi, Mohammad (2010), *Farhang-e Adabiyāt-e Fārsi*, Tehrān: Entešārāt-e Mo'in, pp. 370–371, Gheytanchi, Elham (2011) *Čerāqhā rā man xāmuš mikonam*, Encyclopaedia Iranica Online https://iranicaonline.org/articles/cheragh-ha-ra-man [retrieved 05. 05. 2021].

97 The information of her place of residence after her American publisher's website: oneworld-publications.com [retrieved 10. 05. 2021].

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

several languages, including English and French. In 2009 she was awarded a 2nd prize of French *Courrier International* for the translation of her collection of short stories *Ta'm-e gas-e xormālu* [Acrid taste of persimmon; translated to French as *Le Goût âpre des kakis* by Christophe Balaÿ and published by Zulma].[98]

Pirzād was the second Armenian author to write a novel in Persian (the first being Ālis Ārezumāniān in 1963 with her novel *Hame az yek* [All from one]) (Gheytanchi 2011). Before publishing her debut novel *Čerāqhā rā man xāmuš mikonam* [I turn off the lights] in 2001, she had already been known as the author of three critically acclaimed collections of short stories: (i) *Mesl-e hame-ye asrhā* [Like all the other evenings, 1990], (ii) *Taʿm-e gas-e xormālu* [Acrid taste of persimmon, 1997] and (iii) *Yek ruz mānde be Eyd-e Pāk* [A day before Easter, 1999], then published altogether as one volume under the title *Se ketāb* [Three books, 2003].

Čerāqhā rā man xāmuš mikonam was the most successful among Pirzād's books, awarded with Hušang Golširi Literary Prize [Pers. Ğāyeze-ye Adabi-ye Hušang Golširi], Mehrgān Award [Pers. Ğayeze-ye Mehrgān-e Adab][99] and honorary mention of Yaldā Award [Pers. Ğāyeze-ye Adabi-ye Yaldā] in 2001. It was also chosen a book of the year by the Ministry of Islamic Culture and Guidance [Pers. Ğāyeze-ye Ketāb-e Sāl Ğomhuri-ye Eslāmi-ye Irān] in 2002. The book was praised for its very consistent and flawless, simple yet exact prose, powerful rendition of Armenian dialects in Persian and successful use of dialogue as a storytelling technique (Gheytanchi 2011). The same simple flawless style can be found in *Ādat mikonim* [We'll get used to it], published in 2004. The writer admits herself that the simplicity of language is at the core of her writing:

> What I don't like about Iranian literature is that the characters don't speak like they do in everyday life. When I started to write, the words first presented themselves like this, and I felt that I am close to this way of writing, this is my language. Dialogue is very important, and especially in the Persian language, it can quickly become very heavy. Iranian authors try to write in direct or indirect style. As for me, all my effort is to write neither directly, nor indirectly. As much as I can, I try to bring written language closer to oral. My obsession is to simplify the language. (…) I am looking for simplicity and accuracy. And it's very difficult to write simply.[100] (Pirzād 2009)

98 www.courrierinternational.com/article/2009/10/21/zoya-pirzad-laureate-2009-avec-le-gout-apre-des-kakis [retireved 05.05.2021].

99 This prize was known previously as the Pekā Award [Pers. Ğāyeze-ye Pekā], prize granted by the Institute of Distribution and Development of Iranian Books [Pers. Mo'asese-ye Paxš-o Touse'e-ye Ketāb-e Irān], until 2007/2008, when the Institute ceased its activity.

100 Translation of a quote from an interview with Zoyā Pirzād for *Courrier international* in 2009 [original interview in French] https://www.courrierinternational.com/article/2009/10/30/je-recherche-la-simplicite-et-la-justesse [retrieved 06.11.2021].

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Ādat mikonim (abbreviated as ĀM) is a story of three generations of women in a middle-class family in contemporary Tehrān of two, maybe three decades ago. Main protagonist, Ārezu, is a middle-aged woman. Having divorced her husband, she now lives with her teenage daughter Āye, and she is also responsible for her elderly, widowed mother Māh-Monir. Ārezu runs a real-estate agency inherited after her father and is successful in her professional life. An independent business-woman capable of managing her life on her own on the outside, inside she is entangled in the often conflicted, selfish needs and expectations of her daughter on one hand and her mother on the other. She finds some help and comfort in the relations with her best friend, Širin, another strong female character of the novel, a single middle-aged woman whose fiancée left her after both their mothers were killed in an accident a week before their planned marriage. Yet once a man enters the stage, the balance of this female-dominated world is at risk. When Ārezu meets Sohrāb, a man she is strongly attracted to, the idea of a romance and even remarriage crosses her mind. To her own surprise, she is faced with a lack of understanding, objection and criticism not only from mother and daughter, but even from her best friend. The prejudices and social expectations of a traditional society take precedence over the individual choices and desires and Ārezu is left with a difficult choice between her family on one hand and her heart on the other.

The book was first published in Tehrān in 2004/2005 by Našr-e Markaz and has had 70 reprints so far. For the present analysis, the 14th edition (Tehrān: Markaz, 2006/2007) was used. It consists of 266 pages of approximately 24 lines per page (60 characters per line). Some pages are left half-blank, usually at the beginnings and ends of chapters. In the dialogues, the change of speakers is marked by a line break, but usually the direct speech passages are followed by the description of speaker's actions, thus resulting in longer paragraphs anyway. In the detailed analysis, pages 50–59 were included which approximates 10320 characters (167 full lines).

Ādat mikonim is a realistic third person narrative depicting a contemporary urban middle-class life in Tehrān, the language of the story playing an important role in creating a credible description of the background. The complex compound sentences can hardly be found and the author clearly aims at the simplicity of structure both in terms of language, and in character and plot creation. Development of the plot is often represented in the dialogues rather than described in the narrative parts only. Lexically and syntactically speaking, the dialogues represent the educated Tehrani speech rather than writing, with reversed word order characteristic of the contemporary speech and simple lexical choices. The reversed word order is also frequent in the narrative passages, especially when verbs of movement occur. Yet the phonological characteristics of the spoken idiom can hardly be seen, verbal endings both inside dialogues and outside in the narrative are kept in the MWP standard.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

4.2.5 Rezā Qāsemi, *Hamnavāyi-ye šabāne-ye orkestr-e čubhā* [Nocturnal Harmony of the Orchestra of Woods]

Tehrān: Nilufar, first published in 2001/2002 [1380hš] (17 reprints, 207 pages)

Rezā Qāsemi (Pers. رضا قاسمی) is a novelist, playwright, theatre director and musician. He was born in 1949 in Isfahan and spent first years of his life in southern Iran where his father worked in oil industry. He published his first play *Kasuf* [Eclipse] in a literary magazine *Xuše* [Cluster] in 1969. One year later, he entered the Faculty of Fine Arts of Tehrān University [Pers. Dāneškade-ye Honarhā-ye Zibā], from where he holds a degree in playwriting. After graduating, he worked as a theatre director, published and staged several plays. In 1973 he joined an experimental theatre group and tried his hand as an actor. In the aftermath of the Islamic Revolution of 1979 he decided to move to Paris and has lived there since 1987.[101]

Qāsemi is known to be the first Iranian writer to have turned to the Internet to avoid censorship of his work—the one who took Persian literature to the cyberspace.[102] *Varde-yi ke barrehā mixānanad* [The Spell Chanted by Lambs] was published online on the author's personal blog in 2002 before its traditional publication in 2008 by a Paris-based Persian publisher Našr-e Xāvarān.[103] Another online project entitled *Parvānehā-vo Emkānhā* [Butterflies and Possibilities], an online improvised novel published on Facebook, however, was withdrawn by the author himself right before publication along with his recognition of the cyberspace as dangerous.[104] Qāsemi's website *Davāt* [Inkwell][105] houses his own literary and critical writings, unpublished, raw literary pieces as well as press articles and critical writings by others.

Similarly to Pirzād, Qāsemi also openly expresses his concern in simplifying the language. In the talk mentioned above, he explains how the simpler the

101 Šarifi, Mohammad (1388hš/2010), *Farhang-e Adabiyāt-e Fārsi*, Tehrān: Entešārāt-e Mo'in.

102 Here I refer to the talk given by Rezā Qāsemi during a meeting organized by a Persian-speaking literary organization Raviyun: *Nohomin ğalase-ye Raviun – Rezā Qāsemi va ketāb-e Hamnavāyi-ye šabāne-ye orkestr-e čubhā* [9th Meeting of Raviyun – Rezā Qāsemi and the book *Nocturnal Harmony of the Orchestra of Woods*] (www.youtube.com/watch?v=4smj7j35hUY) [published on 04.04.2018, retrieved 22.04.2021].

103 Its English translation by Erfan Mojib was published in 2013 by by British-Iranian publisher Candle & Fog under the title *The Spell Chanted by Lambs*.

104 Again this is a reference to Qāsemi's own words uttered during the 9th Meeting of Raviyun in 2018 (www.youtube.com/watch?v=4smj7j35hUY) [published on 04.04.2018, retrieved 22.04.2021].

105 Available at www.rezaghassemi.com/davat.htm [retrieved 22.04.2021].

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

sentence, the more influential it becomes and how it is necessary to cut away all the redundant words up until only the essence remains (Qāsemi 2018).[106]

Hamnavāyi-ye šabāne-ye orkestr-e čubhā is Qāsemi's most successful writing. It was first published simultaneously in France and in the United States in 1996 before obtaining a permission to be published in Iran in 2002 (Šarifi 2010: 1114). The book was not only warmly received by the readers, but also critically acclaimed, which resulted in two literary prizes: Press Critics Award [Pers. Ğāyeze-ye Adabi-ye Montaqedān va Nevisandegān-e Matbu'āt] and Hušang Golširi Literary Prize [Pers. Ğāyeze-ye adabi-ye Hušang Golširi] for the Best Novel of 2002.

Hamnavāyi-ye šabāne-ye orkestr-e čubhā (abbreviated as Hn) is a postmodern novel and an example of the migration or expat literature. It is a story of an exiled Iranian man called Yadollāh, living on the sixth floor of an old building in Paris among other immigrants, mostly Iranians. The characters represent a variety of personality types, yet it seems that all have some difficult and disappointing experiences from the past. De-rooted from their homeland, but not able to adapt themselves to the French culture, they mostly turn to themselves and their present life passes in the gloomy, ominous atmosphere in which hardly anything happens. The introduction of a new resident, Profet, disturbs the fragile balance and begins a sequence of events leading to an imminent catastrophe, which in fact is anticipated and signaled from the very beginning of the novel. The story is not told in a linear manner, but rather composed of scattered pieces and protagonist's recollections of the past events, mixed with his internal monologues and posthumous conversations with two angels of death, Munkar and Nakir, known from Islamic eschatology, who are judging his deeds in the afterlife. This way at least three different storylines are layered over one another, and recurring flashbacks and different points of view slowly unfold the actual chain of events, if there is any, since "the ontological despair is taken into account in Reza Ghassemi's *Nocturnal Harmony* and puts the world of story into question" (Shabrang 2020: 26).[107]

As it was already said, the book was first published in 1996 in Los Angeles, California by a Persian publisher Našr-e Ketāb. Its first publication in Iran took

106 *Nohomin ğalase-ye Raviyun – Rezā Qāsemi va ketāb-e Hamnavāyi-ye šabāne-ye orkestr-e čubhā* [9th Meeting of Raviyun – Rezā Qāsemi and the book *Nocturnal Harmony of the Orchestra of Woods*] www.youtube.com/watch?v=4smj7j35hUY [published on 04. 04. 2018, retrieved 22. 04. 2021].

107 *Hamnavāyi-ye šabāne-ye orkestr-e čubhā* has received a wide scholarly and critical interest and there are plenty of papers discussing its structure, its main protagonist as the anti-hero, the use of modernist or postmodern elements and many other aspects. My personal interpretation and understanding of this elaborate and complicated story benefited a lot from Hoda Shabrang (2020) *Every man is an Island: Decanonisation and Fragmentation in Reza Ghassemi's* The Nocturnal Harmony of Wood Orchestra: *Asiatic* vol. 14, no. 1, pp. 23–39.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

place five years later, in 2001/2002, when it was published by Tehrān-based Nilufar. It has had 17 reprints so far and for the present analysis, the 12th edition (Tehrān: Nilufar, 2013/2014) was used. It consists of 207 pages of approximately 25 lines per page (68 characters per line). Most of the pages are fully covered and the few dialogues present are composed within the main body of the text. In the detailed analysis, pages 50–59 were included which approximates 10880 characters (160 full lines).

Hamnavāyi-ye šabāne-ye orkestr-e čubhā is a first-person narrative with many retrospective passages and a complicated structure. The language is careful and precise, and there are no redundant or unnecessary words. According to the authors belief (mentioned above), the shorter and more precise a sentence is, the more effect it will have on the reader. The forms are consistently MWP and only some parts of the dialogues draw on the resources of the spoken idiom in terms of grammar.

4.3 Method of the analysis – an outline

The analysis is modeled to some extent on the multi-dimensional register analysis (Biber 1994, Biber 2004) described in Chapter 2. MD analysis, however, is designed for computational analysis of the digitized data. Therefore certain adjustments need to be made for that tool to be functional for a small, self-designed and undigitized corpus, where data is collected manually from the books and analyzed by the researcher themselves. The **basic assumption** remains the same, though: instead of searching through the material in order to find the pre-listed features, the material is analyzed to find (co-)occurring features and list them afterwards.

The purpose of this analysis is to find MSP features within the texts which should in fact represent MWP. In the light of the basic assumption mentioned above, I decided not to search for the pre-defined MSP features, but rather single out and collect forms which do not represent the textbook standard of MWP in terms of phonology, morphology, syntax and lexis. Such an analysis required a detailed, accurate and in-depth line-by-line reading of the source material. For that reason, the part of text included in the detailed analysis was limited to 10,000 characters for each book, which still proved enough to collect as much as 820 instances (many of them including more than one non-standard feature).

For the collection, description and analysis of the data, an MS Excel spreadsheet was used with separate worksheets for the list of books, detailed analysis of each of the books and the quantitative analysis (altogether).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

4.3.1 First reading

The first step of the analysis, once the list of books was established, was a first reading of the whole book in order to recognize the author's style, see if it is consistent and how do the MSP features surface throughout the text, whether their distribution is regular and repetitive (which proved true for all the analyzed books, allowing to "randomize"[108] the choice of pages for the analysis). General remarks resulting from this stage of the analysis can be found in the preceding section. For each of the books, page 50 was chosen as a starting point and approximately 10,000 characters with spaces were counted to undergo the detailed analysis.

4.3.2 Qualitative analysis

After the first reading of a whole text, the second stage of the analysis was restricted to the above mentioned 10,000 characters from which the forms deviating from MWP were extracted. The in-depth word-by-word analysis of the material was documented in a table including the following parameters:

1) book, page and line number in which the form occurs,
2) form (a word, phrase or sentence) in Persian alphabet,
3) form – transcription,
4) form – translation (into English),
5) grammatical description of the form's components,
6) general grammatical category (phonology/morphology/syntax/lexis)[109] – more than one option available, distinguished by color for readers convenience,
7) detailed grammatical category (e.g., for phonology: alternation *ān>un*, morphology: MSP verbal stem, MSP verbal ending, syntax: MSP word order SVO, lexis: MSP preposition, word or phrase, saying) – at this stage of analysis, the list of detailed categories was open-ended,

108 Randomizing not in the sense of opening the books on a ranom page, but in the sense of not depending on storyline, chapter divisions, presence or absence of dialogues *et cetera* in chosing a certain part.

109 The choice of those general categories was arbitrary. Some of the parameters would possibly fit better into categories such as morphophonology (e.g., verbal endings) or morphosyntax (e.g. suffixed personal pronouns in their syntactic functions). The parameters themselves will be discussed in more detail in later sections of the present chapter. For the purpose of clarity, I decided to stick to the four clear-cut categories of phonology, morphology, syntax and lexis.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

8) MSP vs MWP – a standard written equivalent of the form (i. e. for *mige* 'says' the entry would be *miguyad*),
9) context (narrative/dialogue),
10) range affected (word/sentence) – here 'word' was chosen for those instances where a supposedly MSP word was surrounded by the MWP forms, and 'sentence' for those instances in which more of the elements were in the MSP form,
11) extra comments – a column for storing all the additional information (contextual, grammatical, editorial).

The example of this data table for the book *Dar xarābāte moqān* can be seen below (because of the number of columns and amount of text, the table was cut into three parts).

Page	Line	Form - Persian	Form - Transcription
50	2	جوون ها	ğavun-hā
50	3	نشسس	nešast-an
50	5	به گوش می رسه	be guš mi-res-e
50	5	بر قراره	bar qarār-e
50	6	و خلاصه یه سرزمس جادویی	va xolāse ye sarzamin-e ğādu-yi
50	7	مشری رو	moštari ro
50	7	و رموت رایگانه	[vermut-e] rāygān-e
50	8	خر می کنه	xar mi-kon-e
50	8	اونو	un-o
50	9	وادار می کنه	vādār mi-kon-e
50	10	ماشس رو	māšin ro
50	10	سر و کله تکون می دیم	sar-o-kalle tekun mi-d-im
50	11	از دور پدیدار می شه که	az dur padidār mi-š-e ke
50	12	بدو جف دنبالت می گرده	bo-dou ğef donbāl-et mi-gard-e
50	12	چرا مبایلتو خاموش کرده بودی	čerā mobayl-et-o xāmuš karde bud-i
50	14	داشتم به باخ گوش می کردم	dāšt-am be bāx guš mi-kard-am
50	14	چی شده	či šod-e
50	18	میگه	mi-g-e
50	18	توضیحی نمی دم	touzih-i ne-mi-d-am
50	19	سر میز هشت خرایییه	sar-e miz-e hašt xabar-ā-yi-ye
50	22	بالای دویست تا می شه	bālā-ye devist tā mi-š-e
50	23	معلومه خبر خوسی نیست	ma'alum-e xabar-e xoš-i nist
50	24	بحران های هر روزه ی ماس	bohrān-hā-ye har-ruze-ye mā-s
50	25	که باید باهاشون روبه رو شد	ke bāyad bā-hā-šun ru-be-ru šod
50	25	و رشکسته می شه	var-šekaste mi-š-e
50	26	رو من زیاد حساب می کس	ru man ziyād hesāb mi-kon-an

Figure 1. Qualitative analysis data table, part 1 (source: own research).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Grammatical Description	General Category	Detailed Category
nounPl	phonology	phonological alternation -ān>-un
verb3PlPraet	morphology	verbal ending 3pl
verb3SgPraes	morphology	verbal ending 3sg
verb3SgPraes	morphology	verbal ending (copula) 3sg
conj adv indefArticle noun-ezāfe adj-indef	morphology, phonology	MSP indefinite article yek, final consonant omission
noun-DOmarker	phonology	phonological alternation in DO marker (-rā>-ro)
noun-ezāfe adj-def	morphology	MSP definite article -e
verb3sgPraes	morphology	verbal ending 3Sg
persPron3Pl-DOmarker	phonology	phonological alternation -ān>-un, in DO marker (-rā>-o)
verb3SgPraes	morphology	verbal ending 3sg praes
noun DOmarker	phonology	phonological alternation in DO marker (-rā>-ro)
compoundNoun verb1PlPraes	lexis, phonology, morphology	MSP lexical item, phonological alternation
adv verb3SgPraes conj	morphology, syntax	MSP verbal stem, verbal ending 3sg, ke introducing direct speech
verbImperative2Sg, noun prep-encliticPron2Sg verb3SgPraes	morphology, phonology	preposition + enclPron as IO, verbal ending 3sg
interrog noun-PossEncliticPron2Sg-DOmarker verb2SgPluPerf	phonology	phonological alternation in DO marker (-rā>-o)
auxiliaryVerb1SgPraet prep noun verb1SgImperf	morphology	present progressive tense
interrog pastPart (ommited copula3Sg)	lexis	MSP interrogative
verb3SgPraes	morphology	MSP verbal stem, verbal ending 3sg
nounIndef negVerb1SgPraes	morphology	MSP verbal stem
prep-ezāfe noun number noun-Pl-indef-copula3SgPreas	phonology, lexis	MSP lexical item, phonological alternation (-hā>-ā), verbal ending (copula) 3sg
prep-ezāfe number numerativ verb3SgPraes	morphology	MSP verbal stem, verbal ending 3sg
adj-copula3SgPraes noun-ezāfe adj-indef negCopula3SgPraes	morphology, syntax	verbal ending (copula) 3sgPraes, subordinate clause without conjunction
nounPl-ezāfe adj-ezāfe persPron-copula3Sg	phonology	phonological alternation (deletion of final consonant in copula)
conj part prep-Pl-encliticPron3Pl InfApocop	morphology, phonology	preposition + enclitic pronoun, phonological alternation -ān>-un
verb3SgPraes	morphology	MSP verbal stem, verbal ending 3sg
prep persPron adv verb3PlPraes	lexis, morphology	MSP preposition (no ezāfe), verbal ending 3pl

Figure 2. Qualitative analysis data table, part 2 (source: own research).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Modern Spoken Persian (MSP) vs Modern Written Persian (MWP)	Context	Range Affected	Stylistic Context	Extra Comments
[ğavun-hā] = ğavān-hā	narrative/ semi-monologue	word		plural suffix not affected
[nešast-an] = nešast-and	narrative/ semi-monologue	word		
[mi-res-e] = mi-res-ad	narrative/ semi-monologue	word		
[bar qarār-e] = bar qarār ast	narrative/ semi-monologue	word		
[ye sarzamin-e ğādu-yi] = sarzamin-i ğādu-yi	narrative/ semi-monologue	sentence		
[moštari ro] = moštari rā	narrative/ semi-monologue	sentence		unbound spelling of DO marker
[vermut-e rāygān-e] = in vermut-e rāygān	narrative/ semi-monologue	sentence		
mi-kon-ad	narrative/ semi-monologue	sentence		
ān-rā	narrative/ semi-monologue	sentence		bound spelling of DO marker
vādār mi-kon-ad	narrative/ semi-monologue	sentence		word order and plural suufixes not affected
[māšin ro] = māšin rā	narrative	sentence		unbound spelling of DO marker
[tekun mi-d-im] = tekān mi-deh-im	narrative	sentence		
[padidār mi-š-e] = padidār mi-šav-ad	narrative	sentence		
[donbāl-et mi-gard-e] = donbāl-e to mi-gard-ad	dialogue	sentence		word order not affected
[mobāyl-et-o] = mobāyl-at rā	dialogue	sentence		bound spelling of DO marker
not relevant	dialogue	sentence		progressive tenses do appear in MWP, but they are more common in MSP
[či] = [če]	narrative	sentence		
[mi-g-e] = mi-guy-ad	narrative	sentence		
[ne-mi-d-am] = ne-mi-deh-am	narrative	sentence		
[xabar-ā-yi-ye] = xabar-i hast	dialogue	sentence		*xabar-i-ye* (lit. 'there is some news') is
[mi-š-e] = mi-šav-ad	dialogue	sentence		
[ma'lum-e] = ma'lum ast ke	narrative	sentence		
[mā-s] = mā ast	narrative/ semi-monologue	word		plural suffix not affected
[bā-hā-šun] = bā ānhā	narrative	sentence		
[varšekaste mi-š-e] = varšekaste mi-	narrative	sentence		
[ru man] = ru-ye man, [mi-kon-an] = mi-kon-and	narrative	sentence		

Figure 3. Qualitative analysis data table, part 3 (source: own research).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

At this stage, I was looking for the forms which deviate from MWP and not for those that obviously and evidently represent MSP. More often than not, one form includes more than one deviation and belongs to more than one grammatical category at the same time (consequently, a number of forms is not the same as the number of deviations; for instance, *Dar xarābate moqān* with 282 entries has in fact 475 deviations from the MWP standard in analyzed 10,000 characters). It should be noted that some of the listed forms are not non-standard on the whole and are acceptable in the MWP as well, but either their choice (compared to the other, more neutral options) or their frequency is marked (this is especially valid for the instances of the suffixed personal pronouns in their various syntactic roles). For that reason, such forms were also listed.

4.3.3 Listing parameters found in analyzed forms

During the collection of non-standard forms in the in-depth analysis, the initial open-ended list of parameters soon began to show the repetitive patterns, just as it was expected. There was also a significant co-occurrence of forms, especially those categorized as phonological and morphological (syntactic and lexical forms tended to be more context-independent, that is, appearing both with and without phonological and morphological non-standard surrounding). Collection and unification of the detailed grammatical descriptions from the previous step of the analysis resulted in creating a finite list of non-standard parameters within the four grammatical categories, three of them divided further into subcategories for the purpose of clarity (further division of lexical parameters did not prove necessary). The listed parameters are as follows.

1. Phonology:
   a. Alternation (*ān* > *un* / *-ar* > *-e* / *-e* > *-i* / *rā* > *ro* / *rā* > *o*)
   b. Omission (*h* in any position / final consonant)
   c. Contraction (consonant cluster / enclitic pronoun)
2. Morphology:
   a. Article (definite *-e* / indefinite *yek* / indefinite *yek* + *-i*)
   b. Enclitic pronoun (possessive, partitive[110] / preposition + enclPron as DO/IO / with infinitive (noun of action) as agent or object / with nomVerb as DO / with nomVerb as IO / with nomVerb/preverb as logical subject (quasi impersonal idioms) / with preverb as DO/IO / with verb as DO/IO)

110 This parameter was added during the quantitative analysis. The possessive or partitive enclitic pronouns were not listed at the qualitative stage since their use is by no means a deviation from MWP. On the other hand, a close reading of the source material suggests that a large accumulation of them might actually indicate that a passage inclines towards MSP (this observation would require validation by additional research).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

   c. Particle or conjunction (emphatic ke / temporal ke / emphatic hā / enclitic ham)
   d. Tense (past progressive tense / present progressive tense / two verb construction other than with modal verb)
   e. Verbal ending (3. sg praes / copula 3. sg praes / 2. pl praes/praet / 3. pl praes/praet)
   f. Verbal stem (MSP / alternate)
3. Syntax:
   a. Clause without conjunction (coordinate / subordinate)
   b. Lack of agreement (between subject and verb: quasi-impersonal idioms, logical subject, number disagreement)
   c. Omission (preposition / auxiliary verb)
   d. Word order (SOV > SVO direction, location / SOV > SVO nominal clauses / disrupted)
4. Lexis:
   a. Preposition
   b. Single word / compound
   c. Phrase
   d. Saying

Overall, there are 3 groups of phonological parameters (9 parameters), 6 groups of morphological parameters (24 parameters), 4 groups of syntactic parameters (8 parameters) and 4 lexical parameters (no grouping), which gives a total of 45 parameters. They are discussed in more detail in the next chapter, before the results of the analysis.

4.3.4 Consulting created list with descriptive grammars of Modern Persian

In the previous steps, the occurrence and possible coincidence of the non-standard forms in literary texts were proved. The next step of the analysis would be the attribution of those forms to certain registers of the language, here namely the spoken register. To see if the elicited forms do belong to the characteristics of the spoken rather than written standard of Persian, they were consulted with the chosen descriptive grammars of Modern Persian: 1) G. Lazard, *A Grammar of Contemporary Persian,* transl. by Shirley Lyon, Costa Mesa, California and New York 1992, 2) G. Windfuhr, *Persian Grammar. History and State of its Study*, The Hague 1979, 3) M.R. Majidi, *Strukturelle Grammatik des Neupersischen (Farsi)*, t. I–III, Hamburg 1986, 4) G. Windfuhr and J. Perry, *Persian and Tajik* [in:] G. Windfuhr (ed.), *The Iranian Languages*, London and New York 2009: pp. 416–544, 5) A.K.S. Lambton, *Persian Grammar,* Cambridge 1953. As for the native

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Iranian scholarship, two works were used as the main reference: 6) P.N. Xānlari, *Tārix-e zabān-e fārsi*, Tehrān 1988, 7) H. Ahmadi-Givi and H. Anvari, *Dastur-e zabān-e fārsi* [Persian grammar], Tehrān 2007. Additionally, a linguistic description of the dialect of Tehran by Lazar S. Peisikov (first published in Russian in 1960), entitled *Lahǧe-ye tehrāni* [Tehrāni dialect], transl. by M. Šoǧā'i, Tehrān 2002 and B. Mahmoodi-Bakhtiari's comprehensive list of the diglossic features of Persian (Mahmoodi-Bakhtiari 2018, discussed in chapter 3). Naturally, not all of those grammars included observations of the grammatical features of MSP register (labeled "colloquial," "spoken" or "Tehrani") and in some, those observations were scarce, therefore no reference to them is to be found in the next chapter, where parameters are discussed in detail.

4.3.5 Second reading and quantitative analysis

With a complete list of the 45 parameters characterizing MSP register's literary output, the quantitative analysis of a bigger amount of text was possible.

For this stage, a 30-page piece of each of the books was chosen (pages 50–80) that was equivalent with 45 up to 60 thousand characters. The occurrence of each of the parameters per page was counted but the forms were not stored at this stage. The contextual factor (dialogue/narrative) was excluded. On the other hand, the possessive/partitive pronouns were counted on this stage of the analysis, but had not been listed during the previous stage as their use is perfectly correct in the MWP—but their extensive use seems to be a marked choice, pointing more to the MSP than MWP character of the text. First 10,000 characters were re-analyzed in order to ensure the proper count. The number of forms from each of the categories were entered into a table with category names in rows and page numbers in columns. An additional column for summing up each of the categories was added for each book, and one another for a general sum at the end of the table.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Category	Detailed category	Description	50-56-DXM	57-DXM	58-DXM	59-DXM	60-DXM	61-DXM
Phonology	**alternation**	alternation -ān>-un	70	6	11	15	6	10
Phonology	**alternation**	alternation -e>-i	3	0	2	0	0	0
Phonology	**alternation**	alternation -ar>-e (final consonant omission + -a>-e)	17	1	4	4	0	5
Phonology	**alternation**	DO marker -rā>-ro	41	1	4	2	4	8
Phonology	**alternation**	DO marker -rā>-o	10	3	3	1	1	2
Phonology	**omission**	of h (any position)	7	5	4	1	2	4
Phonology	**omission**	final consonant	20	0	1	4	1	2
Phonology	**contraction**	of enclitic pronoun	8	2	1	4	1	1
Phonology	**contraction**	of consonant cluster / other	5	2	0	0	0	1
Morphology	**verbal ending**	3sg praes	40	5	4	0	2	9
Morphology	**verbal ending**	2pl praes/praet	7	0	1	1	0	0
Morphology	**verbal ending**	3pl praes/praet	53	2	1	4	4	1
Morphology	**verbal ending**	copula 3sg praes	25	1	3	1	3	1
Morphology	**verbal stem**	MSP stem	47	13	5	0	3	12
Morphology	**verbal stem**	alternate stem	3	1	0	0	0	0
Morphology	**tense**	present progressive tense	3	0	2	2	0	0
Morphology	**tense**	past progressive tense	2	0	1	0	0	0
Morphology	**tense**	other	0	0	0	0	0	0
Morphology	**tense**	two verb construction (other than modal)	0	2	1	0	0	1
Morphology	**enclitic pronoun**	possessive / partitive	22	2	4	2	9	5
Morphology	**enclitic pronoun**	with nomVerb as DO	3	0	1	0	0	0
Morphology	**enclitic pronoun**	with nomVerb as IO	1	0	0	1	0	0
Morphology	**enclitic pronoun**	with preverb as DO/IO	0	0	0	0	1	0
Morphology	**enclitic pronoun**	with verb as DO/IO	1	0	0	0	0	0
Morphology	**enclitic pronoun**	preposition + enclPron ie as DO/IO	7	2	0	1	1	1
Morphology	**enclitic pronoun**	with nomVerb/preverb as logical subject (quasi impersonal idioms)	1	1	0	0	0	0
Morphology	**enclitic pronoun**	with infinitive (noun of action) as agent or object	0	0	0	0	0	0
Morphology	**particle or conjunction**	enclitic particle ham	0	1	0	1	0	0
Morphology	**particle or conjunction**	emphatic particle hā	0	0	0	0	0	0
Morphology	**particle or conjunction**	temporal ke	1	0	0	0	0	0
Morphology	**particle or conjunction**	emphatic ke	3	0	1	0	1	0
Morphology	**article**	indefinite article yek	19	0	1	3	2	4
Morphology	**article**	indefinite article yek + -i	1	0	0	0	0	0
Morphology	**article**	definite article -e	2	0	0	0	0	0
Syntax	**word order**	MSP word order SOV>SVO - direction/location	15	1	2	3	1	1
Syntax	**word order**	MSP word order SOV>SVO - nominal clauses	0	0	2	2	1	0
Syntax	**word order**	disrupted word order, SOV>SVO other than two above	3	0	0	0	0	0
Syntax	**omission**	aux verb	1	2	3	1	0	2
Syntax	**omission**	preposition	2	1	0	0	0	0
Syntax	**clause without conjunction**	coordinate	1	3	0	0	0	0
Syntax	**clause without conjunction**	subordinate	2	1	0	1	0	0
Syntax	**lack of agreement**	between subject and verb (quasi impersonal idiom, logical subject, number disagree	2	0	2	0	1	0
Lexis	**MSP lexical item**	preposition MSP	13	2	5	4	0	0
Lexis	**MSP lexical item**	single word, compound word	7	3	4	1	7	2
Lexis	**MSP lexical item**	phrase	4	4	1	3	2	2
Lexis	**MSP lexical item**	saying	1	0	0	0	0	0

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

62-DXM	63-DXM	64-DXM	65-DXM	66-DXM	67-DXM	68-DXM	69-DXM	70-DXM	71-DXM	72-DXM	73-DXM	74-DXM	75-DXM	76-DXM	77-DXM	78-DXM	79-DXM	80-DXM	DXM SUM
6	10	6	10	11	17	6	9	12	10	6	12	14	12	15	14	18	9	11	**326**
1	0	1	4	1	0	0	0	3	4	3	0	2	1	3	0	1	2	0	**31**
0	1	1	3	0	0	0	0	0	3	1	2	1	2	0	1	1	0	4	**51**
1	2	2	1	5	4	4	6	6	3	3	12	0	2	3	4	6	2	3	**129**
2	1	1	0	1	0	2	1	5	3	3	0	2	5	3	2	0	2	2	**55**
3	2	1	2	2	3	0	2	3	2	1	0	1	2	3	1	6	3	2	**62**
4	0	0	4	1	1	1	0	4	0	7	0	0	1	2	1	0	0	0	**54**
2	0	0	1	1	2	0	2	1	0	2	0	1	0	1	0	1	0	3	**34**
0	0	0	2	0	0	0	0	0	0	0	0	0	0	3	1	0	2	0	**16**
18	1	13	18	1	5	5	4	5	4	6	6	3	0	2	0	2	1	12	**166**
0	0	0	7	0	0	0	0	1	0	0	0	0	0	2	0	0	1	0	**20**
1	2	3	3	9	4	2	11	1	6	4	13	15	18	13	9	15	9	4	**207**
4	2	3	3	0	3	0	1	3	4	6	1	1	4	4	0	0	3	2	**78**
14	7	7	17	1	4	1	8	3	4	2	7	5	4	2	1	6	1	9	**183**
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	**5**
0	1	2	0	0	0	0	0	2	0	2	2	1	0	0	0	0	0	0	**17**
0	0	0	0	0	0	0	1	1	0	2	1	0	0	0	0	0	0	0	**8**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	**6**
8	3	2	3	3	2	1	4	5	0	2	0	2	4	1	2	2	2	3	**93**
1	0	0	0	0	0	1	0	0	0	0	1	2	0	0	1	0	0	1	**11**
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	**4**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
5	0	0	1	0	1	0	0	0	0	0	0	0	0	1	1	2	0	1	**24**
1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	**4**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	**1**
0	0	1	2	0	0	0	0	1	0	0	0	0	0	2	0	0	0	0	**8**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**
0	0	0	1	2	0	0	0	1	0	0	1	0	0	1	0	1	0	0	**12**
2	0	1	3	0	0	1	1	4	0	4	0	0	1	2	1	0	0	0	**49**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**
3	0	0	4	2	0	1	0	1	0	1	0	0	1	0	1	0	0	0	**37**
0	0	0	1	2	0	1	0	1	0	0	0	0	0	1	1	2	2	0	**16**
0	0	0	1	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	**7**
3	3	1	4	4	1	0	0	1	1	2	4	4	3	1	2	0	1	1	**45**
0	0	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	0	0	**7**
2	0	0	1	0	0	0	0	0	0	0	0	0	2	2	0	0	0	0	**11**
0	0	0	1	0	0	0	0	0	0	0	0	1	0	3	0	0	0	0	**9**
1	0	1	2	0	0	1	0	0	1	0	1	2	0	0	0	1	0	0	**15**
1	9	3	6	2	4	1	3	3	1	2	0	3	1	0	3	1	1	0	**68**
2	1	1	2	2	0	0	0	1	3	2	0	3	2	3	3	2	3	4	**58**
4	1	3	1	4	2	2	2	2	1	0	0	1	1	1	0	0	0	1	**42**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**

Figure 4. Quantitative analysis data table example (source: own research).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

4.3.6 Quantitative analysis. Presentation of the results

After the quantitative analysis of the source texts was completed and the number of occurrences of each parameter recorded in the table of data, it was possible to analyze the resulting numbers and see if any patterns or tendencies are visible. To facilitate the proceeding of data, it was summarized and presented in the form of smaller and more specified tables as well as charts. The tables allowed for two kinds of observations: on one hand, the tendencies apparent in individual books in regard to the application of MSP patterns in the text and their comparison, and on the other hand, the tendencies of the parameters themselves to appear in literary texts with varying frequency. With acute awareness of the fact that the data sample in this project is too small to allow for a regular statistical analysis, still some observations were too important and too discernible to ignore.

4.3.7 Conclusions

The final, closing stage of the analysis included drawing conclusions from the data presented in the preceding stage and comparing them with the predicted results.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Chapter 5.
Presentation of the results

5.1 Discussion of the parameters

In the section below, the parameters distinguished during analysis are discussed in more detail with reference to their existing grammatical descriptions. Some of them have long been known and described as colloquial, the others might have been associated with local varieties of Persian, especially the dialect of Tehrān. Some may even belong to the written register in general, but in specific contexts or together with other forms become a marker of the spoken register.

Many of the described parameters are rooted in the previous stages of the development of Persian (Middle or even Old Persian) and are not in fact innovative. They might have been eradicted from the written norm and preserved in speech, where they have continuously been in use as one of the alternative expressions of a phenomenon, until they finally reached the time of grammaticalization:

> the possibility, which may have lived through centuries, becomes grammatical only at a definite stage of the language history, perhaps motivated by other simultaneous syntactic changes and becomes a generally obligatory means of expression to substitute for an extinct or changed category (Jeremiás 1984: 286).

And of course, not all of those parameters have been fully grammaticalized as of yet. Many of them are not a completed change, but rather a change in progress.[111]

This description of grammatical features chosen as diagnostic parameters of Modern Spoken Persian is intended to give a synchronically-oriented overview rather than an in-depth analysis of both their synchronic and diachronic aspects, which would be beyond the scope of the present study. As a grammatical reference for the descriptions, the books listed below were mostly used:

111 Such is the case of the grammaticalization of present and past progressives with auxiliary *dāštan* (5.1.2.6), which has lacked a negative form until recently, but the use of negation has been noted in the lyrics of a pop song recently (see the footnotes in the mentioned section).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

1. Gilbert Lazard (1957) *Grammaire du persan contemporain* – over half a century old, this is probably still the most often cited and referred to grammar of Persian language which includes information on the colloquial language. In a preface to the 1992 translation, Ehsan Yarshater calls it "the most detailed, systematic and comprehensive grammar of contemporary Persian written so far" (Lazard 1957/1992: i). Most of the observations made there are still relevant and reliable. In this book, quotes from Lazard come from a 1992 American translation by Shirley Lyons, revised by the author.

