



BIBLIOTHECA
UNIV. JAGELL.
CRACOVENSIS

A 636450

-2

II

PERSO-ARABIC LOANWORDS IN HINDUSTANI

Part II. Linguistic Study



Agnieszka Kuczkiewicz-Fraś

PERSO-ARABIC LOANWORDS IN HINDUSTANI

PART II. LINGUISTIC STUDY

AGNIESZKA KUCZKIEWICZ-FRAŚ

**PERSO-ARABIC LOANWORDS
IN HINDUSTANI**

Part II

LINGUISTIC STUDY



KRAKÓW 2012

Reviewed for publication by:
Dr. hab. Kinga Maciuszak
Prof. Dr. hab. Rahul Peter Das

Cover design:
Igor Stanisławski

Typeset:
Team C&KF

Proofreading and copy editing:
Jan Czekalski

© Agnieszka Kuczkiewicz-Fraś, 2012

The volume was published thanks to the financial support of
the Institute of Oriental Studies, Jagiellonian University in Krakov

ISBN 978-83-7638-294-4

KSIĘGARNIA AKADEMICKA
św. Anny 6, 31-008 Kraków
tel./fax: 12-431-27-43, 12-633-11-67
e-mail: akademicka@akademicka.pl
www.akademicka.pl

Contents

Preface	VII
Acknowledgements	X
Remarks on Transliteration and Transcription	XI
Abbreviations	XV
1. Introduction	1
1.1. Perso-Arabic Loanwords in Hindustani: Preliminary Observations	1
1.2. Hindustani, Hindi, Urdu: a Survey of Approaches and Definitions	5
1.3. Persian and Indo-Aryan: a General Exposition of the Language Contact Situation	34
2. Scripts and Orthography	57
2.1. Writing Systems in Persian and Hindustani	57
2.2. Adaptation of the Persian Writing System in Hindustani/Urdu	68
2.3. Notation of Borrowed Words in Nagari	89
2.4. Tables of Adequacy for the Nagari and Modified Persian Scripts	115
3. Phonology and Phonetics	121
3.1. Phonological Systems of Persian and Hindustani: Similarities and Differences	121
3.2. Phonological Adaptation of the Perso-Arabic Loanwords ..	139
3.3. Quantitative and Qualitative Change of Vowels in the Perso-Arabic Loanwords	153
3.4. Changes in Distribution of Vowel Phonemes in the Borrowed Words	166
3.5. Quantitative and Qualitative Change of Consonants in the Perso-Arabic Loanwords	173
3.6. Changes in Distribution of Consonant Phonemes in the Borrowed Words	181

4. Morphosyntax	187
4.1. General Morphosyntactic Characteristics of Persian and Hindustani: Similarities and Differences	187
4.2. Morphosyntactic Classification of the Perso-Arabic Borrowings in Hindustani	190
4.3. Adaptation of Borrowings to the Hindustani Morphological System	207
4.4. Elements of Arabic and Persian Morphology in Hindustani	214
4.5. Persian and Arabic Morphosyntactic Constructions in Hindustani	219
5. Semantics and Lexicon	227
5.1. Observations on the Morphosemantic Taxonomy of Perso-Arabic Borrowings in Hindustani	227
5.2. Concluding Remarks: Typological Classifications of the Perso-Arabic Lexicon in Hindustani	233
References	249
Index	270

Preface

This book constitutes the second part of the two-volume monograph dedicated to the phenomenon of Perso-Arabic lexical borrowings in Hindustani (by some scholars also named Hindi-Urdu or Hirdu). It is a linguistic study of their adaptation, meant to demonstrate how the analyzed loanwords function in the contemporary system of the Hindustani language, through detailed examination of the adaptation processes, which occurred on all the levels of this system: graphemic, phonetic and phonological, morphological, lexical and semantic. The first part, published in 2008, is a substantial dictionary of borrowings (*Perso-Arabic Loanwords in Hindustani. Part I. Dictionary*), encompassing over 5500 entries, which over the centuries (starting approximately from the 11th-12th c.), entered into the languages and dialects of North India* as a result of the contacts between local communities and Muslim incomers. The lexical material gathered in the *Dictionary* was taken as the corpus for this study.

The analytical part of the book is preceded by a substantial theoretical introduction, in which, along with a brief explanation concerning the aim of the work and its methodical and formal assumptions, two issues, crucial for further considerations, are discussed in a synthetic but at the same time meticulous manner. The first – specifying the language of origin for the studied linguistic material, which was the basis for further analysis – is an attempt to display and scrutinize the ambiguous appellation “Hindustani” in its manifold meanings and scopes – especially in respect to two other terms, with which it is inseparably connected, namely Hindi and Urdu. Discussing the range of their meaning and

demonstrating how they have been understood by particular users, researchers, and various official bodies in different periods, helped to outline the background for both constructing a definition of Hindustani and drawing its boundaries that are in force in the presented study: the linguistic continuum current in the common speech – a *lingua franca*, unofficially, but fully functioning as a link language on the Indian subcontinent. The second problem discussed in the Introduction is the situation of linguistic contact between the Persian and Indo-Arian languages. It is described in a broader context of intercultural encounters, which intensified after the 12th c., when the Muslims took over vast parts of the Indian subcontinent, establishing their domination not only in the sphere of politics and administration but also in that of culture.

In the subsequent part of the study two areas were given special attention: an analysis of graphemic adaptation (Chapter 2) and phonetic and phonological changes (Chapter 3). The reason for this type of approach was the fact that both these areas of the Hindustani linguistic system are of a binary character, i.e. Hindustani employs two writing systems equally – the Nagari and the modified Perso-Arabic script (which was also used for original notation of the borrowed words). This dichotomy in the script can immensely influence the kind and degree of phonological changes and the phonetic realization of given lexemes. All this greatly complicates the process of studying lexical borrowings because, for example, the same lexemes can simultaneously undergo the processes of adaptation and adoption – depending on the language register. In Chapter 4, the analyzed loanwords are described in terms of their adaptation to the Hindustani morphological and syntactic system. This part of the book also includes information on the borrowed morphosyntactic phenomena, introduced into Hindustani during the process of immense lexical borrowing, and their influence on the native system. The last Chapter 5 is an attempt to categorize the discussed lexical material according to

various criteria employed in both diachronic and synchronic linguistics.

The analyses of Perso-Arabic loanwords in Hindustani presented in the book, although comprehensive, did not exhaust the whole extent of the issue – if only because of limiting the studied material to a definite set of lexemes and their sources. However, they provided a solid theoretical basis by showing and explaining the general mechanisms, which rule this phenomenon at all the levels of the language system. It is to be hoped that this work may be an inspiration for further – general or detailed – research, dedicated, for example, to such questions as the contextual functioning of the borrowings (depending on the kind of text/ethnolect/language register), intentional “macaronization” (here: Persianization or Arabicization) or the purification of language in the diachronic and synchronic approaches; internationalisms, doublets and calques (which are only briefly mentioned in the book), etc. The work may also prove useful to researchers engaged in similar studies on other Indo-Aryan languages, as it can be assumed, with a high level of probability, that at least some of the phenomena and processes described in Hindustani are analogous to those occurring in other South Asian linguistic substrata.



* Here and in most of the work, for practical reasons and in accordance with the usage found in most reference texts, “India” is used synonymously with “South Asia”, implying an older meaning of the word, not influenced by present-day circumstances and divisions. However, the name is occasionally used to denote the contemporary state (Republic of India), and in such cases the reader can infer it from the context.

Acknowledgements

I would like to express my gratitude to all those who have extended their help to me in various ways. I am particularly grateful to Prof. Kinga Maciuszak (Jagiellonian University in Krakow) and Prof. Rahul Peter Das (Martin Luther University of Halle-Wittenberg) for reviewing this book for publication and making valuable and useful comments and suggestions; to Prof. Andrzej Pisowicz for his constant and patient support in wading through the twists and turns of Iranian linguistics; to Prof. Ralph Ludwig for our creative and inspiring discussions – I'm still for more. The responsibility for the ideas and views expressed in this book, however, lies solely with me.

Moreover, I would like to thank three anonymous reviewers who, on request of the former State Committee for Scientific Research (Komitet Badań Naukowych), evaluated my research project devoted to the linguistic study of Perso-Arabic loanwords in Hindustani, giving it great recognition and the highest possible score. Thanks to their propitious appraisal I was granted the funds necessary to subsidize a two-year period of research, conducted partly (in 2012) in India (Delhi and Hyderabad).

Last but not least, I would like to express my deep and sincere gratitude to the Genuine Indologists endowed with the only correct knowledge and superior methodology – nothing could be a more effective form of motivation to work hard than the ceaseless repugnance and contempt experienced on a daily basis.

Remarks on Transliteration and Transcription

The Roman transliteration used in this book follows the systems developed for the Nagari script and the modified Arabic script. The systems, presented in the tables below, attempt to be exact, so that the reader may reconstruct the original spelling of transliterated words.

NAGARI

Vowels:

अ *a* आ *ā* इ *i* ई *ī* उ *u* ऊ *ū* ऋ *r̥* ए *e* ऐ *ai* ओ *o* औ *au*

Consonants:

क	<i>ka</i>	ख	<i>kʰa</i>	ग	<i>ga</i>	घ	<i>gʰa</i>	ङ	<i>ṅa</i>
च	<i>ca</i>	छ	<i>cʰa</i>	ज	<i>ja</i>	झ	<i>jʰa</i>	ञ	<i>ña</i>
ट	<i>ṭa</i>	ठ	<i>ṭʰa</i>	ड	<i>ḍa</i>	ढ	<i>ḍʰa</i>	ण	<i>ṇa</i>
त	<i>ta</i>	थ	<i>tʰa</i>	द	<i>da</i>	ध	<i>dʰa</i>	न	<i>na</i>
प	<i>pa</i>	फ	<i>pʰa</i>	ब	<i>ba</i>	भ	<i>bʰa</i>	म	<i>ma</i>
य	<i>ya</i>	र	<i>ra</i>	ल	<i>la</i>	व	<i>va</i>		
श	<i>śa</i>	ष	<i>ṣa</i>	स	<i>sa</i>	ह	<i>ha</i>		

Modifications:

क् *kā* ख *kḥa* ग *gā* ज *zā* ङ *r̥ā* ढ *r̥ʰa* फ *fā*

Nasals: anusvāra ṁ candrabindu ̣ ṁ

ARABIC SCRIPT (BASIC)

FOR HINDUSTANI/URDU, PERSIAN AND ARABIC

ا	A	ص	ṣ
آ	Ā	ض	ẓ
ب	b	ط	ṭ
پ	p	ظ	ẓ
ت	t	ع	ʿ
ٹ	ṭ	غ	ġ
ج	j	ف	f
چ	č	ق	q
ح	ḥ	ک	k
خ	x	گ	g
د	d	ل	l
ذ	ẓ	م	m
ر	r	ن	n
ز	z	و	V
ژ	ž	ه	h
س	s	ی	Y
ش	š	ء	ʾ (hamzah)

General remarks:

- Exact transliteration (whether from non-Roman script into Roman characters or *vice versa*) is always given in parentheses () immediately after a transliterated word/phrase, while the phonetic transcription, consistent with the IPA standards (2005), appears in square brackets [] and the phonemic notation is placed between slashes /. In the English main text of

the book, the simplified phonemic transcription without brackets applies to foreign words and phrases that have become common in English – like geographical and proper names, names of languages (e.g. Farsi, Marathi), cultural or religious terms (e.g. Sufi, Shiah), etc.

