

Modern Spoken Persian in Contemporary Iranian Novels





unipress

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

Advisory Board / Wissenschaftlicher Beirat:

Lorella Bosco (University of Bari, Italy), Leszek Drong (University
of Silesia, Poland), Elizabeth Duclos-Orsello (Salem State
University, USA), Frank Ferguson (University of Ulster, Ireland),
Odile Richard-Pauchet (University of Limoges, France),
Monika Schmitz-Emans (University of Bochum, Germany),
Władysław Witalisz (Jagiellonian University in Kraków, Poland)

The volumes of this series are peer-reviewed.

Die Bände dieser Reihe sind peer-reviewed.

Katarzyna Wąsala

Modern Spoken Persian in Contemporary Iranian Novels

An analysis of selected 21st century novels

With 28 figures

V&R unipress

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.

© 2023 by Brill | V&R unipress, Robert-Bosch-Breite 10, 37079 Göttingen, Germany,
an imprint of the Brill-Group
(Koninklijke Brill NV, Leiden, The Netherlands; Brill USA Inc., Boston MA, USA; Brill Asia Pte Ltd,
Singapore; Brill Deutschland GmbH, Paderborn, Germany; Brill Österreich GmbH, Vienna, Austria)
Koninklijke Brill NV incorporates the imprints Brill, Brill Nijhoff, Brill Schöningh, Brill Fink,
Brill mentis, Brill Wageningen Academic, Vandenhoeck & Ruprecht, Böhlau and V&R unipress.
All rights reserved. No part of this work may be reproduced or utilized in any form or by any means,
electronic or mechanical, including photocopying, recording, or any information storage and
retrieval system, without prior written permission from the publisher.

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 © Alexei, Adobe Stock.

Printed and bound by CPI books GmbH, Birkstraße 10, 25917 Leck, Germany
Printed in the EU.

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

ISSN 2751-8345

ISBN 978-3-8470-1638-0

Contents

Acknowledgements	9
Chapter 1. Introduction	11
1.1 The aim of the research project	14
1.2 Source material	15
1.3 Organization of the book	15
1.4 Notes on transcription and notational conventions	16
Chapter 2. Theoretical framework	19
2.1 Variation in language. Basic definitions	20
2.2 Spoken and written language. Functional variation and the notion of register	25
2.3 Analytical framework for the study of register	26
2.4 Diglossia: a special kind of hierarchized registral variation	29
2.5 Linguistic change in progress	32
2.6 Linguistic variable	34
2.7 Methods applied in present research	35
Chapter 3. State of the research	39
3.1 Defining Modern Persian	39
3.2 Modern New Persian language. An overview	40
3.3 From old to modern: historical development of Persian	41
3.3.1 Old Persian	41
3.3.2 Middle Persian	42
3.3.3 New Persian	43
3.3.4 Modern Persian	46
3.4 The status of spoken and written Persian in contemporary writings in Iran and abroad	48
3.4.1 Early writings on modern Persian (18 th – early 20 th century) . .	48

3.4.2 Early 20 th century writings about Colloquial Persian; descriptive grammars	50
3.5 Contemporary theoretical approaches to the study of Modern Persian	50
3.5.1 Stylistic approach	52
3.5.2 Functionalist approach	56
3.5.3 Diglossic approach	57
3.5.4 Sociolinguistic approach	63
3.5.5 Anthropological approach	64
3.6 Terminological plurality	66
Chapter 4. Analysis of source material	71
4.1 General remarks	71
4.2 Description of the chosen source material	73
4.2.1 Dāryuš Mehrğuyi, <i>Dar xarābāt-e moqān</i> [In the tavern]	74
4.2.2 Maryam Ğa'fari, <i>Šahr-āšub</i> [Femme fatale]	76
4.2.3 Farhād Ğa'fari, <i>Kāfe Piyāno</i> [Café Piano]	79
4.2.4 Zoyā Pīrzād, <i>Ādat mikonim</i> [We'll get used to it]	81
4.2.5 Rezā Qāsemi, <i>Hamnavāyi-ye šabāne-ye orkestr-e čubhā</i> [Nocturnal Harmony of the Orchestra of Woods]	84
4.3 Method of the analysis – an outline	86
4.3.1 First reading	87
4.3.2 Qualitative analysis	87
4.3.3 Listing parameters found in analyzed forms	91
4.3.4 Consulting created list with descriptive grammars of Modern Persian	92
4.3.5 Second reading and quantitative analysis	93
4.3.6 Quantitative analysis. Presentation of the results	96
4.3.7 Conclusions	96
Chapter 5. Presentation of the results	97
5.1 Discussion of the parameters	97
5.1.1 Phonological	100
5.1.2 Morphological	114
5.1.3 Syntactic	143
5.1.4 Lexical	155
5.2 Parameter distribution in analyzed books	160
5.2.1 <i>Dar xarābāt-e moqān</i>	163
5.2.2 <i>Šahr-āšub</i>	165
5.2.3 <i>Kafe Piyāno</i>	166

5.2.4 <i>Ādat mikonim</i>	168
5.2.5 <i>Hamnavāyi-ye šabāne-ye orkestr-e čubhā</i>	170
5.3 Patterns of parameter distribution	172
Chapter 6. Final Conclusions	175
6.1 Revision of the hypotheses. MSP in literature	175
6.2 MSP parameters as linguistic variables, change in progress— historical reference	177
6.3 Ideas for future research	178
List of figures and tables	181
References	183
Appendices	193
Appendix 1. List of the quoted examples with original spelling and expected MWP form	193
Appendix 2. Quantitative analysis data table (summary)	202
Appendix 3. Summary in English	203
Appendix 4. Summary in Polish	206

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 Szcherbowski.

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.

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/toğiki*). 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 *mo-hāvere*, 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.

such as Ğamālzāde (1892–1997), Hedāyat (1903–1951), Ćubak (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 Farrozzād’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 Ğamālzāde (1892–1997), Sādeq Hedāyat (1903–1951) and Sādeq Ćubak (1916–1998) were all important intellectuals and prominent representatives of Persian literature, the founders of modern Persian prose. Ğamālzāde (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 *šē’r *neveštan* ‘to write poetry’ either *šē’r sorudan* ‘to compose poetry’ or *šē’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-šo‘arā’ Bahār described him as a “poet of the common people.” His most important and impressive works are his *tasnifāt* [...]. The *tasnif* 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 Farrozzād (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

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 Ġa’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 Ki-šlowski* [Kieślowski]. transl. by Anna Marciniowska, Tehrān: Qoqnoos (2017); *Hālinā Poš-fiātovskā* [Halina Poświatowska], *Adamak, 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 Wiśława Szymborska’s poetry can be found in: Mārek Smużiński [Marek Smurzyński] (1997) *Hezar nokte-ye bārikar az mu* [in:] *Negah-e nou* 34, Tehrān, pp. 177–183).

Pirzād, i.e., *Ārāghā-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-Exupéry'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 Nağafi's *Dictionary of Colloquial Persian*, Nağafi 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.

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 Ğā'fari (2004) *Šahr-āšub* [Femme-fatale], Tehrān: Šādān, 3) Farhād Ğā'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.

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:

ā – 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: ا – (t)a / ا – (t)e / ا – (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.

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:

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

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.

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.

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.”

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).

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 language¹⁸ (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 semi-otically 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 ideational¹⁹ 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).

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 their 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 registrational 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).

- 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 socio-linguistics, 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).