A description of Lazard's grammar would not be complete without a mention of the fact that could be called **Lazard's paradox:** being treated as a reliable and relevant source of information on the colloquial forms, of which there are numerous examples in that grammar indeed, it actually relies on their **literary versions extracted from modern prose**, the writings of S. Hedāyat, S. Čubak, M.-A. Ğamālzāde (Lazard 1957/1992: 2–3).[112] So taking Lazard's grammar as a source of colloquial forms, we do not base our knowledge on the actual colloquial, but rather on its literary output and those writers' rendering of forms that they have known from their personal experience and decided to include in their books.

2. Gernot L. Windfuhr (1979) *Persian Grammar. History and State of Its Study* – a thorough, comprehensive insight not only into Persian grammar, but also the historical studies up to the day of its publication. This book is concerned more on the literary register of Persian, but holds valuable observations and analyses of the studies of others and often presents grammatical phenomena from more than one point of view. In comparison to (1), which is a typical descriptive grammar, this book is largely analytical.

3. Gernot L. Windfuhr and John Perry (2009) *Persian and Tajik* [in:] G. L. Windfuhr (ed.) *Iranian Languages* – a relatively recent publication that gives an overview of both Persian and Tajik grammar where some observations from (2) can be revised or enriched, but unlike the former, this one is again a descriptive grammar. It could, however, serve as a departure point for some comparison of MSP of Iran with its Tajik counterpart.

4. Lazar Samoeilovič Peisikov (1960) *Tegeranskii dialect* – a comprehensive description of the dialect of Tehran. Published about the same time as (1), it includes descriptions of many features of Tehrani dialect that have become characteristic of the general spoken standard. In this book, instead of the 1960 original version, a Persian translation by Mohsen Šoğā'i entitled *Lahğe-ye tehrāni*, published in 2002 in Tehran is used. It is a deliberate choice: the translation is enriched by the translator's observations and comments, which sometimes alter or even contradict the original observations of the author, and the translator's preface gives some hints on the terminological difficulties of the de-

112 The examples of everyday usage do appear as well, but they are less in number.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

scriptions of contemporary Iranian linguistic situation. An interesting feature of this translation is the spelling of the quoted Tehrani forms: the translator maintains Peisikov's original transcription beside the Persian script, yet the Persian script forms reflect the "proper" Persian spelling rather than what is romanized. For example, the sentence romanized as *tšeraɣo bijar* is spelled as چراغ را بیاور, that is, *čerāq-rā biyāvar* 'bring the lamp' and *piše mane* is spelled as پیش من است *piš-e man ast* (Peisikov 1960/2002: 64). When a change in word order or an additional sound in the final position is the case, they are preserved in Persian spelling as well. Two interpretations of such behavior are possible: on one hand, Persian script could be treated here as a means of showing a regular Persian form in comparison to the dialect. In such a case, however, changes in word order, additions and similar changes should not be applied as well. On the other hand, the conservative (or, even more, preservative) character of Persian script and the matters of correctness might influence the translator and the editor of the book and prevent them from using forms which are generally perceived incorrect.

For the more diachronic insight, some parts draw from Aleksander Chodźko (1883) *Grammaire persane,* known for his observations on the vernacular/colloquial language and Haggard and Le Strange (1882) *The Vazir of Lankarun* in the preface to which the authors aimed at describing Persian language of their day in a way it was actually spoken, which makes it a relevant source for comparison with the MSP of today. The list is by no means finite and additional sources are brought whenever necessary. Not all the descriptive grammars of Persian include information on the features specified in the analysis. This is true especially for those grammars which focus on the written standard and for that reason some previously mentioned books are not referred to in the following sections.

The parameters are described within the four categories: phonological, morphological, syntactic, and lexical. Each of them is divided into subcategories, a number of which differs: there are three groups of phonological parameters (9 parameters on the whole), six groups of morphological parameters (25 parameters), four groups of syntactic parameters (8 parameters) and four groups of lexical parameters (without further subdivisions); the overall number of parameters is 46. Each of the parameters is described in detail and there are examples from the analyzed material provided. More often than not, the examples include more than one parameter and therefore they might appear more than once, even though the effort has been put into bringing a variety of examples from the analyzed corpus.

The overall distribution of parameters can be represented by a following chart (see figure 5 below), of which the implications and further details are discussed in chapter 5.2:

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

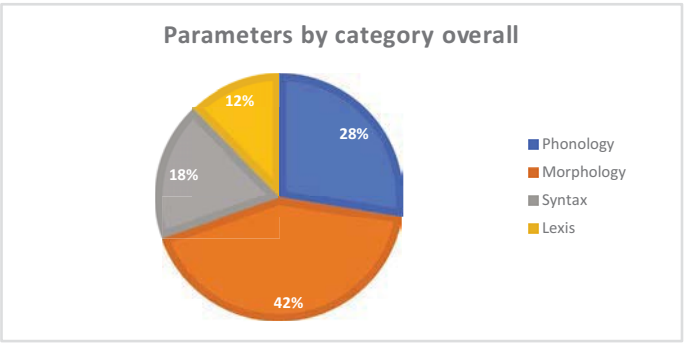

Figure 5. Distribution of MSP parameters in the corpus – proportions (excluding possessive/partitive EPs) (source: own research).

5.1.1 Phonological

There is a tradition of describing most of the MSP (or traditionally colloquial/informal) features as phonological (see Hodge 1957, Jazayery 1970, also Lazard 1957/1992) and their origins are in most cases undoubtedly so (cf. also Jeremiás 1984). Phonological changes which have affected New Persian in the course of its development towards Modern Standard Persian, have often had their effect on the words and morphemes of high frequency such as plural suffix, verbal endings, demonstrative pronouns, or enclitic personal pronouns (also known as personal suffixes). High frequency and grammaticalization of the resulting new phonological forms of those elements allow them to be considered, synchronically speaking, as morphological rather than phonological characteristics. Therefore in the analysis, such elements count as morphological and not phonological. The categories below (alternation, omission, contraction) serve to describe only those features of MSP which can be said to have purely (or almost purely) phonological character.

Naturally, speaking of phonological features in the context of printed source material may raise doubts because no sound analysis is involved. What is meant here by phonological features are those which result in an observable change in the spelling of words, that is, their printed shape. Not all the phonological features of modern spoken standard of Iranian Persian will be visible in print, for two reasons mostly: 1) In Persian script, only the historically long vowels *ā*, *I* and *u* have a stable representation by dedicated characters آ ی و in writing[113]; there are diacritics (called *harakāt*) used to indicate the historically short

113 Those vowels are sometimes also called "stable," in contrast to *a*, *e*, *o*, historically short, which are in that case termed "unstable." The "unstable" vowels undergo certain changes:

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

vowels *a e o*, but their use is by no means obligatory and they are hardly used in popular texts, therefore a change from *a* to *e* will not affect the printed shape of a word in which it occurs; 2) Changes in orthography proceed slowly in comparison to the changes in phonology[114] and it is possible that a new spelling of a word, even common, will nevertheless not be represented in spelling, where an older form will be retained. In my study, I focus on those forms which differ from their regular, standard MWP spelling in a way that indicates their belonging to spoken register. In the analyzed material, three groups of changes can be found that I treat as phonological: alternations, omissions and contractions.

Number of their occurences and their relative distribution is shown in the charts below:

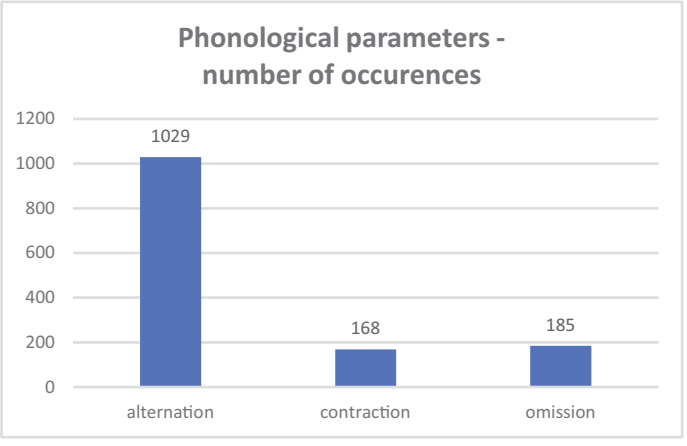

Figure 6. Phonological parameters - number of occurences (source: own research).

assimilation or even complete reduction in rapid speech, whereas the "stable" vowels do not (see, for instance, Pisowicz 1985: 13, Windfuhr and Perry 2009). Although it is not their graphic shape that is beyond this name, I believe the fact that they are recorded in print does have an impact on their stability. The historically short or unstable vowels, which are often invisible in print, may undergo changes completely, so to say, off the record.

114 For a comprehensive study of Persian orthography, see Hashabeiky (2006).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

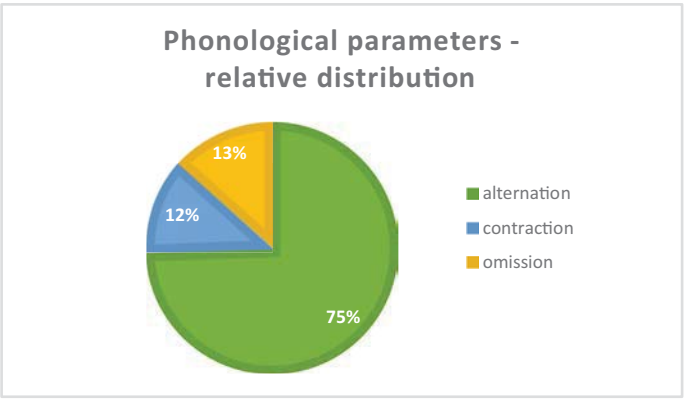

Figure 7. Phonological parameters – relative distribution (source: own research).

5.1.1.1 Alternations

Alternations are the most common of phonological parameters found in the analyzed material. The word 'alternation' is used as a cover term for those cases in which there is a change of quality in the vowel, accompanied or not by other changes. During the analysis, five types of alternations were distinguished:

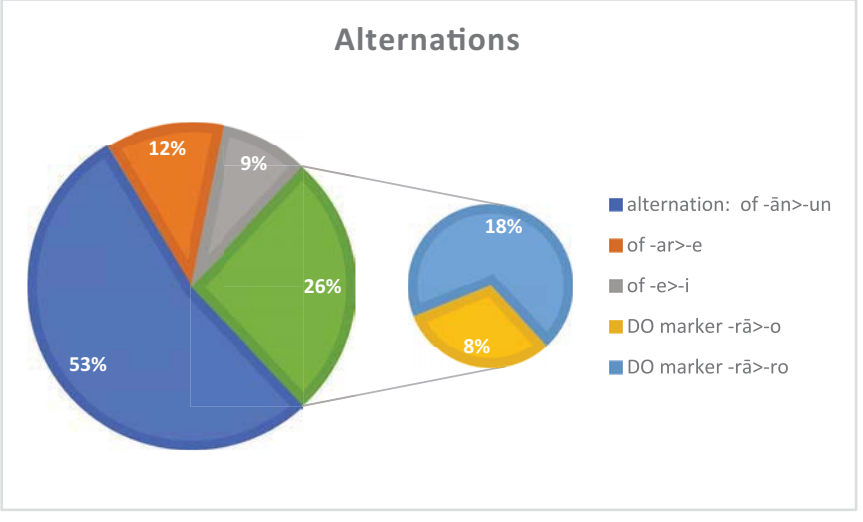

Figure 8. Alternations - relative distribution (source: own research).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

As can be seen in the chart above, more than a half of the recorded alternations were the cases of raising of *ā* to *u* before nasals, a change well-attested and described as typical of the Tehrani dialect (see the description below). The other changes were also present, but much less frequent, the rarest being a change of *-e* to *-i*. Most of the cases of phonological alternations were found in two of the analyzed books, *Dar xarābāt-e moqān* and *Šahr-āšub*.

5.1.1.1.1 Raising of ā *before nasals (*āN > uN*)*

Raising of *ā* to *u* before nasals is perhaps the most recognized phonological feature of the MSP. It is considered one of the distinguishing features of the Tehrani dialect of Persian (Peisikov 1960/2002: 35), that with the growing prestige and usage of this dialect has spread throughout the country and distinguishes spoken register from written. This raising is not a universal process: neither does it affect all groups of words (the words which belong to high register are not affected, even if used in the lower registers), nor is it consequently applied; in fact, it is a feature that can be considered register-sensitive. "In colloquial language, before *n* or *m, ā* changes to *u*. This phenomenon is one of the most characteristic which differentiate the colloquial use of Teheran from the formal language. It is very frequent in colloquial words and may also extend to the most literary words occasionally used in colloquial conversation" (Lazard 1957/1992: 19).

It is worth noting that raising of *ā* before nasals is not new and has already been attested in 19th century. Chodźko (1883: 8) mentions a tendency among the natives of Fars province to pronounce it as *u* (regardless of the environment; but he also mentions the same words being pronounced with *ā* in the heart of Tehran). Haggard and Le Strange (1882: xxiv) also make that observation, but add that "certain words all over Persia are pronounced *colloquially* after this incorrect fashion. Such are نان *nun* 'bread,' and آن *un* 'that:' همان *hamun* for *hamān* is another example of this." Attempting to give the most accurate Tehrani pronunciation of their time, they also mark the pronunciation of long *ā* as in English *ball* typical of the residents of Kāšān (and "a fashion among men of the last generation") (Haggard, Le Strange 1882: xxii). Perry (1996: 274) traces the rounding of *ā* back to as early as Safavid Esfahān and Latin transcriptions of Persian translations of the *Qor'ān* (see also Miller 2012: 162), but calls it a "facultative vowel shift," in fact, not a phonological shift at all since it is selective, facultative and relevant. Miller (2011, 2012) represents a different approach and puts forward a hypothesis that Iranian raising of *ā* to *u* before nasals is the last stage of a phonological process of a chain shift that has never reached completion, with /ā/ as a starting point, then rounded to [ɒ], then raised to [o] and consequently to [u]: "perhaps the underlying forms contain /u/, and these are optionally changed to [ɒ] in higher registers, influenced by the presence of alef ا

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

in orthography" (Miller 2011: 4).[115] The fact that words in which the pre-nasal raising occurs most often are enclitic pronouns and deictics, closed class forms which can resist changes undergone in other parts of language, as noted by Kahn and Bernstein (1981: 138), would seem to support such hypothesis, although it should be noted that this change affects other classes of words frequently as well, even the Arabic loan-words such as *heyvān* > *heyvun* 'animal' or *hammam* > *hammum* 'bath' (Sadeghi 1986: 230).

Raising of *ā* to *u* before nasals is one of the best represented phonological parameters.[116] The total number of its occurrences in the analyzed material is 549 instances (40% of all the phonological parameters), and examples can be found in all the analyzed books except one (Hn). They are extensive in DXM (326 instances) and ŠĀ (189), and observable in KP (28 instances). ĀM has only 6 instances of this parameter, while Hn, as stated, has none.

Examples:

(1) 3pl personal pronoun MWP *ānhā* > MSP *unā* 'they':
unā *šekāyat* *kard-an*
3PL NOMVERB do.PST-3PL
they complaint did
unā šekāyat kardan 'they complained' (DXM 78)

(2) adverb of place MWP *ān-ǧā* > MSP *un-ǧā*[117] 'there'; verb MWP *rānd-e* > MSP *rund-e* 'drove away':
az *un-ǧā* *rund-e*
PREP DEM-place drive.PST-PTCP
from there drove.away
az unǧā runde 'drove away from there' (DXM 79)

(3) pronoun MWP *ān-če* > MSP *un-če* 'that'; verb MWP *mi-dān-id* > MSP *mi-dun-in* 'you know':
un-če *ke* *goft-in* *va* *mi-dun-in*
DEM-PRON RELPRON say.PST-2PL CONJ PRS-know-2PL
that what you.said and you.know
unče ke goftin va midunin 'what you said and know' (ŠĀ 73)

115 According to Miller (2011), the fact that in contemporary Persian (MSP) ā still appears in some words is another phonological change, an example of a change from above (prestige-motivated change introduced by people themselves—in other words, he assumes that there has been a change of *ā* > *u* first, which is then reversed by people to *u* > *ā* again since this form is associated with more prestige.

116 In my opinion, this is due to its clear and instant recognition as a marker of spoken register.

117 In Modern Persian, demonstratives are frequently used to form compounds: the adverb of place *ānǧā/unǧā* 'there' is in fact a compound of demonstrative *ān* 'that' with noun *ǧa* 'place.' A similar process is behind the form from the next example, *ānče*, which is composed of a demonstrative *ān* 'that' with a pronoun *če* 'which.' This demonstrative undergoes raising of *ā* > *u* not only in its independent form, but in the compounds just as well.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(4) noun MWP mehmān > MSP mehmun:
mehmun-e man
guest-EZ 1SG
guest my
mehmun-e man '[be] my guest' (KP 57)

(5) adjective MWP *ğān* > MSP *ğun*
Nosrat-ğun
Nosrat-ADJ
Nosrat-dear
Nosrat ğun ğun 'my dear Nosrat' (AM 55)

5.1.1.1.2 Change of final -ar *to* -e *(final r omission, then a>e)*

Although this change is in fact a result of two subsequent changes: 1) final consonant drop, and 2) raising of *a* to *e*. I decided to combine them together because of the peculiar environment: this alternation is seen especially (if not only) in four words: *digar* 'yet' > *dige; agar* 'if' > *age; magar* 'unless' > *mage*; *āxar* 'in the end, finally, eventually' > *āxe*; yet the frequency of those words is so high that makes the overall perception of this change as a very frequent one.[118]

The first step is a final consonant drop, a change attested for other consonants such as *t* or *k* (not so much for *r* though; in fact, Peisikov (1960/2002: 42) notes that it is one of the consonants that are always preserved), that leaves a word with a final *-a*. And then, since the final vowel *-a* does not appear in MSP pronunciation just as it is not accepted in Tehrani pronunciation (Lazard 1957/1992: 12, 30; Peisikov 1960/2002: 30), the second step is raising of *a* to *e*.[119] According to Perry (1996: 273), this raising is dated as 19th century, although it might have already been in process way in the Safavid period.

The frequency of this parameter is lower than the former (8,54% of all the phonological parameters): DXM has 51 instances, ŠĀ 62, and KP 5. It is not represented at all in ĀM and Hn.

Examples:

(1) adjective MWP *digar* > MSP *dige* 'other':
do se nafar-e dige
NUM NUM person-EZ ADJ
two three person other
do se nafar-e dige 'two [or] three other persons' (DXM 52)

118 The fluctuations of *a* and *e* in other positions and environments both on lexical and morphological level are well attested by Pisowicz (1985: 15).

119 The etymological final *a* also becomes *e*, but returns to *a* after the addition of suffix or clitic, thus *xāne* 'house' (MSP *xune*) + -š 'his/her' > *xuna-š* 'his/her house' (Lazard 1992: 30). According to Pisowicz, this could indicate that in spoken utterances, this vowel is simply treated as non-final, i. e. the form with enclitic pronoun is perceived as a single unit (Pisowicz 1985: 16).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(2) emphatic particle / filler word MWP *āxar* > MSP *āxe* 'oh, but'[120]:

āxe	*mi-tars-am*	*eštebāh*	*kon-i*	*mādar*
EMPH	PRS-fear-1SG	NOMVERB	DO.SUBJ-2SG	NOUN.VOC
oh.but	I.fear	mistake	you.make	mother

āxe mitarsam eštebāh koni mādar 'oh, but I fear you make a mistake, son[121]' (ŠĀ 50)

(3) conjunction MWP *agar* > MSP *age:*

age	*esrār*	*dāšt-am*	*be-r-am*	*ğā-y-i*
CONJ	NOMVERB	have.PST-1SG	SUBJ-go-1SG	place-INDEF
if	insistence	I.had	go	somewhere

age esrār dāštam beram ğāyi 'if I really wanted to go somewhere' (KP 62)

5.1.1.1.3 Raising of e *to* i

This category houses a certain group of frequently used words, especially *če* 'what' which in MSP is often realized as *či.* According to Lazard (1957/1992: 22), *e* does become *i* in Tehrani speech in the neighborhood of prepalatal consonants or syllable with *i.* Such an explanation would suggest that this change might result from vowel harmony, but the actual reason might be rooted deeper in language. Pisowicz (1985: 95–96) notes that this is not a change, but the retainment of the previous pronunciation of classical short *i* and *u*, which in other positions was frequently shifted towards *e* and *o,* respectively. Mahmoodi-Bakhtiari (2018: 192) mentions another group of words in which this change takes place, i. e. *šekam* 'belly' > *šikam*; *ğegar* 'liver' > *ğigar*; *negāh* 'sight' > *nigā(h)*; *šekar* 'sugar' > *šikar*; *tekke* piece' > *tikke*; and *gonğešk* 'sparrow' > *gonğišk*, and calls it a "seemingly regular change" (ibid). Peisikov (1960/2002: 29) gives the examples of imperative and subjunctive forms like *beriz* 'pour' > *biriz; benešin* 'sit' > *bišin*[122] as well as the examples quoted above: *šikar* 'sugar,' *šikam* 'belly,' *nigā* 'sight,' *kučik* 'small' and *ke* > *ki* 'who,' *če* > *či* 'what.' Vowel harmony and the effect of a preceding prepalatal consonant being a sufficient explanation in some of the cases, one should bear in mind that if there is an observable change of *a* to *e* in MSP (see

120 The regular meaning of *āxar* (used as a noun, subject, adjective or adverb) is 'end, final, finally' but when applied as a particle / filler, its meaning changes. When used at the beginning of an utterance, it can express doubt or tentativeness of the speaker.

121 This utterance, a part of a dialogue between mother and son, is an example of address inversion, where mother uses the word *mādar* 'mother' (or *māmān* 'mom') instead of *pesar* 'son' when speaking to her child. Such reverse addressing is characteristic of Persian vernacular language, limited to the close family (mother, father, maternal/paternal uncles and aunts, siblings), and asymmetric (children cannot use reverse addressing towards their parents) (Tavakol, Allami 2013).

122 Here the assimilation is preceded by a middle syllable reduction *benešin* > *bešin* which creates the required phonological neighborhood and leads to assimilating of *e* to *i.*

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

5.1.1.1.2), raising of *e* to *i* seems to be a natural second step in the chain of events.[123]

This parameter appears in all the analyzed books and has a slightly different pattern of frequency: DXM has 31 instances, in ŠĀ there are 36, KP has only 5 instances, but ĀM as much as 18, and Hn 1. Usually those are either the independent pronoun *če* realized as MSP *či* or its compounds, but other forms also appear (see ex. 4 below). Its frequency among other phonological parameters is similar to the former change of *-ar* > *-e,* 8,85% of all the phonological parameters.

Examples:

(1) compound pronoun MWP *har-če* > MSP *har-či* 'anything, whatever':
har-či *hast* *tu* *in* *miz-e*
PRON exist.PRS3SG PREP DEM NOUN-COP3SG
anything there.is in this table.is
harči hast tu in mize 'anything that there is, can be found at this table' (DXM 52)

(2) interrogative pronoun MWP *če* > MSP *či* 'what':
či *mi-xun-in*
INTERROG PRS-read-2PL
what you.read
či mixunin 'what do you read?' (ŠĀ 53)

(3) interrogative pronoun MWP *če* > MSP *či* 'what':
či *šod*
INTERROG happen.PST3SG
what happened
či šod 'what happened?' (ĀM 58)

(4) cardinal number MWP *šeš* > MSP *šiš* 'six':
raft-am *dam-e* *pan-šiš=tā* *šerkat*
go.PST-1SG PREP-EZ NUM-NUM=CL NOUN
I.went to five.six company
raftam dam-e pan-šiš tā šerkat 'I went to five or six companies' (KP 57)

123 This is not in accordance with Labov's Principle II of chain shifting, according to which "In chain shifts, short vowels fall" (Labov 1994: 116). That would suggest a movement in the opposite direction: *i* > *e* > *a* and not raising *a* > *e*> *i*. But even Labov himself (ibid.) notes that there are exceptions to the Principle II. If we take the pronunciation remarks of Haggard and Le Strange (1882: xxii) as legit, then the 19th century short *a* (between English *a* in *bad* and *u* in *bud* or *sun*) was much lower than it is today, fitting the first step of the raising chain well. On the other hand, Windfuhr (1979: 135) describes Iranian Persian vocalic system as such: "Undoubtedly, in classical Persian the vocalic system was marked by a length distinction. In Iranian Persian the system changed, with the lowering of the short high vowels **i, *u* > *e, o* (and the raising of **ē, *ō* to *i, u*)," which would be a perfect example of a chain shift with short vowels falling and long vowels raised. Also the translator and editor of Peisikov's translation, Mohsen Šoğa'i, mentions this change as frequent in the development of Persian and gives the example of *ātaš* > *āteš* > *ātiš* 'fire' (Peisikov 2003: 30, footnotes).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(5) interrogative pronoun MWP *če* > MSP *či* 'what':
manzur-eš *daqiq-an* *či-st*
NOUN-EP3SG ADJ-ADV INTERROG-COP3SGPRS
her.meaning precisely what-is
manzureš daqiqan čist 'what she exactly meant' (Hn 50)

5.1.1.1.4 Change in direct object marker: rā > ro, rā > o

This change is also well attested in the descriptions of modern Persian colloquial and Tehrani speech. The objective enclitic particle *rā* in colloquial language ordinarily changes to *ro* and after a word with final consonant, ordinarily loses its consonant (Lazard 1957/1992: 15, 19) thus *rā* becomes *ro* after a vowel and *o* after a consonant. While it is usually described in terms of a different form of the DO marker, in fact, we could consider it as another example of raising of the long *ā*, this time without nasal neighborhood, but in a very limited, closed-class environment. Windfuhr (1979: 47–57) discusses the functions of *rā* in detail, but does not mention the change in form. Windfuhr and Perry (2009: 443) mention the colloquial forms *-ro* and postconsonantal *-o*. Peisikov mentions *-o* and *-e* in the dialect of Tehran and confesses that he does not find sufficient explanation for which form is used when. The form with retained consonant (*-ro*) is used in more careful speech (Peisikov 1960/2002: 62–63). The analyzed material does not support *-e* for direct object marker, but both forms *-ro* and *-o* are present in three of the books: 184 instances in DXM(129 *-ro*, 55 *-o*), 81 in ŠĀ (57 *-ro*, 24 *-o*), and 5 in KP (1 *-ro*, 4 *-o*). It is the second most frequent phonological parameter, 19,55% of all the recorded phonological parameters belong to this category.

Examples:

(1) post-consonantal *-ro:*
saqf-ro *nešun* *dād-am*
NOUN-DOM NOMVERB give.PST-1SG
ceiling show gave
saqf-ro nešun dādam 'I pointed at the ceiling' (DXM 52)

(2) post-consonantal *-o* after personal pronoun:
man-o *tanhā* *be-zār-in*
PRON1SG-DOM ADV SUBJ-leave.PRS-2PL
me alone leave
man-o tanhā bezārin 'leave me alone' (DXM 53)

(3) post-consonantal *-o* after enclitic personal pronoun:
lebās-et-o *avaz* *kon*
NOUN-EP2SG-DOM NOMVERB do.PRS.IMPSG
your.clothes change do
lebāset-o avaz kon 'change your clothes' (ŠĀ 50)

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(4) post-vowel *-ro* after personal pronoun:

to-ro	*xodā*	*vel-am*	*kon*
PRON2SG-DOM	NOUN	NOMVERB-EP1SG	do.PRS.IMPSG
you	god	go.me	make

to-rā xodā velam kon 'for God's sake, let me go' (KP 64)

(5) post-consonantal *-o:*

ye	*maxsus-eš-o*	*vāse*	*xod-am*	*radif*	*kon-am*
INDEF	ADJ-EP3SG-DOM	PREP	REFL-EP1SG	NOMVERB	make.PRS.SUBJ-1SG
a	special.one	for	myself	queue	make

ye maxsuseš-o vāse xodam radif konam 'I'd prepare one special [of this] for myself' (KP 58)

5.1.1.2 Omissions

5.1.1.2.1 h *omission*

This omission may occur in different positions. One of the examples is the MSP form of the plural suffix *-hā* > *-ā* (Peisikov 1960/2002: 60), although the omission is not obligatory. Another similar case is the omission of *h* in 3rd person plural personal pronoun *ānhā*, often together with raising of the initial *ā* to *u*, so that MWP *ānhā* > MSP *unā.* According to Lazard (1957/1992: 11), "in colloquial speech, *h* tends to disappear in all except initial positions," yet this feature is not very frequent in print. Still, there are some examples from the material. There are 62 instances of *h* omission in DXM, 29 in ŠĀ and 7 in KP. No instances have been recorded in ĀM or Hn.

Examples:

(1) omission of *h* in plural suffix *-hā:*

tā	*kār-ā-m*	*tamum*	*šod*
CONJ	NOUN-PL-EP1SG	NOMVERB	become.PST3SG
until	works.my	finished	became

tā kārām tamum šod 'when my work was finished' (DXM 56)

(2) omission of *h* in plural suffix *-hā,* with demonstrative pronoun MWP *in-hā* > MSP *inā:*

in-ā-ro	*mi-g-i*
DEM-PL-DOM	PRS-say-2SG
all these	you.say

inā-ro migi 'you say all these [things]' (ŠĀ 50)

(3) omission of *h* in plural suffix *-hā* after a noun MWP *pesar-hā* > MSP *pesar-ā:*

pesar-ā-ye	*man*
NOUN-PL-EZ	PRON1SG
sons	my

pesarā-ye man 'my sons' (ŠĀ 50)

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(4) omission of *h* in emphatic enclitic particle *ham:*

fekr-eš-am	*ne-mi-tun-in*	*bo-kon-in*
NOMVERB-EP3SG-EMPH	NEG-PRS-can-2PL	SUBJ-make-2PL
thought.of.it.even	you can't	make

fekreš-am nemitunin bokonin 'you can't even imagine that' (KP 56)

5.1.1.2.2 Final consonant drop

In the preliminary version, this category was designed for all the final consonant omissions except *h* (which has its own category above). However, it should be noted that some of the other frequently dropped consonants were moved to other categories (final *-d* in 5.1.2.3 Verbal endings, final *-r* which occurs only in a few words and is followed by the change of *-a* > *-e* in 5.1.1.2), in some cases the drop of a final consonant is actually a contraction of consonantal group, i. e. *dast* > *das* 'hand' (KP 66) or *dust* > *dus* in the verb *dust dāštan* 'to love' (ŠĀ 61), see both below. So, even though Lazard's observation was of a more general character: "in colloquial language […] some consonants in certain positions or in certain words are dropped" (Lazard 1957/1992: 12–13), in fact most of the examples of final consonant drop are the final *-k* drop in *yek* 'one.' Final consonant drop occurred 54 times in DXM and 26 in KP, but interestingly only 4 times in ŠĀ. There are 3 instances of this feature in ĀM and none in Hn.

Examples:

(1) final *-k* drop in *yek* 'one' in the indefinite article function:

ye	*ğā-ye*	*xalvat*
INDEF	place -EZ	ADJ
a	place	quiet

ye ğā-ye xalvat 'a quiet place' (DXM 56)

(2) final *-ğ* drop in cardinal number in a postconsonantal position:
MWP *panğ* > MSP *pan* 'five'

raft-am	*dam-e*	*pan-šiš=tā*	*šerkat*
go.PST-1SG	PREP-EZ	NUM-NUM=CL	NOUN
I.went	to	five.six	company

raftam dam-e pan-šiš tā šerkat 'I went to five or six companies' (KP 57)

(3) final *-d* drop after a vowel: vocative use of MWP *ostād* > MSP *ustā*

ustā
NOUN.VOC
master

ustā 'master' (ĀM 57)

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

5.1.1.3 Contraction

5.1.1.3.1 Enclitic pronoun

Contraction of the enclitic pronouns takes place after the words which end in a vowel. While in the MWP a glide is inserted between the word-final vowel (*ā* and *u*) and the suffixed pronoun (singular *-y-am, -y-at, -y-aš*; plural *-ye-mān, -ye-tān, -ye-šān*),[124] in the MSP register the suffix may be reduced to the consonant and joined straight to the word-final vowel. This is attested among others by Lazard (1957/1992:33) and Windfuhr and Perry (2009: 435), but it has been mentioned in 19th century already by Haggard and Le Strange (1882: xxxii). For the words ending with *-e*, where no glide is ever inserted, the suffixation of enclitic pronoun has an additional effect: if final *-e* is the etymological final *-a,* it is changed back to its original shape, i. e. *xāne-aš* 'his house' > *xuna-š* (Lazard 1957/1992: 30).

Contraction of enclitic pronouns has a frequency comparable to other phonological parameters except raising of *ā* before nasals (6,3% of all the phonological parameters) and can be found in all analyzed books. There are 34 instances in DXM, 25 in ŠĀ, 19 in KP, 7 in ĀM and 2 in Hn.

For more parameters connected with enclitic pronouns, see 5.1.2.6.

Examples:

(1) contraction of 3rd person singular enclitic pronoun MWP *-aš* > MSP *-š:*

čāqu-š-rā	*gozāšt*	*zir-e*	*galu-ye*	*Seyyed*
NOUN-EP3SG-DOM	put.PST3SG	PREP-EZ	NOUN-EZ	NOUN
his-knife	put	under	throat.of	Seyyed

čāquš-rā gozāšt zir-e galu-ye Seyyed 'he pressed his knife to Seyyed's throat' (Hn 56)

(2) contraction of 1st person singular enclitic pronoun MWP *-am* > MSP *-m*

ǧurāb-hā-m-ro	*ham*	*dar.mi-ār-am*
NOUN-PL-EP1SG-DOM	ADV	PREV.PRS-bring-1SG
my socks	too	I.take.off

ǧurābhām-ro ham dar miyāram 'I take off my socks too' (DXM 56)

Dar āvardan 'to take off' is a compound verb with a preverb that changes the meaning of the basic verb; *āvardan* by itself means 'to bring.'

(3) contraction of 2nd person plural enclitic pronoun MWP *-(ye)tān* > MSP *-tun*

harf-hā-tun-am	*bā*	*ham*	*zad-in*
NOUN-PL-EP2PL-EMPH	PREP	PRON	hit.PST-2SG
your words	with	each.other	you.spoke

harfhātun-am bā ham zadin 'you spoke about it with each other' (ŠĀ 60)

124 It is different for the words with final *-e*, where the enclitic pronouns are joined without a glide, i. e. *xāne-at* 'your house.'

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

In this example, *harf* 'word' has a function of a nominal part of compound verb *harf zadan* 'to talk,' but it is still eligible for taking the plural suffix *-hā* and possessive enclitic pronoun *-tun,* augmented further by an emphatic particle *-ham* with reduced *h* (*ham* appears in this example twice, once as an emphatic particle and then as a reciprocal pronoun).

(4) contraction of 3rd person plural enclitic pronoun MWP *-(ye)šān* > MSP *-šān*
be *hame-šān* *be-deh-ad*
PREP PRON-EP3PL SUBJ-give.PRS-3SG
to all.of.them give
be hamešān bedehad 'give it to all of them' (KP 51)

Here the interesting observation is that contraction is not accompanied by raising of *ā* to *u* (expected form would be *-šun* and not *-šān*).

(5) contraction on 1st person plural enclitic pronoun MWP *-(ye)mān* > MSP *-mān*
āšenāyi-mān
friendship-EP1PL
our.friendship
āšenāyimān 'our friendship (lit. acquaintanceship)' (ĀM 51)

Again, the contraction is not accompanied by raising.

(6) contraction of 3rd person singular enclitic pronoun MWP *-aš* > MSP *-š*
dust-hā-š
NOUN-PL-EP3SG
her.friends
dusthāš 'her friends' (ĀM 50)

5.1.1.3.2 Consonant cluster and other

Under this category, the forms were listed in which a consonantal group is simplified by a drop of one of the elements. This happens usually in the final position and the second or final element is dropped. As Lazard noted, this happens to the groups of occlusive + liquid, spirant + occlusive, nasal + occlusive; the group *-st-* is often assimilated into *-ss-*, the same for *-sl-* in *mesl-e* (Lazard 1957/1992: 28).

Peisikov (1960/2002: 52–54) mentions the consonantal drop (Pers. *hazf-e hamxānhā*) as one of the characteristic features, having impact among others on the verbal system of Tehrani dialect. The most frequent reductions mentioned are: *t* in groups *st, št, ft, qt* and *rt* (it must be noted, however, that t in those groups is said to undergo assimilation to the preceding consonant, so that *st* > *s*, *št* > *š* and *zd* > *z*); *r* in groups *br* and *dr*; voiced dental *d* in *-nd* endings such as 3rd person plural verbal ending (MWP *-and* > MSP *-an*).[125]

125 The last one is included in the 'verbal ending' category.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Within the analyzed materials, certain forms not belonging to MWP in terms of how their spelling reflects their phonological shape were found which cannot be classified as any of the abovementioned groups and so they were added here. One of those is the change in 2nd person singular present verbal ending *-i*, which changes to *-y* in rapid delivery: *mixāyi* > *mixāy* 'you want,' *miyāyi* > *miyāy* 'you come' (Lazard 1957/1992: 13). This contraction is visible in writing and print: it occurs in the postvowel positions such as the above example; the regular *-i* ending is joined to the preceding vowel by a glide *y*, which means that in writing or print it is represented by two adjacent letters *ye* (یی). When the ending is changed from *-i* to *-y*, one of those letters is dropped.

Another very frequent form is the vowel change in the adjective *xub* 'good' to *xob*, used as a particle without a clearly specified meaning that could be roughly translated to English *well, so*. Also, phonological processes affecting consonants, such as assimilation of *č* to *š* before *k* (example 1 below) were included here.

The overall number of contractions of consonant clusters and others is relatively small. They make 5,83% of all the collected phonological parameters, 81 instances distributed as follows: DXM 16 instances, ŠĀ 35, KP 18, ĀM 9, and Hn 3.