- The letters *A*, *Ā*, *Y*, and *V* are used to transliterate ambiguous Arabic characters: ا, آ, ي and و respectively, including situations in which they denote a seat for a *hamzah* (apart from ا which do not have such function). Short vowels are never marked in Roman transliteration from Arabic. To make the pronunciation of a word clear, transliteration is always followed by phonemic or phonetic notation. Roman letters corresponding to consecutive Arabic characters are separated by dashes, e.g.: باء (b-A-r^h) /baʔ^h/, except when the *tašdīd* (doubled consonant) occurs, e.g.: باء (b-ĉĉ-Y) /baʔ^ĉ:i/.
- ا is transliterated as *e* and و is transliterated as *ñ*.
- Vocalisation of words usually follows the original source and short vowels (which are not marked in Arabic script) are most often transcribed in accordance with these sources. The changes in vocalisation may sometimes appear in Persian words (comparing to Steingass's orthography), where the older forms with *i* or *u* have been replaced by forms with *e* or *o* respectively, as they are more compatible with the vocalisation of loanwords in Hindustani.
- Linguistic material quoted after Fergusson (Ferg.) and Gilchrist (OL) follows their own systems of transcription, as neither of these sources gives the original spelling in the Arabic or Nagari script.
- The word-final 'silent' *h* is given as *-h* in pausa. The Arabic grapheme for *tāʾ marbūṭah* (ة) is given as *-i* (following the ISO rule).
- In transliteration from Arabic script *ʾalif maqṣūrah* (ا) is given as *ā* (following the ISO rule).

- In the Roman transliteration the aspirates are represented as a sequence of non-aspirated relevant stops and the letter *h* in superscript, which is the usual IPA method of marking aspiration. This pattern of transliterating the aspirates aims to indicate that they should be perceived and pronounced as unitary sounds, rather than two separate consonants pronounced one after another. It also signals the use of a particular form of *h* in the Perso-Arabic script (cf. 2.2.3.).
- In transliteration from Nagari the inherent *a*, which is not pronounced, is indicated by ' except in word-final position; nasals which have no inherent *a* and non-final consonants of ligatures are not affected by this for obvious reasons.
- In quotations the original spelling and capitalization of the sources has been retained.
- Cyrillic has been Romanized according to the British standard used by the Oxford University Press, which can be found e.g. in: *The Oxford Style Manual* (Ritter 2003: 350).

Abbreviations

A	=	Arabic	n.	=	noun
act.	=	active	NIA	=	New Indo-Aryan
adj.	=	adjective	NP	=	New Persian
adv.	=	adverb	num.	=	numeral
arch.	=	archaic	OIA	=	Old Indo-Aryan
<i>arith.</i>	=	in arithmetic	OP	=	Old Persian
Av.	=	Avestan	orig.	=	original(ly)
B	=	Bengali	P	=	Persian
C	=	consonant	Pah.	=	Pahlavi
Class. P	=	Classical Persian	part.	=	participle
comp.	=	composition	partic.	=	particular(ly)
conj.	=	conjunction	perf.	=	perfectum
conn.	=	connected with	pers.	=	person
E	=	English	pl.	=	plural
f.	=	feminine	poet.	=	poetical
fig.	=	figurative(ly)	ppn.	=	postposition
fut.	=	futurum	Pr.	=	Prakrit
G	=	Greek	pref.	=	prefix
H	=	Hindustani	prep.	=	preposition
IA	=	Indo-Aryan	pron.	=	pronoun
IE	=	Indo-European	S	=	Sanskrit
imperat.	=	imperativus	sg.	=	singular
interj.	=	interjection	sl.	=	slang
inv.	=	invariable	sthg.	=	something
lit.	=	literal(ly)	subj.	=	subjunctive
m.	=	masculine	suf.	=	suffix
<i>math.</i>	=	in mathematics	T	=	Turkic
MIA	=	Middle Indo-Aryan	U	=	Urdu
MP	=	Middle Persian	t.s.	=	the same
<i>mus.</i>	=	in music	V	=	vowel



Fig. 1. A bilingual information board at the entrance to the Labyrinth (Bhul-bhulaiya) in the Barā Imāmbārā complex in Lucknow. The Hindustani name of the place is digraphical here, written in both Perso-Arabic and Nagari scripts. (Photo by the author, 2011)

1. Introduction

1.1. Perso-Arabic Loanwords in Hindustani. Preliminary Observations

*One of the chief fascinations exercised by words
– especially by loan-words – derives from their dual function
as markers of both linguistic and cultural change.*

(Shackle 1978: 73)

1.1.1. Vocabulary of Persian and Arabic origin constitutes a large, extremely important and unalienable part of Hindustani lexical resources. It is a unique heritage acquired as a result of the deep linguistic and cultural contact between the languages and their users that spanned an extended period of approximately eight centuries. Its uniqueness lies mainly in the exceptional vitality and productivity of Perso-Arabic lexemes combined with their enormous capacity to penetrate all the layers of the North Indian social and linguistic environment: there is practically no single walk of life, which has remained untouched by the influence of Perso-Arabic loanwords.

1.1.2. This book presents a complex and detailed analysis of the adaptation of Perso-Arabic lexical borrowings in Hindustani, with a special emphasis on the problem of graphemic, phonetic and phonological adaptation, which, until now, has not been properly scrutinized or has been altogether overlooked by the authors of various studies devoted to Hindustani. The existing works touching upon the question of Persian and Arabic vocabulary in Hindustani (many of them written by Indian and Russian

linguists in the 1950s and 1960s), are of a rather general or even fragmentary nature.¹ In these works, the issue of borrowings, and especially the problem of their adaptation in Hindustani is treated selectively or marginally, their description being often limited to lists of Persian and Arabic lexemes classified into semantic groups. There exist numerous reliable studies on Persian and Arabic borrowings in the languages regarded as “Western” in relation to Persia and Arabia, for example in Turkish,² Spanish,³ or the Balkan⁴ and other European languages⁵. However, the “Eastern” direction of the migration of these borrowings (South Asia and South-East Asia) still leaves much room for extensive research. It is the author’s aim that, in this context, the present work be understood as a groundbreaking study which, hopefully, is to become an important source of information not only for the

¹ Previous works providing some information about Perso-Arabic loanwords in Hindustani may be divided into: (i) general descriptions of the widely understood sociolinguistic situation in South Asia, in its synchronic or diachronic (or both) dimension (e.g. Cardona & Jain 2003; Dimock et al. 1992; Kachru et al. 2008; Shapiro & Schiffman 1981); (ii) older outlines devoted to the broadly understood cultural influence of Islam/Muslims in India, useful but lacking an in-depth formal analysis of linguistic phenomena (e.g. Hamidullah 1954; Bahri 1960; Tivārī 1969); (iii) individual detailed studies restricted to one specific issue (e.g. Apte 1974; Beskrovnýĭ 1960a; Danecki 1971; Kuczkiewicz-Fraś 2013; Zograf 1961).

² E.g.: Doerfer 1963-75; Johanson & Bulut 2006; Perry 2001b; Tietze & Lazard 1967.

³ E.g.: Maíllo Salgado 1991/1998, 1996.

⁴ E.g.: Skaljic 1966; Teufik 2000; Balić 1978; Hemmerdinger 1969; extensive bibliography on Persian lexical borrowings in Greek is available at <http://www.iranicaonline.org/articles/greece-xi-xii>.

⁵ E.g.: Corriente 2008; Cannon 1994; Cannon & Kaye 2001; Turek 2001. On studies of Arabic in contact with other languages cf. Thomason 2006.

specialists in the field of Indo-Arian languages but also for scholars from other fields, e.g. Iranian and Arabic studies.

1.1.3. The linguistic material examined in this book is characterized by a high level of complexity, resulting from: (i) its diachronic character that leads to inconsistencies, sound changes, different forms of the same words etc.; (ii) synchronic digraphia characteristic of Hindustani which uses two scripts – one better for reproducing the independent borrowed lexical units, the other – for representing the same units in the phonological, morphological and syntactic dependencies;⁶ (iii) the multifold parallel process of phonological and morphological nativization of loanwords, depending on such factors as: register of speech, script, historical moment, speaker, and geographical location. Due to all of these features, the methods of dealing with the studied material must also be varied, depending on which aspect of the phenomenon of borrowing is actually analyzed.

1.1.4. The research method applied in this work generally draws from the tradition of structuralism and can be described as a combination of the diachronic method, adequate for researching a historical lexis set within a given time frame and the analytic method – allowing for the thorough and complex description of contact-induced language change. The analyses, illustrated with multiple examples and vastly supplemented with explanatory footnotes, have been additionally put against the solid background of the social and cultural (and, to some extent, also political) context and carried out using well-established and effective theories about linguistic interference (Weinreich 1953), the linguistic contact phenomenon (Thomason & Kauffman 1988; Thomason 2001; Matras 2009) and the mechanisms and the nature of the borrowing processes (Haugen 1950; Myers-Scotton 2006; Romaine 1995).

⁶ Therefore, in this book, in most cases both the graphical forms of words and phrases are given, together with their exact transliterations and phonological transcriptions.

1.1.5. The book primarily addresses issues pertaining to loanwords understood as lexical items characterized by the borrowing of a word form together with its associated meaning (= adoption), which are usually subject to various changes (= adaptation) in the process of adjusting to the matrix of a recipient language. The terms ‘loanword’ (or, rarely, ‘loan’) and ‘(lexical) borrowing’ are used interchangeably in this study, in order to denote – in accordance with Haspelmath’s concise and accurate definition – “a word that at some point in the history of a language entered its lexicon as a result of borrowing” (Haspelmath 2009: 36). The term ‘borrowing’, however, has a broader sense: it can denote something different from a lexical item, transferred from one language to another. Therefore, according to this definition, a concept adopted without lexical importation (Haugen’s ‘semantic loan’) is a borrowing, but not a loanword. Following this logic, the transfer of Perso-Arabic script to Hindustani is also regarded a particular kind of borrowing.⁷

1.1.6. Since most of the sources, which have been referred to in this study, relate – by their authors’ decision – either to ‘Hindi’ or to ‘Urdu’ (even if they concern issues common for both), sometimes the specified names Hindustani/Hindi and Hindustani/Urdu are used to indicate that the matter discussed relates particularly to one of these forms. Such situations occur especially when the variations resulting from the use of different scripts are being characterized. Therefore, to make a long story short, Hindustani/Hindi should be understood as Hindustani depicted in the Nagari characters and, analogously, Hindustani/Urdu – as Hindu-

⁷ Haspelmath also draws attention to the fact that, in its broader sense, borrowing is used to denote the result of all kinds of transfer processes, regardless of the force behind them, i.e. no matter whether they occur due to the choice of the recipient language’s speakers or due to being imposed (e.g. by administrative regulations or political pressure) on a recipient language by the speakers of a donor language (ibid.).

stani written in modified Perso-Arabic script. However, the general and overriding assumption of this study is to both treat and study Hindustani as a single linguistic unit (language, or set of ethnolects), polymorphous and varied, but at the same time having a definite linguistic nucleus that displays all the necessary levels of language structure and vocabulary. A more precise definition of this linguistic entity will be presented in paragraph 1.2.

1.2. Hindustani, Hindi, Urdu: a Survey of Approaches and Definitions

But what is Hindūstānī?

(Ali 1917: 109)

*(...) und man muß zufrieden sein, wenn nur in demselben
Zusammenhange dasselbe Wort immer denselben Sinn hat.*⁸

(Frege 1892: 28)

1.2.1. For a linguist, South Asia remains both a paradise and a hell. A paradise, because the region – having been conceptualized since the fundamental article by Emeneau (1956) as a separate linguistic area⁹ of staggering diversity (Shapiro & Schiffman 1981), characterized by one of the best known, model *Sprachbund* (Trubetzkoy 1928), which could be called paradigmatic thanks to its richness in well-attested data (Masica 2005) – gives endless and inexhaustible research possibilities. A hell, because,

⁸ English translation by Max Kölbel: “(...) and one has to be content if at least in the same context, the same word always has the same Sense” (Byrne & Kölbel 2009: 51).