Examples:

(1) reduction of *-st* to *-s* in final position; MWP *dust* > MSP *dus:*

un	*dus*	*na-dār-e*
DEM	NOMVERB	NEG-have.PRS-3SG
that	love	doesn't.have

un dus nadāre 'he doesn't like it' (ŠĀ 61)

(2) reduction of *-sl-* to *-ss-* before a vowel: MWP *mesl-e* > MSP *mess-e:*

mess-e	*in-ke*
NOUN-EZ	DEM-CONJ
it.seems	that

mess-e inke 'it seems that' (KP 58)

(3) reduction of *n* before *š:* MWP *enšāllāh* > MSP *išālā* 'God willing':

išālā
EXCL
God.willing

išālā 'God willing (I hope so)' (ŠĀ 64)

(4) reduction of *-st-* in the verb *va-istādan* 'to stop, to stand'[126] MSP *vāystād-e* > *vāsāde*

vā-sād-e	*tā*	*resid-e*	*be*	*kešiš*
stand.PST-PTCP	CONJ	arrive.PST-PTCP	PREP	NOUN
stood	until	arrive	to	priest

vāsāde tā reside be kešiš 'he stood until he reached the priest' (DXM 65)

126 This verb is an example of MSP lexical item, it is not present in the standard MWP lexical system.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(5) reduction of copula MWP *ast* > MSP *as*
ruz-e *yekšanbe* *as*
NOUN-EZ NOUN COP3SGPRS
day Sunday is
ruz-e yekšanbe as 'it is Sunday' (DXM 62)

(6) compound pronoun MWP *hič-ki* > MSP *hiš-ki* 'no-one':
hiš-ki *das* *na-zan-e* *lotf-an*
PRON NOMVERB NEG-hit.SUBJ-3SG favor-ADV
no-one hand doesn't.hit please
hiški das nazane lotfan 'nobody touches [it], please' (KP 66)

(7) contracted form of adjective *xub* 'good' > *xob* 'well, so' as a phatic expression:
xob *in* *yek-i* *xatt-e* *Benedikt* *ast*
PART DEM DEMPRON-DEF NOUN-EZ NOUN COP3SGPRS
well this one handwriting Benedict is
xob in yeki xatt-e Benedikt ast 'well, this one is Benedict's handwriting' (Hn 79)

(8) postvowel contraction of 2nd person singular ending connected with rapid delivery MWP *-yi* > MSP *-y:*
to *ne-mi-yā-y* *birun*
PRON2SG NEG-PRS-go-2SG ADV
you don't.go out
to nemiyāy birun 'you don't go out' (ŠĀ 57)

5.1.2 Morphological

Of all the groups of parameters, the morphological ones might seem the most controversial. Indeed, diachronically speaking, many if not all are of the phonological origin, this is also a common method of their description. For the synchronic description of a register of language, however, and for the purpose of clarity, I find it more relevant to speak of those parameters in terms of morphology regardless of their etymological background. An additional argument supporting such a view is the fact that those changes, indisputably phonological in nature, affect only certain groups of words and not others. Therefore it becomes more convenient to speak about a morphological category (such as verbal endings or particles) and group those parameters according to their function, than to choose a phonological point of view and break them down into more numerous subgroups according to the specified phonological phenomena which might occur only in the functionally (and not phonologically) restricted environments. So whenever a categorization was possible that grouped parameters into a certain morphological, functional group, such a categorization was preferred to the phonological. There are two exceptions to this approach, the first being the direct object marker, postposition *-rā* which changes (voluntarily) to *-ro/-o* in the MSP environment and was treated as a phonological parameter

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

despite its restriction to a morphologically defined category (here the reason behind such assignment was the fact that it represents a case of raising of *ā* to *o*, similar to the raising of *ā* to *u* before nasals); the second quasi-exception is the case of functional words *(m)agar/āxar/digar* that change to *(m)age/āxe/dige* (see Kahn and Bernstein 1981), for which it is more convenient to speak of a phonological change than to suggest a morphological closed-class set, even though the group includes only four words and such a phonological change is not witnessed in other words with similar phonetical environment (compare for example MWP *magar* > MSP *mage* 'if not' with MWP *bašar* > **baše* 'human').

5.1.2.1 Articles

5.1.2.1.1 Definite -e

Generally speaking, Persian does not possess a definite article (Lazard 1957/1992: 73–74). Definiteness is marked when a noun or a noun phrase is used in direct object position by postposition *-rā* (Windfuhr and Perry 2009: 442; see also discussion on the use of *-rā* in Windfuhr 1979: 47–57). In other syntactic functions, the nouns or NPs which are not marked as indefinite, are definite by default. Their definiteness can be highlighted by the addition of a demonstrative (*in* 'this' or *ān/un* 'that'). In MSP, in order to mark definiteness of a noun or NP that was previously mentioned, one can also use a facultative stressed suffix *-e* (or *-he* after a vowel). This suffix is present in many dialects of Iran[127] and has been attested in 19th century already, described as definite, demonstrative or determinative (Windfuhr 1979: 40–41). It is also mentioned in Jeremiás (1984) as one of the characteristic features of the L (=low) register of Persian.

Pisowicz observes that there is no mention of this suffix in Peisikov (1960/2002), so it should be exected to be a foreign influence rather than native Tehrani feature[128] (Pisowicz 2003: 179). He contrasts this definite *-e* with a noun-forming

127 For example, in the dialects of Fārs province, instead of *-e* there is a stressed suffix *-u* of the same function. Pisowicz (2003) attributes the origins of the stressed definiteness suffix -e to the influence of Kurdish dialects.

128 It is true that there is no 'definite *-e*' in Peisikov's description of Tehrani dialect. However, an interesting point is made elsewhere, in the discussion of postposition *-rā* as the marker of definiteness in the objective case (Pers. *hālat-e maf'ul*), where beside the form *-o* a possible form *-e* is given. Peisikov notes that there is no exact rule as for when *-o* should be used and when *-e*, but of those two only *-o* appears in script (Peisikov 1960/2002: 62–63). Bearing in mind that forms with *-rā* are definite, this reference could perhaps be relatable to the definiteness marker *-e* discussed here (although some of the examples by Peisikov are attached to a noun with an enclitic pronoun, which is not possible with the definiteness marker).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

suffix *-e* (identical in form), but doubts their common origin[129] and suggests a possible influence of Kurdish dialects (namely some Kurdish-Sorani dialects spoken in Iraq and Mukri dialect spoken in Iran) in which such final *-e* is used as a regular (however complementary) definiteness marker. The Kurdish examples quoted by Pisowicz after a study of Mukri by K. R. Eyyubi and L. A. Smirnova (1968) include mostly forms with a demonstrative: *ew kteb-e* 'this book,' *how dar-e* 'that tree,' though is it mentioned that *-e* can also be used without a demonstrative: *diz-e* 'that thief' (Eyyubi, Smirnova 1968: 22).[130]

It is important to note that stressed *-e* does not mark mere definiteness—it is used to refer to something that has already been mentioned in a conversation and is known to all the involved parties. Therefore perhaps instead of definite or determinative it should be called "referential *-e*." Another interpretation suggests the term *specificity marker* (von Heusinger and Sadeghpoor 2020[131]). This feature is relatively rare because of its peculiarity and the required special context. In the analyzed material, only 11 instances were found (DXM 3, ŠĀ 6, KP 0, ĀM 2, Hn 0; there were more of those forms beyond the range chosen for the analysis). Three of four examples below occur during some recollection of a story, only example (2) is different in that respect, because in there both participants of the dialogue recall the events that both of them took part in few hours before the dialogue.

Examples:

(1) retelling a past event:

ke	*kešiš-e*	*na-fahm-e*
CONJ	NOUN-REF	NEG-understand.SUBJ-3SG
so.that	priest.that	wouldn't.understand

ke kešiše nafahme 'so that the priest would not understand' (DXM 65)

(2) discussing a family situation (in reference to a past event):

pesar-e	*čahārdah-punzdah*	*sāl*
NOUN-REF	NUM-NUM	NOUN
that.boy	14–15	year

129 The noun-deriving *-e* comes from a Middle Persian suffix *-ag* of the same function (which was a continuation of Old Persian *-ka-, -aka-)* (Pisowicz 2003: 178).

130 Those examples resemble another Persian form as well: the demonstrative *-i*, in Persian called *yā-ye ešārat*, or *-i* in a referential or specific function, for example: *ān yek-i-rā did* 'he saw that one' or *kas-i-rā ferestād* 'he sent someone' (see Windfuhr 1979: 37–38 for a detailed discussion of definite and indefinite *-i*). It does not contradict the hypothesis of Kurdish origin of definite *-e* of course, for which Pisowicz (2003) has convincing arguments.

131 Apart from the discussion on *-e*, the paper gives valuable insights into Persian definiteness an indefiniteness marking.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

balke	*bištar*	*az*	*doxtar*	*bozorgtar-e*
CONJ	PRON	PREP	NOUN	ADJ-COP3SGPRS
or	more	from	girl	older.is

pesare čahārdah punzdah sāl balke bištar az doxtar bozorgtare 'that boy is older than the girl by some fourteen-fifteen years or even more' (ŠĀ 65)

(3) mentioning a well-known story:

hekāyat-e	*un*	*doxtar-e*
NOUN-EZ	DEM	NOUN-REF
story	that	that.girl

ke	*tu-ye*	*qal'e-ye*	*div*	*zenduni*	*šode.bud*
RELPRON	PREP-EZ	NOUN-EZ	NOUN	ADJ	become.PLPF3SG
which	in	castle	monster	captive	had become

hekāyat-e un doxtare ke tu-ye qal'e-ye div zenduni šode bud 'the story of that girl who was imprisoned in monster's castle' (ŠĀ 73)

(4) telling a story:

hame-ye	*kas-o-kār-e*	*doxtar-e*	*tu-ye*	*tasādof*	*mord-and*
PRON-EZ	COMPNOUN-EZ	NOUN-REF	PREP-EZ	NOUN	die.PST-3PL
all	relatives	the girl	in	accident	died

hame-ye kas-o-kār-e doxtare tu-ye tasādof mordand 'all of that girl's relatives died in a car accident' (ĀM 70)

5.1.2.1.2 *Indefinite* yek / *indefinite* yek + -i

In written standard Persian, indefiniteness may be marked by the addition of the enclitic *-i*. The numeral *yek* 'one' may also function as indefinite article, slightly less formal then *-i* and more frequent in spoken register (when in that function, it is devoid of its numeral meaning, thus *yek ketāb* might mean either 'a book' or 'one book,' depending on the context). There is also a possibility of combining both markers of indefiniteness, so that *yek* will precede the word and *-i* will be suffixed to it: *yek ketāb-i* (if the noun is complemented by an adjective in the ezāfe construction, indefinite *-i* is suffixed to the adjective rather than the noun). That form is not restricted to colloquial usage, but it is more frequent and more acceptable there (Lazard 1957/1992: 71, Windfuhr 1979:34–38, Windfuhr and Perry 2009: 432) (Lambton 1967: 45; "*Yak* is sometimes used with a noun to which the Indefinite *-i* has been added. Its addition does not materially alter the meaning"). In the MSP environment, final *k* in *yek* is often dropped and the resulting form is *ye* + *-i*.

In the analysis, both *ye(k)* as indefinite article and *ye(k)* + *-i* were included since the proper literary form of expressing indefiniteness would still be suffixation of *-i*.[132] Their appearance was counted separately and more attention was

132 Yadollāh Samare in his 1987 Persian language course *ĀZFĀ* leans towards accepting *yek* + noun (+ adjective) as literary and only *yek* + noun (+ adjective) +*-i* as colloquial (Samare 1987: 48). When indefiniteness is added to the noun + adjective in ezāfe construction, *-i*

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

paid to the circumposed form, which is typical of MSP and not very frequent in MWP. Results of the analysis prove this statement: there were 121 instances of *ye(k)* (DXM 49, ŠĀ 10, KP 48, ĀM 1, Hn 13), but only 20 of *ye(k)* + *-i* (DXM 4, ŠĀ 4, KP 9, ĀM 0, Hn 3).

Examples:

(1) *yek* as indefinite article:

yek	*ğoft-e*	*lab-e*	*barrāq*
INDEF	NOUN-EZ	NOUN-EZ	ADJ
a	pair	lip	shiny

yek ğoft-e lab-e barrāq '[a pair of] shiny lips' (ĀM 54)

(2) *ye* as indefinite article (final *k* dropped):

ye	*bāzğuyi-ye*	*sāde*
INDEF	NOUN-EZ	ADJ
a	investigation	simple

ye bāzğuyi-ye sāde 'a simple investigation' (DXM 84)

(3) *yek* as indefinite article:

vaqti	*ham*	*mi-yā-d*	*eyn-e*	*yek*	*qaribe*	*ast*
CONJ	EMPH	PRS-come-3SG	ADJ-EZ	INDEF	NOUN	COP3SGPRS
when	then	comes	like	a	stranger	is

vaqti ham miyād eyn-e yek qaribe ast 'and when he comes, he is like a stranger' (ŠĀ 52)

(4) *yek* as indefinite article:

yek	*performāns*	*tu-ye*	*Piyāno*	*dāšt-e.bāš-ad*
INDEF	NOUN	PREP-EZ	NOUN	have.PST-PTCP.BE.SUBJ-3SG
a	performance	in	Piano	have

yek performāns tu-ye Piyāno dāšte bāšad '[she would] have a performance in Piano' (KP 50)

(5) *ye* + *-i* as indefinite article (final *k* drop) MWP *čiz-i* > MSP *ye čiz-i*

sobh-eš	*ye*	*čiz-i*	*tarğome*	*kard-e-am*
NOUN-EP3SG	INDEF	NOUN-INDEF	NOMVERB	do.PST-PTCP-PERF1SG
that.morning	a	something	translate	I.did

sobheš ye čizi tarğome karde-am 'I have translated something that morning' (KP 56)

5.1.2.2 Particles and Conjunctions

Among the particles and conjunctions, special attention was given to the hyperproductive particle *ke* which serves a variety of functions and is due to this fact actually devoid of precise independent meaning. In the analysis, two most frequent non-MWP uses of *ke* were taken into account: emphatic *ke* and *ke* as a temporal conjunction. Apart from *ke,* two other emphatic enclitics were in-

can be suffixed to the noun as well, replacing the *-e* of the ezāfe construction (for example, *ketāb-e xub* can either become *ketāb-e xub-i* or *ketāb-i xub*), but such a form is considered very literary.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

cluded: 1) typically colloquial particle *-hā*, used in the sentence-final position to emphasize the verb (sometimes subject to *h* omission it takes the form of *-ā*), 2) enclitic emphatic particle *ham* (*-am* when subject to *h* omission).

5.1.2.2.1 Temporal ke

Ke has multiple functions in MSP and is used to express subordination of different kinds. In the analysis, amongst all the various uses, the temporal was selected because of its frequency in the source texts, but it should be remembered that other options are also possible. As Lazard (1957/1992) pointed out, *ke* as a conjunction is almost universal.

> [Particle *ke* is] "by far the most used of the subordinating link-words. Its uses are very diverse […] it serves to introduce relative clauses; as a proper conjunction it is used above all to introduce object (and subject) clauses. But some subordinates of different kinds, final, consecutive, causal, temporal may also, **especially in colloquial language**, be introduced by the conjunction *ke,* ex. *raftam ke un [ān] ketāb-o [rā] bexaram* 'I went in order to buy (lit. that I buy) that book.' In this case the value of the subordinate clause is not explicitly expressed, but it emerges from the context. The conjunction, devoid of meaning, has no other function than to mark subordination. It behaves then, to a certain extent, as **a kind of universal conjunction**, which may be used instead of a conjunction or conjunctional phrase with more precise meaning, whenever there is no risk of ambiguity" (Lazard 1957/1992: 1992:218–219, my emphasis).

The possible applications of *ke* listed by Lazard include: temporal, adverb clauses of purpose (242), consequence (243, also in conditional sentences 254), cause (245), comparison (249). This multiplicity of applications makes Windfuhr go even further and suggest that this particle is in fact devoid of meaning itself, used only for marking the embedding of a subordinate clause:

> One puzzling question in relation to subordination is the particle *ke.* Already Lumsden (1810 2:92) had recognized that *ke* and *če* are particles and not pronouns, an insight which was not heeded in many subsequent grammars. *Ke* is the most used subordinate 'mot-outil' as Lazard (1957: 211–214) puts it […] The apparent multitude of the functions of *ke* is quite easily explained by the observation that *ke* in contemporary Persian is a 'dummy' which is inserted in the surface string in case of a subordinate clause embedding (Windfuhr 1979: 69).

The temporal use of *ke* in MSP is marked by a specific word order in which (similarly to the sentences with emphatic *ke* described below) the subject, nominal predicate or another noun phrase from the subordinate clause is put at the beginning, followed by *ke* and then the rest of the clause (Lazard 1957/1992: 238), i. e. *dar ke bāz šod* 'when the door opened' (ĀM 52); with compound verbs, it is placed between the nominal and verbal part of the verb (Mahmoodi-Bakhtiari 2018: 205).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Examples:

(1) a sentence with temporal subordinate clause, joined by *ke:*

mi-xāst	*farār*	*kon-e*	*ke*	*gir*	*oftād*
IMPERF-want3SG	NOMVERB	do.SUBJ-3SG	TEMPKE	NOMVERB	fall.PST3SG
wanted	escape	do	when	caught	fell

mixāst farār kone ke gir oftād 'when she wanted to escape, she got caught' (DXM 55)

(2) temporal subordinate clause with *ke:*

zang-e	*dar*	*ke*	*be*	*sedā*	*dar*	*āmad*
NOUN-EZ	NOUN	TEMPKE	PREP	NOUN	PREVERB	come.PST3SG
bell	door	when	to	voice	came	

zang-e dar ke be sedā dar āmad 'when the doorbell rang' (ŠĀ 57)

(3) temporal subordinate clause with *ke* and progressive form:

az	*dar*	*ke*	*dāšt*	*birun*	*mi-zad*
PREP	NOUN	TEMPKE	AUXPST3SG	NOUN	IMPERF-hit.3SG
through	door	when	[he] was	out	going

az dar ke dāšt birun mizad 'when he was going out through the door' (KP 52)

(4) temporal subordinate clause with *ke:*

asabāni	*ke*	*šod*
ADJ	TEMPKE	become.PST3SG
angry	when	she.became

asabāni ke šod 'when she got angry' (ĀM 56)

(5) temporal subordinate clause with *ke:*

ba'd	*ke*	*bar.gašt-am*	*be*	*otāq-e*	*xod-am*
ADV	TEMPKE	return.PST1SG	PREP	NOUN-EZ	REFL-EP1SG
later	when	I.returned	to	room	my

ba'd ke bargaštam be otāq-e xodam 'when I returned to my room later' (Hn 53)

5.1.2.2.2 Emphatic ke

Particle *ke* may also be used as a marker of emphasis in MSP. This usage differs from the typical subordinating conjunction described above. It is not a new phenomenon, it was mentioned in the the 19th century sources such as Chodźko (1957) and Haggard and Le Strange (1882) (see below). Lazard mentions the emphatic use of *ke* as a very frequent feature of colloquial Persian, in which "one (or sometimes several) of the noun phrases of a sentence (subject, nominal predicate, direct object, etc.) is found at the beginning and followed by *ke*, after which comes the rest of the clause. This construction usually has the effect of emphasizing the term thus detached at the beginning, or marks an opposition with what precedes" (Lazard 1957/1992: 255). According to Windfuhr, such a construction is an example of topicalization and does not necessarily occur at the beginning of a sentence, neither is it attached to any specifically located noun phrase, but rather it is able to emphasize any of them, and even the entire clause (e. g. *man pul ke nadāram* or *man pul nadāram ke* 'I have no money'). He sees this

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

construction as probably related to the *nominativus pendens* or anacoluthon known already in Old Persian[133] and mentions that the emphatic function of *ke* is not new and had already been observed by Chodźko in 1852 (Windfuhr 1979: 70–71): "The observation that *ke* may function as an emphatic, most prominently in the colloquial near-standard dialect, is not new; thus Chodźko (1852: 141) suggested the translation-meaning to be 'voilà, mais, eh bien.'" That observation is also repeated by Haggard and Le Strange (1882: xxxiv), even if to a limited extent: "*Kih* is used also to emphasize pronouns, etc. *Tú kih namídáni, "Don't* you *know?"*[original spelling and emphasis]."

Windfuhr quotes different grammatical interpretations of this form, including a special (colloquial) type of relative clause (Jensen 1931: 225–28), but he himself identifies it as an "evident case of topicalization, and thus probably related to what has been called *nominativus pendens* or anacoluthon (well-known already in Old Persian).[134] […] Such topicalization has been called the 'preposition of the psychological subject' of a clause (Jensen 1931: 191); cf. also Homā'i (1959: 145)." (Windfuhr 1979: 71). He describes this construction (after Peisikov 1960, Arends 1941 and Rastorguyeva 1953) altogether with another similar one, in which the psychological subject is different from the grammatical subject of the clause, e. g. *amu zan-aš*, where "*amu* is the emphasized subject while *zan-aš* constitutes the referential part of the predicate in the clause, either as subject or in any other case-function" (Windfuhr 1979: 71).

Peisikov groups particle *ke* together with emphatic *hā* (see below), *āxe* and *dige* (from MWP *āxar* 'end' and *digar* 'yet'), *hey* (a synonym of *hamiše* 'always,' but used in the meaning of 'all the time' and never alone) (Peisikov 1960/2002: 130–131).

Examples:

(1) with emphasis on a place (preceding word):

tu	*kāzino*	*ke*	*ne-mi-š-e*
PREP	NOUN	EMPH	NEG-PRS-become.PRS-3SG
in	casino	(emphatic)	can't

tu kāzino ke nemiše 'it can't be done in [a place such as] a casino' (DXM 56)

This kind of emphasis suggests that the action in question could be performed in another place, but it can't be done in a casino (here the action in question was

133 *Nominativus pendens* occurs when the subject is semantically not connected with the sentence (see Skjærvø 2013: 159), "a 'dangling' nominative that is not affected by the syntax of a sentence" (Skjærvø 2016: 91).

134 This case occurs when the subject is semantically not connected with the sentence (see Skjærvø 2013: 159), "a 'dangling' nominative that is not affected by the syntax of a sentence" (Skjærvø 2016: 91).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

namāz, an act of prayer performed on certain times of a day by a pious Muslim, on a prayer rug and facing Mecca).

(2) at the end of a sentence, emphasizing a whole clause (both person and verb; additional emphasis is put by an exclamation mark):

šoma	*mi-dun-in*	*ke!*
PRON2PL	PRS-know.PRS-2PL	EMPH
you	know	(emphatic)

šoma midunin ke! 'you do know that!' (ŠĀ)

(3) used to both emphasize and contrast the 1st person with the addressee of the clause:

man	*ke*	*pā-ye*	*tork-e*	*xāčik-am*
PRON1SG	EMPH	PREP-EZ	NOUN-EZ	NOUN-COP1SGPRS
I	(emphatic)	by	Turkish	Xachik.am.

man ke pā-ye tork-e xāčik-am 'I do prefer Turkish Khachik[135] [coffee]' (KP 51)

Similar to example (1) above, here the emphatic particle *ke* is used to contrast the speaker (preceding word is a personal pronoun) with the person he is talking to. The context of this utterance is a short exchange between the café owner and a visiting woman, where the owner wants to learn about her taste in coffee and gives her information on his own preference. In a way, this kind of usage reminds of the basic function of *ke* as a relative pronoun (i. e., 'me, who likes Turkish Khachik coffee') but the context is different (i. e., 'me, I like Turkish Khachik, and how about you?').

(4) an example of double emphasis with both particle *ke* and particle *ham:*

ahl-e	*mašquliyāt-e*	*fekr-i*	*ham*	*ke*	*nist*
NOUN-EZ	NOUNPL-EZ	NOUN-ADJ	PART	EMPH	not.be.PRS3SG
the.kind	activities	mental	either	"	isn't

ahl-e mašquliyāt-e fekri ham ke nist 'he is not a person who likes mental activities either' (Hn 55)

In this example, *ham* combines the emphatic article function with the adverbial function similar to English 'either.'

5.1.2.2.3 Emphatic hā

This particle is not mentioned by Lazard (1957/1992), neither does it appear in Windfuhr (1979) or Windfuhr and Perry (2009). Peisikov (1960/2002: 130) mentions two forms: *hā* and *ā*,[136] and describes it as a particle used sentence-

135 "Khachik" is a brand of finely ground coffee available in Iran.

136 The resemblance of this particle and the plural suffix *-hā* (MSP *-hā/-ā*) is interesting. The contemporary plural suffix *-hā* has its roots in the very productive Middle Persian abstract suffix *-ihā*, used to form adverbs (from which it then moved to the category of a non-animate plural suffix) (Skjaervo 2009: 206). Despite the similarity, the two forms do not seem related, especially provided that the emphatic particle *hā* is usually used post-verbally. Perhaps it could rather be associated with another phatic particle *hān*, 'aha, hm.'

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

finally as a kind of warning (*andarzi ya hušdāri* 'advice or warning'): *xafet mikonamā* 'watch out – I'll strangle you.' It can also be used emphatically, for example *midānam-hā* 'I do know.' It is a form associated with Tehrani dialect, but it does appear in the source material so its recognition throughout the MSP might be higher than expected. There are 11 instances of emphatic *hā* (DXM: 0, ŠĀ: 4, KP: 2, ĀM: 5, Hn: 0), which is a frequency similar to referential *-e* (also 11 instances).

Examples:

(1) sentence-final emphatic *hā* with a shade of warning:

ğelou-aš	*in*	*harf-o*	*na-zan-i*	*hā*
PREP-EP3SG	DEM	NOUN-DO	NEG-hit.PRS-2SG	EMPH
in.front.of.him/her	these	word	do.not.say	(emphatic)

ğelouaš in harf-o nazani hā 'don't you dare speak those words in front of him/her' (ŠĀ 51)

(2) sentence-final emphatic *hā* with an auxiliary verb omission in perfect tense (but no shade of warning, rather mocking):

mess-e	*inke*	*emruz*	*kār-o-bār-etun*	*sekke*	*bud-e*	*hā*
NOUN-EZ	CONJ	ADV	COMPNOUN-EP2PL	NOUN	BE.PST-PTCP	EMPH
example	that	today	earnings.your	coins	have.been	

messe inke emruz kār-o-bāretun sekke bude hā 'it seems that you have earned a lot today' (KP 58)

(3) emphatic *hā* after a congratulatory expression, adding a hint of irony to the phrase:

mobārak	*hā*
ADJ	EMPH
congratulations	

mobārak hā 'congratulations' (ĀM 53)

In Persian, this congratulatory phrase is customarily used when somebody obtains a new thing (clothing, haircut, car…). The example comes from a dialogue between mother and daughter. Mother gets a new mobile phone as a gift and asks her daughter not to tell anything to the grandmother in order to avoid questions about the man she received the gift from, which is why the daughter replies ironically.

5.1.2.2.4 Emphatic enclitic ham

Enclitic *ham* is yet another emphatic particle used in the MSP. Its use is frequent, in faster delivery and after a consonant it tends to lose the initial *h* and join the preceding word (also in spelling). According to Lazard (1957/1992: 95) it is "an enclitic particle very frequently used; it serves to emphasize the term on which it bears, with a nuance rather like 'also, even.' […] The enclitic *ham* is usually written separately, but it is sometimes joined in writing to the preceding word." Enclitic particle *ham* should not be confused with the copulative conjunction *ham…ham* and the reciprocal *ham* "that has qualities of a pronoun" (Lazard

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

1957/1992: 118; see example 2, where *ham* is used in both functions in one sentence). Lazard mentions also that it may link coordinate clauses: "the link may be established by the enclitic particle *ham* 'also, even, as to,' more particularly when there is a change of subject" (Lazard 1957/1992: 215). Peisikov mentions *ham* as a reciprocal pronoun only (Peisikov 1960/2002: 74), and not a particle at all. Mahmoodi-Bakhtiari (2018) adds three uses of the clitic *ham* to his list of the spoken language characteristics: first, an "Indifferent *ham* Construction" in which *ham* is inserted between the two repetitions of one verb, for example *rafti ham rafti* '[it is not important] if you go' (lit. you went, you went too)[137] (Mahmoodi Bakhtiari 2018: 206); secondly, two forms attributed to semantic alternations, in the meaning of 'even if': either in a phrase *ham ke šode* or *ham* by itself, e.g., *pesar-e ra'is ham bāši, nemituni beri tu* 'even if you are the boss's son, you may not enter there' (Mahmoodi-Bakhtiari 2018: 208).

Examples:

(1) emphasizing a preceding noun phrase, the meaning similar to English 'even':

yek	*kalame*	*ham*	*na-goft*
NUM/INDEF	NOUN	EMPH	NEG-say.PST3SG
a/one	word	even	didn't.say

yek kalame ham nagoft 'didn't say even a word' (ĀM 77)

(2) emphasizing a preceding pronoun, the meaning similar to English 'at all' (the sentence is completed by a negative present compound verb *farq dāštan* 'to differ' in 3rd person plural, which takes an indefinite *-i* after its nominal part because of the negative form):

ma'ni-yaš	*in*	*nist*	*ke*	*hič*	*ham*	*bā*	*ham*
NOUN-EP3SG	DEM	NEG.be3SG	CONJ	PRON	EMPH	PREP	RCP
its.meaning	that	isn't	that	no	even	with	each.other

ma'ni-yaš in nist ke hič ham bā ham [farqi nadārand] 'it doesn't mean that they [don't differ] from each other at all' (KP 61)

5.1.2.3 Verbal endings

Modern Spoken Persian uses a set of inflectional endings different to that of written standard. The differences result from phonetical reductions similar to those described in the 5.1.1 section: a final consonant drop (3sg *-ad* > *-a*; 3pl *-and* > *-an*), than the change of resulting final *-a* to *-e* when applicable. In the 2nd person plural the *-id* ending is changed to *-in* (here the merely phonological background is doubtful, and the form seems to be a result of assimilation to

137 There is a similar construction with *ke*, e.g. *raft ke raft* '[I don't care that] he went' or 'he went [and did not return],' for which Mahmoodi-Bakhtiari (2018: 205) uses the term 'indifference *Ke*-construction after Sadat-Tehrani (2003).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

3rd person plural, so that 2nd pl and 3rd pl endings both have a final *n* in MSP, such as they both have a final *d* in MWP).[138] The resulting set of endings is as follows:

		MSP		MWP	
		present	**past**	**present**	**past**
singular	1	-am	-am	-am	-am
	2	-i	-i	-i	-i
	3	-e	ø	-ad	ø
plural	1	-im	-im	-im	-im
	2	-in	-in	-id	-id
	3	-an	-an	-and	-and

Table 2. Verbal endings in MWP and MSP, the endings that differ between MSP and MWP are highlighted (source: own research).

Interestingly, different to many (though not all) other parameters, verbal endings are not mentioned in Chodźko (1883) or Haggard and Le Strange (1882), where all the verbal forms quoted (or presented as examples and in paradigms) retain the classical MWP set of endings without phonetical reductions whatsoever.

Within the category "Verbal endings" one more form is included, that is, the MSP form of copula *ast* > *e*. Similarly to the 3rd person singular ending *-ad* > *-e*, where the reduction of final consonant leaves a word-final *-a,* which then according to the phonological rules becomes *-e* in Iranian Persian, also here one could suppose that the final consonant cluster is reduced (*ast* > *ass* > *a*) and then the resulting final *-a* is changed to *-e* (unstressed copula becomes a clitic and can be sensed a part of the preceding word). This seems to be a long-existent feature of spoken Persian which, according to Jeremiás (2012), was attested as early as 13th century by Nasir ad-Din Tusi (1201–1274), who would notice in his discussion of a sentence *Hasan dabīr-i* that "in some *ʿajam* (non-Arab, that is, Persian) dialects the *ḥaraka*[139] on the letter *rā* in the word *dabīr* may fulfill the function of the copula. He seems to mean that the copula was expressed as *ast* in formal Persian, but expressed by the *ḥaraka* elsewhere. […] He might have referred to an existing colloquial pronunciation of the word *ast* > *e* (marked with *kasra*)" (Jeremiás 2012: 137). Today the MSP copula *e* is not spelled with *kasre* but (if spelled at all, since it is in principle a spoken form) with the letter ه *he*, the same one that is used to represent a word-final *-e* (Peisikov 1960/2002: 30).

138 Lazard mentions the 3rd person (both singular and plural) endings as resulting from phonetical reductions. As for the 2nd person plural, he only mentions the *-in* ending as "the very much used" in Tehrani Persian (1992: 135–136). Windfuhr (1979) does not discuss inflectional endings at all.

139 Arabic *ḥaraka* (plural *ḥarakāt*) are the diacritics representing short vowels, which are otherwise not represented in script. Kasra is the diacritic representing *e*.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Examples:

(1) MSP 3pl ending with the imperfect form of the compound verb *bāvar kardan* 'to believe':
bāvar *ne-mi-kard-an*
NOMVERB NEG-DUR-do.PST-3PL
believe were.not-making
bāvar nemikardan 'they weren't believing' (DXM 54)

(2) MSP 3sg ending with the present form of the compound verb *šoru' kardan* 'to start':
šoru' *mi-š-e*
NOMVERB DUR-become. PRS-3SG
beginning become
šoru' miše 'it starts' (KP 50)

5.1.2.4 Verbal stems

In Persian verbal system, every verb has two stems, traditionally called present and past, from which tenses are formed: present stem for present tenses and present and active participles and past stem for past tenses, past and passive participles. The past stem of the regular verbs is formed by adding *-id* suffix to the present stem (i. e. *ras-* > *rasid-* 'arrive'); the stems of the irregular verbs cannot be derived from one another (i. e. *sāz-* > *sāxt-* 'build'). The regular/irregular distinction is related to the stems relationship only. Among the irregular verbs, there are small groups with irregular formants or those where an underlying morphophonemic change is still visible. The most irregular verbs are those most frequent like *rav-* > *raft-* 'go,' *kon-* > *kard-* 'do,' *gu(y)-* > *goft-* 'say,' *bin-* > *did-* 'see,' *šav-* > *šod-* 'become,' *dār-* > *dāšt-* 'have' (Windfuhr and Perry 2009: 447).

Some of the most frequent verbs (which often have irregular present stem) differ in their present stem between MWP and MSP. While there is usually a phonological change underlying the difference, those MSP stems are so well grounded in language that their origin probably goes unnoticed. Lazard (1957/1992: 143) mentions the verbs *xāstan/xāh-*[140] 'to want,' *āmadan/ā(y)-* 'to come,' *raftan/rav-* 'to go,' *šodan/šav-* 'to become,' *dādan/deh-* 'to give,' *goftan/gu(y)-* 'to say,' to be reduced by contraction in their present stem forms[141] and indeed, some of them are contracted to the extreme (single consonant); their present stems are presented in Table 3 below.

140 Lazard's (1957/1992) way of giving verbal forms is different from that of Windfuhr and Perry (2009); instead of present/past stem he gives the infinitive/present stem. Past stem can be easily derived from the infinitive by removing the suffix *-an*, but the order of forms may be confusing.

141 He also explains contracted verbal stems of verb such as *raftan (rav-)* 'to go,' *dādan (deh-)* 'to give' or *xāstan (xāh-)* 'to want' in terms of final *-h*, *-y* or *-v* drop (thus *miravam* > *miram*, *mideham* > *midam*, *mixāham* > *mixām*) (Lazard 1992: 35).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

infinitive	meaning	past stem	MWP		MSP	
			present stem	1 sg praes	present stem	1 sg praes
xāstan	'to want'	*xāst-*	*xāh-*	*mi**xāh**am*	*xā(y)-*	*mi**x**ām*
āmadan	'to come'	*āmad-*	*ā(y)-*	*mi(y)**āy**ām*	*ā-*	*mi(y)**ā**m*
raftan	'to go'	*raft-*	*rav-*	*mi**rav**am*	*r-*	*mi**r**am*
šodan	'to become'	*šod-*	*šav-*	*miš**av**am*	*š-*	*miš**a**m*
dādan	'to give'	*dād-*	*deh-*	*mi**deh**am*	*d-*	*mi**d**am*
goftan	'to say'	*goft-*	*gu(y)-*	*mi**guy**am*	*g-*	*mi**g**am*

Table 3. Reduction in MSP verbal stems, part 1 (based on Lazard 1957/1992).

Another group are the verbs with longer present stems, in which there is a reduction of a vowel (and, consequently, a syllable), sometimes accompanied by a reduction of adjacent consonant. In some of those verbs, a short (i. e., unstable) vowel of the first syllable tends to be reduced (see Table 3 above). This change, however, is not visible is the script since the short vowels are not represented. While there is a possibility of representing them with a *haraka* called *sokun* (used when there is no vowel at all) to indicate that it is a contracted form, this hardly happens in literary texts. It is different when the stem begins with a vowel: with bound spelling of *mi-* (a present tense prefix; its bound or unbound spelling is optional), the opening *alef* representing that vowel can be omitted and *mi-* joined straight to the adjacent consonant, which would suggest to the reader that the vowel is omitted, although not in an absolutely definite way.

infinitive	meaning	past stem	MWP		MSP	
			present stem	1 sg praes	present stem	1 sg praes
gozāštan	'to put; to let'	*(go)zāšt-*[142]	*gozār-*	*mi**gozār**am*	*zār-*	*mi**zār**am*
āvordan	'to bring'	*āvord-*	*āvar-*	*mi(y)**āvar**am*	*ār-*	*mi(y)**ār**am*
nešastan	'to sit'	*n(e)šast-*	*nešin-*	*mine**šin**am*	*šin-*	*miš**in**am*
šekastan	'to break'	*šekast-*	*šekan-*	*miš**ekan**am*	*škan-*	*miš**kan**am*
šenidan[143]	'to hear'	*šenid-*	*šenav-*	*miš**enav**am*	*šnav-/ šnou-*	*miš**nav**am/ miš**nou**am*
sepordan	'to deposit'	*sepord-*	*sepār-*	*mi**sepār**am*	*spār-*	*mi**spār**am*
andāxtan	'to throw'	*andāxt-*	*andāz-*	*mi(y)**andāz**am*	*ndāz-*	*mi**ndāz**am*

142 This verb can have its stem reduced also in the past: *gozāšt-* (MWP) > *zāšt-* (MSP): 1 sg praet *migozāštam* vs *mizāštam.*

143 *Šenidan/šenav-* 'to hear' has an irregular, colloquial form *šenoftan* (Lazard 1992: 135), of which the present stem remains unchanged.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(Continued)

infinitive	meaning	past stem	MWP		MSP	
			present stem	1 sg praes	present stem	1 sg praes
foruxtan	'to sell'	*foruxt-*	*foruš-*	*mi**forušam***	*fruš-*	*mi**frušam***

Table 4. Reduction in MSP verbal stems, part 2 (source: own research, based on data from Lazard 1957/1992).

In the analysis, only the forms from the first group, where the reduction is undoubtedly visible in print, were tagged as 'MSP stem.'

Examples:

(1) contracted, MSP stem of the verb *āvardan* 'to bring' (MWP: āvar-, MSP: ār-); in this example, present tense is used to express the future, and an interesting feature is the lack of postposition *-rā* after *Parviz* to mark its function as a definite direct object:
Parviz ham mi(y)-ār-am
NOUN PART DUR-bring.PRSMSP-1SG
Parviz too I.bring
Parviz ham miyāram 'I will bring Parviz too' (ŠĀ 77)

(2) contracted stem of the verb *gozāštan* 'to put' (MWP: *gozār-*, MSP: *zār-*; interestingly, spelled with the letter ږ instead of regular ږ, which might be a typographic mistake, though, since the same verb is spelled properly on the same page just two lines above); it is also worth noticing that in this example there are two phonological alternations typical for MSP: *hič-kas* 'nobody' become *hiš-ki* and *dast* 'hand' as a nominal part of the compound verb *dast zadan* 'to touch' is contracted to *das* with final consonant drop (it is probably *dass* with compensatory doubling of the final *s*, however it is not indicated in spelling), also the verbal ending in 3sg praes is changed from MWP -ad to MSP -e:

xod-am	*mi-zār-am*	*hiš-ki*	*das*	*na-zan-e*	*lotfan*
REFL-1SG	DUR-put.PRS-1SG	PRON	NOMVERB	NEG-HIT.PRS-3SG	ADV
I.myself	put	nobody	hand	does.not.touch	please

xodam mizāram, hiški das nazane lotfan 'I'll put it [there] myself; no one touches it please' (KP 66)

5.1.2.5 Tense

Under this heading, two periphrastic verbal forms are included: the present and past progressive with auxiliary *dāštan*. Lazard (1957/1992) describes them altogether as a progressive paraphrase: "In order to indicate that an action is being accomplished, that it is still progressing, colloquial language frequently uses a periphrase constituted with the help of the verb *dāštan/dār-* 'to have, hold,' followed by a personal form (in the indicative) of the main verb. The two verbal forms may immediately follow each other or may be separated by one or more words. For action in progress at the present moment, the two verbs are in the

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

present. […] In the past, the verb *dāštan* is in the preterite, the main verb in the imperfect. […] In the 'completed past,' the verb *dāštan* is in the perfect, the significant verb in the compound imperfect." (Lazard 1957/1992:161). Both of them appear also in Windfuhr (1979: 102–103; among aspectual auxiliaries).

The tenses with auxiliary *dāštan* are described by Peisikov as forms which were first frequent in spoken idiom, but recently (the book was published in 1960s, so that means some 1950s) have entered literary language and are now common in literature, press, political and scientific writings (Peisikov 1960/2002: 104; section 93).

One interesting feature of the progressive forms, which is not mentioned in those three sources, is its lack of negative form. The fact that a regular negation is lacking might indicate that the form has not been fully grammaticalized as a tense yet.[144,145]

Examples:

(1) past progressive of the compound verb *touzih dādan* 'to explain'—it is possible to put additional modifiers between the auxiliary (*dāšt*) and the imperfect form of the verb (*touzih midād*), as exemplified by the adverb *bivaqfe* 'continuously':

dāšt	*bi-vaqfe*	*touzih*	*mi-dād*
have.PST3SG	ADV	NOMVERB	DUR-give.PST3SG
(auxiliary)	continuously	explanation	was.giving

dāšt bi-vaqfe touzih midād 'she was explaining continuously' (KP 50)

(2) present progressive of the compound verb *šuxi kardan* 'to joke':

dār-am	*šuxi*	*mi-kon-am*
have.PRS-1SG	NOMVERB	DUR-do-1SG
(auxiliary)	joke	I.am.making

dāram šuxi mikonam 'I am joking' (ŠĀ 54)

5.1.2.6 Two-verb constructions

In the analyzed material, the constructions of more than one verb were found that could not be satisfactorily described neither in terms of tenses not clause coordination or subordination, though the verbs in question were not considered regular auxiliaries used in composing other verbal forms. This kind of phenomenon is discussed in Windfuhr (1979) as "aspectual auxiliaries," the forms

144 However, as Taleghani (2008: 134) observes, the lack of direct negation in those forms (which she describes as Aspectual Complex Predicates) might be related to semantic factors (as no morpho-syntactic factor is involved).

145 It is possible that the grammaticalization of negative present progressive form has begun, as Nematollahi (2018) pointed that a form *dāram nafas* ***ne****mikešam* 'I am not breathing' appears in lyrics of a pop song by Iranian singer Benyamin Bahadori *Ye bār-e dige eštebāh kon* (published in 2016 and available in streaming services such as YouTube, Aparat or Spotify: https://open.spotify.com/album/5v9GxbtT0F1Moxg4lAE0df).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

which are treated as marginal and little researched in his opinion, especially when the colloquial forms are concerned: "This is especially true for the 'marginal' colloquial auxiliaries such as *gereftam hesābi xordam* 'I took a good bite' as well as for the 'marginal' functions of the better known auxiliaries such as the inchoative use of *xāh* (e. g. *mixād šoru' beše* 'is about to begin') […],[146] the 'potential' use of *šav* (e. g. *dide nemiše* 'cannot be seen')" (Windfuhr 1979: 102; contrary to Windfuhr, I have decided to include the progressive forms with *dāštan* in the morphology > tense category).

Windfuhr (1979) mentions a few aspectual auxiliaries in his book: *gereftan* 'to take' perfective inchoative (idem: 103–104), *xāstan* 'to want' (idem: 104–105), *šodan* 'to become' (idem: 105–106). To that list, *raftan* as inchoative could possibly be added (see example (4) below).

Differentiating between a real two-verb construction with aspectual (or other) auxiliary and verbs denoting a chain of events as coordinate clauses without conjunction might be difficult at times, yet in the source material two constructions not mentioned in Lazard (1957/1992) or Windfuhr (1979) were found which are clearly distinguishable: one with auxiliary *dādan* 'to give' and the other with auxiliary *kardan* 'to do' which are syntactically distinct from the regular usage of those verbs. They seem to be relatively rare: only a few instances of the construction with *dādan* were found in the selected parts of *Kāfe Piyāno*, *Dar Xarābat-e Moqān* and *Šahr-Āšub*.[147] The two verb construction with *dādan* as auxiliary seems to carry a causative meaning; almost all the examples have a similar context of ordering something to be done (poster design, clothwash, construction of a cabin); there is one in which *dādan* is used in the meaning of giving permission (see example (3)). The construction with *kardan*, on the other hand, seems to be associated with those situations, in which the subject refuses to act in order to perform the action denoted by the second verb (cleaning, seeing the unwanted guests off, surrendering to good manners).

146 I have excluded from this list "the 'obligatory' use of *dār* (e. g. *bā ke dāram harf bezanam* 'whom have I (here) to talk with' or 'with whom do I have to talk)." The commentary to this form on the adjacent page cites the example in a different way: *bā ke harf dāram bezanam,* which is a quote from a play by Āxundov, cited by Christensen in 1934 as an example of progressive subjunctive, repeated in Lazard (1957) without reference and corrected by Yarshater in his review of the Lazard's grammar to be a regular use of the verb *dāštan* 'to have.' The change in word order might be intentional or not and with the lack of other examples, the existence of this form remains doubtful to me.

147 The small number of those constructions can be of course explained by the small size of analyzed corpus. The fact that they are present (even if scarce) in three out of the five books suggests that this is not an individual feature of one author's style. Those constructions definitely require a more in-depth research.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Examples:

(1) *dādan* as auxiliary with a meaning similar to English causative *have sth done:*

be-deh-am	*tarrāhi*	*kon-and*
SUBJ-give.PRS-1SG	NOMVERB	make.PRS-SUBJ3PL
I.would.give	design	[they]make

bedeham tarrāhi konand 'I'll have [it] designed [lit. I would give they would design]' (KP 52)

(2) *dādan* as auxiliary, again in the causative function:

dād-am	*ye*	*ālunak-e*	*kučulu*	*barā-m*	*dorost*	*kard-an*
give.PST-1SG	INDEF	NOUN-EZ	ADJ	PREP-EP1SG	NOMVERB	make.PST-3PL
I.gave	a/one	cabin	little	for.me	made	

dādam ye ālunak-e kuchulu barām dorost kardan 'I had a small cabin built for me; I ordered a little cabin [lit. I gave a little cabin for me they built]' (DXM 56)

(3) *dādan* as auxiliary, but with a different meaning (giving permission):

ne-mi-deh-am	*pip-am-ro*	*baqiye*	*be-gozār-and*	*guše-ye*	*lab-ešun*
NEG-give.PRS-1SG	NOUN-EP1SG-DO	NOUN	SUBJ-put.PRS-3PL	NOUN-EZ	NOUN-EP3PL
I.do.not.give	my.pipe	others	little	corner	their.lips

nemideham pipam-rā baqiye begozārand guše-ye labešān 'I don't let others pu my pipe in the corner of their lips [lit. I don't give others put my pipe in the corner of their lips]' (KP 59)

(4) inchoative *raftan* denoting (and emphasizing) the beginning of an action or process:

raft	*yahudi*	*šod*
go.PST3SG	ADJ	become.PST3SG
a/one	word	even

raft yahudi šod 'he became a Jew [i. e. converted to Judaism; lit. went and became a Jew]' (DXM 57)

(5) *kardan* as auxiliary with inverted word order in the subordinate clause, denoting a situation in which the subject did not make any effort towards the action denoted by the second verb:

na-kard	*lā-aqal*	*tā*	*dam-e*	*dar*	*bi-yā-d*	*badraqe-mun*
NEG-make.PST3SG	ADV	PREP	NOUN-EZ	NOUN	SUBJ-come.PRS-3SG	NOUN-EP1PL
not.did	at.least	to	threshold	door	come	to.see.us.

nakard lā-aqal tā dam-e dar biyād badraqemun 'he didn't even come to see us off at the threshold [lit. didn't make it to come]' (ŠĀ 70)

(6) again *kardan* as auxiliary with a similar meaning intended, lack of any effort towards the action denoted by the second verb:

na-kard	*tamiz-eš*	*kon-ad*
NEG-make.PST3SG	ADJ-EP3SG	make.PRSSUBJ-3SG
not.did	clean.it	make

nakard tamiz-eš konad 'he didn't clean it [lit. didn't make it to to clean it]' (KP 55)

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

5.1.2.7 Enclitic pronouns

The possibility of replacing the regular, full-form personal pronouns with their enclitic, suffixed counterparts is one of the characteristic features of Modern Persian. Those enclitic personal pronouns, Persian *zamir-e šaxsi-ye peyvaste* (Xānlari 1970: 199), *zamirhā-ye pey-časb* (Peisikov 1960/2: 74) are also known as personal suffixes (Lazard 1957/1992) or personal enclitics (Mahmoodi-Bakhtiari 2018) and they are not new to the grammar of Persian. Jeremiás (1984) points out that "their use, function and position in preclassical texts (10^{th}–11^{th} c.) are the same as those in today's informal variety. They continued to be used in the classical poetry, while they were less frequently applied in prose texts" (Jeremiás 1984: 284).

The use of enclitic pronouns does not mark spoken register by itself, as they are used both in MSP and MWP. Their usage in MWP is more restricted though (mainly as a possessive pronoun, less frequently as a direct or indirect object). A shift from written to spoken register correlates with their growing frequency, enough to make Lazard state that "the frequency of the suffix *-aš/-eš/-š* is one of the most characteristic features of the colloquial language" (Lazard 1957/1992: 113). Those enclitic pronouns can serve a variety of syntactic functions depending on what words are they affixed to (they do not have a strictly defined position): possessive pronoun, direct or indirect object, and even subject. Of those functions, the possessive/partitive function as typical of MWP and MSP likewise, was not included in the first stage of the analysis. So **it is not their use but their frequency that is a marked linguistic choice.** Also, they are so frequent in comparison to the other functions that their number distorts the interpretation of the analysis results, not only on the level of morphological parameters, but also in the general distribution of the grammatical categories. As it is visible in the chart below, with possessive function included, enclitic pronouns constitute almost 50% of all morphological parameters (66% of which and 31% of all morphological parameters are the possessive EPs).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

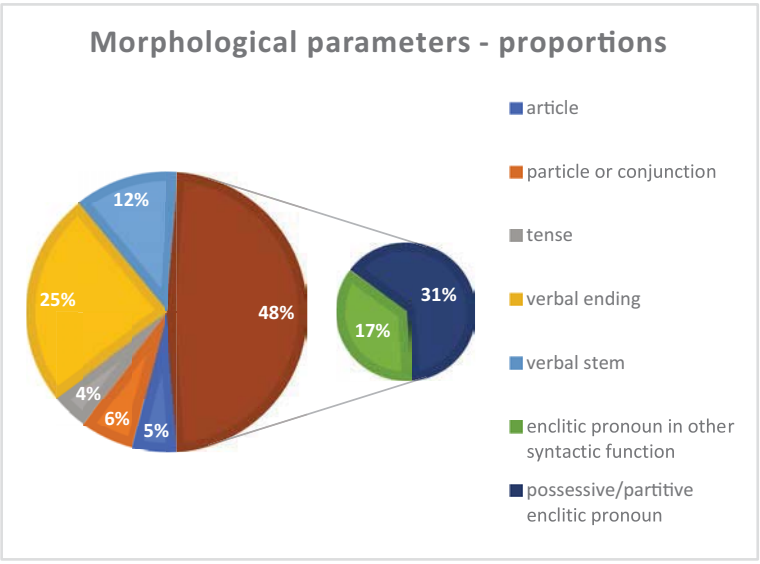

Figure 9. Morphological parameters with possessive/partitive EPs (source: own research).

Those proportions change significantly once we exclude the possessive/partitive EPs:

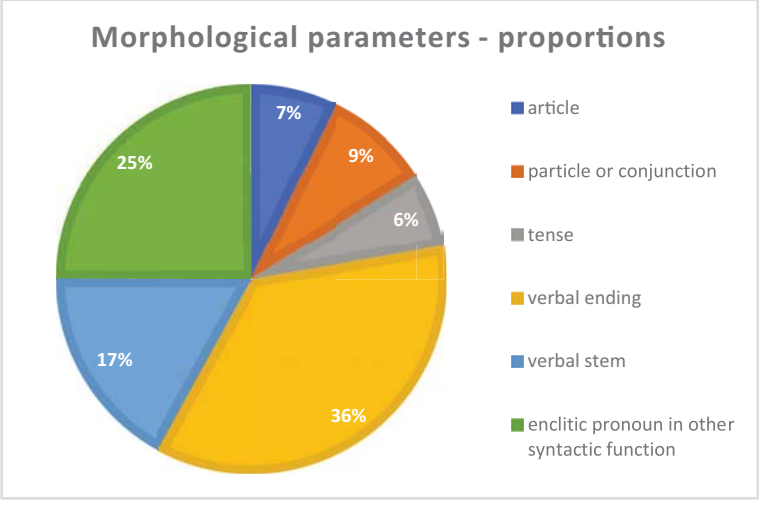

Figure 10. Morphological parameters without possessive/partitive EPs (source: own research).

Such a difference is also, naturally, reflected in the proportion of morphological parameters among the other grammatical categories.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

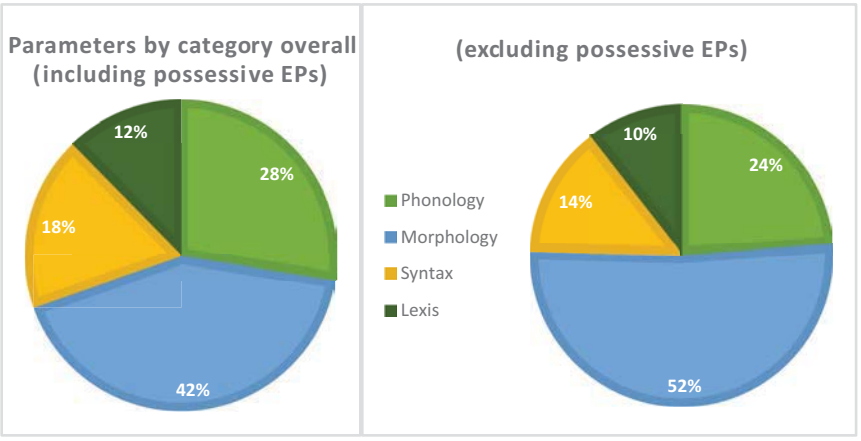

Figure 11. Parameters by category overall: 1) excluding and 2) including possessive/partitive EPs (source: own research).

Including possessive/partitive pronouns, which are not an inherent MSP marker, might result in a distorted interpretation of the proportions, giving morphological features a significance 10% bigger than it actually is.

Considering all arguments for and against, it seems reasonable **not to include** possessive/partitive pronouns in the discussion of the general distribution of forms, but rather treat them as an **additional marker** and compare their frequency in the analyzed books without affecting the general results.

As for the other syntactic functions of enclitic pronouns, the direct/indirect object function, was further split into subcategories according to what the suffix is attached to: a verb, a nominal/preverbial part of a compound verb, a preposition. These subcategories are described in the subsequent sections along with corresponding examples.

Enclitic pronouns undergo the phonological processes typical for MSP such as the *āN* > *uN* alternation in plural forms and *a* > *e* alternation in singular (which is in fact invisible in print since none of those two vowels is marked in writing). After a vowel, the literary forms are joined to the word by a glide *y*, while the MSP forms are reduced to a single consonant and attached to the final vowel itself (cf. Peisikov 1960/2002: 75). Those changes, however, are understood and counted as phonological and not morphological parameters (morphological parameter is the occurrence of the suffixed pronoun in specified position regardless of its phonological form).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

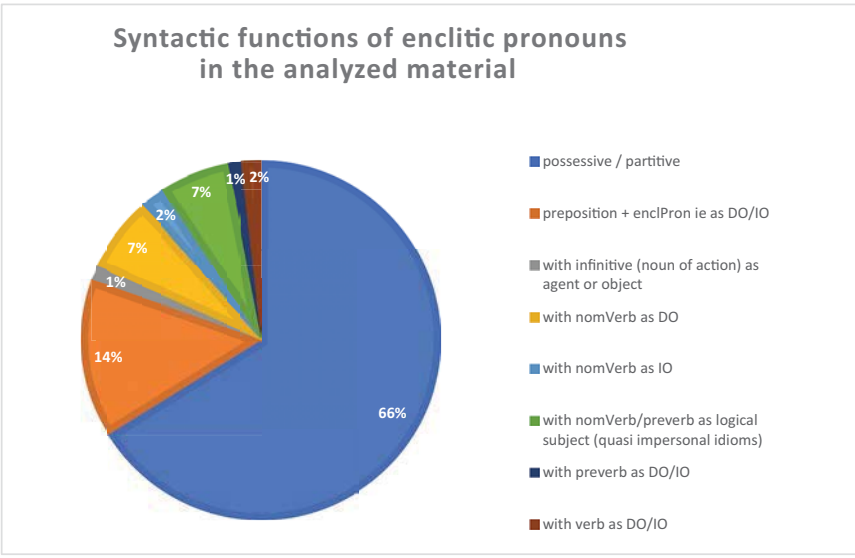

Figure 12. Syntactic functions of enclitic pronouns in the analyzed material (source: own research).

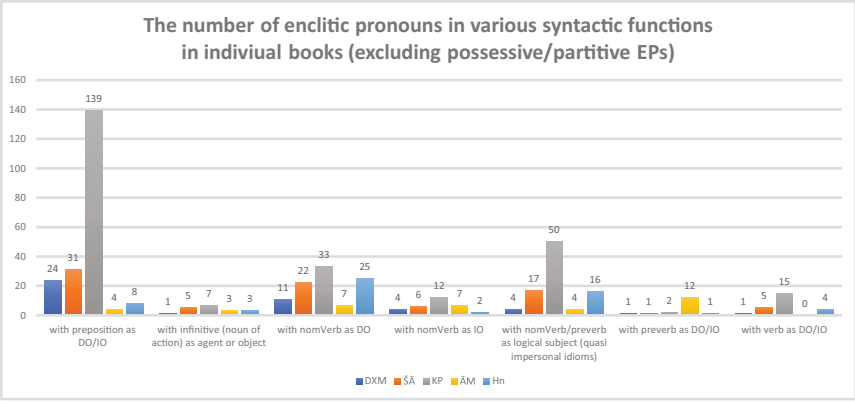

Figure 13. The number of enclitic pronouns in various syntactic functions in individual books (excluding possessive/partitive EPs) (source: own research).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

5.1.2.7.1 Possessive, partitive

The basic and most important function of the enclitic pronouns is the expression of possession. In the qualitative analysis, such a usage (possessive and partitive) was not elicited from the analyzed texts nor listed, since their regular use is by no means a deviation from MWP. During the quantitative analysis, however, it showed that their frequency does coincide with an inclination of the analyzed passage towards MSP. This is in line with an observation recorded by Lazard:

> Personal suffixes are much used in colloquial language. They are much less used in literary language which, generally speaking (except in poetry), tends to avoid them and to prefer the personal independent pronouns: they are hardly ever found except in the uses indicated below under numbers 1 [joined to a noun (or an infinitive), a pronoun, a numerical phrase, the suffix is equal to the construction ezāfe + personal pronoun – with all the values of this construction] and 3 [affixed to the personal forms of the verb, assuming the function of direct object, the suffixes are equivalent to personal pronouns followed by *rā*] (Lazard 1957/1992:109–110)

According to Xānlari, eclitic pronoun adjoined to verb will have the function of its object (*maf'ul*), and when adjoined to noun, it can be either a complement (*motammem*) or express possession (*mozāfun-aleyh*). Other syntactic functions are not mentioned (Xānlari 1970: 199).

The genitive relation expressed by enclitic pronouns is a hallmark of Persian language to the extent in which giving any examples might seem redundant since they are obviously present in all of the analyzed books. Their frequency is not uniform, though. From the analyzed books, *Kāfe Piyāno*, is a particularly rich source of this kind of forms, which seems to be a conscious choice of the author, helping to create a feeling of relaxed, loose style of a familiar conversation. It is well reflected in numbers: 38% of all the occurences of possessive/partitive enclitic pronouns come from *Kāfe Piyāno* (357 out of 951; to compare, there are 246 such forms in *Šahr-Āšub*, 178 in *Hamnavāyi*, but less than a hundred in both *Dar Xarābāt-e Moqān* and *Ādat mikonim*).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

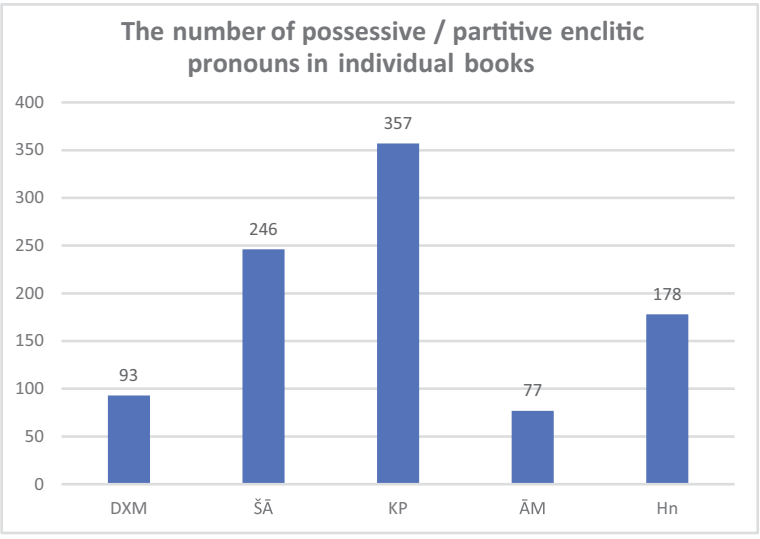

Figure 14. The number of possessive/partitive enclitic pronouns in individual books (source: own research).

Examples:

(1) possessive enclitic pronoun in a clause with reversed word order and omitted auxiliary verb in perfect tense:

mesl-e	*hamiše*	*fešār-et*	*āmad-e*	*pāyin*
NOUN-EZ	ADV	NOUN-EP2SG	come.PST-PTCP	ADV
as	always	pressure.your	has.come	low

mesl-e hamiše fešāret āmade pāyin 'as always, your blood pressure has fallen' (ĀM 59)

(2) possessive enclitic pronoun added to a noun (body part):

čašm-ān-aš	*bāz*	*ham*	*qarq-e*	*ašk*	*mi-šod*
NOUN-PL-EP3SG	ADV	EMPH	ADJ-EZ	NOUN	DUR-become.PST3SG
eyes.her	again	(emph)	full.of	tear	were.becoming

čašmānaš bāz ham qarq-e ašk mišod 'her eyes were filling up with tears all over again' (ŠĀ 79)

(3) partitive enclitic pronoun joined to *hame* 'all':

hame-aš-rā	*be*	*sarnevešt-e*	*nekbat-emān*	*nesbat*	*mi-deh-im*
PRON-EP3SG-DO	PREP	NOUN-EZ	ADJ-EP1PL	NOMVERB	DUR-give.PRS-1PL
all.of.that	to	fate	wretched.our	connection	we.give

hame-aš-rā ham be sarnevešt-e nekbatemān nesbat midehim 'for all of that, we blame our wretched fate' (KP 55)

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(4) partitive enclitic pronoun added to *yek-i* 'one':

yek-i-aš-rā	*ham-in*	*diruz*	*doxtarak*	*zad*	*be*	*goldān*
NUM-INDEF-EP3SG-DO	EMPH-DEM	ADV	NOUN	hit.PST3SG	PREP	NOUN
one.of	that	yesterday	girl	she.hit	to	vase

yeki-aš-ro hamin diruz doxtarak […] zad be goldān 'a little girl hit one of those against a vase just yesterday' (KP 63)

5.1.2.7.2 With preposition, as direct or indirect object

Personal suffixes can be affixed to prepositions (or prepositional phrases with ezāfe) to express direct or indirect object. This usage is typical for (but not restricted to) MSP register nowadays, though in the observations of Jeremiás (1984: 284) and Lazard (1992: 110) it was acceptable only in colloquial language. It is also the second most popular application of the enclitic pronouns in the analyzed material: while possessive/partitive EPs constitute 66% of all of their uses, EPs with prepositions make 14%; the third group are EPs joined to the nominal part of the compound verbs, which make about another 14% (7% as DO and 7% as logical subject); and the rest is evenly distributed among other categories. In case of the prepositional phrases with ezafe, the ezafe is often dropped and enclitic pronoun affixed to the preposition itself. The prepositions *bā* and *be* might take epenthetic *h* to avoid hiatus, thus becoming *bā-h-ā*-enclitic pronoun and *be-h-e*-enclitic pronoun. In this case, the unstable vowel of the enclitic pronoun itself becomes subject to vocal harmony and is changed to *ā* in the first example and *e* in the second:

bā + *-at* > *bā-h-aš* > *bāhāt* 'with you'

be + *-at* > *be-h-at* > *behet* 'to you'

The might also be joined without the epenthetic *h*, in which case the enclitic pronoun's vowel is reduced:

bāhāt > *bāt* 'with you'

behet > *bet* 'to you'

Both cases are well attested in the analyzed material.

The prepositional phrase *barā-ye* loses its ezāfe and the enclitic pronoun is joined to the preposition itself, as indicated above, in two possible patterns: either with a glide *y* to avoid hiatus, or with a reduction of enclitic pronoun's vowel:

(1) *barā-y-* + enclitic pronoun: *barā-y-at* 'for you'

or (2) *barā-* + enclitic pronoun with vowel reduction: *barā-t* 'for you'

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

and the MSP preposition *vāse* 'for' that is used to replace *barā-ye*, when combined with enclitic pronouns, loses its final *-e* and is joined to the EP itself:

vāse + enclitic pronoun: *vāse* + *-at* > *vās-at* 'for you'

A complete paradigm for the prepositions *be* 'to,' *barā-ye* and *vase* 'for' is given in Peisikov (1960/2002: 77).

Examples:

(1) with preposition *bā* 'with' (with epenthetic *hā*):

xod-am	*bā-hā-š*	*harf*	*mi-zan-am*
REFL-EP1SG	PREP-EP3SG	NOMVERB	PRS-hit.PRS-1SG
myself	with.him	word	hit[148]

xodam bāhāš harf mizanam 'I'll talk to him myself' (ĀM 57)

(2) with preposition *be* 'to' (with epenthetic *h*):

faqat	*be-h-em*	*be-gu*
ADV	PREP-EP1SG	SUBJ-tell.IMP2SG
just	to.me	tell

faqat behem begu 'just tell me' (ŠĀ 74)

(3) with preposition *barā-ye* 'for' (with vowel reduction):

nazar-et	*barā-m*	*mohemm-e*
PREP	PREP-EP1SG	NOMVERB-COP3SGPRS
opinion.your	for.me	important.is

nazaret barām mohemme 'your opinion is important to me' (ŠĀ 74)

5.1.2.7.3 With infinitive (noun of action) as agent or object

Enclitic pronouns can be affixed to different nouns, also nouns of action (which are, in fact, Persian infinitives). In that case, the enclitic defines the agent (intransitive verbs) or the object (transitive verbs), such as a regular personal pronoun adjoined to the verb in an ezafe construction would do: *raftan-eš* 'his/her leaving' but *didan-eš* 'seeing him/her.' According to Lazard, "joined to a noun (or an infinitive), a pronoun, a numerical phrase, the suffix is equal to the construction ezāfe + personal pronoun – with all the values of this construction" (Lazard 1957/1992: 110).

Examples:

(1) enclitic pronoun added to the infinitive *raftan* 'to go/going' to indicate the subject of that action:

az	*Āye*	*va*	*Farānse*	*raftan-eš*	*harf*	*be-zan-im*
PREP	NOUN	CONJ	NOUN	go.INF-EP3SG	NOMVERB	SUBJ-hit.PRS-1PL
of	Āye	and	France	going.her	words	let's.hit

az Āye va Farānse raftaneš harf bezanim 'let's talk about Āye and her going to France' (ĀM 77)

148 Compound verb *harf zadan* 'to talk' literally means 'to hit words.'

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(2) enclitic pronoun (again 3rd person singular) added to the infinitive *budan* 'to be/being' to indicate the subject of a state:

Zimens	*budan-eš*	*ma'lum*	*bāš-ad*
NOUN	be.INF-EP3SG	NOMVERB	be.PRSSUBJ-3SG
Siemens	being.it	obvious	would.be

Zimens budan-eš ma'lum bāš-ad '[so that] it being a Siemens [appliance] is obvious' (KP 63)

(3) enclitic pronoun (1st person singular) added to the infinitive *didan* 'to see' to denote the object of the action:

engār	*entezār-e*	*didan-am-ro*	*na-dāšt-i*
ADV	NOUN-EZ	see.INF-EP1SG	NEG-have.PST-2SG
seemingly	expectation	seeing.me	you.didn't.have

engār entezār-e didan-am-ro nadāšti 'it seems you didn't expect to see me' (ŠĀ 74)

5.1.2.7.4 With preverb or nominal part of a compound verb as psychological subject (quasi-impersonal idioms)

One of the frequent constructions with enclitic pronouns in contemporary Persian are the so-called impersonal idioms (also impersonal phrases, semi-personal phrases, quasi-impersonal idioms[149] or quasi-impersonal structures). In those phrases, the verb (obligatorily a compound one) is in 3rd person singular, and the enclitic pronoun that is affixed to the nominal part of the verb indicates who is affected by the action or state. "In impersonal or semi-personal phrases, the suffix, usually joined to the non-verbal element, represents the interested person (the one found in the condition or submitting to the event indicated by the phrase) ex. *xoš-am miāyad* 'I like it' (lit. it pleases me)" (Lazard 1957/1992: 111). Windfuhr calls them indirect verbs and quotes the interpretation of Xānlari (from 1970s) that "the surface object of those constructions is in fact subject (*nehād*), e. g. classical *ma-rā xoš āmad* = modern *xoš-am āmad* 'it came me well' – 'I liked it.' Here is the expanding-construction type of contemporary Persian, the underlying subject is surfaced as dative (not as subject) which obligatorily has to be a personal suffix" (Windfuhr 1979: 51). In Peisikov (1960/2002: 79–80) those phrases are called *ebārathā-ye qeyr-e šaxsi* 'impersonal phrases,' but in fact the term 'semi-impersonal' suggested by Lazard (or 'quasi-impersonal' as in Mace 2003 and Yousef 2018) is more accurate, since the person is always indicated by the enclitic pronoun.

In the definition by Windfuhr, the underlying subject surfaces as dative, replacing the classical dative construction with *-rā* (see above), but it is does not

149 The term *quasi-impersonal idioms* can be found in a relatively recent grammar of contemporary Persian by Saeed Yousef (2018), where they are defined as a construction which can hardly be called *impersonal* since the person is included, only not at its usual place (i. e. the conjugational ending of the verb) but attached as a possessive or dative suffix to the 'grammatical subject' of the verb. A list of such expressions is given (Yousef 2018: 309).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

show the complete picture. It is true for some of those constructions, but not for all—in some of them, the relation between the underlying and surface subject is of genitive nature. Compare: *sard-am ast* < *ma-rā sard ast* 'I am cold (lit. it is cold to me)' or *xāb-am bord* < *ma-rā xob bord* 'I fell asleep (lit. sleep took me)' with *bāvar-am mišavad* 'I believe (lit. it becomes my belief), *del-am mixāhad* 'I want (lit. my heart wants), *sar-am mišavad* 'I understand (lit. it gets to my head). Still, the logical subject of all those phrases is 1st person singular *man.*

Examples:

(1) with MSP verb *qeyb-EP zadan* 'to disappear':
qeyb-eš *zad-e.bud*
NOUN hit.PST-PTCP.be.PST3SG
absence.him had.hit
qeyb-eš zade bud 's/he had disappeared' (Hn 55)

The literal meaning is similar to 'be hit by absence' with absence as the subject and the disappearing object represented by the attached enclitic pronoun.

(2) with verb *bāvar-EP šodan* 'to believe':
hanuz-am *bāvar-am* *ne-mi-š-e*
ADV-EMPH NOMVERB-EP3SG NEG-DUR-become.PRS-3SG
still believe.me doesn't.come
hanuzam bāvaram nemiše 'I still can't believe' (ŠĀ 50)

The literal meaning would be similar to 'it doesn't come to my belief.' This case is slightly different because of the main verb used. In the previous example, *zadan* 'to hit' was a typical transitive verb of action requiring an argument (object). Here, *šodan* is an intransitive verb denoting state, i. e., it does not take a direct object. Thus the psychological subject cannot surface as an object of the sentence and has to take another function: indirect object.

(3) with verb *sar-EP šodan* 'to understand':
enqadr-hā *sar-am* *mi-šav-ad*
ADV-PL NOUN-EP1SG DUR-become.PRS-3SG
this.much head.my becomes
enqadrhā saram mišavad 'this much I understand' (KP 50)

In this example, again the intransitive *šodan* 'to become' is used, thus the underlying subject surfaces in a possessive construction with the nominal part of the verb (idiom).

5.1.2.7.5 With preverb or nominal part of a compound verb as direct or indirect object
In the data count, the preverb and nominal part of a compound verb are treated separately to see which of those attract suffixes more, but mechanism is the same. The suffix can be added both to the preverb/nominal part or to the verb itself: "In

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

the case of a verb with a preverb, the suffix may be affixed to the preverb or to the verbal form, ex. *bar-eš dār/bar dār-eš* 'lift it' [=ān-rā/u-rā bar dār]. Similarly in the case of a compound verb, the suffix may be affixed to the nominal element or to the verbal form, ex. *dust-et dāram/ dust dāram-et* 'I love you' [=to-rā dust dāram]" (Lazard 1957/1992:111) but in the data count, only the forms with preverb + enclitic pronoun and nominal part of the verb + enclitic pronoun were taken into account. The function of such pronoun in those constructions depends on the verb: *komak-et mikonam* 'I will help you' (indirect object; equivalent: *be to komak mikonam*) or *negāh-eš kardam* 'I looked at him/her' (indirect object, equivalent: *be u negāh kardam*) but *bāz-eš kardam* 'I opened it' (direct object; equivalent *ān-rā bāz kardam*). The forms in which pronoun is attached to the verb itself were included in the section below.

Examples:

(1) with nominal part of the verb *surāx kardan* 'to tear':

age	*surāx-eš*	*kon-i*
CONJ	NOMVERB-EP3SG	make.PRS.SUBJ-2SG
if	tear.them	you.will

age surāxeš koni 'if you tear them' (ŠĀ 58)

(2) with preverb, in a quasi-impersonal idiom *zouq (kasi-rā) bar dāštan* 'to get excited' (literally, it is the excitement that takes over somebody'):

zouq	*bar-eš*	*dāšt*
NOUN	PREVERB-EP3SG	have.PST3SG
excitement	over.her	had

zouq bar-eš dāšt 's/he got excited' (KP 50)

However described as a quasi-impersonal idiom, this one differs from the examples in 5.1.2.7.4 above: first, the enclitic pronoun can be replaced here by a full form of *kasi-rā* 'someone,' and second, the enclitic pronoun is not attached to the nominal part of the idiom (i. e. *zouq*) but to the preverbial part of the compound verb (i. e. *bar*). Changing the place of the enclitic pronoun would alter the meaning of the whole phrase by changing the relation from dative to genitive.

5.1.2.7.6 With verb as direct or indirect object (or subject)

The enclitic pronouns can also be affixed to the simple verbs themselves. The function of such pronoun in those constructions depends on the verb: *mibinam-et* 'I will see you' (direct object = *to-rā mibinam*) but *goftam-eš* 'I told him' (indirect object; equivalent with *be u goftam*). Lazard mentioned only the direct object function: "Affixed to the personal forms of the verb, assuming the function of direct object, the suffixes are equivalent to personal pronouns followed by *rā*, ex. *didam-et* 'I saw you' [=to-rā didam]" (idem 1992: 110) He did not mark such usage as colloquial and it is not remarkable enough to mark MSP register by itself,

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

but a higher frequency of this feature should be associated with an inclination towards the spoken rather that written standard.

There is one more form where an enclitic pronoun of the 3rd person singular is affixed to the 3rd person singular verb. It is attested by Lazard: "A personal suffix of the third person affixed to a personal verbal form may represent the subject. This use is frequent in colloquial language; it sometimes indicates a slight insistence on the identity of the subject, ex. *umad-eš* 'he came' [=u āmad]" (Lazard 1957/1992:112). Peisikov (1960/2002: 78) also mentions such forms as a representation of subject, but this observation is corrected by Šoǧā'i. According to his comment, it is a continuation of ergative construction in Middle Persian. However it was used only with a limited number of verbs, so is the discussed form nowadays (ibidem).

Examples:

(1) 3sg subjunctive form of the verb *didan* 'to see' with 1sg enclitic pronoun *-am* as direct object:

xāste.bud	*be-bin-ad-am*
will.PLPF3SG	SUBJ-see-3SG-EP1SG
had.wanted	would.see.me

xāste bud bebinadam 'he had wanted to see me' (Hn 48)

(2) 3sg past form of the causative verb *xābāndan* 'to lay [somebody/something]' with 3sg enclitic pronoun as direct object:

xābānd-im-eš	*ru-ye*	*taxt*
lay.PST-3PL-EP3SG	PREP-EZ	NOUN
we.laid.him	on	bed

xābāndimeš ru-ye taxt 'we laid him on the bed' (Hn 51)

(3) here the enclitic pronoun is unexpectedly joined to the copula instead of the nominal part of the verb, constituted by a participial form of a causative verb *pušāndan* 'to cover' (a more common form of incorporating object into this kind of forms would be *pušānde-aš bud*, similar to the examples from 5.1.7.2.5):

dud-e	*yeknavāxt-e*	*qaliz-i*	*pušānd-e*	*bud-eš*
NOUN-EZ	PREP-EZ	ADJ-INDEF	cover.PST-PTCP	be.PST3SG
smoke	uniform	thick	covering	was.that

dud-e yeknavāxt-e qalizi pušānde budeš 'a uniformly thick smoked had covered it' (KP 63)

5.1.3 Syntactic

Although syntactic parameters make up only as much as 18 percent of the overall number, their importance for the results of the analysis lies in their ever-presence —it is the most universally found group of parameters, and also the one that is the most often incorporated into otherwise "proper" MWP passages. Two basic groups can be distinguished within this category: the aberrations of the word

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

order and the omissions (of conjuctions in compound clauses, of auxiliary verbs mainly in perfect tense forms and prepositions in locative phrases); there is also a small group of parameters representing the so-called 'lack of agreement,' that is, phrases with psychological verbs and semi-impersonal idioms where the grammatical agreement, usually between subject and verb, is compromised.