⁹ Or “convergence area” as proposed first by Weinreich (1958); cf. also Hock 1991 and Subbārāo 2012.

no matter how hard you try, nothing can be established with absolute certainty and once and for all. The confusion starts at the very basic level of taxonomy, and is brought about not only by the multiplicity and diversity of languages and dialects belonging to four different linguistic families, but also by using many different names for one linguistic entity and, conversely, by assigning one name to different individual tongues. The case of Hindi-Urdu-Hindustani is the most glaring example of such a convoluted situation (Masica calls it “the ultimate anomaly”¹⁰). It is further complicated by the vast spectrum of factors which are extra-linguistic in their nature: ranging from religious, cultural, and ethnic, sociological and economic to geographical or political. For decades, or even centuries, these factors have had an equal influence on the sphere of official and academic normalization and the speakers’ perception of and attitude towards the language they speak. Yet, as Sommerstein rightly observed, when analysing particular linguistic material or linguistic features, one must know what language one is interested in (Sommerstein 1977: 16). Therefore specifying the object of research, or to put the matter more simply, defining what language is and what language is not the source of analyzed research material, remains the crucial – though not always easy – requirement of every linguistic study.

1.2.2. However, the simplest answer to the aforementioned condition that comes to mind, evoked by a term used in the very title of this book, namely Hindustani, not only fails to clarify the issue but, on the contrary, makes the matter much more complicated than in the case of naming any other South Asian language – except maybe Hindi and probably also Urdu. Throughout history the scope and understanding of the term Hindustani varied significantly, usually relying more on socio-political circum-

¹⁰ Masica 1991: 27.

stances than on purely linguistic determinants. For that reason, before explaining in detail how Hindustani is understood in this study, and what constitutes the linguistic entity which remains our object of research, it seems indispensable to sketch the history of the formation and the use of this term at different times and by different authors.

1.2.3. Dozens of books and hundreds of articles have been written in reference to questions directly or indirectly connected with Hindustani. Among the most crucial of these questions are: the dilemma of India's *lingua franca* or link language and advantages and disadvantages of Hindi/Hindustani/English as the pan-Indian official and/or national language (e.g. Chand 1944 (esp. Chapter I); Chatterji 1949; Deshpande 2000a; Gandhi 1965; Lelyveld 2001; Mohan 2000; Nayar 1968; Rai 2005; Rao 1978); Hindi-Urdu controversy and its linguistic and non-linguistic aspects (e.g. Beg 1996; Hasnain & Rajyashree 2004; King 1974, 1989, 1994; Matin et al. 2001; Rai 1984; Russel 1996; Trivedi 2004); the role of the British in shaping the linguistic situation in India (e.g. Cohn 1985; Friedlander 2006; Khan 2006;¹¹ Kidwai 1972; Lelyveld 1993; Siddiqi 1963; Steadman-Jones 2007); common linguistic and literary heritage of Hindi-Urdu-Hindustani (e.g. Barannikov & Barannikov 1956; Barz & Yadav 1993; Beskrovnŭi 1960b; Everaert 2010; Faruqi 2003; Khan 1999; Martynyuk 2003; Orsini 2010; Rahman 2011; Shackle & Snell 1990; Singh 1995; Trivedi 2013; Zograf 1961) etc. It is worth

¹¹ This book (*Urdu/Hindi: An Artificial Divide* by Abdul Jamil Khan), labelled by the author himself as "iconoclastic", and full of peculiar statements and linguistically groundless assumptions (for example, "Urdu is one of the earliest known languages" or "Hindi came from pre-Aryan Dravidian and Austric-Munda (...) not from Aryan Sanskrit") is quoted here only for the record and for the sake of the analysis performed by its author, concerning the role played by British colonialism in splitting the alleged linguistic unity of North India.

noting that not all these works are focused on sociolinguistic or philological dimensions of the analyzed phenomena, the nature of which is genuinely and undeniably associated with, speaking broadly, the science of language. But since language remains also one of the most important constitutive factors of the human national identity (e.g. Fishman 1973; Gellner 1983; Kymlicka 1995), the role played by Hindustani (and Hindi, and Urdu) – mainly in the late 19th and early 20th centuries – in the construction of the contending Hindu and Muslim nationalisms, engaged in mutual combat, has become a marginal facet, often examined by authors of studies belonging to the realm of history, political sciences or sociology (e.g. Aneesh 2010; Ayres 2009; Brass 2005; Orsini 2002; Rai 2007; Van der Veer 1994; Zavos 2000). It is not our goal, however, to picture or discuss research approaches and scholars' concurrences and discrepancies which over the years have accumulated around all these issues; much less is it an attempt at resolving problems that for decades have remained unsolvable (and are likely to remain so for many years to come). Let us focus instead on displaying and analyzing the ambiguous appellation "Hindustani" in its manifold meanings and scopes – certainly set together with two other terms, inseparably connected with it, namely Hindi and Urdu. Discussing the range of their meaning and demonstrating how they have been understood by particular users, researchers and various official bodies in different periods will help to outline the background for both constructing the definition and drawing the boundaries of Hindustani that are in force in this study.

1.2.4. These two words, Hindi and Hindustani, being actually adjectives (viz. 'of India; belonging to the country of Hind/Hindustan') but used substantively, mostly as proper names of languages, are themselves of Iranian origin and can be traced back hundreds of years. The *Hobson-Jobson Dictionary* confirms the name of the language in the Westernized form "Indostan" as

early as 1616.¹² A similar designation (“industanes”) appears in a Jesuit account from the times of Akbar (r. 1556-1605). It was written in French du Jarric’s *Histoire...*, completed in 1611, where we find the following passus:

La nuict de Noel on fit ioüer vne Pastorale en l’âgue Persique, sur le mystere de la Natiuité de N. Seigneur, y entre-meslant quelques dictons ou sentences **industanes**.¹³

Of course, we cannot be sure which local language du Jarric was referring to. We can only assume that it was an Indian language used at the Akbar’s court, but such a guess still leaves us with at least a few possibilities. It might have been a North-Indian vernacular based on K^haṛī bolī speech (cf. *infra* 1.2.5.), which was widely adopted by the Muslim newcomers as a common language of interaction with the local population (Kachru 2008: 82),

¹² The *Hobson-Jobson Dictionary* itself proves great confusion over explaining the terms Hindi/Hindustani and mutual relations between them. We read, on the one hand, that “**Hindee** (...) is most commonly **used for those dialects of Hindustani speech** which are less modified by P.[ersian] vocables than the usual Hindustani and which are spoken by the rural population of the N.W. Provinces and its outskirts” (Yule & Burnell 1996: 417), and on the other that “Hindustanee, s. *Hindūstānī* (...) [is] **the language** (...) **developed out of the Hindi dialect** of the Doab chiefly, and of the territory round Agra and Delhi, with a mixture of Persian vocables and phrases, and a rediness to adopt other foreign words. It is also called Oordoo (...)” (ibid.). [All the emphases are mine.] However, depending on how the terms Hindi and Hindustani are defined, both the statements, though contradictory at first glance, may be considered to be correct. On reciprocal dependencies and shared identity of Hindi and Hindustani *vide infra*, 1.2.7. ff.

¹³ Du Jarric 1611: 673. In an English translation by Payne: “On Christmas night, they played a pastorelle in the Persian language, illustrating the mystery of the Nativity of our Saviour, introducing into it various Hindustani phrases and sayings” (du Jarric 1926: 41). It seems that this is the earliest known mention of the name Hindustani by the non-Persian speaking author referring to an Indian vernacular.

and ultimately developed into Hindi and Urdu. Alternatively, it might have been Braj b^hāṣā,¹⁴ the vernacular of the Braj district around Agra (Akbar's capital), described by Allison Busch as the "other literary language of the Mughal Court" (Busch 2010: 267). It might have even been Sanskrit.¹⁵ The derivational base of the word itself, however, viz. the Iranian toponym in its various forms (Av. *hindav-*, *hendav-* (not certain), OP *Hiⁿdu-*, Pah. *Hind* etc.), referring to the (lower) Indus country, dates as far back as the ancient times of Younger Avesta (Karttunen 1995: 152). The fact that both the words, *hindī* and *hindūstānī*, are derivatives

¹⁴ Cf.: "'Brajbhasha' had many names in the premodern period, including 'Bhasha', 'Hindi', and 'Gwaliyari' (among others)" (Busch 2010: 268). From roughly the 16th until the 19th century Braj b^hāṣā was a literary dialect of Hindi, widely used for composing poetry. Religious (Kṛṣṇaite) devotion and courtly patronage contributed to the growth of its meaning and, consequently, until the end of the 19th c. it dominated the Hindi scene to such an extent that it was not the neighbouring K^haṛī bolī, but Braj b^hāṣā which was essentially meant by the designation 'Hindi' (cf. Shackle & Snell 1990: 13-14). Over the centuries Braj b^hāṣā vocabulary penetrated Hindustani, which itself was essentially a colloquial register and never became the language of sublime literature or high culture (Kachru 2008: 81; on registers of Hindustani see paragraph 1.2.14.). The features of Hindustani and Braj b^hāṣā mixed in works of many writers (McGregor 1999: *passim*), leading to a certain amount of convergence and, ultimately, to the emergence of later Standard Literary Urdu and Modern Standard Hindi. The complicated matters of mutual relations and dependencies between K^haṛī bolī, Braj b^hāṣā and other local varieties have been analyzed e.g. in Nespital 1994.

¹⁵ Using the term 'Hindi' when referring to Sanskrit was not an uncommon practice in Indo-Persian texts. I am indebted to Audrey Truschke for indicating relevant examples, like those of Firištah (1560-1629), who mentions translation of the *Mahābhārata* from *hindī* (evidently referring to Sanskrit), or Abū 'l-Faẓl (1551-1602), whose term *hindī*, employed by him repeatedly, especially in the final part of the *Ā'in-i Akbarī* devoted to Indian intellectual traditions, "has a range of meanings, including Hindi, Sanskrit, and Indian" (cf. Truschke 2012: 238 and 262 respectively).

created by means of the highly productive Persian adjectival suffix *-ī*, induce the belief that they must have been in common use long before they were noted down (see also Rahman 2011: 32). Additionally, the logical assumption is that in the sense of “the Indian language/the language of Hind” they were not used by the Indian native-speakers, but rather by the foreign newcomers, who contrasted their own non-Indian tongues with the local “Indian” speech.¹⁶ According to Tara Chand, the name Hindi/Hindvī was assigned to an “Indo-Aryan dialect” as early as in the 13th century. This dialect was also sporadically used as a medium for writing poetry (Chand 1944: Ch. III). Faruqi points out the diversity and fortuity of nomenclature referring to this linguistic entity over the years. He enumerates such names as: Hindvī, Hindī, Dihlavī, Gujrī, Dakanī and Reḡhtah, Moors, Hindoostanic, Hindoostanee and Indostans¹⁷ (Faruqi 2001: 22). A similar range is presented by Rahman – the names he specifies and unravels one after the other are: Hindi/Hindvi/Hindui, Dehlavi, Gujri/Gojri/Gujarati, Dakhni, Indostan/Moors, Rekhta and Hindustani, the latter being examined by the author in great detail (Rahman 2011: 31-41).¹⁸ Leaving aside sheer variety of terms and their

¹⁶ Cf. Carl Ernst’s remarks on the terms ‘Hindu’ and ‘Hinduism’, which in his opinion represent, in fact, foreigners’ concept of India and do not correspond – either in terms of geography, or religion – to any indigenous Indian notion (Ernst 1992: 22-23). I am grateful to Audrey Truschke for pointing me toward this book.

¹⁷ All forms of the names quoted here retain the original spelling of the sources. Cf. also various names discussed in Bailey 1938: 7-10.

¹⁸ Note that the attempts of naming or defining vernaculars used in vast areas of North India for everyday communication and later also for literary and administrative purposes, have a very long history. Such information, including remarks on linguistic diversity and phonological changes occurring between the eastern (Prācyā) and western (Udīcyā) forms of Aryan speech current in North India already appeared in the Brāhmaṇas (e.g. *Kauṣītaki* and *Pañcaviṃśa*, ca. 800-500 BC, dates after Macdonell

exact explanations which might be found in the books referred to, let us now take under closer scrutiny the one which is central to this study.