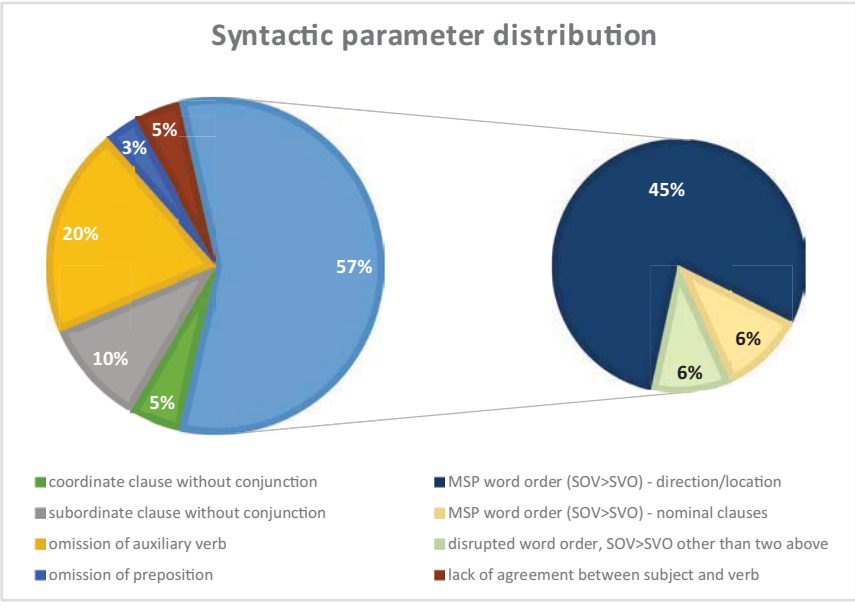

Figure 15. Syntactic parameter distribution (source: own research).

5.1.3.1 Clause without conjunction (coordinate/subordinate)

In the analyzed material, one of the most frequent deviations from MWP in terms of syntax is the omission of conjunctions. This is true both for coordinate and subordinate clauses and has been observed already in 19th century. In 1882, Haggard and Le Strange noted that "there is in Persian generally a great looseness of style, a state of things fostered by their **idiomatic disregard of conjunctions** (such as *and, then, if, in order that,* and the like), and the very unprecise nature of the relative pronoun" (Haggard, Le Strange 1882: xxix, my emphasis). While the problems connected with the relative pronoun *ke*, which serves a variety of other syntactic functions as well, are described in 5.1.2.2, the abovementioned disregard of conjunctions can still be seen. 15,61% of all the syntactic parameters recorded are clauses without conjunction (5,35% coordinate and 10,26% subordinate), which makes it the third most frequent syntactic MSP feature after reversed word order and auxiliary verb omission in perfect tenses. Jeremiás

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(1984: 285) notes that "changing subordinate clauses into co-ordinate through the omission of the conjunctions […] is also a characteristic of the informal variety."

5.1.3.1.1 Coordinate clauses

Coordinate clauses without conjunction appear usually in those passages in which the events take place immediately one after another. It could be said that this way of joining sentences without any unnecessary disruption speeds up the pace with which events occur for the reader. According to Lazard, "in colloquial language, it frequently happens that two clauses or two verbs follow one another without appreciable pause when describing actions which succeed each other immediately and tend to be conceived as forming one event" (Lazard 1957/1992: 216). This happens especially often when the first verb is a verb of movement such as *raftan* 'to go' or *āmadan* 'to come,' as can be seen in the examples below.

Examples:

(1) a coordinate clause after a modal verb *xāstan* 'to want':

mi-xāst-i	*negah-eš.dār-i*	*torši.be-ndāz-i*
DUR-want.PST-2SG	NOMVERB-EP3SG.have.PRSSUBJ-2SG	NOMVERB-EP3SG.SUBJ-throw.PRS-2SG
you.wanted	keep.her	make.her.sour

mixāsti negah-eš dāri torši bendāzi 'you wanted to keep her and make her an old maid [lit. make her sour]' (ŠĀ 51)

The expected conjunction *va* 'and' would fall after the compound verb *negah-eš dāri: mixāsti negah-eš dāri* [va] *torši bendāzi.*

(2) a set of three subjunctive clauses with more than one interpretation possible:

tāksi	*be-gir-ad*	*xāne*	*be-rav-ad*	*duš*	*be-gir-ad*
NOUN	SUBJ-take.PRS-3SG	NOUN	SUBJ-go.PRS-3SG	NOUN	SUBJ-take.PRS-3SG
cab	take	home	go	shower	take

tāksi begirad beravad xāne duš begirad 'take a cab, go home, take a shower' (KP 52)

First possibility is to use *va* 'and' twice and have three coordinate clauses: *tāksi begirad* [va] *beravad xāne* [va] *duš begirad* 'take a cab, go home and have a shower.'

Second option is to use *va* 'and' first and *tā* 'to' second (first two clauses coordinate and the third subordinate to second): *tāksi begirad* [va] *beravad xāne* [tā] *duš begirad* 'take a cab and go home to have a shower.'

The third options uses the same conjunctions but in reversed order: first *tā* 'to' and later *va* 'and' (as a result, second and third coordinate with each other and they are together subordinate to the first): *tāksi begirad* [tā] *beravad xāne* [va] *duš begirad* 'take a cab to go home and have a shower.'

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(3) coordinate clauses with verb of movement (coming/going):

āmad	*āmpul-e.fešār*	*zad*
come.PST3SG	NOUN-EZ.NOUN	hit.PST3SG
came	blood.pressure.injection	gave

āmad āmpul-e fešār zad '[the doctor] came gave [her] the injection for blood pressure' (ĀM 56)

Verbs of movement, especially *raftan* 'to go,' *āmadan* 'to come,' *bar gaštan* 'to come back, to turn back' are commonly found in the coordinate clauses without conjunction, especially at the beginning. Here both verbs are in the simple past tense, thus interpretations other than coordination with *va* 'and' are not possible (subordinate *tā* 'to' would require a subjunctive mode in the second clause): *āmad* [va] *āmpul-e fešār zad* 'came and gave the injection.'

(4) coordinate clauses with *bar gaštan* 'to turn':

bar.gašt	*borāq.šod*	*be*	*Ārezu*
PREVERB.turn.PST3SG	NOMVERB.become.PST3SG	PREP	NOUN
turned.back	got.angry	at	Ārezu

bar gašt borāq šod be Ārezu 'he turned back got angry at Ārezu' (ĀM 57)

As in the former example, also here only one interpretation is possible, with *va* 'and' after *bar gašt* 'he turned.' The verb *borāq šodan* means 'to become angry, to show anger towards somebody.'

5.1.3.1.2 Subordinate clauses

With subordinate clauses, we return to the *idiomatic disregard of conjunctions*. As Lazard puts it, "in colloquial language subordinate words are often omitted. The clauses are purely and simply juxtaposed (parataxis); the relation of subordination emerges from the context and from the intonation and rhythm of the sentence. The nature of subordination (temporal, final, conditional, etc. …) is shown only by the context, that is, by the semantic relationship between the main clauses and the subordinate clause" (Lazard 1957/1992: 220). I would go even one step further: the nature of relation between clauses might be shown only by the context, for in some cases the absent conjunction could be a copulative *va* 'and' just as well as a subordinative *tā* 'so that.' While in general, the difference will surface in verbal forms (the subordinated clause introduced by *tā*, present or absent, would rather use subjunctive forms), there are cases just as the example (2) above in 5.1.3.1.1, where all the verbs are in fact subjunctive and the decision of reading them as coordinate or subordinate is left to the reader. Relative clauses are, to some extent, an exception to this rule, since they are obligatorily introduced by *ke;* unless they belong to a "particular sub-class of generalized relative clauses introduced by *har* 'each, every,' where *ke* is optional" (Windfuhr 1979:67; see example (4) below). Windfuhr adds that there are no specified rules for the application or omission of *ke* (Windfuhr 1979: 69). The examples collected

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

in my analysis show that *tā* 'to; until' (replaceable, naturally, by *ke*) is one of those conjunctions that are mostly prone to be omitted, just as it is in the first two examples below.

Examples:

(1) omitted *tā* or *ke:*

boland.šod-am	*be-rav-am*
NOMVERB.become.PST-1SG	SUBJ-go.PRS-1SG
I.stood.up	I.would.go

boland šodam beravam 'I stood up [so that] I go' or 'I stood up to go' (Hn 59)

(2) omitted *tā* or *ke:*

mi-r-am	*čāy*	*biy-ār-am*
DUR.go.PRS-1SG	NOUN	SUBJ-bring.PRS-1SG
I.go	tea	I.would.bring

miram [tā] *čāy biyāram* 'I'll go [to] bring some tea' (ŠĀ 51)

(3) omitted *ke* 'that' introducing a declarative content clause:

be-h-eš	*etminān.dād-am*	*man*	*mesl-e*	*ānhā*	*fekr.ne-mi-kon-am*
PREP-EP3SG	NOMVERB.give.PST-1SG	PRON	ADV-EZ	PRON	NOMVERB.NEG-DUR-make.PRS-1SG
to.him	I.made.sure	I	like	them	don't.think

beheš etminān dādam [ke] *man* […] *mesl-e ānhā fekr nemikonam* 'I ensured him [that] I don't think like the others' (KP 57)

Here quoting the whole sentence gives a hint as to why *ke* is omitted:

> *beheš etminān dādam [ke] man yeki*[150] *ke mesl-e ānhā fekr mikonam va be nazaram; az xeyli ādamhā-ye digari ke mišenāsam ādam-e ma'qultari-st*
>
> 'I made sure I don't think like the others and in my opinion, he is more reasonable than many of the people I know'

The underlined parts are two *ke*-introduced relative clauses, which usually are not subject to the omission of conjunctions (see quote from Windfuhr 1979 above). Therefore the only *ke* which is not necessary in the context, is omitted in order to avoid repetition.

(4) a special kind of relative clause, in which the use of compound pronoun with *har* makes it possible to omit *ke* (Windfuhr 1979: 67, quoted above):

har-či	*dam-e*	*dast-am*	*resid*	*bo-xor-am*
PRON-PRON	NOUN-EZ	NOUN-EP1SG	ADV-EZ	SUBJ-eat.PRS-1SG
everything	to	my.hands	came	I.would.eat

harči [ke] *dam-e dastam resid boxoram* 'I would eat anything that came to my hands' (ĀM 59)

150 The omission of copula in this sentence is not a regular grammatical feature of Persian; it might be a result of editorial mistake or an element of author's stylization.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

5.1.3.2 Lack of agreement (between subject and verb: quasi-impersonal idioms, logical subject, number disagreement)

In this category, forms were listed in which a grammatical agreement between elements is compromised. It includes the impersonal constructions of the *xoš-am miyāyad* 'I like it (lit. [it] comes pleasant to me),' (I refer to them as quasi-impersonal idioms) in which the psychological subject of the verb is represented by an enclitic personal pronoun (also called personal suffix) attached to the nominal part of the verb, while the grammatical subject of the predicate is the psychological object of the action (or, when there is none, a 3rd person singular form is used). According to Lazard, those are "certain expressions designating psychic or physical states […] of an impersonal style, ex. *sard-am ast* 'I am cold' [lit. [it] is cold to me]" (Lazard 1957/1992: 177). Windfuhr (1979: 126–128) describes those forms as *indirect (middle) verbs*, he also mentions a Persian term *nā-gozar* (lit. non-passing, i.e., intransitive) used by Xānlari for the verbs in which the formal object is in fact the logical subject (1970: 1972).

The 'logical subject' category refers to the clauses where the grammatical subject of the predicate is different from the psychological subject of the clause. The function of such constructions is usually emphatic and they seem to have been known in Persian grammar since long ago. According to Lazard:

> "The following turn of phrase is frequent, especially in colloquial language. The psychological subject (the word representing the person or the thing for which one want to affirm something), which is not the grammatical subject of the verb, is placed at the beginning of the sentence, then cross-referenced in the body of the sentence by a personal pronoun (most often a suffix) used in the syntactical function required by the construction, so that the entire sentence appears as the psychological predicate of the word thus detached at the beginning. This latter is followed by a pause more or less perceptible, ex. *xānum, hāl-ešun [-ešān] četowr-e [ast]* 'how is Madame?' (lit. 'Madame, her condition is how?')" (Lazard 1957/1992: 182).

Windfuhr sees it as a "rare register-marked variant of anacoluthon constructions" and relates it to the topicalization with emphatic *ke* (see 5.1.2.2 above) as a "'preposition of the psychological subject' of a clause (Jensen 1931: 191; cf. also Homā'i 1959: 45)." (Windfuhr 1979: 52, 71–72). He also mentions that such a construction was found in traditional grammars, described as a kind of *badal*, an introduction of the first member of a sentence in order to draw listeners' attention to it (Windfuhr 1979: 72; after Lumsden 1810: 228–231), and adds another point of view by Rastorguyeva (1953), who treats *amu zan-aš* as a unit, equivalent to ezāfe construction *zan-e amu* (according to Windfuhr, that understanding is based on a Tajiki evidence. Peisikov (1960/2002) describes the construction in detail, but does not give a satisfying description. Not accepting Rastorguyeva's interpretation that equals the construction with ezāfe, he does not give sufficient

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

evidence to counter it. So for Peisikov (1960/2002, after Arends [1941: 32–34]) in *amu zan-aš* 'uncle, his wife' *amu* is the emphasized subject and *zan-aš* a referential part of the predicate (Peisikov 1960/2002: 187–196, 380–387; after Windfuhr 1979: 71). Pisowicz, on the other hand, attributes this construction to a Turkish influence ("Turkish izafet" like *pesar ketāb-eš* instead of *ketāb-e pesar*) (Pisowicz 1985: 91, 2003: 179).

The least represented subcategory were those sentences in which the forms of address do not correspond with the verb. This happens quite often in the dialogues, when the issues of politeness are involved. The second person singular or plural pronouns are sometimes used interchangeably, such as the singular and plural verbal endings. This phenomenon is mentioned by Lazard: "In colloquial language, the use of *to* 'you' (coll.) and of *šomā* 'you' (polite form) and the verbal forms of second pers. sing. and of second pers. plural frequently alternate at short distances. Sometimes *šomā* is followed by a verb in the second pers. sing., particularly in the imperative" (Lazard 1957/1992: 181).

Examples:

(1) an example of the anacoluthon construction discussed above:

in	*ğavān*	*xun-aš*	*be*	*ğuš*	*āmad-e*
DEM	NOUN	NOUN-EP3SG	PREP	NOUN	come.PST-PTCP
this	young.person	his.blood	to	boil	has.come

in ğavān xunaš be ğuš āmade 'this young person has become angry [lit. this young person, his/her blood has come to boil]' (Hn 56)

In this example, the young person (a young man living in the same building as the narrator of the story) is the logical (psychological) subject of the sentence. Yet his state is described by a quasi-impersonal idiomatic phrase where the psychological subject is realized as the possessor of the grammatical subject (enclitic pronoun). To clarify whose blood has boiled, the psychological subject is introduced at the beginning, without a grammatical bond with the rest of the clause. Interestingly, the complimentary coordinate clause is: *va xaste ast* 'and is tired' (where *xaste* 'tired' is an adjective and *ast* is a 3sg present copula), so it goes back to the original subject (i. e. psychological subject from the previous clause) without repeating it or signalizing it by means of a pronoun. The whole sentence would translate literally into: "the young person, his blood has come to boil, and is tired," where the subject of the second clause could be read as ambiguous (even if the pronoun was used, both *xun-aš* 'his blood' and *in ğavān* 'this young person' would use the same 3rd person pronoun *u*). It should also be noted that the perfect form *be ğuš āmad-e* 'has come to boil' has the auxiliary verb omitted (see examples in the section 5.1.3.3.1 below).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(2) another example of the enclitic pronoun denoting the psychological (and logical) subject of the clause; here in fact the construction is ambiguous:

Ārezu	*yād-eš*	*bud*
NOUN	NOMVERB-EP3SG	be.PST3SG
Ārezu (name)	memory.her	was

Ārezu yādeš bud 'Ārezu remembered [lit. Ārezu, her memory was]' (ĀM 55)

In this example, *Ārezu* (a personal name) might denote either the psychological subject (i.e. 'Ārezu remembered') or the object of the clause (i.e. 's/he remembered Ārezu'). In speech, the difference could be expressed by intonation: in the former case, word stress would fall on the nominal part of the compound verb, that is, on *yād-eš,* while in the latter, it would be on the object, that is, *Ārezu.* In writing or print, this pattern of stress is sometimes expressed by punctuation (a psychological subject is followed by a comma), but it is not obligatory; no comma is used in the quoted example. Here, the context comes to reader's aid: the quote is a part of a longer sentence which reads: *az vaqti ke ārezu yad-eš bud lebāshā-ye Nosrat goldār bud* 'from the time Ārezu remembered Nosrat's clothing was always [covered with] flowery patterns' and since the clothing is not capable of remembering, there is no other candidate for the psychological subject in the sentence except Ārezu.

5.1.3.3 Omission (of preposition or auxiliary verb)

5.1.3.3.1 Auxiliary verb

In this category, the sentences with analytic verbal forms were listed in which the finite form of verb (the auxiliary) is missing. In the source material, that is in fact restricted to the perfect tense or at least those forms which surface as a perfect form with missing auxiliary (although they might represent different underlying categories with the same surface form). A discussion on the constructions with finite verb missing can be found i.e. in Windfuhr (1979: 74) under "participial clauses." He lists two kinds of those, subordinate (*fe'l-e vasfi,* in which the participial form of verb is used) and coordinate (*hazf-e fe'l,* where the ellipsis of finite verb marking is found) and declares them as "generally confined to the literary dialect," known and found in earlier Persian already.

> Coordinate participial clauses imply a literary rule by which all but the last verb form of coordinate clauses are surfaced as a (perfect) participle; e.g.: *vazir bar gašte, va esterāhat xāhand nemud* 'the vazir returned and will relax' or 'the vazir will return and relax.'
>
> Subordinate participial clauses participialize the verb of the subordinate clause; e.g.: *vazir bar gašte esterāhat xāhand nemud* 'the vazir having returned will relax.'

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

> Both constructions generally require identical subjects, but not obligatorily so. Both constructions are known in early Persian already, where the finite form was not obligatorily the final one. […] As can be seen from the two examples above, the differentiation between the two constructions is not always self-evident (Windfuhr 1979: 75).

Haggard and Le Strange (1882: xxxvii) also describe the participial clauses of that kind, especially in historic narration when, as they put it: "an entire page or more may be composed with a succession of subordinate clauses, each with its past participle: then, at the end of all things, and far removed from its subject, come the one principal verb."

Lazard (1957/1992) describes such forms under gerund (=past participle), which "coordinated to a personal verb which follows it, [it] serves as substitute to a personal verb form of the same type as that verb; it may be preceded by complements in the same manner as a personal verb" (Lazard 1957/1992: 170).

As can be seen in the results of the analysis, the ellipsis of finite verb marking is very frequent in dialogues where a perfect form of verb is used. Or, in other words, the gerund or past participle forms (past verbal stem with -e ending, i. e. *rafte* 'gone,' *dide* 'seen') in dialogues are hardly accompanied by a form of finite auxiliary verb *budan* 'to be,' regardless of the presence of any other finite verb in the sentence, by which feature they differ from the participial clauses described by Windfuhr and also by Lazard as they are not coordinated to any other verb. All the clauses with missing finite verb found in the source material are 3rd person singular, therefore lack of finite verb with verbal ending specifying person is not significant. The question arises whether those are the remnants of the participial clauses discussed above, or rather a new tendency of omitting auxiliary personal verb after perfect forms.

For the perfective forms, regardless of the auxiliary being present or omitted, there is one more important note: the notion of 'non-evident' or 'reported' past, denoting actions in the past, in which the speaker was not directly involved, e. g. *tarğome mikarde-and* 'they were translating.' This usage is seemingly restricted to the third person (singular or plural) and it was traditionally called 'continuous perfect,' (the continuous aspect denoted by the use of durative prefix *mi-)*, but according to Windfuhr, it is not in fact the difference of aspect or tense, but the evidentiality. Thus the forms belonging to the present perfect have double function: they express the event experienced in the past and anterior to the present, or the 'reported' past events, not experienced directly by the speaker (1979: 85, 88–89). Perry (2000: 229) notes that in some forms of standard Persian the evidential (indirective or epistemic, in his words) meaning is predominant: past perfect formed by a past participle with the past participle of the auxiliary verb and a finite form of the auxiliary (for example, *karde bude-ast*) and what he

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

calls a durative perfect (*mi-karde-ast*). He also adds that in all varieties of Persian, the final auxiliary might be omitted.

In the analysis, all the perfective verbal forms without auxiliary were treated altogether under the category of auxiliary verb omission, regardless of their belonging to the actual perfect, the reported past or other. Further study on the subject would be necessary to determine whether there is another function of the seemingly perfect forms without auxiliary, that is however beyond the scope of this research project.

The category proved very frequent, covering 20% of all the collected syntactic parameters (DXM: 45, ŠĀ: 33, KP: 57, ĀM: 44, Hn:6); numerous examples are to be found in the clauses quoted above at another parameters.

Examples:

(1) auxiliary verb omission after a regular use of a perfect tense:

tā	*inǧā*	*čid-e*	*āmad-e*	*bālā*
PREP	PRON	put.PST-PTCP	come.PST-PTCP	ADV
until	here	has.put	has.come	up

tā inǧā čide āmade bālā '[he] has laid [bricks] up to here and has gone up further' (ĀM 56)

In this example, a housekeeper describes the progress of wall construction to the house owner's daughter at the time when the wall has not been finished yet. Apart from the omission of auxiliary verb, the sentence exemplifies the omission of coordinate conjunction and a change in word order (MWP rendering of this sentence: *tā inǧā čide va bālā āmade ast*).

(2) a 3rd person "durative perfect" used as the epistemic verb:

ne-mi-tavānest-e	*in*	*kār-rā*	*bo-kon-ad*
NEG-DUR-can.PST-PTCP	DEM	NOUN-DO	SUBJ-do.PRS-3SG
he.couldn't.have	this	thing	done

nemitavāneste in kār rā bokonad 'he couldn't have done this' (KP 56)

The epistemic meaning is suggested by the context: *xob, ma'lum ast agar mariz bud nemiyāmad pišet benšinad. ya'ni nemitavāneste in kār rā bokonad* 'well, it is clear that if he was sick, he wouldn't come to sit beside you; I mean, he wouldn't be able to do that'—the sentence does not state an observable fact. It is rather the speaker's interpretation and presupposition beyond his personal experience.

5.1.3.3.2 Preposition

The omission of preposition is strongly connected with the reversed word order, especially in sentences with adverb or noun phrase of direction, place or origin. The preposition (usually *be* 'to' or *dar* 'in') can also be omitted in a regular word order, so a simple sentence in MWP: *be Tehrān miravam* 'I go to Tehrān' can become *Tehrān mira(va)m* or *mira(va)m Tehrān* without affecting the meaning

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(*Tehrān* can of course submit to raising and become *Tehrun*, if the clause is supposed to become MSP to the extreme). *Tehrān-am,* name with a personal form of 'to be,' is enough to state 'I am in Tehrān.' Prepositions other than *be* and *dar* are not omitted that readily, even though the following passage from Lazard (1957/1992) would suggest that.

> Objects other than direct object are for the most part introduced by a preposition […]. In colloquial language, the prepositions are frequently omitted. In the sentence—as a general rule—when the function of the noun (or pronoun) object is made clear by its own meaning and/or by the context, the preposition is omitted, ex. *borrow manzel* 'go [to] the house.' […] This happens especially with the prepositions *be* 'to, with' (direction, manner), *dar* 'to, in' (place), *bar* 'on' (place, direction) and sometimes also with *az* 'from' (origin). Prepositions provided with a more "concrete" meaning, *bā* 'with' (accompaniment, instrument), *bi* 'without,' *tā* 'until,' *ǧoz* 'except,' *čun* 'like' are not omitted (Lazard 1957/1992: 194).

Omission of preposition is one of those features which are common in MSP, but not represented extensively in print, contrary to the reversed word order. The analysis recorded only 30 examples within the chosen material: an average of 8 instances in DXM, ŠĀ, KP and ĀM—and none in Hn.

Examples:

(1) omission of *be* 'to' with reversed word order and omission of auxiliary verb in perfect tense:

mağbur	*bud-e*	*be-rav-ad*	*dastšuyi*
NOMVERB	be.PST-PTCP	SUBJ-go-3SG	NOUN
obliged	was	would.go	bathroom

mağbur bude beravad dastšuyi 's/he had to go to the bathroom' (KP 56)

(2) omission of *dar* 'in' without reversed word order:

tehrān	*be-mān-ad*
NOUN(LOC)	SUBJ-stay-3SG
Tehran	would.stay

Tehrān bemānad 's/he would stay in Tehrān' (KP 61)

(3) omission of *be* 'to' with reversed word order:

Āye-rā	*mi-bord*	*mahd-e*	*kudak*
NOUN-DO	DUR-take.PST3SG	NOUN-EZ	NOUN
Āye	s/he.was.taking	kindergarten	

Āye-rā mi-bord mahd-e kudak 'she was taking Āye to kindergarten' (ĀM 54)

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

5.1.3.4 Word Order (SOV > SVO direction, location / SOV > SVO nominal clauses / disrupted)

Modern Written Persian is characterized by relatively regular order of sentence constituents. Grammars do agree that the basic, neutral word order in Persian sentence is SOV, with the finite verb falling at the end of the sentence (Dabir-Moghaddam 2018: 44). It does not mean that there are no exceptions, in fact, as Lazard put it,

> "deviations from the normal order are not rare, especially in colloquial language; but they are not lacking, either, in formal language. They are owing sometimes to reasons of rhythm (relative length of different terms), and sometimes to the desire to emphasize a term. They are unequally frequent according to their nature: an adverbial phrase is displace rather easily, even after the verb; on the other hand, the inversion of the direct object and of the verb, and especially that of the subject and of the verb, has a marked stylistic value" (Lazard 1957/1992: 208)

The observations below were made by Lazard on the syntax of colloquial language: 1) indirect object or adverb is often placed after the verb, 2) direct object follows rather frequently the verb (Lazard 1957/1992: 210–211), and (this being the most frequent one) 3) the adverb of direction is very often without preposition; in this case it is usually placed after the verb (Lazard 1957/1992: 197). Jeremiás attributes the looser word order of the informal variety to the change or omission of other grammatical elements (such as the omission of prepositions) and mentions the suprasegmental features of the speech (stress, intonation, pitch) that might take over and fulfil the syntactic functions of the grammatical morphemes typical of the written variety (Jeremiás 1984: 285). According to Mahmoodi-Bakhtiari, the movement of modifier to the end of the sentence (after the verb) results from the deletion of preposition (2018:201), but the analyzed data does not prove that assumption: indeed it is often the case that a modifier is moved along with the preposition, or the deletion of preposition does not trigger the change of word order.

The syntax of nominal clauses is more flexible in terms of verb placement also in MWP, which is why they were treated separately in the analysis. "In noun-predicate clauses, when the predicate is a relatively long group of words, a part is often thrown after the verbal copula in such a way that the latter finds itself embedded within the predicate. This construction is usual, even in formal language" (Lazard 1957/1992: 208).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Examples:

(1) reversed word order SVO > SOV (direction/ location):

nešast-e.bud	*ru-ye*	*labe-ye*	*taxt*
SIT.PST-PTCP.BE.PST3SG	PREP-EZ	NOUN-EZ	NOUN
had sat	on	edge.of	bed

nešaste bud ru-ye labe-ye taxt 'she was sitting on the edge of the bed' (Hn 53)

The verbs *istādan* 'to stand,' *nešastan* 'to sit' and *xābidan* 'to lay down' in Persian describe the action od standing up, sitting down or laying down, respectively. To express the state of standing, sitting or laying, a perfect form (i. e. *istāde ast* 'has stood up,' *nešaste ast* 'has sat' and *xābide ast* 'has laid') is used, thus the pluperfect form in the example (the event is taking place in the past).

(2) reversed word order SVO > SOV:

šoru'	*kard-am*	*be*	*sanǧeš-e*	*ouzā'*
NOMVERB	make.PST-1SG	PREP	NOUN-EZ	NOUN
beginning	made.I	to	examination.of	situation

šoru' kardam be sangeš-e ouzā' 'I began to examine the situation' (DXM 53)

This kind of sentence is not marked as typical MSP and can occur in MWP as well.

(3) reversed word order SVO>SOV with prepositional phrase (preposition with enclitic pronoun) in a postverbal position:

mi-mir-am	*barā-ye-šān*
DUR-die.PRS-1SG	PREP-EZ-EP3PL
I.die	for.them

mimiram barāyešān 'I'd die for them (expression of liking something)' (KP 61)

5.1.4 Lexical

The items listed in this group are perhaps the least controversial in terms of their belonging. They are grouped into four categories according to length and syntactic function. If any of the phonological, morphological or syntactic features described above were found in forms from this category, they were listed separately under the appropriate headings and counted there as well.

The distribution of lexical parameters is presented in the charts below.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

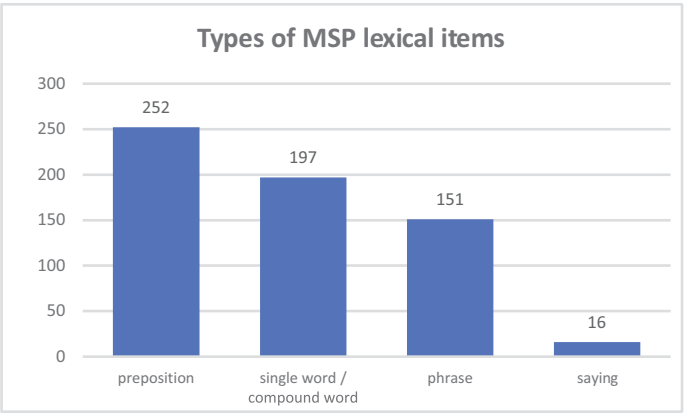

Figure 16. Types of MSP lexical items and number of occurences (source: own research).

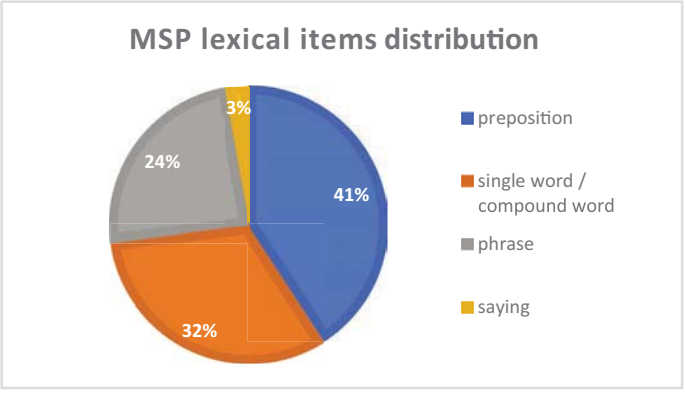

Figure 17. Individual lexical parameter distribution (source: own research).

5.1.4.1 Preposition

In fact, only two prepositions were counted as MSP: *tu* 'in' and *vāse* 'for.' In the analyzed source material, especially in the dialogue parts, the preposition *dar* 'in' hardly ever appears and is replaced by *tu* in most of the cases, with or without ezāfe. This is in line with the observation made by Lazard:

> "In colloquial language, the prepositions *dar* and *bar* are practically not used in their locative meaning. When there is a need to specify the position 'inside of' or 'on' something, one uses the prepositional phrases *tu-ye* 'inside of' and *ru-ye* '[lit. on] the face of,' almost always without a proper preposition; these expressions often lose the ezāfe […]. Very frequently, in colloquial language, the locative phrase is without any preposition." (Lazard 1957/1992: 196)

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

As stated, *tu* 'in' (with or without ezafe) is very frequent to the extent that it almost completely replaces *dar* (which was difficult to find within the analyzed corpus). *Vāse* is not that omnipresent, but it is worth saying that it does replace *barā* in a variety of contexts, not only in the actual benefactive meaning, but also causative: MWP *barā-ye hamin ast* > MSP *vāse hamin-e* 'because of that' (DXM 51, DXM 56). The popularity of those two prepositions is visible in their relative frequency: they appear in the corpus 252 times, which makes 40,91% of all lexical parameters (616 appearances).

Examples:

(1) *tu* 'in' without ezafe, replacing MWP *dar* 'in' or *dāxel-e* 'into':

parid-am	*tu*	*māšin*
jump.PST-1SG	PREP	NOUN
I.jumped	in	car

paridam tu māšin 'I jumped into the car' (DXM 56)

(2) *tu* 'in' with ezafe, replacing MWP *dar* 'in':

xod-am	*be-mān-am*	*tu-ye*	*xāne*	*be-pus-am*
REFLEX-EP1SG	SUBJ-stay.PRS-1SG	PREP-EZ	NOUN	SUBJ-rot. PRS-1SG
myself	I.should.stay	in	house	I.should.rot

xodam bemānam tu-ye xāne bepusam 'I should stay alone and rot in this house' (ĀM 59)[151]

(3) *vāse* without ezāfe replacing *barā-ye* 'for':

ye	*maxsus-eš-o*	*vāse*	*xod-am*	*radif*	*kon-am*
INDEF	ADJ-EP3SG-DOM	PREP	REFL-EP1SG	NOMVERB	make.PRS.SUBJ-1SG
a	special.one	for	myself	queue	make

ye maxsuseš-o vāse xodam radif konam 'I'd prepare one special [of this] for myself' (KP 58).

5.1.4.2 Single word / Compound word

In this category the words are included which are marked as colloquial in the dictionary (the dictionaries of Dehxodā and Soxan were used as a reference) as well as those with a less concrete meaning that have a special usage in MSP different than their regular standard use. The example of such functional words might be *digar*, which "in colloquial language […] is used very frequently (at the end of a sentence), as a simple particle supporting an interrogation, an exclamation, an injunction etc." (Lazard 1957/1992: 128). Apart from those, the copulative compounds of two words with similar meaning connected with -o- were also included.

151 Using English *should* in the translation is more relevant to the context than *would*, which I give otherwise for that grammatical form. It should be remembered that those are only rough translations. Persian subjunctive forms can be translated in more than one way, depending on their original context.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Examples:

(1) a compound adjective formed by a double imperative:
bo-dou- *-bo-dou*
IMP-run.PRS- IMP-run.PRS
run run
bodou-bodou 'fast (lit. running)' (DXM 55)

(2) a compound noun made of two past verbal stems:
sāxt- *-o-* *-pāxt*
BUILD.PRS COPCONJ COOK.PRS
build and cook
sāxt-o-pāxt 'realized plan' (DXM 53)

(3) a compound noun made of two nouns with a possessive enclitic pronoun:
kār- *-o-* *-bār-* *-etun*
NOUN COPCONJ NOUN EP2PL
work and weight yours
kār-o-bāretun 'your affairs' (KP 58).

5.1.4.3 Phrase

In this section, idiomatic and set phrases (Pers. *estelāhāt*) associated with colloquial usage were listed. In the traditional understanding, it is exactly the *estelāhat* that express the colloquial belonging of a passage of text, it is also them that should be preserved[152] (which is why they were very frequently employed in the prose writings of the early 20th century). Even these days, according to my field experience, a bookseller asked about a book with colloquial language would naturally turn toward those packed with *estelāhāt* rather than those with grammatical characteristics of *mohāvere*.[153] Among the idiomatic and set phrases no further divisions were introduced. Despite their obvious existence (the most striking is probably the difference between idiomatic phrases with unexpected meanings and more regular phrases which include colloquial elements), separation of those two groups would not be significant for the present study or alter the results of the analysis.

152 Their special place in language and literary canon is explicitly described by Mohammad-Ali Ğamālzāde in the introduction to the first edition of his short stories *Yeki bud, yeki nabud* (Mohammad-Ali Ğamālzāde, *Dibāče* [Introduction] [in:] idem, *Yeki bud, yeki nabud*, 7th edition, Tehrān 1966, pp. 3–21; 1st edition, Berlin 1921).

153 This is my personal experience which has repeated itself over numerous visits to Persian bookstores especially in Tehran (2012, 2015), but also Esfahan (2012), Shiraz (2012) and even Berlin (2019). While such an experience cannot naturally serve as any kind of evidence for the study, it is worth a mention.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Examples:

(1) a typical copulative compound noun: two synonyms, one literary and the other colloquial:

sar-o-kalle	*tekun*	*mi-d-im*
compNoun	nomVerb	prs-give.prs-1pl
head-and-head	move	we.give

sar-o-kalle tekun midim 'we nod heads' (DXM 50)

(2) a colloquial expression *xabar-i-ye* 'there are news' augmented by the use of plural *xabar-ā* instead of singular *xabar* 'news':

sar-e	*miz-e*	*hašt*	*xabar-ā-yi-ye*
prep-ez	noun-ez	subj-go-3sg	noun-pl-indef-cop3sgPrs
at	table	eighth	news.some.are

sar-e miz-e hašt xabarāyiye 'something is going on at table 8' (DXM 53)

(3) an idiomatic phrase:

sang	*ru*	*yax-am*	*kon-e*
noun	prep	noun-ep1sg	make.prs.subj-3sg
stone	on	ice.me	make

sang ru yaxam kone 'he would humiliate me (lit. make me a stone on ice)' (ŠĀ 51)

(4) another idiomatic phrase:

sar-e	*man-o*	*be-kub-i*	*be*	*tāq*
noun-ez	pron-do	subj-hit.prs-2sg	prep	noun
head	my	hit	to	niche

sar-e mano bekubi be tāq 'you would distract me (lit. hit my head against a niche)' (ŠĀ 52)

(5) an idiomatic verbal phrase:

be-zan-am	*tu-ye*	*zouq-eš*
subj-hit.prs-1sg	prep-ez	noun-ep3sg
I.would.hit	into	his/her.enthusiasm

bezanam tu-ye zouqeš '[I'd] dampen (lit. hit into) his enthusiasm' (KP 57)

(6) another idiomatic phrase:

āsmān-rismān	*bāft*
noun-noun	plait.pst3sg
sky-rope	plaited

āsmān-rismān bāft 'she talked nonsense (lit. she plaited sky with a rope)' (ĀM 65)

(7) a colloquial verb:

vel-am	*kon*
nomVerb-ep1sg	imperSg
loose.me	make

velam kon 'leave me alone (lit. let me loose)' (ĀM 77)

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

5.1.4.4 Saying

This category lists proverbs and sayings (Pers. *zarb-ol-masal*) used as independent clauses or sentences which are usually brought into the dialogues/ narrative to give it a more colloquial, conversational feeling. It should be noted that sometimes distinction between a saying and a colloquial phrase might be difficult and arbitrary. Here the distinction was made between those phrases and sayings which can be used as independent syntactic units (like proverbs), which were counted as sayings, and those idiomatic expressions which imply a meaning different to the literal meaning of their elements, but operate within context in functions similar to other syntactic units (i. e. as a compound verb, complement etc.), which were counted as phrases (see the category above). The use of syntactically independent clauses counted as sayings is of course a marginal phenomenon, which is well reflected in their low frequency within the analysis results (a total of 17 occurrences in the whole analyzed material).

Examples:

(1) used in 1st person singular:

bid-i	*nist-am*	*ke*	*be*	*in*	*bād-hā*	*be-larz-am*
NOUN-REL	be.not.PRS-1SG	CONJ	PREP	DEM	NOUN-PL	SUBJ-tremble.PRS-1SG
willow	am.not	that	at	those	winds	would tremble

[*man*] *bidi nistam ke bā in bādhā belarzam* 'I do not fear that (lit. I am not a willow to tremble at such winds)'(ŠĀ 61)

(2) used impersonally, with emphatic *ke*, omission of preposition and auxiliary verb in perfect tense:

āsmān	*ke*	*zamin*	*na-yāmad-e*
NOUN	EMPH	NOUN	NEG-come.PST-PTCP
sky	[that]	ground	hasn't come

āsmān ke zamin nayāmade 'this is not the end of the world (lit. heaven hasn't come to earth yet)' (ĀM 77)

(3) used in dialogue, as consolation/ advice:

xodā-ro	*če*	*did-i*	*kār-e*	*na-šod*	*na-dār-e*
NOUN-DO	PRON	see.PST-2SG	NOUN-EZ	NEG-become.PST3SG	NEG-have.PRS-3SG
god	what	did.you.see	work	did.not.happen	does.not.have

xodā-ro če didi? [*Be qoul-e māmān*] *kār-e našod nadāre* 'what have you seen of God? [As mom says,] nothing is impossible for him' (ŠĀ 61).