1.2.5. Having gained currency as a common name of the language sometime at the turn of the 18th century,¹⁹ the label Hindustani has been used in several senses. Its broad definition may be found in *Encyclopaedia Britannica*,²⁰ where Hindustani is described as India's most common *lingua franca*, mixed speech, which started developing during the 13th century and was initially used as a contact language to facilitate interaction between the speakers of K^haṛī bolī, a regional dialect current in and around the Indian cities of Delhi and Meerut, that developed out of Śaurasenī Apabhraṃś (cf. Bailey 1926), and the speakers of Persian, Turkic, Pashto and Arabic, who migrated in great number to North India after the establishment of Muslim hegemony in the early 13th century (cf. also Bhatia & Machida 2008: 1). In this sense the term was employed by the earliest authors of grammars and dictionaries, viz. Jean Josua Ketelaar (1698), Benjamin Schulze (1745), Georges Hadley (1772), John Fergusson (1773) and John Bortwick Gilchrist (1787-90; 1796; 1798) etc. There also seems to be no doubt that at least some of these authors were aware of existing variants of the "Indian language" and distinguished between the vernacular, which played the role of the common tongue and the genuine language of the Hindus.

1900), and is then repeated in Patañjali's *Mahābhāṣya* (2nd c. BC). Secular Sanskrit, codified in the 5th c. BC by Pāṇini, was also a vernacular, based on the spoken Western dialect. It was then adopted as a link language by the entire Brahmanical world, and spread its supra-regional influence far beyond India. Cf. Chatterji 1960: 62ff; also see Hock 1991: 473-474.

¹⁹ More on the late 17th and 18th c. sources see also Grierson 1916: 3-12.

²⁰ <http://www.britannica.com/EBchecked/topic/266468/Hindustani-language> (accessed 2012-10-17).

In the short preface (voor Reeden) to his grammar²¹ Ketelaar observed:

The language used by the Hindustānī, Gujarathi and Bengali Muhammadans for the most part depending on Persian. The language of pagans, whose countries are now being ruled by the Moors has been blended with Persian. The resulting mixed language which is being used by the Moors nowadays is mostly called 'Hindustanki boeli' or 'Hindoestani', although it rightly might be styled degenerate Hindustānī, for the tongue of the Moorish Muhammadans is most different from the speech of the true Hindustani or pagans. So, in order to make a distinction between them, they call the Moorish language as stated above, 'Hindoestanki boeli'.²²

1.2.6. Schultze clearly defined 'lingua Hindostanica' as the language used by the inhabitants of the land, known under the name 'Hindostan' (which he considered to be the specific and the most appropriate name of this country), who call themselves 'Hindostanos'.²³ This language, being a vernacular in the realm of the Great Moghul, was prevalent far and wide together with its

²¹ *Instructie off onderwijnsinge der Hindoustanse, en Persiaanse Talen, nevens hare declinatie en conjugatie, als mede vergeleykinge der hindoustanse med de hollandse maat en gewichten mitsgaders beduydingh eenieger moorse namen etc.*, door Joan Josua Ketelaar, Elbingensem en gecopieert door Isaacq van der Hoeve, van Uytreght. Tot Leckenaauw A° 1698. [tr.: *Instruction or Tuition in the Hindūstānī and Persian languages, besides their declension and conjugation, together with a comparison of the Hindūstānī with the Dutch weights and measures, likewise the significance of sundry 'Moorish' names, etc.*, by Joan Josua Ketelaar, Elbingensem, and copied by Isaac van der Hoeve of Utrecht. At Lucknow, A° 1698] (Vogel 1936: 820-821).

²² Translated from Dutch by Bodewitz (Bodewitz 1994-95), cited after Bhatia & Machida 2008: 37.

²³ Cf.: "Verumenimvero nomen ejus proprium ac rectissimum, Hindostan vocatur (...) Ast incolae hujus terrae se metipsos Hindostanos, & linguam, qua utuntur, Hindostanicam appellare solent" (Schultze 1745: 1).

several varieties.²⁴ One of these varieties the existence of which Schultze was aware of, was Dakkani. He called the language ‘Dacni’ and described it in one of his Mission Reports (he was a missionary of the Danish-Halle mission in Tranquebar in South India) as the Moorish speech current in the province Deccan, into which many Arabic words have been imported, especially in the theological style, but by nature very different from the Arabic, and a bit more approximate to the Persian, with which it has many words in common.²⁵

1.2.7. Gilchrist, on the other hand, uses the names Hindustani and Hindi interchangeably (for one linguistic entity) in his writings. We can see in the following fragment from his *Dialogues, English and Hindoostanee* that he labels the same language *hindoostanee* (or *hinddee*) in English, but *hinddee* in the transliterated text:²⁶

*pray, sir, in your opinion, whether is the **hindoostanee** or persian language the most difficult?*

kuho şahib, toomharee danist meñ **hinddee** zuban ki farsee kuonsee ziyadu mooshkil hue?

*as to the difficulty of the **hinddee** there can be no doubt, but it is more necessary than the persian; we therefore are striving to learn it. can you teach us?*

²⁴ Cf.: “Verumenimvero cum Magni Mogulis territorium, in qua Hindostanica lingua Vernacula est, se longe lateque extendat, ejusdem linguae aliquot dialectos quoque dari necesse est” (ibid.: 3).

²⁵ Cf.: “Die Mohrische Sprache wird von der Provinz Decan genannt Dacni (...) In diese Sprache sind zwar viele arabische Wörter gebracht worden, sonderlich in Stilum Theologicum, sie ist aber an sich von der Arabischen sehr unterschieden. Der Persischen kommt sie etwas näher, als mit der sie viele Wörter gemein hat” (qtd. in Pelikan 1987: 15).

²⁶ Friedlander 2006: 44 also draws attention to this practice.

hindee ke ishkal meṇ koochḥ shoobuh nuheen, lekin farsee se ziyadu durkar, isee waste hum muqdoor bḥur seekḥte huen. toom hum ko sikhla sukoge?

*do say, in your idea, for the person who has transactions of all sorts, with both the high and the low, throughout hindoostan – of these two languages, viz. the **hindoostanee** and persian, which is the most requisite?*

bḥula toomhare khiyal meṇ kuho to oos shukḥs ke waste jo moo,amilu ya suro kar rukḥta hue, kya awam kya khuwaṣ se, hindoostan bḥur meṇ, in dono zubanoṇ meṇ se, yuṇe farsee uor **hindee** kuonsee poorzuroor hue?²⁷

At the same time Gilchrist tells apart Hindoostanee/Hindee and “the proper Hinduwee (...) written and read from left to right, in a character called *Naguree*” (Gilchrist 1796: 4), which he occasionally calles Bhakha²⁸ or “rustic tongue”.²⁹ It also seems that Gilchrist saw the “proper Hinduwee” as a source of Hindustani, equall in this regard to Arabic and Persian.³⁰ The distinction was in line with the linguistic policy of the British, who promoted Hindustani extensively during the 19th c. and initiated efforts at its standardization and separation. As Vasudha Dalmia points out, the whole subsequent notion of Hindi and Urdu as two autono-

²⁷ Gilchrist 1826: 136-137. The emphases are mine.

²⁸ Cf.: “In the *B. hak ha* or pure *Hinduwee* there are still many elegant poems (...)” (Gilchrist 1796: 335).

²⁹ Cf.: “the fox and stork in the rustic tongue, or *bhukhu*” (Gilchrist 1826: iv).

³⁰ Cf.: “let us now try the celebrated moral ode [...] which i shall here translate as literally as possible, and at the same time agreeably to the rules of construction, in order to instruct the reader effectually how to do so hisself, with any piece of writing in this dialect [viz. Hindustani], that he may hereafter wish to attend to, as a student, translator, or oriental linguist as such, however, should he aim at a very distinguished place, he will find it advisable to dip also occasionally in the sources theirselves, viz. the arabic, persian, and hinduwee tongues” (Gilchrist 1826: 255).

mous linguistic entities was set into motion and cemented in the early 19th c., when two separate linguistic departments were established at the College of Fort William: one for Hindustani, which later was made tantamount to Urdu, and the other for Hindi, later known as Hindi (Dalmia 2003: 1378).

1.2.8. Over a centenary after Gilchrist, Suniti Kumar Chatterji followed in his footsteps, and equated Hindi with Hindustani, declaring in 1940:

I use both the names *Hindi* and *Hindusthani*³¹ to mean in a general way the great Indo-Aryan ‘Lingua Franca’ which may be described as ‘the *kā-mē-par-sē*, *is-us-jis-kis* and *nā-tā-ā-gā* Speech’ (taking into note its characteristic post-positions and inflexions for the Noun, the Pronoun and the Verb) and which forms the basis of the two cultured and literary languages, the Hindu *High-Hindi* (or *Nāgarī Hindi*) and the Muhammadan *Urdū*. *Hindī* is the oldest and simplest of names for the current speech of Northern India (from the East of the Panjab to Bengal) after the Turki conquest in the 12th-13th centuries, and I use it in its old connotation which is still present among the masses. (Chatterji 1960: 138).

Chatterji’s great predecessor, George Abraham Grierson, who for thirty years (1898-1928) conducted the Linguistic Survey of In-

³¹ In his series of lectures, delivered in 1940 in Ahmedabad, Chatterji differentiates between Hindūsthānī “with *-sthān* from Sanskrit *-sthāna* as the second element”, a strange hybrid neologism which he uses to denote the “popular North Indian ‘Lingua’ Franca”, and Hindustani, “a much later and a more cumbrous formation”, synonymous, according to him, with Urdu (cf. Chatterji 1960: 138-139). However, neither in his earlier works (e.g. Chatterji 1931) nor in writings published after 1940 does he again make such a distinction and in his other texts we find only the name Hindustani, usually contrasted with High Hindi, on the one hand, or with Urdu, on the other. Cf. e.g.: “The Hindi or Hindustani speech (...) is the standard speech for some 40 millions (...) who use it (in either of its two forms High Hindi and Urdu) as their language of education and literature” (Chatterji 1949: 121).

dia (henceforth: LSI), described Hindustani as a dialect of Western Hindi, recognized by him as being “the most important, both politically and in point of numbers” (Grierson 1916: xiii) among all the Indian vernaculars. But he also emphasized the fact that Hindustani “is not a typical dialect of a language” (ibid.) and is responsible for Western Hindi spreading throughout the whole of India:

[Hindi] is not only a local vernacular, but in one of its forms, “Hindōstānī”, it is spoken over the whole of the north and west of continental India as a *lingua franca* employed alike in the court and in the market-place by everyone with any claim to education. (...) in the days of the early Muṃl sovereignty of India it was the common speech of the *bāzār* of Delhi (...). From Delhi it was carried all over India by the Muṃl armies. (Grierson 1918: 51)

It seems that Grierson, despite his denials, followed the perception of Hindi common among the Europeans of his times and used this term in two mutually contradictory senses (cf. also definitions given in the *Hobson-Jobson Dictionary*, q.v. fn. 12): either as a collective term, to indicate all the vernaculars spoken in the vast area between Bengal and Punjab, divided into two main big groups of dialects: Eastern Hindi and Western Hindi (which understood in such a way is the source of Hindustani, perceived as one of its forms, *vide supra*); or to indicate the Sanskritized variant of Hindustani which in this sense consists of three main linguistic varieties, namely:

Literary Hindōstānī proper, employed by both Musalmāns and Hindus for literary purposes and as *lingua franca*; Urdū,³² employed chiefly by Musalmāns and by Hindūs who have adopted

³² Urdu being itself differentiated into two separate varieties: “the standard literary form of Delhi and Lucknow, and the Dakhinī, spoken, and used as a literary medium, by Musalmāns of Southern India” (Grierson 1916: 1).

the Musalmān system of education, and a modern development, called Hindī, employed only by Hindūs who have been educated on a Hindū system. (Grierson 1916: 1)