5.2 Parameter distribution in analyzed books

Having presented the parameters in detail, it is worth seeing now how are they distributed in the analyzed material. Preliminary observations on the language of each of the books were presented in Chapter 4.2, however, they were made *a priori*

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

and their relevance to the results of analysis are yet to be seen. The interesting question is, if those observations can be related to or proven by the numbers. In the sections below, each of the books will be presented with a set of charts representing the MSP parameter distribution in the analyzed part of the text to facilitate a later comparison, but before, some general remarks should be made.

In the preceding section of this chapter, the MSP parameters elicited from the texts were discussed in terms of their grammatical features, possible genesis and meaning. For some of them, the number of occurences was given to exemplify their higher or lower frequency, but in most of the cases without much further inquiry; and some of the parameters were not given the statistics at all.

The charts, tables and statistics should be considered of informative and illustratory character. Although the size of the analyzed corpus does not allow for drawing general conclusions valid for all the contemporary literature based on the statistical results, yet the patterns observable on that small and restricted span of text are repetitive.[154]

First, the overall distribution of parameters in the whole corpus should be described. The gross number of collected forms was almost five thousand (5028 forms)[155] with a relative distribution in the categories as represented by two charts below:

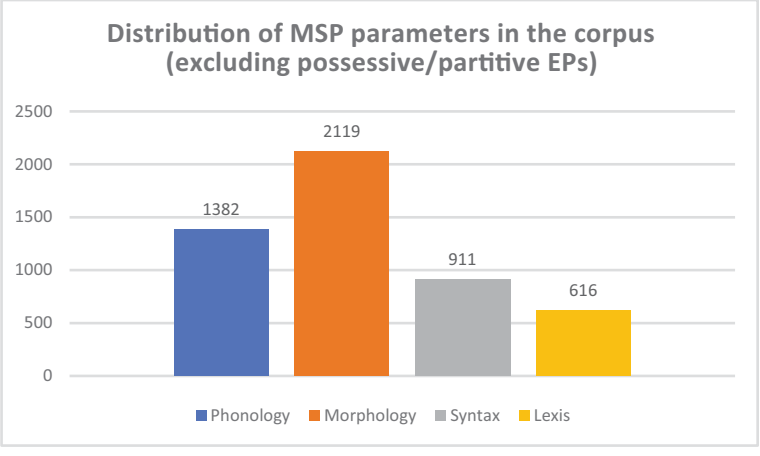

Figure 18. Overall distribution of MSP parameters in the corpus (excluding possessive/partitive EPs) (source: own research).

154 It was visible during the collection of non-standard forms in the first stage of the analysis—even within the chosen part of the texts, the types of non-standard forms and their application was repetitive (see chapter 4, sections 4.2 and 4.3).

155 5979 if we include possessive/partitive enclitic pronouns (951 forms).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

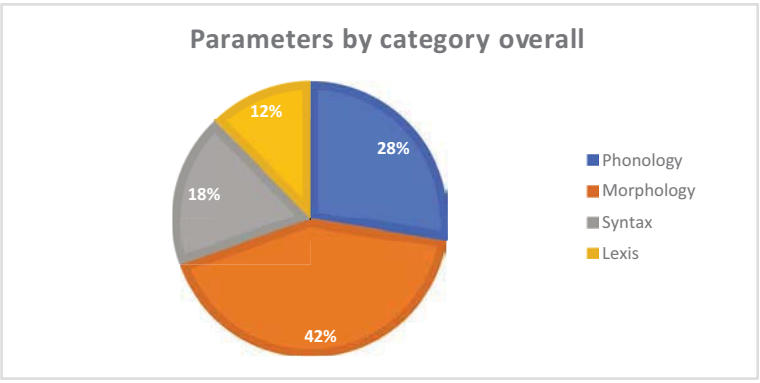

Figure 19. Distribution of MSP parameters in the corpus – proportions (excluding possessive/partitive EPs) (source: own research).

As for the exclusion of possessive/partitive enclitic pronouns from the general charts, the explanation is given in section 5.1.2.7 of the present chapter. The addition of those 951 forms to the morphological parameters would naturally result in a greater predominance of the forms of this category over the others.

The charts show an averaged result of the analysis of individual books, which differ greatly between themselves in regard to number of MSP forms and their distribution between categories. The predominance of morphological features is clear. Lexical forms, a typical marker of a colloquial style of text for a layman, are in fact a minority. The relatively large amount of phonological features is in reality due to two books (DXM and ŠĀ) which use them freely either in the whole text (DXM) or in the dialogues (ŠĀ). The other authors are more careful in applying the MSP spelling of words even within dialogues, which is visible in the diagrams of individual books in the sections below. Syntactic features represent only 18% of all collected forms, but here also significant differences between the books can be noticed. To present the differences between individual books in regard to the relative distribution of categories, another chart can be useful:

This chart clearly presents that the distribution and overall number of MSP forms is not equal among the books. While the differences are relatively small for lexical or syntactic parameters, they are paramount as far as phonology and morphology are concerned. The chart shows also the trends in distribution of categories in the individual books which will be discussed in the sections below.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

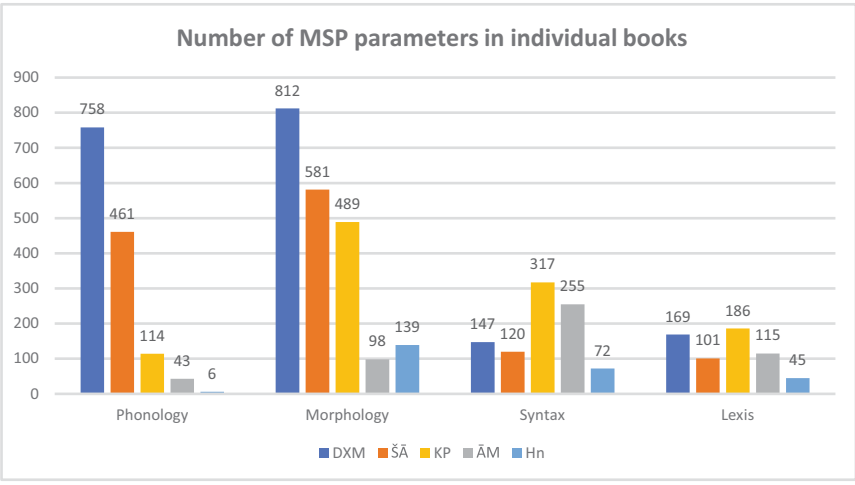

Figure 20. Number of MSP parameters in individual books (source: own research).

5.2.1 *Dar xarābāt-e moqān*

Dar xarābāt-e moqān is the leading book as far as the amount of MSP features is concerned (1886 forms). A striking feature of this book is the application of phonological features of MSP to all the text, regardless of it being a dialogue or part of a narrative. It is justified by narrating a whole novel as a first person retrospective, thus a kind of a monologue. In his relation of the past events, the narrator mixes an action story with spiritual and philosophical reflections. While in the action parts the dynamics of the story allow for syntactic "shortcuts" such as the omission of conjunctions (20) or auxiliaries (45) and changes in word order (60), the philosophical and spiritual content requires different means of expression and is responsible for a relatively small number of lexical parameters in the text (169 forms), mostly prepositions (68) and single words (58).

An interesting feature of DXM is a small number of enclitic pronouns, whether in possessive/partitive function (only 93) or in another syntactic functions (46). The morphological MSP features found in DXM are mostly verbal endings (471 out of 905) and verbal stems (188). The reader can immediately recognize the elements of the MSP repertoire and interpretes the language of the book as colloquial because of the spelling representing MSP phonological elements, which however does not affect the overall cohesion and tidiness of the text (they could be at risk when enclitic pronouns or and disruptions in word order are abundant). The chart representing the distribution of MSP parameters in DXM can be found below:

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

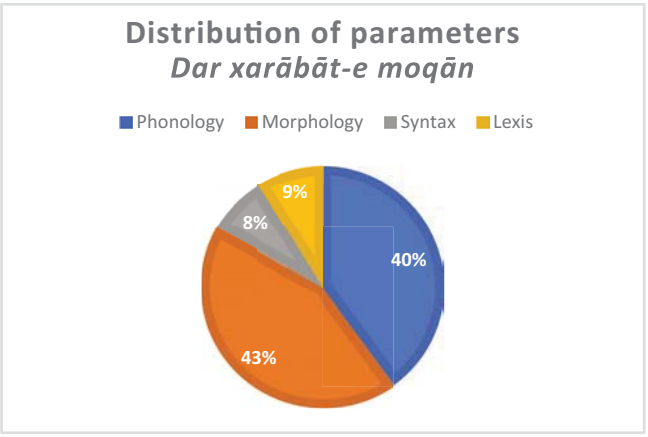

Figure 21. Distribution of parameters (DXM) (source: own research).

Unlike the general trend visible in Fig. 14, here the morphological parameters are represented almost equally to the phonological ones, while the syntactic parameters are definitely underrepresented, although still found in the text.

One of the possible interpretations of such distribution is that the text was shaped rather in a process of conscious stylization than less conscious application of everyday communicative patterns. In other words, as if the author intended to write close to MSP, so he used the most visible markers (i. e., the spelling affected by MSP phonology, verbal stems and endings). Yet at the same time, the syntax underneath reveals an adherence to MWP rules greater than expected. The application of MSP phonological features regardless of context (narrative or dialogue) is definitely innovative in Persian literature and there are not many examples of such a strategy.[156] At the same time, the professional background of the author should be taken into consideration. Dāriuš Mehrǧuyi is an experienced film director and screenwriter. While the language of books can be said to be restrictive and conservative, the language of screenplays is supposed to represent the natural speech much more. As a screenwriter and director, Mehrǧuyi is thus used to writing down the spoken forms of words and sentences —an experience which might simply be prolonged into his literary works. Moreover, his professional status might influence the readership and make the introduced MSP forms more acceptable.

156 Of the novels written entirely with the use of features of spoken register, *Sang-e sabur* [The Patient Stone] by Sādeq Čubak should be mentioned. Published in Iran in 1966, it is a critique of the Iranian society of that time, narrated by the characters one by one in a technique resembling, but not entirely adhering to, the stream of consciousness. The reproduction of a colloquial and sometimes obscene language stirred conflicting reactions of the critics; some praised it, others criticized or even condemned it (Ferdowsi 2011).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

5.2.2 *Šahr-āšub*

In terms of the amount of MSP forms, *Šahr-āšub* ranks second after DXM, however, it does so with a 33% difference in their total number (1263 compared to 1886).

Unlike DXM, in ŠĀ most of the recorded occurences of MSP forms is in dialogues, while the language of the narrative parts is almost overly adjusted to the norms of MWP, especially as far as the syntactic rules (word order) are concerned. This tendency is well visible in the table of forms from the qualitative analysis, where entries are marked by a page and line number. Many of the dialogues in the analyzed part were short exchanges of a sentence or two, usually accompanied by an introductory phrase; introductory phrases were usually free of MSP features, which were on the contrary abundantly present in the dialogues themselves. It is not surprising then, that the collected forms often come from every other line of a page.

Among the 461 collected phonological parameters, 368 are alternations, especially raising of āN>uN (189). Alternation of *-ar>-e* is also frequent (62, more than other books), which is due to the frequent use of functional words *dige* (MWP *digar*) and *āxe* (MWP *āxar*) in their phatic function (devoid of their traditional meaning).

As far as the lexical parameters are concerned, ŠĀ includes colloquial phrases and sayings in amounts bigger than other analyzed books. Despite the overall number of lexical parameters (101) not exceeding their number in the other books (see Fig. 13), the number of sayings is twice the number from others (11 compared to 0–4 in other books). There are also frequent colloquial phrases and words, which however do not stand out among the general results.

The relative distribution of MSP parameters is represented by the chart below (Fig. 22).

Such a pattern of parameter distribution is almost strikingly similar to what can be seen in DXM (compare Fig. 19 in the previous section). There are small differences of distribution between syntactic and lexical features, but the overall scheme remains very close. The difference lays in the distribution of parameters throughout the text. While in DXM they can be found both in the narrative parts (1st person narrative allows for a direct recollection of the protagonist's experience, which in consequence make it possible to create all the narrative as a quasi-monologue, enriched by dialogues from time to time), in ŠĀ the MSP parameters (except the lexical intrusions: idiomatic and colloquial phrases) are limited to the dialogues. In this respect, it can be said that ŠĀ is not innovative at all—introducing the features of spoken register in the dialogues is a technique familiar to Persian writers since Hedāyat, Čubak or Ǧamālzāde.[157] Narrative parts

157 Already mentioned in the introductory chapter.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

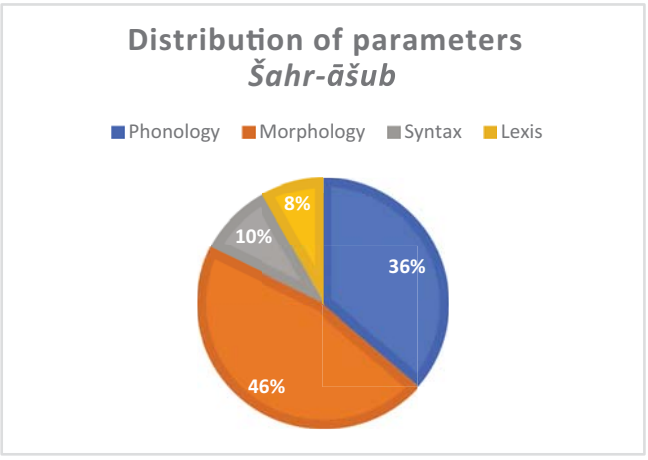

Figure 22. Distribution of parameters (ŠĀ) (source: own research).

are grammaticaly and syntactically conservative. It is even more interesting and meaningful then to see the correspondence of patterns of distribution between DXM and ŠĀ: both authors replicate the spoken register with the same means. It becomes clear then that for both authors, there exists a certain set of parameters with a certain relative frequency which is characteristic of MSP.

5.2.3 *Kafe Piyāno*

The number of MSP forms collected from *Kāfe Piyāno* is not significantly smaller than in *Šahr-āšub* (1106 compared to 1263), yet the style in which MSP forms are employed in those two books is radically different. While ŠĀ is consequently marking the difference between the narrative and the dialogue by the use of MWP and MSP syntactic features, in KP that difference is not to be felt. It is, of course, closely related to the narrative techniques applied: in ŠĀ, the reader is merely an observer, who learns the story from a distanced background information given by a third person narrator and overhears the dialogues of the characters; in KP the distance between the first person narrator and the reader is shortened, and the role of the reader is rather that of an active listener than a passive observant. This effect is achieved by the application of certain linguistic means, namely the MSP parameters which bring the language of the book closer to the actual everyday speech. Yet instead of exploiting the rich repertoire of the MSP register's phonological features (just as it was in the two previous books), KP employs the strategies characteristic of spoken language in general: lexical sparsity and grammatical intricacy typical for speech (Halliday 1985: 97, Roberts and Street

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

1998: 11). The role of phonological parameters is very limited (114 entries): they appear in the dialogues and (surprisingly) in chapter titles; it is the morphological and syntactic parameters that are the most frequent, as represented in the chart below (Fig. 23).

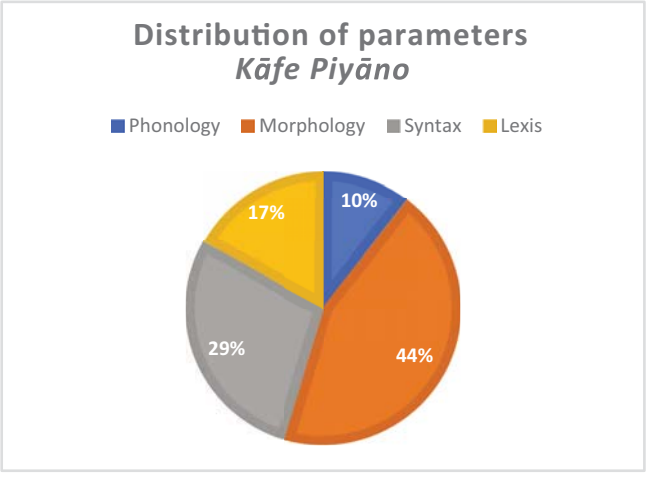

Figure 23. Distribution of parameters (KP) (source: own research).

Among the morphological parameters in question, the interesting and significant one—that has been excluded from the general classification—are the possessive/partitive enclitic pronouns, the number of which clearly exceeds the value for other books. Indeed, the use of EPs in almost all syntactic functions is one of the hallmark features of the KP's style, and their representation in all the categories exceeds that of other books, which is well represented by the chart below (figure 13 in chapter 5, section 5.1.2.7, repeated here for readers' convenience).

As far as the other morphological parameters are concerned, there is an outstanding number of progressives (22 past, 17 present; compare DXM 8/17, ŠĀ 7/16, ĀM 7/1, Hn 6/2) and two verb constructions (17, including the constructions with *dādan* and *kardan* as auxiliaries, see 5.1.2.6; compare DXM 6, ŠĀ 4, ĀM 4, Hn 3). The number of MSP verbal stems (27) and endings (51) is definitely lower that in DXM (188/471) and ŠĀ (145/239), although still higher than ĀM (4/2) and Hn (0/1). The number of MSP articles (especially indefinite *yek*) is also higher than elsewhere.

Lexically speaking, KP has a largest number of MSP prepositions (104, shared between *tu* for MWP *dar* 'in' and *vāse* for MWP *barā-ye* 'for; to compare, there are only 68 of those in DXM, 50 in ĀM, 25 in ŠĀ and just 5 in Hn). There are some colloquial words and phrases, but their number does not stand out from that of the other books. It is the enclitic pronouns, extensive use of determinatives,

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

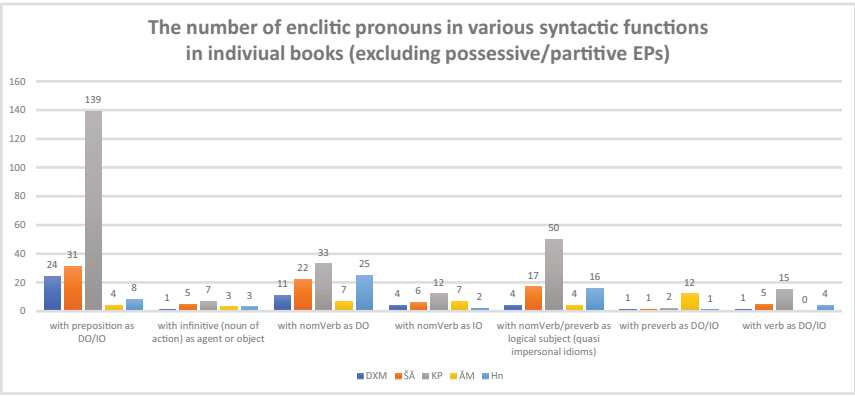

Figure 24. Syntactic functions of EPs in individual books (source: own research).

introduction of new auxiliaries (categorized as morphology>two-verb construction in the analysis), omission of auxiliaries and changes from the regular SOV word order that are responsible for the overall spoken character (grammatical intricacy) of the text.

It could be said that *Kāfe Piyāno*, in comparison with DXM and ŠĀ, does not simply employ the features of spoken register of Persian; rather than that, it is composed in a way which would represent spoken features in any language. While the MSP parameters of DXM and ŠĀ would pose a great difficulty in case of translation (and perhaps would be entirely lost in the process because of the lack of adequate measures in another language), KP should be able to retain its conversational character because of the universal character of the means employed there.[158]

5.2.4 *Ādat mikonim*

Ādat mikonim represents a third approach to the application of MSP characteristics into the literary text. It is neither phonological (as DXM and ŠĀ) not morphological (as KP); here the spoken register features surface mostly in syntax (255 parameters in comparison with 43 phonological, 175 morphological including 77 possessive/partitive EPs and 115 lexical ones).

As far as the representation of MSP phonology is concerned, Zoyā Pirzād's book is rather conservative. There are some dialogues in which there appear

158 Unfortunately, it is not possible to prove or disprove such hypothesis at present since none of those three books has a published English translation so far; and naturally other factors would be involved such as the skill of the translator. The hypothesis is of a purely linguistic background.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

words spelled according to their MSP pronunciation, moreover, there are also colloquial forms in those dialogues which belong to the *āmiyāne* style rather than to the MSP/*mohāvereyi* register. But those can be found in a dialogue of the middle-class protagonist with a lower-class construction worker. The dialogues between the protagonist and her mother (who sees herself as a representative of bourgeoisie) or best friend (middle-class) are usually free of the phonological features. When a teenage daughter of the protagonist has a say, she usually employs morphological and syntactic features of MSP more than the phonological ones as well. In general, 43 occurences of phonological parameters were listed, most of them being alternations of *-e>-i* (18), and contractions (16). In comparison to the previous books, ĀM represents a relatively small number of the instances of āN>uN raising (only 6), there is also no instance of a phonological alternation in DO marker *-rā*.

Among the morphological parameters, those connected to syntax appear much more frequently: thus there are almost none MSP verbal endings (2 overall) and stems (4) or articles (3). On the other hand, there is a frequent use of *ke* as an emphatic particle (12) or temporal conjunction (15, the biggest number among all the books). Also the emphatic particle *hā* was recorded 5 times, which is more than in any other book from the analysis.

As far as the syntax is concerned, the substantial part of the 255 parameters are changes of the word order (171 cases), especially in the locative meaning (133); omission of the auxiliaries in perfect tenses is also frequent (44 instances), and there are 30 clauses with omitted conjunctions, mostly coordinate (22 compared to 8 subordinate).

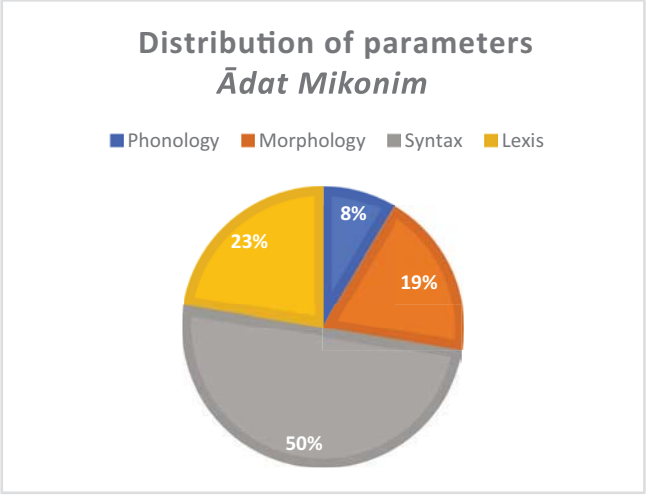

Figure 25. Distribution of parameters (ĀM) (source: own research).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

What is important in ĀM is the fact that syntactic features of MSP are distributed evenly throughout the text. There are frequent omissions of auxiliaries and SOV word order is often compromised, especially when the verbal phrase denotes movement or there is a prepositional phrase denoting location or direction—but not only there; there are also the nominal clauses which exhibit aberrations of the typical MWP word order.

As a result, as the first hand impression the reader might feel that s/he is in contact with a typical literary text in Modern Persian, neat and tidy in composition, adhering to all the expected norms. The MSP experience unfolds oly after a longer contact with the text, and it might pose certain difficulty to the reader to point where exactly does MSP surface. It is entirely different to the experience with DXM or ŠĀ, where pointing at MSP features in the text does not even require reading—it is visible at the first sight.

To summarize, ĀM could be said to represent a strategy opposite to DXM (see 5.2.1). In this book, the MSP phonology is almost invisible in spelling except for some dialogue parts (not even all the dialogues; compare 758 forms in DXM and 43 in ĀM). At the same time, it has the second rank in terms of forms representing syntactic deviations from MWP (255, compared with 147 in DXM). Although the relative number of lexical parameters is high, the language does not feel colloquialised (as it does in KP). The general interpretation of such distribution of parameters could be that the author intended to avoid the overt application of MSP patterns and adhere to the proper language of books (i.e., MWP), whereas the underlying layer of syntactic structures reveals that the language is not as much MWP as it seems at the first impression. It is not possible to judge whether that is the result of actual everyday language's influence or of an intricate and well-designed stylization, nevertheless the resulting style preserves the neat character of literary Persian.

5.2.5 *Hamnavāyi-ye šabāne-ye orkestr-e čubhā*

Among the analyzed five books, *Hamnavāyi-ye šabāne-ye orkestr-e čubhā* is undoubtedly the one with the least traceable influences of the MSP register. There are 262 MSP parameters overall. Only 6 of them are phonological, and those are the contractions of enclitic pronouns (2) and consonant clusters (3), plus one alternation of final *-e>-i.*

Among the 139 morphological parameters (the statistics shown in the chart below do not include the 178 possessive/partitive EPs), enclitic pronouns in the direct and indirect object function are frequent. There are almost no instances of MSP verbal endings or stems. On the other hand, the number of MSP particles and conjunctions (52) is higher than DXM (23), KP (39) or ĀM (40); only in ŠĀ

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

there are more (61). The use of progressives is on a level similar to ĀM (8 instances). The syntactic parameters are mostly the omission of conjunctions in subordinate sentences (17) and the reversed word order in sentences with a locative meaning (38; it is more than in DXM or ŠĀ, but less than KP and ĀM). 45 recorded lexical parameters are distributed between the MSP phrases (24), prepositions (5) and single words (16).

The distribution of parameters within categories is represented by the chart below:

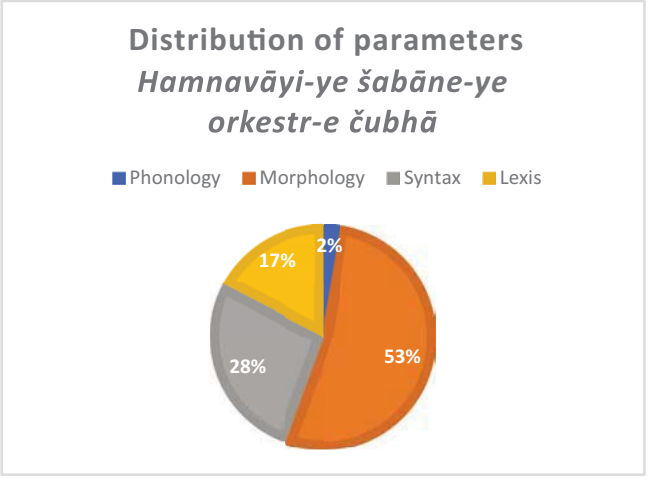

Figure 26. Distribution of parameters (Hn) (source: own research).

Interestingly, even though the differences in numbers are substantial, the distribution of the collected forms among grammatical categories is in fact similar to their distribution in *Kāfe Piyāno*, which was described in 5.2.3 as largely influenced by MSP. Similarly to the case of DXM and ŠĀ, where a coincidence of distributive patterns was observed in spite of the differences in style and application of MSP forms, also here it can be pointed out to the certain characteristics shared by those two books, seemingly very distant from each other.

Hamnavāyi is very conservative in terms of spelling and there are hardly any forms that are influenced by the way they are actually pronounced. On the other hand, the syntax is definitely influenced by that of spoken register, and also the enclitic pronouns do serve a variety of syntactic functions. Therefore, similar to KP, the general properties of spoken language such as grammatical intricacy appear here. The grammatical intricacy corresponds well to the generally intricate plot of the book. Even though it is told in 1st person and describes personal experience in a kind of monologue, Qāsemi prefers to use morphological and

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

syntactic means of suggesting that.[159] The gross number of MSP parameters in Hn is 262; it is almost three times less than just the phonological parameters (not to mention other categories) in DXM, so the difference is profound and without the analysis, one could suppose that Hn is free of the spoken register influence. After the analysis, it becomes clear that even in the case of a very limited influence, there are categories which are sensitive to MSP patterns and those are enclitic pronouns and changes in word order.

5.3 Patterns of parameter distribution

Analyzing the distribution of MSP parameters in the analyzed books, it becomes clear that it is by no means accidental. There might be various sources of influence, and authors' background and their specific purposes cannot be excluded.

Among the five books chosen for the analysis, a hierarchy of MSP-ization can be observed: each of them represents a different level, that is, a different number of MSP forms to be extracted from the text. Judging by the gross number of parameters, the books could be hierarchized as follows:

MSP

Dar xarābat-e moqan (1886)

Šahr-āšub (1263)

Kāfe Piyāno (1106)

Ādat mikonim (511)

Hamnavayi (262)

MWP

where DXM is the closest to MSP and Hn is the most conservative, the closest to the MWP standard. But, as we saw in the previous section, the way in which MSP parameters appear in those books is not uniform and their distribution between grammatical categories is not even. Therefore, a more complex approach is needed. Moreover, the distribution of parameters in the individual books shows certain similarities not connected with their overall numbers. The general observations are:

159 Rezā Qāsemi's background is tied to theatre, so it woud not be surprising to see him reproduce spoken forms more freely, such as Mehrǧuyi does in DXM. He does experiment with spoken register in his later works (see 4.2.5), but not in *Hamnavāyi*.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

a) the most frequent parameters are morphological,
b) there are two groups of morphological parameters: those influenced by phonology (verbal endings, verbal stems) and those influencing syntax (particles and conjunctions, enclitic pronouns in various syntactic functions, tenses),
c) high frequency of phonological parameters correlates with a low frequency of the syntactic ones, and high frequency of syntactic parameters correlates with a low frequency of the phonological ones,
d) when there is a general domination of phonological parameters, the number of the phonologically-influenced morphological parameters is higher, and the higher frequency of the syntactically-influenced parameters is observable in those books where phonological parameters are scarce and syntactic parameters dominate.

The lexical parameters behave differently—their number bears more relevance to the number of parameters in general than to their categorization.

The distribution of parameters between categories in individual books was presented on a chart in section 5.2 (Fig. 18), as well as in the pie charts in the subsequent sections. Compared with each other, they show three distinct patterns that can be distinguished in the analyzed material despite sometimes substantial differences in gross number of parameters. Two pairs undoubtedly share a distributional pattern of the parameters between categories; one book stands out and marks a pattern of its own.

Pattern 1 (DXM and ŠĀ) is characterized by the abundance of phonological parameters. The most of them are phonological alternations, and among alternations the most frequent is raising of *ā* before nasals. This observation is not surprising, as such a phonological environment is observable in the extremely frequent demonstrative *ān* (MSP *un*; not only on its own, but also in compounds). Changes in other frequent functional words and morphemes are observable: in the postposition *-rā* (MSP *-ro/-o*), words such as *digar* (MSP *dige*) or *āxar* (MSP *āxe*). The tendency to favorize phonology is visible also among the morphological parameters, where the most frequent are those which are formed as a result of phonological alternations too: verbal endings, verbal stems *et cetera*. A sharp decline is visible with the movement towards syntax and lexis, in which categories the number of parameters is almost equal.

This pattern could be labeled as "phonological," with an inclinations towards morphology.

Pattern 2 (KP and Hn) has a much lower share of phonological parameters, while the morphological parameters stand out; then there is a visible decline towards the syntactic parameters, continued into lexis. The difference between this pattern and the previous one lies in a sharp increase between phonology and

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

morphology, missing in DXM and ŠĀ. In both books, the frequency of enclitic pronouns is noticeable: the are abundant both in their basic function of expressing possession and in other syntactic functions, especially as direct and indirect object. Syntactic parameters such as a changed word order and omission of conjunctions are also frequent.

This pattern could be labeled as "morphological," with an inclination towards syntax.

Pattern 3, represented by *Ādat mikonim* only, is mainly based on the syntactic parameters. The phonological ones are scarce, and the morphological ones with a phonological background are similarly not represented too much. A growing tendency is visible from phonology through morphology until syntax, and then there is a decrease at lexis. Therefore, this pattern could be labeled "syntactic."

Interestingly, none of the analyzed books represented a pattern that could be labeled as "lexical" (i. e., with a domination of lexical parameters). It indicates that in fact, what lies beneath an act of MSP-ization of book's language are not colloquial phrases and idiomatic expressions, but the small yet significant grammatical decisions. This is an argument that supports differentiating between MSP register, which I decided to call with a Persian term *mohāvereyi*, and colloquial style, for which I suggest adapting Persian *āmiyāne*.

	Pattern 1	Pattern 2	Pattern 3
Visibility	overt	overt/covert	covert
Dominating category	morphology	morphology	syntax
Distinguishing category	phonology	syntax	syntax

Table 5. Patterns of MSP application (source: own research).

The described patterns express different attitudes towards understanding MSP. Pattern 1 uses the most conspicuous means, that is, the graphic representation of the MSP phonology, and matches it with a less innovative syntax. Pattern 2 is closer to the reproduction of the features of speech in general, albeit with the means available in Modern Spoken Persian repertoire. Both pattern 2 and pattern 3 avoid overt and evident declaration of phonological features, restraining themselves to the more subtle (i. e. less overt) syntactic means. It does not mean that the books in which patterns 2 and 3 are employed are somehow less MSP-influenced—in fact, it might be the opposite, since the MSP parameters found in them are rooted deeper than the superficial changes of spelling on the surface level. Whether introduced by the authors on purpose or unconsciously, those are more difficult to spot at the first sight or easily transform back to the "literary proper" MWP (although such a process is usually possible to some extent).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Chapter 6.
Final Conclusions

6.1 Revision of the hypotheses. MSP in literature

Repeating the conclusion from 5.3, the analysis showed that the presence of Modern Spoken Persian in literary texts can be proved and it is far from being accidental. There are patterns of application, which may be employed by the authors consciously or unconsciously, but in either case, they open the door for the "spoken" forms to enter the literary language.

The collected forms are not very innovative in character. Most of the parameters extracted from the analyzed books have already been described in the 1950s by Peisikov or Lazard, although in some of the cases forms described by them as singularities are slowly grammaticalized, just as it happens with the periphrastic progressive tenses. An interesting observation is the fact that some of the forms and patterns present in MSP (and not present in MWP, which is supposed to be not only a more conservative, but also an older form) strikingly resemble those of Middle Persian or Early New Persian (eradicated during the development of New Persian in 13th–19th century), which raises some doubt on their innovative character and would definitely require more study.

This goes in line with a general observation: those forms of MSP which are recorded in literature are mainly the forms already well familiarized and accepted by the users, grounded in their imaginary picture of the language. Therefore those are rather the core features than brand new, fresh changes, which will be more likely to be encountered in song lyrics (just as the case of negative periphrastic progressive in the Benyamin's song) or short captions somewhere in the digital sphere (for example Internet memes). Those core features, familiarized even more by their presentation in literature, which shapes people's perception of language, have a chance of being 'promoted' into the standard and seen as correct by the users.

The analysis was conducted with certain questions in mind and the results do provide the expected answers. The characteristic features of spoken Persian, i. e., the Modern Spoken Persian register, do appear in literary texts and are not

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

limited to the specified parts of it. Although it has long been customary to write dialogues in a style resembling actual speech, nowadays the MSP features can be found in the narrative parts as well. According to the research results, it is the grammatical category of the features that has more influence on their literary application than the surrounding context. Written rendering of phonological features is more likely to appear in dialogues, while syntactic features are those which can be found in narrative parts with a similar probability. The morphological features are more difficult to address, because there are two kinds of them: those with a phonological background (verbal endings, verbal stems) and those that are closer to the syntactic ones (the use of new auxiliaries and periphrastic constructions); and they behave similarly to the phonological and syntactic parametrs, respectively. Hence, it is more the parameters themselves than the textual circumstances which condition the appearance of MSP in modern prose.

The assumption that author's background plays a role in choosing a pattern of MSP application into the text seems plausible, but it would require further investigation to be proven. The example of Dariuš Mehrǧuyi and his *Dar xarābāt-e moqān* serves as a perfect example, since the author is an experienced film director and in the film industry, the reproduction of actual spoken language in screenwriting is a natural process. Such a process is less typical for the theatre, which might be one of the reasons behind a more conservative language of the novel *Hamnavāyi-ye šabāne-ye orkestr-e čubhā*, written by Rezā Qāsemi, experienced in theatrical activities. A journalistic experience of Farhād Ǧa'fari and his activity on a blog has definitely influenced the loose style of his *Kāfe Piyāno*.

One can get an impression that little has been said about the orthography of the MSP forms found in the material. It is because in the analyzed texts, it is rather consistent regardless of the author. If there are differences, they are usually connected with the bound and unbound spelling of certain morphological elements such as the durative prefix *mi-*, enclitic pronouns and particles, which are not limited to the MSP—the same problems are discussed in the MWP register (see Hashabeiky 2008). On a high level of generality, the fact that phonological characteristics of the MSP register are represented is an orthographical matter and the fact that they are found more in some books and less in the others is linked to the author's more or less restricted approach to traditional, standard spelling.

All the described observations of the roles of MSP register in literary writings and patterns of its use are significant. No less important is the classification of the varieties of modern Persian as dialects, styles and registers, in which it has been shown how the functional character of the differences between spoken and written varieties supports its treatment as a register on a high level of generality and not as an element of the style continuum. The most significant result of the analysis and the essential conclusion, however, is the fact that MSP

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

register can be diagnosed by the co-existence of certain parameters, and those parameters can be elicited from the text rather than assumed in advance. Therefore, by searching for the co-occurrence of those parameters in the texts other than those analyzed for this project, their level and pattern of MSP presence could be assessed. This opens the possibility for a wider study and a comparison of tendencies over time.

6.2 MSP parameters as linguistic variables, change in progress—historical reference

The history of Persian from Old Persian period until today is a history of developing diglossic relationships, resolved by the end of each period by the appearance of a new form, sharing the qualities of both diglossic varieties in one, then splitting into a diglossic relationship again. Of course, it is a far-reaching simplification, which should only be discussed with regard to the historical and sociolinguistic circumstances: the processes leading to the formation of Middle Persian after the Old Persian period were not the same as those in power when the New Persian emerged. Nevertheless, they share some common traits, as was explained in chapter 3.3. In both cases, some late writings (inscriptions) were discovered which included the alleged errors, and in both cases a significant historical event led to (or at least co-incided with) a transition to the next period in the history of language.

At the present moment, the differences between spoken and written registers are wide again. The registers do not change this much themselves, the characteristics of MSP register recorded in the scholarship of 1950s are still correct to a great extent. What does change is the relationship of the registers and their mixing environments. The parameters of the spoken register are observable in literature, and many of them are clearly a violation of rules and norms of the MWP register, considered a standard version of 21st century Modern Persian. Of course, the circumstances are completely different than those of the previous transitional periods: the wide access to education and low level of illiteracy enable both passive and active contact with the written word to the majority of society (although, naturally, to a varying extent). It is also facilitated by the digital means of communication available in the present times.

The studies of Labov and other variationists show that a phonological linguistic change can be studied while still in progress and suggest that similar studies might be undertaken for the morphological and syntactic linguistic change. As can be seen in the results of the analysis, such morphological and syntactic changes are in fact observable for Modern Persian. Therefore as it has

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

been stated in chapter 3, it should be possible to treat those equivalents as linguistic variables (that is, find a MWP equivalent for each MSP parameter included in the analysis) and compare their frequencies by the means of LVC methodology. An analysis of the occurrence of variables in certain contexts could support the hypothesis of spoken register gaining more power over written register over the course of time. The MSP parameters visible in the literary texts could serve a function parallel to the "mistakes" found in the late Old Persian and late Middle Persian texts and mark a step forward in a transitional period towards a next stage in the development of Persian.