But the varieties marked out by Grierson are only apparently equivalent, and – similarly to Chatterji – he also considered Hindustani as a broader phenomenon, not a counterpart of the other two forms but a combination of both, which is clearly visible in the concluding definition of the three terms that he propounded in the LSI:

Hindōstānī is primarily the language of the Upper Gangetic Doab, and is also the *lingua franca* of India, capable of being written in both Persian and Dēva-nāgarī characters, and without purism, avoiding alike the excessive use of either Persian or Sanskrit words when employed for literature. The name ‘Urdū’ can then be confined to that special variety of Hindōstānī in which Persian words are of frequent occurrence, and which hence can be only written in the Persian character, and, similarly, ‘Hindī’ can be confined to the form of Hindōstānī in which Sanskrit words abound, and which hence can only be written in the Dēva-nāgarī character. (Grierson 1916: 47)

1.2.9. The end of the 19th c. and the first decades of the 20th c. witnessed the widening of the gap between Hindi and Urdu, both sociolinguistically marked – the former as the Hindu language, the latter as connected with Muslims and Islam. Against this background, religiously neutral Hindustani flourished as a common speech of the British Raj, and even though it enjoyed lesser literary prestige than (High) Hindi or Urdu, it was learned widely throughout India, even in Bengal and the Dravidian south (Hock 1991: 475). However, the difference between Hindi and Urdu and the basis for distinguishing the two characterized forms were, in most cases, limited merely to the spheres of script and lexis (as in the Grierson’s definition, *vide supra*). Such a commonly accepted assumption corresponded well with

two popular convictions. The first, remaining a rather European point of view, was that the Hindi language (as well as Urdu prose) was invented by the teachers of the College of Fort William in a desire “to create a Hindustānī for the use of Hindūs” (Barannikov 1936: 374³³). Hindi, as perceived in this way, was subject to widespread disapproval,³⁴ an example of which can be found in de Tassy’s 1874 report, where the author supports his own opinion with a citation from a letter by Lieutenant-Colonel J. Chambers, Professor of Hindustani at the University of Oxford:

Je suis charmé de voir que vous et beaucoup de savants hindoustaniens soutiennent l’usage de l’urdu contre celui de l’hindi moderne. Je n’aime pas ce dernier dialecte comme étant un mélange désagréable de persan, d’arabe, d’anglais et de termes locaux avec un grand nombre de composés sanscrits nulle part à ma connaissance en usage dans l’Inde (...) ³⁵

³³ Barannikov quotes, i.a., Grierson’s opinion on the matter and popular sources, such as *Encyclopaedia Britannica*. According to Grierson, literary Hindi, or High Hindi, was (re)created by the teachers of the College of Fort William by “taking Urdū, the only form then known, as a basis, ejecting therefrom all words of Persian or Arabic origin, and substituting in their place words borrowed from, or derived from, the indigenous Sanskrit” (Grierson 1918: 52).

³⁴ Views of this kind are not unfamiliar also to more contemporary scholars, cf.: “The language called ‘Hindi’ was until very recently only an artificial, public style of what most linguists call Hindustani or, following Kelkar (1968), Hirdu” (Singh 1995: 1).

³⁵ De Tassy 1875: 3-4. In English translation by S. Kamal Abdali: “I am delighted to learn that you and many scholars of Hindustani support the use of Urdu against that of modern Hindi. I am not fond of the latter, which is an unpleasant mixture of Persian, Arabic, English, and local phrases with a large number of Sanskrit compounds, none of which, to my knowledge, are used in India (...)” (de Tassy 2011: 142-143).

The other belief arose among the Hindu scholars, as a result of a not necessarily unprejudiced revision of the previous opinion and stated, in turn, the purely artificial origin of Hindustani.³⁶ According to these scholars, the process was initiated by the English, entirely “for political reasons”. Commenting on such theories Barannikov wrote with a great deal of criticism:

These scholars imagine the affair to have been enacted in the following way: the English selected out of the Urdū and Hindī words common to both languages, kept the Hindī grammar and in this way invented a new language. Such a mode of explaining the formation of the idiom which the Hindus themselves call Hindustānī is quite mechanical and contradicts their own words, as the Hindū scholars declare that the Hindustānī is just the form of language *jise sab log bolcāl mē kām mē lāte haī* “which is used by all in conversation. (Barannikov 1936: 389)

But regardless of such opinions, the aim of which was to boycott Hindustani by presenting it as the invention of the colonial authorities, the “British language of command” (Cohn 1985), and pronouncing its defence as anti-national (Rai 2007), many scholars still advocated Hindustani as the most representative language of India, remaining in common use as the medium of literature and culture “in the United Provinces (...), Panjab, Rajputana, Central India, Bihar, and the Central Provinces as well, and (...) understood everywhere in the rest of Aryan India” (Chatterji 1922: 186). For the secular nationalists, like Gandhi,³⁷ Nehru or Maulana Abul Kalam Azad, Hindustani, being a compromise between Urdu and High Hindi (Chatterji 1960: 168), was the perfect form of a literary language which thanks to its inclusive quality, could unite the Hindus and Muslims and stop the grow-

³⁶ Out of three literary forms set up on the basis of K^haī bolī, viz. High Hindi, Urdu, and Hindustani (Barannikov 1936: 389).

³⁷ On Gandhi's views on the linguistic questions and his attitude towards Hindustani cf. Lelyveld 2001.

ing division between these two religious communities (cf. Aneesh 2010: 98). In 1947, on the eve of independence, the Language Sub-Committee of the Constituent Assembly, supported by many national leaders with a pluralist orientation, recommended that Hindustani, “written (...) either in Devanagari or the Persian script shall, as the national language, be the first official language of the Union” (qtd. in Rai 2007: 111). However, two years later this recommendation was ultimately not adopted by the Constituent Assembly in reaction to the partition of the sub-continent and Pakistan’s declaration of Urdu as its national language.³⁸ The common tongue has been “divided against itself because of religious and political differences” (Hock & Joseph 1996: 61). Hindi, written exclusively in the Nagari script, was proclaimed the official language of the Republic of India; the idea of Hindustani slowly sank into oblivion and faded away.³⁹

1.2.10. In post-independence India, Hindustani has practically disappeared from the linguistic map created on the basis of data gathered during the consecutive decennial censuses.⁴⁰ It is

³⁸ More informative and detailed analyses concerning the battle for the national language of India can be found e.g. in: Aneesh 2010; Hasnain & Rajashree 2004; King 1994; Rai 2007 and Trivedi 2004.

³⁹ “Partition killed Hindustani” – declared Granville Austin (Austin 1966: 277). The idea of Hindustani as the factor uniting the two distinct religious communities had become, in the moment of partition, unacceptable not only for Hindu but also for Muslim nationalists. In the opinion of the Muslim League “Urdu and Urdu alone is «the common language» of both Muslims and Hindus in North India. Replacing it with Hindi, «even under the pseudonym of Hindustani», is «an attempt to suppress those parts of Indian culture, which are not exclusively Hindu but in the making of which Muslims too have had a hand»” (qtd. in Lelyveld 1993: 677-678).

⁴⁰ For example, in Uttar Pradesh the percentage of speakers who declared Hindustani as their mother tongue decreased from 99,75% in 1921 and 99,68% in 1931 to 10,67% in 1951 and 0,01% in 1961. Consequently, in the 1961 census, Hindi was declared as the mother tongue by 85,39%

not enumerated in the documents of the 2001 census (recent available data), either among 22 scheduled languages (viz. those which are specified in the Eight Schedule), or in the group of “other languages and mother tongues with speakers’ strength of 10,000 and above at the all India level, grouped under each language”. There are a total of 122 languages and 234 mother tongues in the census documents – Hindustani is non-existent in this medley.⁴¹ But the name itself is still used and the linguistic phenomenon it designates still analyzed, remaining the subject of not only academic, but also public discourse.⁴² The recurring discussion, raising the matter of “India’s *lingua franca*”, or “link language”, occasionally flares up in the press devoted to socio-political and cultural issues, often featuring academic articles by top scholars, like the daily *Jan’satta*, published in Hindi, from Delhi, the English-medium daily *The Hindu* (Chennai) or *Economic and Political Weekly* (Mumbai).⁴³

and Urdu – by 10,70% of the respondents. These data, reflecting the effects of the intensified pro-Hindi and pro-Urdu movements in the 1950s, confirm Brass’s opinion that “the language censuses in North India are political, not philological, documents” (Brass 2005: 180), and as such they cannot be treated as a reliable source of information or a representation of the actual linguistic situation.

⁴¹ http://www.censusindia.gov.in/Census_Data_2001/Census_Data_Online/Language/Statement1.htm (accessed 2012-01-12).

⁴² Linguistic questions concerning Hindustani have been very often discussed against the ethnic or religious factors or as the manifestation of various political or nationalistic forces (e.g. Amritavalli & Jayaseelan 2007; Aneesh 2010; Rai 2005; Windmiller 1954).

⁴³ See, for example, Nayar’s article ‘Hindi as Link Language’ (Nayar 1968) and the subsequent polemics: Rao 1968; Karve 1968; Datt 1968, or three decades later ‘Hindustani in India’ by Deshpande (2000a) and following commentaries: Mohan 2000; Deshpande 2000b; Srivastava 2000. The vivid debate concerning all-India link language took place recently in the

1.2.11. Although Hindustani, to which an alternative appellation Hindi-Urdu is often attributed by contemporary researchers and users, is commonly believed to be the *lingua franca* of India, it is not officially recognized (e.g. it is not enumerated on the Ethnologue list of South-Asian languages⁴⁴) and its status as a vernacular renders it difficult to measure its precise number of speakers. However, in its *bāzār* or marketplace varieties it is used as a *lingua franca* also on the peripheries of the “Hindi belt”, in the areas where other Indo-Aryan languages are prevalent. Usually it has a reduced form, which may be classified as pidginized (e.g. Shapiro & Schiffman 1981; Apte 1974), characterized by numerous simplifications.⁴⁵ Nevertheless, specialists, both lin-

pages of *The Hindu*, provoked by the text published by Markandey Katyu, the Chairman of the Press Council of India (Katy 2012).

⁴⁴ Cf.: “«Hindustani», though not listed separately in India, refers here to the unofficial lingua franca of northwest India. Has a lexical mixture in varying proportions of Hindi (vocabulary derived from Sanskrit) and Urdu (vocabulary derived from Persian or Arabic).” (http://www.ethnologue.com/show_country.asp?name=IN; accessed 2012-09-03). According to the Ethnologue, the name Hindustani formally applies to the several quite different languages, based on various dialects of Hindi, that are spoken outside of the South Asia: Caribbean Hindustani – in Suriname, Guyana or Trinidad and Tobago (http://www.ethnologue.com/show_language.asp?code=hns; accessed 2012-09-03), and Fijian/Fiji Hindustani in Fiji (http://www.ethnologue.com/show_country.asp?name=FJ; accessed 2012-09-03).

⁴⁵ S. K. Chatterji, describing Hindustani “of the streets and *bāzārs*” as a jargon dialect of Calcutta, second in this area only to Bengali, “very living speech” in which “a great deal of transactions of life are carried on”, enumerates such simplifications as: elimination of grammatical gender and passive and ergative constructions or unification of future tense conjugation forms (Chatterji 1931). On the other hand, both Chernyšhev and Apte, characterizing the Bombay variety of Hindustani (Apte calls it Hindi-Urdu) underline the strong influence of Marathi and other regional languages on this vernacular, manifesting themselves e.g. in lacking aspiration of consonants and loss of case, number, gender and person in nouns and verbs

guists and non-linguists (regardless of the official rosters), agree on the fact of its well established existence:

We might call the latter ‘Hindi-Urdu’, ‘Urdu-Hindi’, ‘Hirdu’, ‘Urdu’, ‘Hindustani’, or ‘Hindi’, but whatever the name, what is shared by Hindi and Urdu has become the true national link language, the language best known through ‘Hindi’ films which are popular all over India and in many places abroad. (King 1994: 175)

Apart from the differences in terminology, most researchers explain it in a similar way, as the language which is

neither the highly Sanskritized Hindī nor the highly Perso-Arabicized Urdū of the purists, but rather the ‘common core’ language generally spoken by educated Hindus and Muslims in an area of the South Asian subcontinent that includes the states of Delhi, Uttar Pradesh, Madhya Pradesh, and Bihar, as well as parts of the Punjab and Himachal Pradesh, and in such distant urban centers as Karachi and Calcutta, Delhi and Dacca, Hyderabad and Lahore. (Narang & Becker 1971: 646 fn. 1)

Of course, definitions vary in details; Brass, for example, following Austin’s reflections on the Indian Constitution (Austin 1966), describes Hindustani as “the common spoken language of north India, written in both the Devanagari and Persian-Arabic scripts, and drawing its vocabulary from English and Urdu, more than from Sanskrit”, differentiating it clearly, at the same time, from both Hindi and Urdu (Brass 1994: 158). Even those who deny its present existence can hardly avoid using its name, like Friedlander, who predicates that the term Hindustani “is no longer used as a name to designate a living language, but is associated with the British colonial period and various failed attempts to

which – together with elimination of the ergative construction – cause the increasing and paramount role of postpositions in syntax (cf. Chernyšhev 1971; Apte 1974).

foster a compromise between Hindi and Urdu”⁴⁶ (Friedlander 2006: 40), yet in his study he treats all three terms (Hindustani, Hindi and Urdu) synonymously to a large degree, using mostly the term Hindustani when speaking about the linguistic entity they designate.

1.2.12. One of the most stereotypical definitions of Hindustani can be found in a German handbook, published in 1945 by Otto von Spies and Ernst Bannerth. It reads as follows:

Unter Hindūstānī verstehen wir eigentlich zwei Sprachen Indiens: Urdū und Hindī. Das Urdū wird meist von den Mohammedanern, das Hindī ausschließlich von den Hindus gebraucht. Grammatik und Aufbau sind in den beiden Sprachen gleich: beträchtliche Unterschiede bestehen nur in bezug auf den Wortschatz und die Schriftzeichen.⁴⁷ (Spies & Bannerth 1945: 1)

⁴⁶ Incidentally, such a categorical statement would be rather difficult to justify and carries some additional risk. It is an irrefutable fact that languages change, some even die and disappear, and other languages replace them. But how would it be possible, that the language of the well established, centuries-old tradition, used as a spoken and to some extent also literary medium by millions, taught and learnt in the subcontinent and beyond – had this language vanished all of a sudden in a puff of smoke, having been replaced by another language(s), and only because it had not been given official recognition when the Constitutional Assembly opted for Hindi against Hindustani? Obviously, the question is provocative, since Friedlander meant “the term” not “the language” which was not used anymore. But lack of acknowledgment of the name can appear only as a result of the denial of the existence of its reference (cf. Frege 1892). And here lies the main danger, because, consequently, the non-existence of the name in public awareness may (and most probably will) lead to doubts about, and, as a result, to the denial of the existence of the reference itself. Fortunately, we can safely abandon further ontological considerations, since in the case of Hindustani both these denials are groundless.

⁴⁷ “Under the term Hindustani we mean two languages of India: Urdu and Hindi. Urdu is used mostly by the Muslims, Hindi exclusively by the

Further in the introduction the authors specify Hindustani as a general slang, or colloquial language, that employs Arabic and Persian loanwords rather than Sanskrit lexemes, which are used exclusively to name religious and cultural concepts or objects connected with Hinduism.⁴⁸ However, a closer look at the above-given description reveals two problems that require more scrutiny. The first concerns the commonly prevailing view, that Hindi, written in the Nagari script, is the language of the Hindus, and Urdu, using the Perso-Arabic characters – the language of the Muslims. This notion dates back to the times of the College of Fort William and since that time, it has been spreading among the researchers as well as the speakers. However, several recent studies (e.g. Abbi et al. 2004; Ahmad 2011, 2012; Trivedi 2013) clearly undermine the validity of this opinion and prove that both Hindi and Urdu exceed the limits imposed by these stereotypical views, both in terms of their users' religious classification⁴⁹ and of the script employed.⁵⁰ The second issue is of a taxonomical

Hindus. Grammar and structure are the same in both the languages: considerable differences exist only in relation to the vocabulary and the script".

⁴⁸ "Die allgemeine Umgangssprache gebraucht heute mit Vorliebe die heimisch gewordenen arabischen und persischen Lehnwörter in begrenztem Maße und verwendet Sanskrit-Wörter nur für religiöse und kulturelle Begriffe und Sachen des Hinduismus. Diese Sprache nennen wir heute Hindūstānī..." (Spies & Bannerth 1945: 2).

⁴⁹ This oversimplified opinion is questioned, though not very often, also in public discourse, cf. e.g. the statement by the Muslim intellectual and politician Syed Shahabuddin: "Language has no religion. Nor is any language synonymous with any particular religion or vice versa. Urdu is not the mother tongue of all Muslim Indians (...) nor has Urdu become the language of Muslims only (...) To give an example, the entire debate on the Punjabi Suba between the Hindus and the Sikhs was carried out in Urdu." (Shahabuddin 2001)

⁵⁰ There are many examples of daily proof confirming the absurdity of such a religion-based distinction between Hindi and Urdu. One of the most

nature and results from very unclear and vague boundaries between such entities as language, dialect and subdialect, which is characteristic of the South Asian linguistic area.⁵¹ Researchers point out the lack of generally accepted criteria for distinguishing between these phenomena and their corresponding terms, as well as the fact, that these terms are often used, both popularly and by scholars, in several conflicting ways (Masica 1991; cf. also Gumperz & Naim 1960; Gumperz 1961; Shapiro & Schiffman 1981).⁵² Masica most authoritatively states that Urdu and Modern

spectacular comes in the form of Hindu nationalist demonstrations directed against Muslims and Pakistan, during which the protesters, often carrying visible nationalistic symbols and banners in Nagari, shout slogans in Urdu, such as “Pākistān murdābād! Hindūstān zindābād!” (transl. “Death to Pakistan! Long live India!”).

⁵¹ Four of the eight aspects of language use, enumerated by Ferguson, which characterize South Asia, remain closely connected with this problem, viz. multilingualism, dialect variation, register variation and diverse scripts (Ferguson 1992: 27). The insightful discussion of the terms ‘language’ and ‘dialect’ and the relationships between them may be found in Haugen 1966.

⁵² This generally tangled situation is additionally complicated by the influence of regional and subregional vernaculars, further differentiated (often on the basis of their literary or non-literary status) into more standard acrolects and less standard (or non-standard) patois; jargons peculiar to the members of various social groups (like trade languages or caste languages); local varieties spoken around home and farm etc. (more cf. Gumperz 1958; Gumperz & Naim 1960; Shapiro 2003: 251-253). This immense variety of often overlapping literary and oral linguistic entities, historically and currently in use in North and Central South Asia, is difficult, or maybe even impossible, to categorize or systematize. Hence the common and inveterate tendency of covering this linguistic mosaic with the vague and capacious blanket term ‘Hindi-Urdu’/‘Hindustani’, which allows to create an impression of some taxonomic and genetic order without giving details concerning the enveloped objects (tongues) and the relationships between them. But this is a palliative, not a solution to the problem and does not meet the need

Standard Hindi, which are counted as separate languages officially and socioculturally (i.e. both possess a “written standard” and a literature), are actually not even different dialects or subdialects (i.e. subvarieties of a larger unit). According to the American linguist, they are “different *literary styles* based on the *same* linguistically defined subdialect” (Masica 1991: 23, 27). Zograf (1961: 3-5), similarly, defines Hindi and Urdu as two literary styles (*literaturnaya forma*, *literaturnyi stil'*) of Hindustani, which he describes as the interregional vernacular (*mezhhoblastnoi razgovornyi yazyk*) that has emerged on the basis of a local dialect current in the vicinity of Delhi, i.e. K^haṛī bolī. Yet, at the same time, he remarks that since Hindi and Urdu have been chosen as the national languages of newly independent India and Pakistan respectively, their separate development, backed up by the use of different writing systems, made them so distant, that they are often labelled as separate languages (*ikh neredko schitayut raznyimi yazykami*). A strong supporter of the two-language theory was, among others, Ralph Russell, who in one of his articles advocates with full conviction:

For all practical purposes (...) Urdu and Hindi are two separate languages and should be described as such, despite their almost completely common structure and less completely common stock of everyday words. (Russell 1996: 204)

On the other hand, Russell's attitude towards the existence and function of Hindustani seemed to be somewhat reluctant and he avoided using this name, even though he agreed that

The link language of everyday communication in India continues to be, as it was before independence, one which is as much Hindi as it is Urdu. (Russell 1999: 44)

for meticulous survey of ‘Hindi-Urdu’/‘Hindustani’ (which also is not the subject of this study, anyway).

The fact that Russell changed the primordial title of his manual – from *A New Course in Hindustani for Learners in Britain* (1st publ. in 1980) to *A New Course in Urdu and Spoken Hindi...*, published in 1986 solidifies the aforementioned approach of the great scholar.

1.2.13. The above illustrations represent only an exemplary sample of the prolonged and inconclusive discussion on the problem of mutual connections and dependencies between Hindustani, Hindi and Urdu. But even such a fragmentary recapitulation allows one to see the dimension and complexity of the issues that have arisen, over the centuries, in connection to these terms and their referents. Finally, having presented various facts, opinions and formulations there comes a moment when the term Hindustani, as it is understood and used in this study, should be clarified.

1.2.14. In the sense which has been assumed for the purpose of this analysis, the term Hindustani refers to the linguistic continuum current in the common speech – a *lingua franca*, unofficially, but fully functioning as a link language in the Indian subcontinent. It denotes the common form of a medium of communication – the vernacular widespread in the vast regions of contemporary Northern India, Pakistan and Bangladesh,⁵³ intelligible for the most part to users of both Hindi and Urdu, and often to the speakers of other South Asian languages as well. Kachru identifies Hindustani, by many called also Hindi-Urdu,⁵⁴ as a

⁵³ As was aptly noted by Kelkar, from among the modern languages (which he terms *koinés*) creating, in his opinion, the IA linguistic continuum of the Northern and Central South Asia, only Hindustani lacks the homeland in the sense of a clearly defined area over which a given language (*koiné*) dominates (Kelkar 1968: 2). This fact speaks in favour of understanding Hindustani as the regionally unmarked, supra-regional South Asian *lingua franca*.

⁵⁴ E.g.: Apte 1974; Gumperz & Naim 1960; Narang & Becker 1971. The hyphenated form implies that the features and processes described by

common form shared by Hindi and Urdu, and remarks that it is essentially a colloquial language (Kachru 2008: 81). Ohala uses the term ‘Hindi’ for the same variety and defines it as the dialect spoken by educated urban people, in everyday casual conversation, in cities such as Varanasi, Lucknow, Delhi etc. (Ohala 1983: xiii; Ohala 1999: 100). But this ‘colloquial language’ or ‘spoken dialect’ is not restricted to use in speech only. It can also be found in popular literature, everyday media, advertizing, cinema and film songs, and its products (texts) may be oral as well as written in either Nagari or Perso-Arabic script.⁵⁵ Kelkar (1968: 6-10) distinguished several registers of Hindustani (spanning various spoken as well as written scopes), ranging from the very colloquial *bāzārī* variety to the highly formalized separate literary styles of ‘High’ Hindi and ‘High’ Urdu. They have been schematically depicted on p. 31.

the authors are common for both Hindi and Urdu and that the language which is in their focus is neither the ‘high Hindi’ (highly Sanskritized and also known as the literary style of Hindi), nor the ‘high Urdu’ (strongly Perso-Arabicized and labeled as the literary style of Urdu). Kelkar (1968) coined the cover-name ‘Hirdu’ for Hindi-Urdu and uses both the forms in his book.