6.3 Ideas for future research

This study was designed to assess the presence of spoken register forms in the contemporary Persian prose on a limited sample of five novels. Setting such a limit allowed for a close, in-depth analysis of the language employed in the studied material and enabled the elicitation of spoken register marking features "from below," instead of presupposing a list of features beforehand, "from above." Those features, here labeled "parameters," were common in all the analyzed material, they were also coherent with the features of this register described in the grammars and linguistic analyses of Modern Persian. Therefore they could serve as a tool in further research of the interplay between Modern Spoken and Written Persian.

It would definitely be interesting to repeat a quantitative study on a bigger corpus, including a variety of genres, and see if there is a correlation between genre and the pattern of parameter distribution. Such a study would not have to be limited to the literary language—the results would benefit greatly from an addition of press articles or even samples of the everyday electronic communication (text messages, social media posts, blog entries), where no special education or skill is required in order to create and publish any content. Moreover, a comparative study of the literary language from earlier (e. g. 1990s) and later (2020s) periods would shed some light on the development of the researched phenomenon, thus contributing to the research of potential change in progress. The ongoing development of technology should facilitate a similar study covering a wider text sample and drawing from a bigger corpus.

Apart from a general study of the changing relations between spoken and written registers of Modern Persian, many of the parameters would benefit greatly from an individualized study of their development over time. Among the interesting topics, there are new auxiliaries, changes in syntax, grammaticalization of periphrastic progressives, all of which would require further study. Also the phonological changes observable during the transformation of MWP forms

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

into MSP and back could be analyzed separately, especially in the vein of the LVC methodology. Further investigation of any of those topics would advance the study of contemporary Persian and help capture changes in the language right now, exactly when they are taking place, therefore creating a rare opportunity to see a language of which all three previous stages of development are known, during the period which with much probability is in fact a transition to another stage.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

List of figures and tables

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

References

Source texts

Mehrǧuyi, Dāryuš. 2010. *Dar xarābāt-e moqān* [In the tavern], 3rd edition. Tehrān: Našr-e Qatre.

Ǧa'fari, Mariam. 2004. *Šahr-āšub* [Femme-fatale], 2nd edition. Tehrān: Šādān.

Ǧa'fari, Farhād. 2007. *Kāfe Piyāno* [Café Piano], 35th edition. Tehrān: Našr-e Češme.

Pirzād, Zoyā. 2004. *Ādat mikonim* [We'll get used to it], 14th edition. Tehrān: Našr-e Markaz.

Qāsemi, Rezā. 2001. *Hamnavāyi-ye šabāne-ye orkestr-e čubhā* [Nocturnal harmony of the orchestra of woods], 12th edition. Tehrān: Nilufar.

Theoretical references, critical literature and others

EI = Encyclopaedia Iranica

– *Encyclopædia Iranica*, vol. I–, London, 1982–.

– *Encyclopædia Iranica*, online edition, New York, 1996–present.

OED = The Oxford English Dictionary

Afxami, Ali and Ziyā ad-Din Qāsemi. 2010. Sabkhā-ye zabāni dar zabān-e fārsi va namud-e adabi [Styles of Persian language and their literary examples]. *Pažuheš-e zabānhā-ye xāreǧi,* 55(2010), pp. 5–17.

Ahmadi-Givi, Hasan and Hasan Anvari. 1385h.š. [=2007]. *Dastur-e zabān-e fārsi* [Persian grammar]. Tehrān: Entešārāt-e Fātemi.

Alexiou, Margaret. 1982. Diglossia in Greece [in:] William Haas (ed.) *Standard Languages: Spoken and Written*, Manchester: Manchester University Press, pp. 156–192.

Anvari, Hasan. 2007. *Farhang-e fešorde-ye soxan.* [The concise Persian dictionary]. Tehran: Soxan.

Beeman, William O. 1976. Status, Style and Strategy in Iranian Interaction. *Anthropological Linguistics*, vol. 18, no. 7 (Oct.), pp. 305–322.

Idem. 1986. *Language, Status and Power in Iran.* Bloomington: Indiana University Press.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Idem. 2005. *Persian, Dari and Tajik in Central Asia.* Seattle: The National Council for Eurasian and East European Research, Title VIII Program. Online: http://www.ucis.pitt.edu/nceeer/2005_817_17g_Beeman.pdf (retrieved 29. 10. 2021).

Biber, Douglas. 2004. Conversation Text Types: A Multi-dimensional Analysis. *Journées internationales d'Analyse statistique des Données Textuelles,* no. 7.

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

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

Bloomfield, Leonard. 1933. *Language.* Chicago: University of Chicago Press.

Boyle, John Andrew. 1966. *Grammar of Modern Persian.* Wiesbaden: Otto Harrassowitz.

Britto, Francis. 1986. *Diglossia: A Study of the Theory with Application to Tamil.* Washington, D.C.: Georgetown University Press.

Browne, Edward G. 1908. *A Literary History of Persia.* London: T. Fisher Unwin.

Calmard, Jean. 1991. Chodźko, Aleksander Borejko [in:] *EI*, vol. V, fasc. 5, pp. 502–504.

Chodźko, Aleksander. 1852. *Grammaire persane ou principes de l'iranien modern.* Paris: Imprimerie nationale.

Chomsky, Noam. 1957. *Syntactic Structures.* The Hague: Mouton.

Chwilczyńska-Wawrzyniak, Monika. 1998. *Język perski* [Persian language]. Warszawa: Dialog.

Clyne, Michael. 1997. Multilingualism [in:] Florian Coulmas (ed.) *The Handbook of Sociolinguistics.* Oxford: Blackwell, pp: 301–314, DOI: 10.1002/9781405166256.ch18.

Coulmas, Florian (ed.). 1997. *The Handbook of Sociolinguistics.* Oxford: Blackwell, DOI: 10.1002/9781405166256.

Crystal, David. 2008. *A Dictionary of Linguistics and Phonology* (6th edition). Oxford: Blackwell.

Crystal, David and Derek Davy. 1969. *Investigating English Style.* New York: Longman.

Dabir-Moghaddam, Mohammad. 2012. Linguistic Typology: An Iranian Perspective. *Journal of Universal Language* 13(1), pp. 31–70.

Idem. 2018. Typological Approaches and Dialects [in:] A. Sedighi and P. Shabani-Jadidi (eds.) *The Oxford Handbook of Persian Linguistics.* Oxford: Oxford University Press, pp. 43–138.

Dehqān, Ali and Fāteme Hamāmči. 2016. Zabān-e 'āmiyāne dar āsār-e Ali Mohammad Afqāni [Colloquial language in the writings of Ali Mohammad Afqāni]. *Faslnāme-ye taxassosi-ye zabān-o-adabiyāt-e fārsi-ye Dānešgāh-e Āzād-e Eslāmi-ye Mašhad,* no. 3/1394, pp. 51–73.

Denison Ross, Edward. 1931. Modern Persian Colloquial. *Bulletin of the School of Oriental Studies, University of London*, vol. 6, no. 3, pp. 601–615, available online at http://www.jstor.org/stable/607196 [retrieved 17. 04. 2018].

Eckert, Penelope and John R. Rickford (eds.). 2001. *Style and Sociolinguistic Variation.* Cambridge: Cambridge University Press.

Ellis, J.M. 1970. Linguistics, Literature, and the Concept of Style. *Word*, vol. 26 no.1, pp. 65–78.

Elwell-Sutton, Lawrence P. 1941. *Colloquial Persian.* London: Routledge.

Enkvist, Nils E. 1986. What Has Discourse Linguistics Done to Stylistics? [in:] Simon P. X. Battestini (ed.) *Developments in Linguistics and Semiotics: Language Teaching and*

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Learning, Communication Across Cultures (GURT 1986). Washington, DC: Georgetown University Press, pp. 19–36.

Ervin-Tripp, Susan. 1972. *On Sociolinguistic Rules: Alternation and Co-Occurrence* [in:] John Gumperz and Dell Hymes (eds.) *Directions in Sociolinguistics.* New York: Holt, Rinehart and Winston, pp. 213–50.

Eslāmzāde, Sa'id. 1387h.š. [= 2009]. Zabānšenāsi va guyeshā-ye irāni dar goftogu bā doktor Irān-e Kalbāsi [Linguistics and Iranian dialects in an interview with dr Irān Kalbāsi]. *Pažuhešgarān,* no.14–15, pp. 29–37.

Eyyubi, Karim R. and Iraida A. Smirnova, *Kurdskiy dialect mukri.* Leningrad 1968.

Ferguson, Charles. 1958. Word Stress in Persian. *Language*, vol. 33 no. 2, pp. 123–135.

Idem. 1959. Diglossia. *Word,* no. 15, pp. 325–340.

Idem. 1971. *Language Structure and Language Use: Essays.* Stanford: Stanford University Press.

Idem. 1973. Language Problems of Variation and Repertoire. *Daedalus*, vol. 102, no. 3 *Language as a Human Problem*, pp. 37–46.

Idem. 1994. Dialect, Register, and Genre: Working Assumptions About Conventionalization [in:] D. Biber and E. Finegan (eds.) *Sociolinguistic Perspectives on Register.* New York and Oxford: Oxford University Press, pp. 15–30.

Idem. 1996. *Sociolinguistic Perspectives*, ed. by Thom Huebner, New York and Oxford: Oxford University Press.

Idem. 1996. Epilogue: Diglossia Revisited [in:] A. Elgibali, Alaa (ed.) *Understanding Arabic: Essays in Contemporary Arabic Linguistics in Honor of El-Said Badawi.* Cairo: American University in Cairo Press, pp. 49–67.

Ferguson, Charles and Afia Dil. 1979. The Sociolinguistic Variable (s) in Bengali: A Sound Change in Progress? [in:] Charles Ferguson. 1996. *Sociolinguistic Perspectives*, ed. by Thom Huebner, New York and Oxford: Oxford University Press, pp. 181–188.

Ferdowsi, Ali. 2011. Sang-e sabur [in:] *EI Online,* available at: https://iranicaonline.org/articles/sang-e-sabur [retrieved 26.12.2021].

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

Finegan, Edward and Douglas Biber. 2001. Register Variation and Social Dialect Variation: the Register Axiom [in:] Penelope Eckert and John R. Rickford (eds.) *Style and Sociolinguistic Variation.* Cambridge: Cambridge University Press, pp. 235–267.

Fishman, Joshua. 1967. Bilingualism with and without Diglossia; Diglossia with and without Bilingualism. *Journal of Social Issues*, vol. 23(2), pp. 29–38.

Idem. 1980. Bilingualism and Biculturism as Individual and as Societal Phenomena. *Journal of Multilingual and Multicultural Development*, vol. 1, pp. 3–15.

de Fouchécour, Charles-Henri. 2006. *Classical Persian Literature* [in:] *EI*, vol. XIII, fasc. 4, pp. 414–432, available online at: https://www.iranicaonline.org/articles/iran-viii2-classical-persian-literature [retrieved 26.11.202].

Fowler, Roger. 1988. Studying Literature as Language. *Revista alicantina de estudios ingleses*, no. 1, pp. 81–90.

Ğamālzāde, Mohammad-Ali. 1921/1966. *Dibāče* [Introduction] [in:] idem, *Yeki bud, yeki nabud*, 7th edition. Tehrān 1966, pp. 3–21; 1st edition, Berlin 1921.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Ghayoomi, Masood, Momtazi, Saeedeh and Mahmood Bijankhan. 2010. A Study of Corpus Developoment for Persian. *International Journal on Asian Language Processing,* vol. 20 no. 1, pp. 17–33.

Gheytanchi, Elham. 2011. *Čerāqhā rā man xāmuš mikonam* [in:] *EI*, available online at https://iranicaonline.org/articles/cheragh-ha-ra-man [retrieved 05. 05. 2021].

Gignoux, Philippe. 1984. *Le Livre d'Arda Viraz. Translittération, transcription et traduction du texte pehlevi.* Paris: Recherche sur les Civilisations.

Gumperz, John. 1962. Types of Linguistic Communities. *Anthropological Linguistics* 4(1), pp. 28–40.

Hāfez. 1384h.š. [=2006]. *Divān* [A collection of poetry]. Tehrān: Ǧeyhun.

Haggard, William H.D. and Guy Le Strange. 1882. *The Vazir of Lankurān. A Persian Play. A text-book of Modern Colloquial Persian for the Use of European Travellers, Residents in Persia, and Students in India.* London: Trübner & Co., available online in public domain at https://hdl.handle.net/2027/nyp.33433081879946 [retrieved: 08. 11. 2021].

Halliday, Michael A.K. 1971. Linguistic Function and Literary Style: An Inquiry into the Language of William Golding's "The Inheritors" [in:] S. Chatman (ed.) *Literary Style: A Symposium.* London and New York: Oxford University Press, pp. 330–365.

Idem. 1989. *Spoken and Written Language.* Oxford: Oxford University Press.

Hashabeiky, Forogh. 2005. Persian Orthography. Modification or Changeover? (1850–2000). *Acta Universitatis Upsaliensis. Studia Iranica Upsaliensia 7,* Uppsala: Acta Universitatis Upsaliensis.

Haugen, Einar. 1966. Dialect, Language, Nation. *American Anthropologist,* vol. 68, no. 4, pp. 922–935.

von Heusinger, Klaus and Roya Sadeghpoor. 2020. The specificity marker *-e* with indefinite noun phrases in Modern Colloquial Persian [in:] K. Balogh, A. Latrouite and R. D. Van Valin Jr. (eds.), *Nominal anchoring: Specificity, definiteness and article systems across languages.* Berlin: Language Science Press, pp. 115–147. DOI: 10.5281/zenodo.4049685

Hodge, Carleton T. 1957. Some Aspects of Persian Style. *Language*, vol. 33 no. 3, pp. 355–369.

Horn, Paul. 1895–1901. *Neupersischen Schriftsprache* [in:] W. Geiger and E. Kuhn (eds.) *Grundriss der iranischen Philologie*, Bd. 1, Abt. 2, Strassburg: Verlag von Karl J. Trübner, pp. 1–200.

Hudson, Alan. 1994). *Diglossia as a Special Case of Register Variation* [in:] D. Biber and E. Finegan (eds.) *Sociolinguistic Perspectives on Register.* Oxford: Oxford University Press, pp. 294–314.

Idem. 2002. Outline of a Theory of Diglossia. *International Journal of Society and Language,* no. 157, pp. 1–48.

Hymes, Dell. 1974. *Foundations in Sociolinguistics: An Ethnographic Approach.* Philadelphia: University of Pennsylvania Press.

Ibrahim, Muhammad Mirza. 1841. *A Grammar of the Persian Language.* London.

Irvine, Judith T. 2001. "Style" as distinctiveness: The culture and ideology of linguistic differentiation [in:] Penelope Eckert and John R. Rickford (eds.) *Style and sociolinguistic variation.* Cambridge: Cambridge University Press, pp. 21–43.

Jahangiri, Nader. 1980. *A Sociolinguistic Study of Tehrani Persian.* (Doctoral Dissertation, University of London).

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Jakobson, Roman. 1960. Closing Statements: Linguistics and Poetics [in:] Thomas A. Sebeok (ed.) *Style In Language.* Cambridge Massachusetts: MIT Press, pp. 350–377.

Jazayery, Mohammad Ali. 1966. Western influence in Contemporary Persian: A General View. *Bulletin of the School of Oriental and African Studies,* vol. 29, no. 1, London: Cambridge University Press, pp. 79–96.

Idem. 1970. Observations on Stylistic Variation in Persian [in:] *Actes du Xe Congres International des Linguistes.* vol. III. Bucharest: Editions de l'Academie de la Republique Socialiste de Roumaine, pp. 447–457.

Jeremiás, Éva. 1984. Diglossia in Persian. *Acta Linguistica Academiar Scientiarum Hungaricae*, tomus 34 (3–4), pp. 271–281.

Eadem. 2012. Matthew Lumsden's 'Persian Grammar' (Calcutta, 1810). Part I. *Iran*, vol. 50 (2012), pp. 129–140).

Jones, William. 1771. *Grammar of the Persian Language.* London: W. & J. Richardson.

Kalbāsi, Irān. 1380 h.š. [=2002]. Fārsiye goftāri va neveštāri [Spoken and written Persian]. *Farhang*, no. 37–38, pp. 49–68.

Eadem. 1381h.š. [=2003]. Lahǧe-ye tehrāni, fārsi-ye goftāri, yā fārsi-ye 'āmiyāne? [Tehrani dialect, spoken Persian or colloquial Persian?]. *Ǧahān-e Ketāb*, no. 15–16/7, p. 10.

Kamshad, Hassan and Nahid Mozaffari. 2016. Jamalzadeh-Mohammad-Ali ii, [in:] *EI*, vol. XIV, fasc. 4, pp. 444–448 and vol. XIV, fasc. 5, pp. 449–452.

Kerswill, Paul. 2000. Koineization and accommodation [in:] J. K. Chambers, P. Trudgill, N. Schilling-Estes (eds.) *The Handbook of Language Variation and Change.* Oxford: Blackwell, pp. 669–702.

Kieślowski, Krzysztof. 2017. *Zendegi-am hame čiz-e man ast: mağmu'e-ye goftegu-hā bā Kišlufski*, transl. by A. Marcinowska. Tehrān: Qoqnoos.

Labov, William. 1966. *The Social Stratification of English in New York City.* Cambridge: Cambridge University Press.

Idem. 1972. *Sociolinguistic Patterns.* Philadelphia: University of Pennsylvania Press.

Idem. 1994 [vol. I], 2001 [vol. 2], 2010 [vol. III]. *Principles of Linguistic Change.* Oxford: Blackwell.

Idem. 2001. The Anatomy of Style-shifting [in:] Penelope Eckert and John R. Rickford (eds.) *Style and Sociolinguistic Variation.* Cambridge: Cambridge University Press, pp. 85–108.

Lambton, Ann K.S. 1953. *Persian Grammar.* Cambridge: Cambridge University Press.

Latham, J. Derek. 1997. Ebn Al-Moqaffaʿ, Abū Moḥammad ʿAbd-Allāh Rōzbeh [in:] *EI*, vol. VIII, fasc. 1, pp. 39–43, available online at http://www.iranicaonline.org/articles/ebn-al-moqaffa [retrieved: 08. 11. 2021].

Lazard, Gilbert. 1963. *La Langue des plus anciens monuments de la prose persane.* Paris: Klincksieck.

Idem. 1975. Rise of New Persian Language [in:] *Cambridge History of Iran*, vol. IV. Cambridge: Cambridge University Press, pp. 595–632.

Idem. 1993. *The Origins of Literary Persian,* Washington, D.C.: The Foundation for Iranian Studies.

Idem. 2003. Du pehlevi au persan: diachronie ou diatopie? [in:] Paul, Ludwig (ed.) *Persian Origins – Early Judaeo-Persian and the Emergence of New Persian. Collected Papers of the Symposium, Göttingen 1999.* Wiesbaden: Harrasowitz Verlag, pp. 95–102.

Lazard, Gilbert and M. Ghavam-Nejad. 1990. *Dictionnaire persan-français.* Leiden.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Levy, Reuben. 1951. *The Persian Language.* New York: Philosophical Library.
Mace, John. 2003. *Persian Grammar.* London and New York: Routlege Curzon.
Maggi, Mauro and Paola Orsatti. 2018. From Old to New Persian [in:] A. Sedighi and P. Shabani-Jadidi (eds.) *The Oxford Handbook of Persian Linguistics.* Oxford: Oxford Univeristy Press, available online, DOI: 10.1093/oxfordhb/9780198736745.013.2 [retrieved 10. 01. 2021].
Mahmoodi-Bakhtiari, Behrooz. 2018. Spoken vs. Written Persian: Is Persian Diglossic? [in:] A. Korangy and C. Miller (eds.) *Trends in Iranian and Persian Linguistics.* Berlin/ Boston: The Gruyter Mouton, pp. 183–212.
Majidi, Mohammad Reza. 1986. *Strukturelle Grammatik des Neupersischen (Fārsi),* vol. I–III. Hamburg.
Malinowski, Bronislaw. 1923. The Problem of Meaning in Primitive Languages. Supplementary essay [in:] C. K. Ogden and I. A. Richards (eds.) *The Meaning of Meaning.* New York: Harcourt, Brace, pp. 296–336.
Mann, William C. and Christian Matthiessen. 1991. Functions of language in two frameworks. *Word,* no. 42 vol. 3, pp. 231–249.
Matini, Jalal and Margaret Caton. 1986. ʿĀref Qazvini [in:] *EI*, vol. II, fasc. 4, pp. 391–392, available online at: https://www.iranicaonline.org/articles/aref-qazvini-poet [retrieved 20. 12. 2021].
Megerdoomian, Karine. 2018. Computational Linguistics [in:] A. Sedighi and P. Shabani-Jadidi (eds.) *The Oxford Handbook of Persian Linguistics.* Oxford University Press: Oxford, Chapter 19, pp. 461–480.
Mękarska, Barbara. 2005. Odmiany języka perskiego uwarunkowane historycznie i regionalnie. *Bulletin de la Société Polonaise de Linguistique,* fasc. LXI, pp. 5–11.
Milroy, James and Leslie Milroy. 1997. Varieties and Variations [in:] F. Coulmas (ed.) *Handbook of Sociolinguistics.* Oxford: Blackwell, pp. 45–64, DOI: https://doi.org/10.1002/9781405166256.ch3.
Modarresi, Yahya. 2018. Sociolinguistics [in:] A. Sedighi and P. Shabani-Jadidi (eds.) *The Oxford Handbook of Persian Linguistics.* Oxford University Press: Oxford, Chapter 13, pp. 329–346.
Moises, Edward. 1792. *The Persian Interpreter: in Three Parts.* Newcastle: Printed by L. Hodgson.
Naǧafi, Abol-Hassan. 1378 h.š. [=2000]. *Farhang-e fārsi-ye ʿāmiyāne* [Dictionary of colloquial Persian]. Tehrān: Nilufar.
Nematollahi, Narges. 2018. HAVE-progressive in Persian: a case of pattern replication? *Diachronica* 35:1, pp. 144–156, available online at: https://doi.org/10.1075/dia.17023.nem [retrieved 20. 12. 2021].
Ong, Walter J. 2002. *Orality and Literacy. The Technologizing of the World.* London and New York: Routledge (first published in 1982).
Orsatti, Paola. 2015. Spoken Features in Classical Persian Texts: subordinate conditional clauses without a conjunction [in:] I. Szántó (ed.) *From Aṣl to Zā'id: Essays in Honour of Éva M. Jeremiás.* Piliscsaba, pp. 155–165.
Paul, Ludwig. 2000/2013. Persian Language. i. Early New Persian [in:] *EI Online*, available online at: https://iranicaonline.org/articles/persian-language-1-early-new-persian [retrieved 10. 01. 2022].

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Idem. 2002. A Linguist's Fresh View on 'Classical Persian' [in:] M. Szuppe (ed.) *Iran. Questions et connaissances. Actes du IVe Congrès Européen des Études Iraniennes, organisé par la Societas Iranologica Europaea, Paris, 6–10 Septembre 1999 II. Périodes médiévale et moderne.* Paris, pp. 21–34.

Idem. 2003. Early Judaeo-Persian in a Historical Perspective: The Case of the Prepositions bē, o, pa(d), and the Suffix -rā [in:] Ludwig Paul (ed.) *Persian Origins – Early Judaeo-Persian and the Emergence of New Persian. Collected Papers of the Symposium, Göttingen 1999.* Wiesbaden, pp. 177–94.

Peisikov, Lazar. 1380h.š. [=2002]. *Lahǧe-ye tehrāni* [Tehrani dialect], transl. by Mohsen Šoǧā'i. Tehrān: Farhangestān-e Zabān-o Adab-e Fārsi.

Pellò, Stefano. 2016. A Linguistic Conversion: Mirza Muhammad Hasan Qatil and the Varieties of Persian (ca. 1790) [in:] Stefano Pellò (ed.) *Borders: Itineraries on the Edges of Iran.* Venezia: Edizioni Ca' Foscari – Digital Publishing, pp. 203–240. DOI: 10.14277/6969-100-3/EUR-5-9 [retrieved: 02. 11. 2021].

Perry, John R. 1999. Comparative Perspectives on Language Planning in Iran and Tajikistan [in:] Yasir Suleiman (ed.) *Language and Society in the Middle East and North Africa. Studies in Variation and Identity.* Curzon, pp. 154–174.

Idem. 2000. Epistemic verb forms in Persian of Iran, Afghanistan and Tajikistan [in:] L. Johanson and B. Utas (eds.) *Evidentials. Turkis, Iran and Neighbouring Languages.* Berlin, New York: Mouton de Gryuter, pp. 229–257.

Idem. 2003. Persian as a Homoglossic Language. *Studia Iranica*, no. 27, pp. 11–28.

Idem. 2012. New Persian: Expansion, Standardization, and Inclusivity [in:] B. Spooner and W. Hanaway (eds.) *Literacy in the Persianate World: Writing and the Social Order.* Philadelphia: Museum Publications, pp. 70–94.

Phillot, Douglas Craven. 1919. *Higher Persian Grammar for the Use of Calcutta University, Showing Differences Between Afghan and Modern Persian, with Notes on Rhetoric.* Calcutta: Baptist Mission Press.

Pirzād, Zoyā. 2009. "Je recherche la simplicité et la justesse." Interview by Hamdam Mostafavi. *Courrier International*, 03 November 2009, https://www.courrierinternational.com/article/2009/10/30/je-recherche-la-simplicite-et-la-justesse [retrieved: 06. 11. 2021].

Pisowicz, Andrzej. 1985. *Origins of the New and Middle Persian Phonological Systems*, transl. by Barbara Buja-Stanek. Kraków: WUJ.

Idem. 2003. The possibile influence of Kurdish dialects on colloquial Persian (the definiteness suffix -é). *Folia Orientalia* 39, pp. 177–180.

Polański, Kazimierz (ed.). 1993. *Encyklopedia językoznawstwa ogólnego* [Encyclopaedia of general linguistics]. Wrocław–Warszawa–Kraków: Ossolineum.

Poświatowska, Halina. 2015. *Ādamak, matarsak va āšeq*, transl. by K. Wąsala and Z. Qāsemi.Tehrān: Sarzamin-e Āhurāyi.

Qāsemi, Rezā. 2018. *Nohomin ǧalase-ye Raviun – Rezā Qāsemi va ketāb-e* Hamnavāyi-ye šabāne-ye orkestr-e čubhā [9th Meeting of Raviyun – Rezā Qāsemi and the book *Nocturnal Harmony of the Orchestra of Woods*], available at: www.youtube.com/watch?v=4smj7j35hUY [published 04. 04. 2018, retrieved 22. 04. 2021].

Rafiee, Abdi. 2001. *Colloquial Persian. The Complete Course for Beginners.* London, New York: Routledge.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Rahimi, Bahman. 2003. *Šāyad fardā, be gune-yi digar… Moqaddame* [Maybe tomorrow, in a different way… Preface], [in:] Ğa'fari, Maryam. 2003. *Šahr-āšub* [Femme fatale]. Tehrān: Šādān, pp. 5–6.

Rastorgueva, Vera S. 1953. Kratkij očerk grammatiki persidskogo âzyka [in:] Boris V. Miller (ed.) *Persidsko-russkij slovar.* Moscow: Gosudarstvennoe Izdatel'stvo Inostrannyh i Nacional'nyh Slovarej.

Idem. 1966. *Sredne-persidskij jazyk.* Moscow: Izdatel'stvo Nauka.

Rezāyi, Vāli. 1384h.š. [=2006]. Zabān-e me'yār čist va če vižegihāyi dārad? [What is standard language and what characterizes it?]. *Nāme-ye Farhangestān,* no. 3/6, pp. 20–36.

Roberts, Celia and Brian Street. 1998. Spoken and Written Language [in:] F. Coulmas (ed.) *The Handbook of Sociolinguistics.* Oxford: Blackwell, DOI: 10.1002/9781405166256.ch10.

Romaine, Suzanne. 1995. *Bilingualism.* Oxford, Cambridge: Wiley-Blackwell.

Rossi, Adriano V. 2015. Diglossia in Persian [in:] I. Szántó (ed.) *From Aṣl to Zā'id: Essays in Honour of Éva Jeremiás.* Piliscsaba: The Avicenna Institute of Middle Eastern Studies, pp. 211–219.

Rzehak, Lutz. 2001. *Vom Persischen zum Tadschikischen. Sprachliches Handeln und Sprachplanung in Transoxanien zwischen Tradition, Moderne und Sowjetmacht (1900–1956).* Wiesbaden: Reichert Verlag.

Idem. 2015. *Dari – the Persian Language of Afghanistan.* A lecture delivered at the Department of Iranian Studies, Jagiellonian University, Cracow on 5th October 2015.

Idem. 2016. Recent Developments in Afghanistan [in:] H. Hock and E. Bashir (eds.), *The Languages and Linguistics of South Asia.* Berlin/Boston: De Gruyter, pp. 292–299.

Sādeqi, Ali Ašraf. 1356h.š. [=1978]. *Takvin-e zabān-e fārsi* [The emergence of Persian language]. Tehrān: Dānešgāh-e Āzād.

Idem. 1362h.š. [=1984]. Zabān-e me'yār [Standard language]. *Našr-e dāneš*, no. 4/3, pp. 16–21.

Idem. 1986. Arabic Language. i. Arabic elements in Persian [in:] *EI* vol. II, fasc. 3, pp. 229–231.

Sārli, Nāser Qoli (et al.). 1396h.š. [=2017]. Moqāyese-ye me'yār-sāzi-ye zabān-e fārsi dar dourehā-ye piš o pas az Mašrute [A comparison of standardization of Persian before and after the Constitutional Period]. *Ğostārhā-ye zabāni,* no. 5(40)/8, pp. 205–234.

Sadat-Tehrani, Nima. 2003. The indifference *ke* construction in modern Persian. *Linguistica Atlantica* 21, pp. 137–151.

Samā'i, Mehdi. 1382h.š. [=2003]. *Farhang-e loqāt-e zabān-e maxfi* [A Persian Dictionary of Argot]. Tehrān: Našr-e Markaz.

Samare, Yadollāh. 1966h.š. [=1987/88] *ĀZFĀ. Āmuzeš-e zabān-e fārsi* [Persian Language Teaching]. 5 volumes. Tehrān: International Relations Department.

Sami'i, Ahmad. 1378h.š. [=2000]. *Negāreš va virāyeš* [Writing and editing].Tehrān: Samt.

Sankoff, David and Pierrette Thibault. 1981. Weak complementarity: Tense and Aspect in Montreal French [in:] J. B. Strong and D. Strong (eds.), *Syntactic Change.* Ann Arbor: University of Michigan, pp. 206–216.

de Sausurre, Ferdinand. 1959. *Course in General Linguistics*, ed. by Charles Bally and Albert Sechehaye, transl. by Wade Baskin. New York: Philosophical Library.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Idem. 1996. Saussure's First Course of Lectures on General Linguistics (1907). From the notebooks of Albert Riedlinger, ed. and transl. by Eisuke Komatsu and George Wolf. Oxford, New York, Seoul, Tokio: Pergamon.

Sayyedyazdi, Zahrā and Farzāne Hakimi. 1395h.š. [=2016]. Zabān-e 'āmiyāne dar *Firuzšāh-nāme* [Colloquial language in *Firuzšāh-nāme*]. *Motāle'āt-e zabāni-bolāqi,* no. 13/7, pp. 98–114.

Schiffman, Harold. 1997. *Diglossia as a Sociolinguistic Situation* [in:] F. Coulmas (ed.) *The Handbook of Sociolinguistics*, Oxford: Blackwell, pp. 205–216, DOI: 10.1002/9781405166256.ch12.

Sedighi, Anousha. 2010. *Agreement Restrictions in Persian.* Leiden: Leiden University Press.

Sedighi, Anousha and Pouneh Shabani-Jadidi (eds.). 2018. *The Oxford Handbook of Persian Linguistics.* Oxford: Oxford University Press.

Seraji, Mojgan. 2015. Morphosyntactic Corpora and Tools for Persian. *Studia Linguistica Upsaliensia 16.* Uppsala: Acta Universitatis Upsaliensis.

Shabrang, Hoda. 2020. Every man is an Island: Decanonisation and Fragmentation in Reza Ghassemi's "The Nocturnal Harmony of Wood Orchestra." *Asiatic* vol. 14, no. 1, pp. 23–39.

Simpson A.J. and E.S.C. Weiner (eds.). 1989. *The Oxford English Dictionary.* Oxford: Clarendon Press; Oxford; New York: Oxford University Press.

Skalmowski, Wojciech. 1986. Języki nowoirańskie [New Iranian Languages] [in:] Leszek Bednarczuk (ed.) *Jezyki indoeuropejskie* [Indoeuropean Languages], vol. 1. Warszawa: Państwowe Wydawnictwo Naukowe, pp. 163–215.

Składanek, Bogdan. 1997. *Wprowadzenie do gramatyki języka perskiego* [An Introduction to Persian Grammar]. Warszawa: Dialog.

Skjærvø, Prods Oktor. 2006. Iran. VI. Iranian Languages and Scripts [in:] *EI*, vol. XIII, fasc. 4, pp. 344–377.

Idem. 2009. Middle West Iranian [in:] G. L. Windfuhr (ed.), *The Iranian Languages.* London and New York: Routlegde, pp. 196–278.

Smurzyński, Marek. 1997. Hezar nokte-ye bariktar az mu. *Negah-e nou* 34, Tehrān, pp. 177–183.

Spooner, Brian. 2012. Persian, Farsi, Dari, Tajiki: Language Names and Language Policies. [in:] H. Schiffman (ed.) *Language Policy and Language Conflict in Afghanistan and Its Neighbors: The Changing Politics of Language Choice.* Leiden, Boston: Brill, pp. 89–117.

Statistical Center of Iran. 2011. *National population and Housing Census 2011 (1390): Selected Findings.* Tehrān. Available online: https://irandataportal.syr.edu/wp-content/uploads/2011_iran_census_national-population-and-housing-census.pdf [retrieved 31.10.2021].

Stolper, Matthew and Jean Tavernier. 2007. From the Persepolis Fortification Archive Project, 1: An Old Persian Administrative Tablet from the Persepolis Fortification. *Achaemenid Research on Texts and Archaeology,* vol. 2007, no.1, Paris: Achemenet.

Šarifi, Mohammad. 2010. *Farhang-e Adabiyāt-e Fārsi* [A Dictionary of Persian Literature]. Tehrān: Entešārāt-e Mo'in.

Tagliamonte, Sali A. 2006. *Analyzing Sociolinguistic Variation.* Cambridge: Cambridge University Press.

Taleghani, Azita H. 2008. *Modality, Aspect and Negation in Persian.* Amsterdam and Philadephia: John Benjamins Publishing.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Tavakol, Mahbube and Hamid Allami. 2013. Reverse Addressing in Modern Persian. *International Journal of Society, Culture & Language*, 2(1), 2014, pp. 37–52.

Tisdall, William St. Clair. 1902. *Modern Persian conversation-grammar: with reading lessons, English-Persian vocabulary and Persian letters.* London, New York, Heidelberg: Julius Groos.

Trudgill, Peter. 1992. *Introducing Language and Society.* London: Penguin.

Utas, Bo. 2009. Iranian Languages as Gardens. *Iranian Journal of Applied Language Studies,* no.1(1), pp. 142–160.

Wakeman, John. 1988. *World Film Directors*, vol. 2. The H. W. Wilson Company.

Wąsala, Katarzyna. 2015. How the characters speak for themselves: colloquial language as a mean of expressing identity in *Čerāqhā rā man xāmuš mikonam*, a novel by Zoyā Pirzād [in:] M. Michalak and M. Rodziewicz (eds.) *In Quest of Identity. Studies on Persianate World.* Warszawa: Wydawnictwo Akademickie Dialog, pp. 193–209.

Eadem. 2018. The Little Traveller's adventures in the land of colloquiality. On the Persian translations of de Saint-Exupéry's *Le Petit Prince* [in:] R. Krupa and I. Piechnik (eds.) *Saint-Exupéry relu et traduit.* Kraków: Biblioteka Jagiellońska, pp. 180–190.

Eadem. 2019. *The case of Persian diglossia: do we really need an answer?* Conference presentation: European Conference of Iranian Studies 9, Freie Universität Berlin, 9–13 September 2019.

Weinreich, Uriel. 1954. Is a Structural Dialectology Possible? *WORD*, 10: 2–3, pp. 388–400.

Wexler, Paul. 1971. Diglossia, Language Standardization and Purism: Parameters for a Typology of Literary Languages. *Lingua*, no. 27, pp. 330–354.

Windfuhr, Gernot L. 1979. *Persian Grammar. History and State of its Study.* The Hague: Mouton.

Windfuhr, Gernot L. and John Perry. 2009. *Persian and Tajik* [in:] G. L. Windfuhr (ed.) *The Iranian Languages.* London and New York: Routlegde, pp. 416–544.

Xānlari, Parviz Nātel. 1970. *Dastur-e zabān-e fārsi* [Grammar of Persian Language]. Tehrān: Entešārāt-e Tus.

Idem. 1987. *Tārix-e zabān-e fārsi* [History of Persian Language]. Tehrān: Entešārāt-e Nou.

Yousef, Saeed. 2018. *Persian: A Comprehensive Grammar.* New York: Routledge.

Zarrinkub, Abdolhossein. 1957. *Do qarn sokut* [Two centuries of Silence]. Tehrān: Amir Kabir.

Zhukovsky, Valentin. 1922. Материалы для изучения персидских наречей [Materialy dlya izucheniya persidskikh narechiĭ – Materials for Studying Persian Dialects], A. A. Romaskewicz (ed.). Petrograd. (Жуковский, Валентин Алексеевич. 1888–1922. Материалы для изучения персидских наречий. Санкт-Петербург: тип. Акад. наук.)