⁵⁵ In recent years, the practice of using Roman script for Hindustani has become increasingly popular. This trend has been spreading mainly due to the development and globalizing influence of the visual mass media and the overpowering influence of English. It is becoming a mass phenomenon in advertisement and on various websites, frequently resulting in written multilingual texts, characterized by language mixing and code-switching (on written multilingualism cf. Sebba et al. 2012).

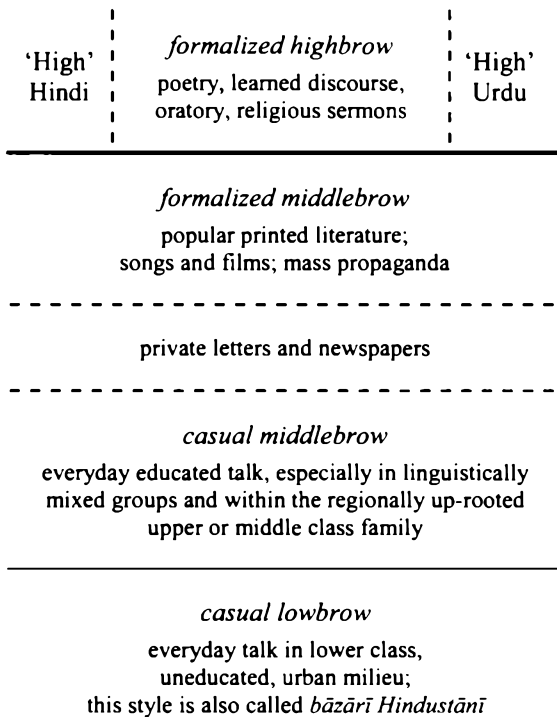


Fig. 2. Schematic illustration of Hindustani registers (based on Kelkar's description, 1968).

1.2.15. The linguistic material (Perso-Arabic loanwords in Hindustani) which forms the basis for the research in this study (Kuczkiewicz-Fraś 2008), generally comes from middlebrow registers (both casual and formalized). However, some parts of it can also appear in the inventories of the lowest and the highest strata. It must be noted, that the middlebrow registers (connected with the educated urban population) have always been that form

of a language, in which the entire process of enriching one speech by the elements of another can be easily observed. From the very beginning of its existence Hindustani has been subject to foreign and local influences and has borrowed equally freely from both the superstrata (previously mainly from Persian, and through it – from Arabic, nowadays – from English) and from the adstrata or substrata (various indigenous languages and dialects, such as Panjabi, Braj b^hāṣā, Bhojpurī, Avadhī etc.).⁵⁶ In this study, however, Hindi and Urdu will be treated as two forms of Hindustani, differing most noticeably in their highly literary and intellectual varieties but lacking any visible distinctions in their everyday forms of use.

1.2.15. I would like to conclude this chapter with some personal observations made during one of my recent trips to India. In April 2012, I travelled to Hyderabad and nearby places to gather linguistic and historical data connected with my academic work. Well aware of Hyderabad being a great centre of Urdu language and culture, second only to Lucknow, I tried my best to speak “Urdu” not “Hindi” to the very first person I encountered, who happened to be a taxi driver driving me from the airport to my hotel. After a short casual conversation concerning the long distance to the city and the city’s vast extent, the man turned to me, smiling, and exclaimed: *Yah pahlī bār hai ki main hindī bol-nevālī aqnābī ko dek^h rahā hūn!* [“It’s the first time I’ve seen a foreigner speaking Hindi!”]. A bit mazed I asked him what language he speaks at home. *Ghar meñ ham log urdū bolte haiñ* [“At home we speak Urdu”], he answered and seeing my puzzled face added quickly: *Lekin hindī-urdū meñ koi farq nahīñ, madam. Yah sab ek hī zabān hai. Sab hindūstānī hai* [“But there

⁵⁶ Cf. Nespital 1994: 308-310. On strata linguistics, i.e. different types of relative social status of languages participating in contact situation see e.g. Hock 1991: 409-411; Hock & Joseph 1996: 274; McMahon 1994: 254ff; Tristram 2007.

is no difference in Hindi-Urdu. All this is one language. Everything is Hindustani].⁵⁷ It was, however, only in Hyderabad that the language I spoke to people was assessed as Hindi, not Urdu. In other places of the region, like Bijapur or Gulbarga, I was always praised for speaking Urdu (and never Hindi), regardless of whether they were residents of the place or visitors, like me. I still cannot solve this sociolinguistic conundrum.



Fig. 3. A one pie coin minted by the East India Company in Calcutta (Bengal Presidency) in 1831. Denomination is marked in four scripts/languages – obverse: English ONE PIE / Bengali এক পাই (= *ek pāi*); reverse: Hindustani/Urdu ایک پائی (= *ek pāi*) / Nagari script ऐक पाई (= *aik pāi*). (Source: Pridmore 1975: 215)

⁵⁷ This situation reminded me of the anecdote told by David Matthews about his experience in Chennai, when, having ordered tea in Urdu, he was instructed by a Tamil waiter, that people there did not like Hindi (Matthews 2002: 166).

1.3. Persian and Indo-Aryan: a General Exposition of the Language Contact Situation

Every Moosulman and Hindoo, who would assume the office of a moonshee (prop. a writer, secretary), or teacher of Hindoostanee, can read Persian.

(Gilchrist 1796: 4).

Without a knowledge of Persian our sentiments cannot be correctly expressed nor can we correctly achieve polished conversation

(Sharar 1975: 101)

1.3.1. In the introductory remarks to his paradigmatic book, Uriel Weinreich stresses the fact that language contact is often viewed as an aspect of cultural contact and language interference⁵⁸ as a facet of cultural diffusion and acculturation. According to the American linguist, extra-linguistic factors in the sociocultural context of the linguistic contact situation are as important for determining the phenomenon of interference as the

⁵⁸ I use this term as defined by Weinreich, viz. “the rearrangement of patterns that result from the introduction of foreign elements into the more highly structured domains of language”, such as the phonemic system, morphology, syntax or lexis (Weinreich 1953: 1). Weinreich underlines the difference between interference (phonemic and grammatical) and borrowing, which he qualifies as “mere additions to the vocabulary”. The former entails “deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language, i.e. as a result of language contact”, assuming, that two or more languages are said to be in contact if they are used alternately by the same persons (ibid.; cf. also Haugen 1950). The latter is less restricted to the bilingual representants of a language community (Weinreich 1953: 56). On bilingualism as a factor essential for the language contact *vide infra* paragraph 1.3.4.

very structure of the languages in contact (Weinreich 1953). Weinreich's views concerning the interdependency between linguistic and sociocultural domains were obviously based on former observations expressed by other scholars, like those made by Sapir or Bloomfield, who a few decades earlier wrote:

Languages, like cultures, are rarely sufficient unto themselves. The necessities of intercourse bring the speakers of one language into direct or indirect contact with those of neighboring or culturally dominant languages. (...) The language of a people that is looked upon as a center of culture is naturally far more likely to exert an appreciable influence on other languages spoken in its vicinity than to be influenced by them. (...) It is generally assumed that the nature and extent of borrowing depend entirely on the historical facts of culture relation. (Sapir 1921: 205, 208)

Every speech-community learns from its neighbors. Objects, both natural and manufactured, pass from one community to the other, and so do patterns of action, such as technical procedures, warlike practices, religious rites, or fashions of individual conduct (...) Along with objects or practices, the speech-forms by which these are named often pass from people to people. (Bloomfield 1933: 445)

Weinreich was, however, the first to realize, and to effectively demonstrate, that both internal (or structural) and sociocultural factors play an equally important role in shaping the linguistic results of contact. This opinion is now universally accepted⁵⁹ and the concepts and principles formulated by him still pave the way

⁵⁹ For example by William Labov in the monumental *Principles of Linguistic Change* (1994, 2001, 2010), Gillian Sankoff in her sociolinguistic works (e.g. Sankoff 1980, Sankoff 2002) or by Sarah Thomason and Terrence Kaufman in their fundamental study of various types of language contact, in which they went so far as to argue that "it is the sociolinguistic history of the speakers, and not the structure of their language, that is the primary determinant of the linguistic outcome of language contact" (Thomason & Kaufman 1988: 35).

for research on multilingual communities and language contact. Such phenomena as “borrowing”, “interference”, “bilingualism”, “code-switching” and “mixing the languages”, perceived as the results of, broadly speaking, a relationship between language and culture, have attracted the interest of researchers in many fields – especially those in anthropology, linguistics and sociology. They have pointed out the parallel between the linguistic and cultural contact situations, and have described how one can influence the other, namely, how the acquisition of language is tied to social and cultural context, and to what extent the acquisition of culture occurs through language.⁶⁰ In this chapter, the situation of language contact between the Persian⁶¹ and the Indo-Aryan lan-

⁶⁰ In 1975 Benjamin T'sou determined five stages of linguistic assimilation, viz. linguistic importation, linguistic substitution, code-switching, bilingualism, and residual interference, and made an attempt at correlating them with five phases of cultural assimilation, viz. sojourner phase, congregated phase, transitional phase, acculturated phase, and assimilated phase (T'sou 1975: 455). However, this comparison requires two remarks: firstly, elements of both the processes are not mutually exclusive but progressively overlapping, which the author himself accentuated. Secondly, T'sou based his hypothesis on data obtained primarily from a study of a small group of Cantonese immigrants in the U.S., and his theory might be relevant in many similar situations, i.e. when a given group of newcomers, characterized by a distinct culture and language, is subjected to a process of acculturation to the “bigger” and “stronger” linguistic and cultural environment of the place of destination. But, as we will see further in this chapter, this was not exactly the case of those who brought the Persian culture and the Persian language to India. Cf. also paragraph 1.3.10, where the applicability of the hypothesis presented by T'sou in the context of the Indo-Persian contact situation is further discussed.

⁶¹ I write ‘Persian’, but with the notion kept in mind, that this term should be understood very broadly, since it also includes elements of Arabic and Turkic languages. These elements were mainly lexical, but to a limited extent also grammatical, and had been absorbed and assimilated by Persian before it was introduced to South Asia. The phenomenon of Arabic

guages will be discussed and exposed against the background of the Indo-Persian cross-cultural relations. On this basis I will attempt to determine how cultural adaptation, acculturation, and cultural change affected the ways the language(s) was/were used by members of the community and, *vice versa*, how these cultural changes were themselves affected by the language(s). I will also refer to one of the most vital problems posed by sociolinguists: how various social phenomena manipulated essentially neutral linguistic processes into a system which reflected and reinforced social and power difference (cf. Sankoff 1980: 5).

1.3.2. In the Islamic world, Persian achieved high status as a language of administration long before its introduction in South Asia. It was used for recording business transactions and revenue accounts as well as for expressing and transmitting refined and philosophical ideas. In the Umayyad period, Persian language and culture continued to spread throughout the territories of Persia, Syria, Caucasus, Transoxiana and Sind – engulfed by rapidly spreading Islam – but it was in the middle of the 8th c., with the Abbasids' rise to power, when their full blossoming started. For the following century the Persian ascendancy in the Caliphate was overwhelming, including Baghdad and other leading centres of Abbasid supremacy (Syed 2012). This was the moment in history when the formation of the Perso-Islamic culture, “the second great cultural nexus of the Islamic world” (Robinson 1991), was initiated, spreading, at its furthest reach, from the Aegean and the Eu-

and Turkic impact on Persian has been analyzed attentively by John Perry (1984; 1992; 2001b; 2002; 2005) and also by a few other authors, e.g. Canfield 1991; Doerfer 1963-75; Jazayery 1970; Johanson & Bulut 2006; Utas 2006. On Arabic in South Asia *vide infra* fn. 75; on Turkic in South Asia *vide* fn. 76; cf. also paragraph 5.2.2. A general survey of the use of Arabic, Persian and Turkic in India, supported by many interesting examples of language contact may be found in Chaudhary 2009: 135-162. Unfortunately, his examples are quoted mostly after secondary sources.

phrates in the West, to Chinese Turkestan and the Bay of Bengal in the East, and from the Russian Steppe in the North, to the Indian Ocean in the South. The Persian language was the first and crucial pillar of this culture, the common denominator, which, having established its supremacy over other languages used in the Iranian region (both Iranian and Turkic) in the 13th c., remained not only the major native and contact vernacular of the area, but also the interregional written language of choice for the next six hundred years. Persian replaced Arabic in the sphere of Islam “for all but scriptural and liturgical purposes” (Perry 2001a: 191), it became a language of poetry and belles-lettres, and a universal tool of communication in the administration and diplomacy of Iran, Afghanistan, Central Asia, and Northern and Central India. Persian literary models and cultural ideals were widely cultivated, additionally pervading and affecting the languages spoken on the fringes of the areas under the direct influence of Persian tradition – in Anatolia, the South Caucasus, the Persian Gulf, South India and Southeast Asia (ibid.).