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Appendices

Appendix 1. List of the quoted examples with original spelling and expected MWP form

No.	Persian (romanized)	Persian (original spelling)	Expected MWP form	English translation	Source
1	*unā šekāyat kard-an*	اونا شکایت کردن	ānhā šekāyat kard-and	they complained	DXM 78
2	*az unğā runde*	از اونجا رونده	az ānğā rānde	drove away from there	DXM 79
3	*unče ke goftin va midunin*	اونچه که گفتین و می دونمین	ānče ke goftid va midānid	what you said and [what you] know	ŠĀ 73
4	*mehmun-e man*	مهمون من	mehmān-e man	[be] my guest	KP 57
5	*Nosrat ğun ğun*	نصرت جون جون	Nosrat ğān	my dear Nosrat	AM 55
6	*do se nafar-e dige*	دو سه نفر دیگه	do se nafar-e digar	two [or] three other persons	DXM 52
7	*āxe mitarsam eštebāh koni mādar*	آخه می‌ترسم اشتباه کنی مادر	[āxar] mitarsam ke eštebāh koni, mādar	oh, but I fear you make a mistake, son (lit. mother)	ŠĀ 50
8	*age esrār dāš-tam beram ğāyi*	اگه اصرار داشتم برم جایی	agar esrār dāštam ke be ğāyi ber-avam	if I really wanted to go somewhere	KP 62
9	*harči hast tu in mize*	هرچی هست تو این میزه	harče ke hast dar ān miz ast	anything that there is, can be found at this table	DXM 52
10	*či mixunin*	چی می‌خونین	če mixānid	what do you read	ŠĀ 53
11	*či šod*	چی شد	če šod	what happened	ĀM 58

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(Fortsetzung)

No.	Persian (romanized)	Persian (original spelling)	Expected MWP form	English translation	Source
12	*raftam dam-e pan-šiš tā šerkat*	رفتم دم پن شش تا شرکت	dam-e dar-e panǧ-šeš šerkat raftam	I went to five or six companies	KP 57
13	*manzureš daqiqan čist*	منظورش دقیقا چیست	manzuraš daqiqan če ast	what she exactly meant	Hn 50
14	*saqf-ro nešun dādam*	سقف رو نشون دادم	saqf-rā nešān dādam	I pointed at the ceiling	DXM 52
15	*man-o tanhā bezārin*	منو تنها بذارین	ma-rā tanhā be-gozārid	leave me alone	DXM 53
16	*lebāset-o avaz kon*	لباستو عوض کن	lebāset-rā avaz kon	change your clothes	ŠĀ 50
17	*to-rā xodā velam kon*	تورا خدا ولم کن	to-rā xodā velam kon	for God's sake, let me go	KP 64
18	*ye maxsuseš-o vāse xodam radif konam*	یه مخصوصشو واسه خودم ردیف کنم	yek [fenǧān-e/šo-kālāt-e] maxsus barā-ye xodam radif konam	I'd prepare one special [of this] for myself	KP 58
19	*tā kārām tamum šod*	تا کارم تموم شد	tā kāram tamām šavad	when my work was finished	DXM 56
20	*inā-ro migi*	اینارو میگی	ānhā-rā miguyi	you say all these [things]	ŠĀ 50
21	*pesarā-ye man*	پسرای من	pesarhāye man	my sons	ŠĀ 50
22	*fekreš-am nemitunin bokonin*	فکرشم نمی تونین بکنین	fekr-e ān-rā ham nemitavānid bokonid	you can't even imagine that	KP 56
23	*ye ǧā-ye xalvat*	یه جای خلوت	yek ǧā-ye xalvat	a quiet place	DXM 56
24	*ustā*	اوستا	ostād	master	ĀM 57
25	*čāquš-rā gozāšt zir-e galu-ye Seyyed*	چاقوش را گذاشت زیر گلوی سید	čāquyaš-rā zir-e galu-ye Seyyed gozāšt	he pressed his knife to Seyyed's throat	Hn 56
26	*ǧurābhām-ro ham dar miyāram*	جوراب‌هام رو هم در می آرم	ǧurābhāyam-rā ham dar miyāvaram	I take off my socks too	DXM 56
27	*harfhātun-am bā ham zadin*	حرفاتونم با هم زدین	harfhāyetān-rā bā ham zadid	you spoke about it with each other	ŠĀ 60
28	*be hamešān bedehad*	به همه‌شان بدهد	be hameyešān bedehad	give it to all of them	KP 51
29	*āšenāyimān*	آشنایی‌مان	āšenāyiyemān	our acquaintanceship	ĀM 51

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(Fortsetzung)

No.	Persian (romanized)	Persian (original spelling)	Expected MWP form	English translation	Source
30	*dusthāš*	دوست‌هاش	dusthāyaš	her friends	ĀM 50
31	*un dus nadāre*	اون دوس نداره	ān dust nadārad	he doesn't like it	ŠĀ 61
32	*mess-e inke*	مث اینکه	mesl-e inke	it seems that	KP 58
33	*išālā*	اشالا	enšāllāh	God willing	ŠĀ 64
34	*vāsāde tā reside be kešiš*	واساده تا رسیده به کشیش	montazer bud tā be kešiš beresad	he stood until he reached the priest	DXM 65
35	*ruz-e yekšanbe as*	روز یکشنبه اس	ruz-e yekšanbe ast	it is Sunday	DXM 62
36	*hiški das na-zane lotfan*	هیشکی در نزنه لطفا	hičkas dast naza-nad lotfan	nobody touches [it], please	KP 66
37	*xob in yeki xatt-e Benedikt ast*	خب این یکی خط بندیکت است	xub, in xatt-e Bendikt ast	well, this one is Benedict's handwriting	Hn 79
38	*to nemiyāy birun*	تو نمیای بیرون	to birun nemiyāyi	you don't go out	ŠĀ 57
39	*ke kešiše na-fahme*	که کشیشه نفهمه	ke ān kešiš nafah-mad	so that the priest would not under-stand	DXM 65
40	*pesare čahār-dah punzdah sāl balke bištar az doxtar bo-zorgtare*	پسره چهارده پونزده سال بلکه بیشتر از دختر بزرگتره	ān pesar chahār-dah pānzdah sāl yā hattā bistar az doxtar bozorgtar ast	that boy is fourteen-fifteen years older than the girl, if not more	ŠĀ 65
41	*hekāyat-e un doxtare ke tu-ye qal'e-ye div zenduni šode bud*	حکایت اون دختره که توی قلعه دیو زندونی شده بود	hekāyat-e ān dox-tar-i ke dar qal'e-ye div zendāni sode bud	the story of that girl who was imprisoned in monster's castle	ŠĀ 73
42	*hame-ye kas-o-kār-e doxtare tu-ye tasādof mordand*	همه‌ی کس و کار دختره توی تصادف مردند	hame-ye ba-stegān-e ān doxtar dar tasādof mor-dand	all of that girl's rela-tives died in a car ac-cident	ĀM 70 (p. 103)
43	*yek ğoft-e lab-e barrāq*	یک جفت لب براق	yek ğoft/ğofti lab-e barrāq	(lit. a pair of) shiny lips	ĀM 54
44	*ye bāzğuyi-ye sāde*	یه بازجویی ساده	bāzğuyi-ye sāde-yi	a simple investigation	DXM 84

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(Fortsetzung)

No.	Persian (romanized)	Persian (original spelling)	Expected MWP form	English translation	Source
45	*vaqti ham miyād eyn-e yek qaribe ast*	وقتی هم میاد عین یک غریبه [است]	vaqti ke miyāyad, eyn-e qaribe-yi-st	and when he comes, he is like a stranger	ŠĀ 52
46	*yek performāns tu-ye Piyāno dāšte bāšad*	یک پرفورمانس توی پیانو داشته باشد	dar Piyāno performānsi dāšte bāšad	'[she would] have a performance in Piano'	KP 50
47	*sobheš ye čizi tarğome karde-am*	صبحش یه چیزی ترجمه کرده ام	sobh-e ān ruz čizi tarğome karde-am	I have translated something that morning	KP 56
48	*mixāst farār kone ke gir oftād*	می خواست فرار کنه که گیر افتاد	vaqti ke mixāst farār konad, gir oftād	when she wanted to escape, she got caught	DXM 55
49	*zang-e dar ke be sedā dar āmad*	زنگ در که به صدا در آمد	vaqti ke zang-e dar be sedā dar āmad	when the doorbell rang	ŠĀ 57
50	*az dar ke dāšt birun mizad*	از در که داشت بیرون می زد	vaqti ke az dar birun miyāmad	when he was going out through the door	KP 52
51	*asabāni ke šod*	عصبانی که شد	vaqti ke asabāni šod	when she got angry	ĀM 56
52	*ba'd ke bargaštam be otāq-e xodam*	بعد که برگشتم به اتاق خودم	ba'd vaqti ke be otāq-e xodam bar gaštam	when I returned to my room later	Hn 53
53	*tu kāzino ke nemiše*	تو کازینو که نمیشه	dar kāzino nemišavad	it can't be done in [a place such as] a casino	DXM 56
54	*šoma midunin ke!*	شما می دونین که!	šomā midānid	you do know that!	ŠĀ
55	*man ke pā-ye tork-e xāčik-am*	من که پایه ی ترک خاچیک ام	man pā-ye tork-e xāčikam	I do prefer Turkish Khachik [coffee brand]	KP 51
56	*ahl-e mašquliyāt-e fekri ham ke nist*	اهل مشغولیت فکری هم که نیست	ahl-e mašquliyat-e fekri ham nist	he is not a person who likes mental activities either	Hn 55
57	*ğelouaš in harf-o nazani hā*	جلوش این حرفو نزنی ها	ğelou-e u in harf-hā nazanid	don't you dare speak those words in front of him/her	ŠĀ 51
58	*messe inke emruz kār-o-bāretun sekke bude hā*	مث این که امروز کار و بارتون سکه بوده ها	mesl-e inke emruz kar-o-bār-e shomā sekke bude ast	it seems that you have earned a lot today	KP 58

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(Fortsetzung)

No.	Persian (romanized)	Persian (original spelling)	Expected MWP form	English translation	Source
59	*mobārak hā*	مبارک ها	*mobārak bāšad* [without expressing the irony]	congratulations [ironic]	ĀM 53
60	*yek kalame ham nagoft*	یک کلمه هم نگفت	hattā yek kalame nagoft	didn't say even a word	ĀM 77
61	*ma'ni-yaš in nist ke hič ham bā ham farqi nadārand*	معنی‌اش این نیست که هیچ هم با هم فرقی ندارند	ma'ni-ye ān in nist ke bā ham hič farqi nadārand	it doesn't mean that they don't differ from each other at all	KP 61
62	*bāvar nemikardan*	باور نمی کردن	bāvar nemikardand	they were not believing	DXM 54
63	*šoru' miše*	شروع می‌شه	šoru' mišavad	it starts	KP 50
64	*Parviz ham miyāram*	پرویز هم میارم	Parviz-rā ham miyāvaram	I'll bring Parviz [with me] too	ŠĀ 52
65	*xodam mizāram; hiški das nazane lotfan*	خودم می زارم هیشکی دس نزنه لطفا	xodam migozāram; lotfan hičkas dast nazanad	I'll put it [there] myself; no one touches it please	KP 66
66	*dāšt bivaqfe touzih midād*	داشت بی‌وقفه توضیح می داد	bivaqfe touzih midād	she was explaining continuously	KP 50
67	*dāram šuxi mikonam*	دارم شوخی می‌کنم	šuxi mikonam	I am joking	ŠĀ 54
68	*bedeham tarrāhi konand*	بدهم طراحی کنند	[sefāreš] bedeham [tā/ke] tarrāhi konand	I'll have [it] designed [lit. I would give they would design]	KP 52
69	*dādam ye ālunak-e kuchulu barām dorost kardan*	دادم یه آلونک کوچولو برام درست کردن	[sefāreš] dād-am [tā/ke]yek ālunak-e kuchek barā-ye man dorost kardand	I had a small cabin built for me; I ordered a little cabin [lit. I gave a little cabin for me they built]	DXM 56
70	*nemideham pipam-rā baqiye begozārand guše-ye labešān*	نمی دهم پیپم را بقیه بگذارند گوشه ی لبشان	nemigozāram ke baqiye pipam-rā dar guše-ye labešān begozārand	I don't let others pu my pipe in the corner of their lips [lit. I don't give others put my pipe in the corner of their lips]	KP 59
71	*raft yahudi šod*	رفت یهودی شد	yahudi šod	he became a Jew [i. e. converted to Judaism; lit. went and became a Jew]	DXM 57

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(Fortsetzung)

No.	Persian (romanized)	Persian (original spelling)	Expected MWP form	English translation	Source
72	*nakard lā-aqal tā dam-e dar biyād badraqe-mun*	نکرد لا اقل تا دم در بیاد بدرقه‌مون	zahmat nakešid ke dam-e dar ba-draqeyemān biyāyad	he didn't even come to see us off at the threshold [lit. didn't make it to come]	ŠĀ 70
73	*nakard tamizeš konad*	نکرد تمیزش کند	zahmat nakešid ke ān-rā tamiz konad	he didn't clean it [lit. didn't make it to to clean it]	KP 55
74	*mesl-e hamiše fešāret āmade pāyin*	مثل همیشه فشارت آمده پایین	fešāre to mesl-e hamiše pāyin āmade ast	as always, your blood pressure has fallen'	ĀM 59
75	*čašmānaš bāz ham qarq-e ašk mišod*	چشمانش باز هم قرق اشک می‌شد	čašmānaš bāz qarq-e ašq mišod	her eyes were filling up with tears all over again	ŠĀ 79
76	*hame-aš-rā ham be sarn-evešt-e nek-batemān nesbat midehim*	همه‌اش را هم به سرنوشت نکبتمان نسبت می‌دهیم	mā hame-ye ān-rā be sarnevešt-e nekbatemān nes-bat midahim	for all of that, we blame our wretched fate	KP 55
77	*yeki-aš-ro hamin diruz doxtarak […] zad be goldān*	یکی‌اش رو همین دیروز دخترک [] زد به گلدان	hamin diruz dox-tarak bā yeki az ānhā be goldān zad	a little girl hit one of those against a vase just yesterday	KP 63
78	*xodam bāhāš harf mizanam*	خودم باهاش حرف می‌زنم	xodam bā u harf mizanam	I'll talk to him myself	ĀM 57
79	*faqat behem begu*	فقط بهم بگو	faqat be man begu	just tell me	ŠĀ 74
80	*nazaret barām mohemme*	نظرت برام مهمه	nazar-e to barā-ye man mohemm ast	your opinion is im-portant for me	ŠĀ 74
81	*az Āye va Fa-rānse raftaneš harf bezanim*	از آیه و فرانسه رفتنش حرف بزنیم	az Āye va raftan-e u be Farānse harf bezanim	let's talk about Āye and her going to France	ĀM 77
82	*Zimens buda-neš ma'lum bāšad*	زیمنس بودنش معلوم باشد	tā ma'lum bāšad ke Zimens ast	[so that] it being a Siemens [appliance] is obvious	KP 63
83	*engār entezār-e didanam-ro na-dāšti*	انگار انتظار دیدنم رو نداشتی	engār entezār-e didan-e ma-rā na-dāšti	it seems you didn't expect to see me	ŠĀ 74
84	*qeybeš zade bud*	قیبش زده بود	u gom šod	s/he had dissapeared	Hn 55
85	*hanuzam bā-varam nemiše*	هنوزم باورم نمی‌شه	hanuz ham bā-varam nemišavad	I still can't believe	ŠĀ 50

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(Fortsetzung)

No.	Persian (romanized)	Persian (original spelling)	Expected MWP form	English translation	Source
86	*enqadrhā saram mišavad*	انقدرها یرم می‌شود	inqadr mitavānam befahmam	this much I understand	KP 50
87	*age surāxeš koni*	اگه سوراخش کنی	agar ān-rāsurāx koni	if you tear them	ŠĀ 58
88	*zouq bar-eš dāšt*	ذوق برش داشت		s/he got excited	KP 50
89	*xāste bud bebinadam*	خواسته بود ببیندم	xāste bud ma-rā bebinad	he had wanted to see me	Hn 48
90	*xābāndimeš ru-ye taxt*	خواباندیمش روی تخت	u-rā ru-ye taxt xābāndim	we laid him on the bed	Hn 51
91	*dud-e yeknavāxt-e qalizi pušānde budeš*	دود یکنواخت غلیظی پوشانده بودش	dud-e yeknavāxt-e qalizi ān-rā pušānde bud	a uniformly thick smoked had covered it	KP 63
92	*mixāsti negaheš dāri torši bendāzi*	می‌خواستی نگهش داری ترشی بندازی	mixāsti u-rā negah dāri va torši biyandāzi	you wanted to keep her and make her an old maid [lit. make her sour]	ŠĀ 51
93	*tāksi begirad beravad xāne duš begirad*	تاکسی بگیرد برود خانه دوش بگیرد	tāksi begirad va be xāne beravad tā duš begirad	take a cab, go home, take a shower	KP 52
94	*āmad āmpul-e fešār zad*	آمد آمپول فشار زد	āmad va āmpul-e fešār zad	[the doctor] came gave [her] the injection for blood pressure	ĀM 56
95	*bar gašt borāq šod be Ārezu*	بر گشت براق شد به آرزو	bar gašt va be Ārezu borāq šod	he turned back, got angry at Ārezu	ĀM 57
96	*boland šodam beravam*	بلند شدم بروم	boland šodam tā beravam	I stood up [so that] I go / I stood up to go	Hn 59
97	*miram čāy biyāram*	میرم چای بیارم	miravam tā čāy biyāvaram	I'll go [to] bring some tea	ŠĀ 51
98	*beheš etminān dādam* [ke] *man* […] *mesl-e ānhā fekr nemikonam*	بهش اطمینان دادم من مثل آنها فکر نمی کنم	be u etminān dādam ke man mesl-e ānhā fekr nemikonam	I ensured him [that] I don't think like the others	KP 57
99	*harči* [ke] *dam-e dastam resid boxoram*	هرچی که دم دستم رسید بخورم	harče ke dam-e dastam resid, boxoram	I would eat anything that came to my hands	ĀM 59

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(Fortsetzung)

No.	Persian (romanized)	Persian (original spelling)	Expected MWP form	English translation	Source
100	*in ğavān xunaš be ğuš āmade*	این جوان خونش به جوش آمده	in ğavān asabāni šod	this young person has become angry (lit. this young person, his/her blood has come to boil)	Hn 56
101	*Ārezu yādeš bud*	آرزو یادش بود	Ārezu be yād dāšt / yād-e Ārezu bud	Ārezu remembered (lit. Ārezu, her memory was)	ĀM 55
102	*tā inğā čide āmade bālā*	تا اینجا چیده آمده بالا	vaqti tā inğā čid, āmad bālā	[he] has laid [bricks] up to here and has gone up further	ĀM 56
103	*nemitavāneste in kār rā bokonad*	نمی‌توانسته این کار را بکند	nemitavānest in kār rā bokonad	he couldn't have done this	KP 56
104	*mağbur bude beravad dastšuyi*	مجبور بوده برود	mağbur bud be dastšuyi beravad	s/he had to go to the bathroom	KP 56
105	*Tehrān bemānad*	تهران بماند	dar Tehrān bemānad	s/he would stay in Tehrān	KP 61
106	*Āye-rā mi-bord mahd-e kudak*	آیه را می‌برد مهد کودک	Āye-rā be mahd-e kudak mibord	she was taking Āye to kindergarten	ĀM 54
107	*nešaste bud ru-ye labe-ye taxt*	نشسته بود روی لبه‌ی تخت	ru-ye labe-ye taxt nešaste bud	she was sitting on the edge of the bed	Hn 53
108	*šoru' kardam be sangeš-e ouzā'*	شروع کردم به سنجش اوضاع	be sangeš-e ouzā' šoru' kardam / sangeš-e ouzā'-rā šoru' kardam	I began examining the situation	DXM 53
109	*mimiram barāyešān*	می‌میرم برایشان	barā-ye ānhā mimiram	I'd die for them (expression of liking something)	KP 61
110	*paridam tu māšin*	پریدم تو ماشین	dāxel-e māšin paridam	I jumped into the car	DXM 56
111	*xodam bemānam tu-ye xāne bepusam*	خودم بمانم توی خانه بپوسم	xodam dar xāne bemānam va be-pusam	I should stay alone and rot in this house	ĀM 59
112	*ye maxsuseš-o vāse xodam radif konam*	یه مخصوصشو واسه خودم ردیف کنم	yek [šokālāt-e] maxsus-rā barā-ye xodam radif konam	I'd prepare one special [of this] for myself	KP 58
113	*bodou-bodou*	بدو بدو	tond	fast [lit. run-run]	DXM 55

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

(Fortsetzung)

No.	Persian (romanized)	Persian (original spelling)	Expected MWP form	English translation	Source
114	*sāxt-o-pāxt*	ساخت و پاخت	/idiomatic phrase not used in MWP form /	realized plan	DXM 53
115	*kār-o-bāretun*	کاروبارتون	kārhāyetān	your affairs	KP 58
116	*sar-o-kalle tekun midim*	سروکله تکون می‌دیم	sar tekān midahim	we nod heads	DXM 50
117	*sar-e miz-e hašt xabarāyiye*	سر میز هشت خبراییه	sar-e miz-e hast čizi ettefāq miy-ofte	something is going on at table eight	DXM 50
118	*sang ru yaxam kone*	سنگ رو یخم کنه	/idiomatic phrase not used in MWP form /	he would humiliate me (lit. make me a stone on ice)	ŠĀ 51
119	*sar-e mano bekubi be tāq*	سر منو بکوبی به طاق	sar-e ma-rā be tāq bekubi /idiomatic phrase not used in MWP form/	you would distract me (lit. hit my head against a niche)	ŠĀ 52
120	*bezanam tu-ye zouqeš*	بزنم توی ذوقش	/idiomatic phrase not used in MWP form /	[I'd] dampen (lit. hit into) his enthusiasm	KP 57
121	*āsmān-rismān bāft*	آسمان ریسمان بافت	/idiomatic phrase not used in MWP form /	she talked nonsense (lit. she plaited sky with a rope)	ĀM 65
122	*velam kon*	ولم کن	ma-rā vel kon /idiomatic phrase/	leave me alone (lit. let me loose)	ĀM 77
123	[*man*] *bidi nistam ke bā in bādhā belarzam*	بیدی نیستم که با این بادها بلرزم	/a popular saying, no difference between forms/	I do not fear that (lit. I am not a willow to tremble at such winds)	ŠĀ 61
124	*āsmān ke zamin nayāmade*	آسمان که زمین نیامده	āsmān be zamin nayāmade ast /idiomatic phrase not used in MWP form /	this is not the end of the world (lit. heaven hasn't come to earth yet)	ĀM 77
125	*xodā-ro če didi?* [*Be qoul-e māmān*] *kār-e našod nadāre*	خدا رو چه دیدی؟ به قول مامان کار نشد نداره	xodā-rā če didi? [Be qoul-e māmān] kār-e našod nadārad /idiomatic phrase/	what have you seen of God? [As mom says,] nothing is impossible for him	ŠĀ 61

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Appendix 2. Quantitative analysis data table (summary)

Parameter occurrence according to category					
Category	Dar Xarābāt-e Moqān	Šahr-āšub	Kāfe Piyāno	Ādat Mikonim	Hamnavāyi...
Phonology	**758**	**461**	**114**	**43**	**6**
alternation	**592**	**368**	**44**	**24**	**1**
alternation -ān>-un	326	189	28	6	0
alternation -ar>-e (final consonant omission + -a>-e)	51	62	6	0	0
alternation -e>-i	31	36	5	18	1
DO marker -rā>-o	55	24	4	0	0
DO marker -rā>-ro	129	57	1	0	0
contraction	**50**	**60**	**37**	**16**	**5**
of consonant cluster / other	16	35	18	9	3
of enclitic pronoun	34	25	19	7	2
omission	**116**	**33**	**33**	**3**	**0**
final consonant	54	4	26	3	0
of h (any position)	62	29	7	0	0
Morphology	**905**	**827**	**846**	**175**	**317**
article	**53**	**21**	**57**	**3**	**16**
definite article -e	3	6	0	2	0
indefinite article yek	49	11	48	1	13
indefinite article yek + -i	1	4	9	0	3
enclitic pronoun	**139**	**333**	**615**	**114**	**237**
possessive / partitive	93	246	357	77	178
preposition + enclPron eg as DO/IO	24	31	139	4	8
with infinitive (noun of action) as agent or object	1	5	7	3	3
with nomVerb as DO	11	22	33	7	25
with nomVerb as IO	4	6	12	7	2
with nomVerb/preverb as logical subject (quasi impersonal idioms)	4	17	50	4	16
with preverb as DO/IO	1	1	2	12	1
with verb as DO/IO	1	5	15	0	4
particle or conjunction	**23**	**61**	**39**	**40**	**52**
emphatic ke	12	29	8	12	13
emphatic particle hā	0	4	2	5	0
enclitic particle ham	8	23	18	8	25
temporal ke	3	5	11	15	14
tense	**31**	**28**	**56**	**12**	**11**
other	0	1	0	0	0
past progressive tense	8	7	22	7	6
present progressive tense	17	16	17	1	2
two verb construction (other than modal)	6	4	17	4	3
verbal ending	**471**	**239**	**51**	**2**	**1**
2pl praes/praet	20	49	9	0	0
3pl praes/praet	207	13	7	0	0
3sg praes	166	110	20	2	0
copula 3sg praes	78	67	15	0	1
verbal stem	**188**	**145**	**28**	**4**	**0**
alternate stem	5	2	1	2	0
MSP stem	183	143	27	2	0
Syntax	**147**	**120**	**317**	**255**	**72**
clause without conjunction	**20**	**32**	**40**	**30**	**17**
coordinate	11	8	4	22	0
subordinate	9	24	36	8	17
lack of agreement	**15**	**4**	**17**	**3**	**1**
between subject and verb (quasi impersonal idiom, logical subject, number disagreer	15	4	17	3	1
omission	**52**	**42**	**63**	**51**	**6**
aux verb	45	33	56	44	6
preposition	7	9	7	7	0
word order	**60**	**42**	**197**	**171**	**48**
disrupted word order, SOV>SVO other than two above	7	8	17	14	7
MSP word order SOV>SVO - direction/location	37	30	171	133	38
MSP word order SOV>SVO - nominal clauses	16	4	9	24	3
Lexis	**169**	**101**	**186**	**115**	**45**
MSP lexical item	**169**	**101**	**186**	**115**	**45**
phrase	42	27	33	25	24
preposition MSP	68	25	104	50	5
saying	1	11	0	4	0
single word, compound word	58	38	49	36	16
Sum	**1979**	**1509**	**1463**	**588**	**440**

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Appendix 3. Summary in English

"Modern Spoken Persian in contemporary literary texts on the example of selected 21st century novels"

One of the characteristic features of the linguistic situation of contemporary Iran is the coexistence of two standards of Modern Persian: the spoken standard and the written standard. While the written standard, strongly rooted in the Persian literary tradition and drawing patterns from it, does not raise any doubts, the existence and status of the spoken standard are still sometimes considered controversial.

The book is concerned with the issue of the occurrence of forms typical for the spoken variety of the modern Persian language in contemporary Persian literature. This phenomenon, present in Persian literature since around the middle of the last century, has not been thoroughly described so far, and the very fact of the co-occurrence of two equally important variants of the language has received numerous studies in several research paradigms, but without decisive conclusions that would allow to assess what exactly the character of the relationship between these variants is.

The aim of this study was, above all, to indicate exactly what forms of the spoken language variety penetrate into the written variety and how it is conditioned. For this purpose, an analysis of five contemporary novels, published in the first decade of the 21st century and meeting certain conditions (including the diversification of authors in terms of gender, place of birth or residence, or non-writing professional experience) was carried out. The second, equally important aspect was the reference to the existing studies on the issue of the diversity of the written and spoken language varieties, their analysis and an attempt to indicate the research methodology that best describes them.

The book is divided into six chapters, the first three of which constitute the theoretical part, while the remaining three are devoted to the practical analysis of the material and the conclusions drawn from it. The first chapter is introductory. It presents a general picture of Iran's linguistic situation, the aims of the work and the material selected for analysis. Moreover, this chapter explains the Persian transcription method used in the described research project and the abbreviations used in the text. The second chapter discusses the theoretical approaches that influenced the applied theory, and also indicates how exactly it will be applied. This chapter presents, among other things, various approaches to concepts such as 'dialect,' 'style' and 'register,' necessary in the discussion of the linguistic situation of contemporary Iran. The third chapter is devoted to the Persian language: its historical development, division into spoken and written forms, and the descriptions of this phenomenon in Iranian linguistic works both

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

in Iran and abroad, within five methodological approaches: stylistic, functional, diglossic, sociolinguistic, and anthropological. By comparing several theoretical approaches and confronting them with the existing descriptions of the linguistic situation of today's Iran, it has been shown that the written and spoken forms are most accurately described using the notion of a register, and their interdependence shows the features of a diglossic relationship. The difference between the written and spoken varieties and the stylistic diversity of the language was also shown, which in Persian terms was marked by the separation of the Persian-language terms *mohāvereyi* 'colloquial [= spoken]' and *āmiyāne* 'colloquial [= colloquial],' assigning the former to the register and the latter to style. These conclusions summarize the first theoretical part of the work.

The fourth chapter is an introduction to the analysis of the material. It includes some general remarks, a discussion of the selected material (for each of the selected novels it is the author's biography, plot summary and a brief description of the style) and a detailed description of the methods used. In the theoretical chapters, it was shown that the spoken variety of the modern Persian language (MSP) and its written form (MWP) are in fact registers of this language, therefore the method of analysis was based on the methods of register analysis, with the necessary modifications.

The first stage of the analysis was qualitative. Selected fragments of books were analysed in detail in order to distinguish non-standard forms. It was evident at an early stage that these non-standard forms are repetitive, both within and between individual books. On the basis of the distinguished forms, a catalog of 45 non-standard parameters was created, which most often co-occurred with each other within sentences or paragraphs. The parameters were divided into four groups according to grammatical categories: phonological, morphological, syntactic and lexical parameters (phonological parameters are those for which an alternative spelling, reflecting their actual pronunciation in the spoken register, is used instead of a regular, dictionary form based on a written standard).

Detailed description of the indicated parameters, presented in the first part of the fifth chapter, explains their grammatical characteristics, and where possible, apart from the synchronous description, also refers to diachrony. For each parameter, there are also examples taken from the analyzed material, often with additional comments. Creating a list of parameters as a result of qualitative analysis served as the starting point for the second stage, i. e. quantitative analysis, in which larger batches of selected novels were analyzed in terms of parameter's presence and frequency. The frequency of the selected parameters in the analyzed texts, within four categories (phonology, morphology, syntax and lexis) for each of the books, revealed the presence of three patterns, within which the MSP parameters appear in literary texts: phonological, morphological and syntactic; those patterns could be related to the characteristics of the books in

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

which they appeared. The description of patterns is in fact a conclusion of the second stage of the analytical part.

Chapter six summarizes the research carried out. Additionally, it proposes directions for further research and the possible use of the results obtained thanks to the analysis carried out in this paper for further research projects.

In conclusion, thanks to the conducted analysis, it was determined that the relationship of the spoken and written variety of Persian is of a functional character, and these varieties can be classified as registers of a high level of generality. The typical features of the spoken register of contemporary Persian language were indicated, and it was confirmed that the phenomenon of their penetration into literature does indeed take place and is diversified in terms of its intensity and type of forms, which is described in the form of three patterns of the occurrence of MSP forms in literary texts.

The occurrence of spoken forms in written texts is not a new phenomenon in the history of the Persian language. Research on the transitional periods between Old Persian and Middle Persian, and then Middle and New Persian, shows that in these periods there was also a gap between the language variety used in writing and the one used in speech, and the texts created in their final stages were penetrated by the elements closer to the new stage. At first they were seen as scribes' mistakes, but today they are believed to be symptoms of a transitional period and of an impending change, not sudden, but gradual. Contemporarily, incorporation of forms typical of the spoken Persian into literary texts also evokes ambivalent feelings. The written variety, deeply rooted in literature, is much more conservative than the spoken variety, in which progressive changes can be observed. Perhaps this means that we are witnessing another transitional period in the history of the Persian language—exceptional in that this time we have the opportunity to observe changes in progress, just as they are happening.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

Appendix 4. Summary in Polish

"Odmiana mówiona języka perskiego we współczesnych tekstach literackich na podstawie wybranych powieści z XXI wieku"

Jedną z cech charakterystycznych sytuacji językowej współczesnego Iranu jest współwystępowanie dwóch standardów języka perskiego: standardu mówionego oraz standardu pisanego. O ile standard pisany, silnie zakorzeniony w perskiej tradycji literackiej i czerpiący z niej wzorce, nie budzi wątpliwości, o tyle istnienie i status standardu mówionego wciąż bywają uznawane za kwestie kontrowersyjne.

Niniejsza praca poświęcona została zagadnieniu występowania form typowych dla mówionej odmiany współczesnego języka perskiego we współczesnej perskiej literaturze. Zjawisko to, obecne w literaturze perskiej od około połowy ubiegłego wieku, nie zostało dotąd gruntownie opisane, a sam fakt współwystępowania we współczesnej perszczyźnie dwóch równorzędnych wariantów języka doczekał się licznych opracowań w kilku paradygmatach badawczych, lecz bez decydujących wniosków, które pozwalałyby ocenić, jaki dokładnie charakter ma relacja pomiędzy tymi wariantami.

Celem pracy było przede wszystkim wskazanie, jakie dokładnie formy mówionej odmiany języka przenikają do odmiany pisanej i czym jest to warunkowane. Do tego celu przeprowadzono analizę pięciu współczesnych powieści, wydanych w pierwszym dziesięcioleciu XXI wieku i spełniających określone warunki (w tym zróżnicowanie autorów pod względem płci, miejsca urodzenia bądź zamieszkania czy też pozapisarskiego doświadczenia zawodowego). Drugim, równie ważnym aspektem było przywołanie istniejących opracowań zagadnienia zróżnicowania pisanej i mówionej odmiany języka, ich analiza i próba wskazania najlepiej opisującej je metodologii badawczej.

Praca podzielona została na sześć rozdziałów, przy czym pierwsze trzy z nich stanowią część teoretyczną, pozostałe trzy natomiast są poświęcone praktycznej analizie materiału i wypływającym z niej wnioskom. Rozdział pierwszy ma charakter wprowadzający. Jest w nim zarysowany ogólny obraz sytuacji językowej Iranu, cele pracy i wybrany do analizy materiał. Rozdział ten objaśnia także zastosowaną transkrypcję języka perskiego oraz występujące w tekście skróty. Rozdział drugi omawia podejścia teoretyczne, które wpłynęły na metodologię zastosowaną w pracy, a także wskazuje, w jaki sposób będzie ona wykorzystana w analizie materiału. W tym rozdziale dyskutowane są między innymi różne podejścia do pojęć takich jak 'dialekt,' 'styl' i 'rejestr,' niezbędnych w dyskusji nad sytuacją językową dzisiejszego Iranu. Rozdział trzeci poświęcony został językowi perskiemu: jego rozwojowi historycznemu, podziałowi na odmianę mówioną i pisaną, a także opisom tego zjawiska w iranistycznych pracach językoznawczych

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

zarówno w Iranie, jak i poza jego granicami, w obrębie pięciu grup metodologicznych: stylistycznej, funkcjonalnej, diglosyjnej, socjolingwistycznej oraz antropologicznej.

Dzięki porównaniu kilku podejść teoretycznych oraz skonfrontowaniu ich z istniejącymi opisami sytuacji językowej dzisiejszego Iranu wskazano, że odmiana pisana i mówiona najtrafniej dają się opisać przy użyciu pojęcia rejestru, a ich współzależność wykazuje cechy relacji diglosyjnej. Pokazano także odrębność relacji odmian pisanej i mówionej od zróżnicowania stylistycznego języka, co na gruncie perskim zaznaczono poprzez rozdzielenie stosowanych w pracach perskojęzycznych terminów *mohāvereyi* 'potoczny [=mówiony]' oraz *āmiyāne* 'potoczny [=kolokwialny],' wprowadzając pierwszy z nich do kategorii rejestru, drugi natomiast do kategorii stylu. Te wnioski stanowią podsumowanie pierwszej, teoretycznej części pracy.

Rozdział czwarty jest wprowadzeniem do analizy materiału. Obejmuje on ogólne wprowadzenie, omówienie wybranego materiału (dla każdej z wybranych powieści jest to biografia autora, streszczenie fabuły oraz krótka charakterystyka stylu) i szczegółowy opis poszczególnych etapów analizy i zastosowanych metod. W części teoretycznej wykazano, że odmiana mówiona współczesnego języka perskiego (MSP) i jego odmiana pisana (MWP) są w istocie rejestrami tego języka, dlatego też metoda analizy materiału została oparta na metodach analizy rejestrów, z niezbędnymi modyfikacjami.

Pierwsza część analizy miała charakter jakościowy. Wybrane fragmenty książek poddano szczegółowej analizie zastosowanych form w celu wyodrębnienia form niestandardowych. Już na wczesnym etapie uwidoczniono, że owe niestandardowe formy powtarzają się zarówno w obrębie poszczególnych książek, jak i pomiędzy nimi. Na podstawie wyodrębnionych form powstał katalog 45 parametrów niestandardowych, najczęściej współwystępujących ze sobą w obrębie zdań czy akapitów, podzielonych według kategorii gramatycznych na cztery grupy: parametry fonologiczne, morfologiczne, składniowe i leksykalne (za parametry fonologiczne uważane są te formy, których reprezentacja graficzna odzwierciedla ich formę stosowaną w rejestrze mówionym, a nie formę słownikową, opartą na standardzie pisanym).

Szczegółowy opis wskazanych parametrów, przedstawiony w pierwszej części rozdziału piątego, objaśnia ich charakterystykę gramatyczną, a tam, gdzie to możliwe, oprócz opisu synchronicznego nawiązuje również do diachronii. Dla każdego parametru podane są też przykłady zaczerpnięte z przeanalizowanego materiału, często z dodatkowymi komentarzami.

Stworzenie listy parametrów w rezultacie analizy jakościowej było punktem wyjścia do drugiego etapu, czyli analizy ilościowej, w której większe partie wybranych powieści zostały przeanalizowane pod kątem ich obecności. Frekwencja wyodrębnionych parametrów w analizowanych tekstach, w obrębie czterech

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384

kategorii (fonologii, morfologii, składni i leksyki) dla każdej z książek, ujawniła występowanie w badanym materiale trzech schematów występowania parametrów MSP w tekstach literackich: fonologicznego, morfologicznego i składniowego, powiązanych z wcześniej opisaną charakterystyką książek, w których wystąpiły. Opis owych schematów stanowi podsumowanie drugiego etapu części analitycznej.

Rozdział szósty stanowi podsumowanie przeprowadzonych badań, zawiera także propozycje kierunków dalszych badań i możliwego wykorzystania rezultatów uzyskanych dzięki przeprowadzonej w niniejszej pracy analizie.

Podsumowując, dzięki przeprowadzonej analizie określono, iż relacja odmiany mówionej i pisanej języka perskiego ma charakter funkcjonalny, a odmiany te można zaklasyfikować jako rejestry o wysokim poziomie ogólności. Wskazano, jakie są typowe cechy mówionego rejestru współczesnej perszczyzny, a także potwierdzono, że zjawisko ich przenikania do literatury rzeczywiście ma miejsce i jest zróżnicowane pod względem skali oraz typu występujących form, co opisano w formie trzech schematów występowania form MSP w tekstach literackich.

Zjawisko występowania w tekstach pisanych form pochodzących z odmiany mówionej nie jest zjawiskiem nowym w historii języka perskiego. Badania nad okresami przejściowymi pomiędzy językiem staroperskim i średnioperskim, a następnie średnio- i nowoperskim pokazują, że w okresach tych również występował rozziew pomiędzy odmianą języka stosowaną w piśmie a używaną w mowie, a do tekstów powstających w ich końcowych fazach przenikały elementy bliższe już nowemu etapowi. Z początku były one postrzegane jako błędy skrybów, dziś jednak uważa się, że były raczej symptomami okresu przejściowego i nadchodzącej zmiany, która wszak nie następowała nagle, ale stopniowo.

Współczesne włączanie do tekstów literackich form typowych dla mówionej odmiany języka również dziś bywa postrzegane ambiwalentnie. Odmiana pisana, zakorzeniona w literaturze, jest znacznie bardziej zachowawcza od odmiany mówionej, w której można zaobserwować postępujące zmiany. Być może oznacza to, że jesteśmy świadkami kolejnego okresu przejściowego w historii języka perskiego – wyjątkowego o tyle, że tym razem mamy możliwość prowadzenia obserwacji zmian na bieżąco.

ISBN Print: 9783847116387 – ISBN E-Lib: 9783737016384