1.3.3. The influence of Persian language and culture on the Indian subcontinent dates back to the times of the Ghaznavids, who, despite being themselves of Turkic origin, have inherited Samanid administrative, political, and cultural traditions and were the first to introduce the Persianate⁶² way of life in Northern India. As a result of repeated invasions undertaken by Mahmud of Ghazna (Maḥmūd-e Ġaznavī, *r.* 998-1030), a large number of people of Turkic, Persian and Afghan descent were brought to India. Soldiers of fortune, merchants, adventurers, scholars and

⁶² Analogously to the term ‘Islamicate’, coined to indicate the cultural traditions of the Muslim civilization at large, but not restricted in terms of its reference to the Islamic religion (Hodgson 1974), the term ‘Persianate’ refers to civilizational traditions strongly influenced by Persian language, literature, and culture, but separate ethnically and geographically from Persia itself (cf. Hakala 2010; Truschke 2012).

religious men in Islam, artists and craftsmen – they left their homelands and settled down in the Punjab, around Delhi and possibly in some areas of Gujarat. As a result of the political stabilization brought by the rule of the Ghurids (Ġorī, *r.* 1148-1215), who are credited with laying the foundations of Muslim domination in India that lasted for several centuries, and the Slave Dynasty, ruling India's Delhi Sultanate for almost the entire 13th c. (1206-1290), these strangers started perceiving India as their own country. Patronised by the Muslim, and sometimes even by the Hindu, rulers, they helped to enrich the mediaeval culture of the subcontinent.⁶³ India thus became a part of the Perso-Islamic world, not different in this respect from its other parts, like Transoxiana, Ghazna or Ghur. Delhi and Lahore enjoyed a reputation of great centres of Perso-Islamic culture, equal to Bukhara, Herat or Tirmiz (Alam 2003: 134). Cultural integration with a coherent Perso-Islamic identity, which had emerged in the 13th c., started bearing its fruits in the form of syncretic Indo-Persian culture (described also as Persianate – or, broadly speaking – Islamicate culture of South Asia). It reached its peak in the Mughal era⁶⁴ and manifested itself in many different areas of life: from architecture, garden planning, painting, music, literature, philosophical and religious ideas, to dressing style and cuisine. This culture, pertaining to the elites as well as the common man, flourished mostly in cities and towns. From there, it

⁶³ For detailed and more comprehensive accounts of the earliest period of mutual relations between South Asian and Perso-Islamic traditions see e.g.: Abidi & Gargesh 2008; Ahmad 2002; Alam 2003; Eaton 2003; Ernst 1992; Kuczkiewicz-Fraś 2003b; Syed 2012; Wink 1999.

⁶⁴ The Mughal (Moġūl) rule in India lasted for over three hundred years (1526-1858), but the golden age of Indo-Persian culture spans the reign of Akbar (Jalāl ud-Dīn Muḥammad Akbar, *r.* 1556-1605) and his two successors, Jahangir (Nūruddīn Salīm Jahāngīr, *r.* 1605-1627) and Shah Jahan (Šahābuddīn Muḥammad Šahjahān, *r.* 1628-1658).

gradually penetrated into small towns and rural centres, and, widely disseminated as a result of military campaigns and waves of migration, it made a lasting impression on most local traditions of the subcontinent.

1.3.4. Contact breeds imitation and imitation breeds linguistic convergence.⁶⁵ Cultural integration fostered integration in the sphere of speech, which, on the one hand, became a communication must, but on the other, was in many cases a matter of more or less conscious choice. It was not a new phenomenon – since antiquity, bilingual⁶⁶ competence has been a central factor in the everyday linguistic reality of much of the South Asian population,⁶⁷ and Persian, having appeared on the linguistic scene of the subcontinent, became just one of many elements (although a very important one) of this jigsaw puzzle. Language contact relies on bilingualism.⁶⁸ Thus, as a consequence of the natural tendency

⁶⁵ André Martinet in his preface to Weinreich 1953: VIII.

⁶⁶ Henceforth I will use the term “bilingualism” in both the senses, viz. bi- and multilingualism.

⁶⁷ On bilingualism in South Asia see: Bhatia & Ritchie 2004; Hock 1991: 498-504; Jain 2003: 54-56; Kachru et al. 2008: part 4; Pattanayak 2001; Shapiro & Schiffman 1981: Ch. 6; Thomason 2001: 42-46.

⁶⁸ Scholars agree that some minimum of bilingual proficiency is a prerequisite for linguistic borrowing and language interference. Such an approach follows Hermann Paul's opinion, expressed by the German linguist as early as the end of the 19th c.: “Die meiste Veranlassung zur Mischung ist gegeben, wo es Individuen gibt, die doppelsprachig sind, mehrere Sprachen neben einander sprechen oder mindestens eine andere neben ihrer Muttersprache verstehen. Ein gewisses Minimum von Verständnis einer fremden Sprache ist unter allen Umständen erforderlich.” [Most occasions for mixing occur when there are individuals who are bilingual, who can speak several languages side by side, or at least understand one other beside their mother tongue. A certain minimum understanding of a foreign language is required in all circumstances.], Paul 1886: 338; cf. also Haugen

toward bilingualism, prevailing among the people of India, the language contact between South-Asian languages and Persian became a fruitful reality from the very beginning of its presence in the subcontinent.⁶⁹ As early as the 11th c. there appeared the first poets born in India who composed their verse in Persian. Some of these authors, like Mas'ūd Sa'd Salmān (1046-1121) or Amīr Xusrau Dihlavī⁷⁰ (1253-1325), wrote not only in Persian, but also in the North Indian vernacular, commonly designated as Hindavi.⁷¹ Annemarie Schimmel commented that “Even re-

1950; Hoffmann 1991: Ch. 5; McMahon 1994: Ch. 8; Romaine 1995: Ch. 2; Weinreich 1953.

⁶⁹ Muḥammad Ḥusain Āzād made the following observation: “As soon as Islam came, it laid the foundation for the mixing of words” (Āzād 2001: 64), and then, commenting on the language (called by him “Hindi”) of Cand Bar'dāī's heroic poem *Prṭ'vīrāj-rasau* (ca. 1193), he added: “It is astonishing (...) to see how quickly the language accepted the influence of Arabic and Persian; on every page there are quite a number of words”. Thirteen examples of such words are listed (*ibid.*).

⁷⁰ Amīr Xusrau is regarded as the greatest Persian-writing poet of mediaeval India, to whom the authorship of the poetic lexicon *Xāliq bārī* (or at least a part of it) is also attributed (discussion of the authorship of this work cf. Zaidi 1993: 23-24; also Hakala 2010: Ch. 4). This glossary-poem, reported to have originally contained 170 verses, was the first Persian-Hindavi dictionary, on which, in the following centuries, many similar works were modelled (Samiuddin 2007: 189).

⁷¹ In fact, the Persian adjective “Hindavī”, meaning “Indian”, was used by Muslims in this sense to denote every local language, all over the subcontinent. For example, in the 16th c., the sultan of Golkonda, 'Abdullāh Quṭb Šāh, instituted a special office, called *dabīrī-yi farāmīn-i Hindvī*, whose task it was to prepare royal edicts in Telugu (Alam 2003: 157) – a Dravidian language which has nothing in common with the Indo-Aryan *lingua franca* of North India. In Bijapur the term “Hindvī” was used in contradistinction to Persian and assigned to the local language written in Perso-Arabic characters, characterized by Prakrit structure and a strong Marathi influence (Sherwani 1985: 296-297). Cf. also paragraph 1.2.4.

nowned poets were tempted to use the popular language in their works in order to prove that their literary talent could display itself through any linguistic medium” (Schimmel 1975: 128). These poets became very prominent and influential “loci of the contact” (as Weinreich put it), since their verses, in which they mixed local and Persian idioms, gained immense popularity and influenced subsequent artists, who often drew inspiration from their works.⁷² But the contact between languages was not confined solely to the realm of literature. Persian names were given to various products, offices, and even towns and places.⁷³ New idioms of Persian became current, mixing with rapidly developing local dialects. The process of cultural borrowing⁷⁴ entailed borrowing of new, foreign (Persian, Turkic or Arabic⁷⁵) words,

⁷² For information about poetry and prose under the early sultans of Northern India see Alam 2003: 135ff and Schimmel 1973: 2-22.

⁷³ See e.g. Kuczkiewicz-Fraś 2013 on Delhi urban toponyms of Perso-Arabic origin.

⁷⁴ The notion of cultural borrowing was first formulated by Sapir, cf.: “The simplest kind of influence that one language may exert on another is the *borrowing* of words. When there is cultural borrowing there is always the likelihood that the associated words may be borrowed too” (Sapir 1921: 206). Later on the idea was developed by Bloomfield, who distinguished between “*dialect borrowing*, where the borrowed features come from within the same speech-area (...) and *cultural borrowing*, where the borrowed features come from a different language” (Bloomfield 1933: 444).

⁷⁵ It must be noted that the first Arab sailors and traders, and with them also the Arabic language, reached the ports of South India as early as in the beginning of the first millennium CE. Then, in the 8th c., Sind – the western borderland of India – was invaded and conquered by the Muslim Arabs led by Muḥammad bin al-Qāsim and the Arabs began to settle there, establishing India’s first considerable contacts with Islam and with the Arabic language. Later on, Arabic was present in the realms ruled by all the succeeding Muslim dynasties of South Asia, but its role was generally confined to the sphere of the Islamic religion and, unlike Persian, it has never achieved the status of being the secular language of culture or ad-

especially in areas dominated by Muslim rule,⁷⁶ i.e. on the extensive and fertile terrain of the Indo-Gangetic Plain and afterwards also in Deccan. These earliest loanwords, pertaining predominantly to administration and various amenities of urban life, could be found in most New Indo-Aryan vocabularies and above all in Hindavi.⁷⁷

1.3.5. Initial “transmission channels” for permeating new words must have been created by those domains of life, in which communication between the residents and the newcomers was unavoidable and, consequently, the interactions between them were most intense: trade, offices and the army. Thus, it was need (Hock 1991: 408), sheer necessity (McMahon 1994: 201), that

ministration. However, its occurrence in the language contact situation, especially in South India, although limited, cannot be denied – some Arabic borrowings in Indian languages undoubtedly resulted from a direct influence. For more information on Arabic in India cf. Qutbuddin 2007.

⁷⁶ Many of the Muslim nobles and military commanders who came from the Afghan dynasties Sūr or Lodīs and established their kingdoms in Gujarat, Malva, the Deccan, Bihar, and Bengal (Rizvi 2000: 197), did not speak (or could speak very little) Persian. Even though they mostly adopted Persianate customs and court etiquette, many continued using their own tongue (often some form of Turkic or Pashto) to communicate with their kinsmen, or shifted to local vernaculars while dealing with their subjects. Bābur (whose mother tongue, and the medium he chose for his writing, was not Persian either, but Chagatai of the Turkic family) recorded in his memoir that in order to address one of the mightiest nobles, Daulat Xān Lodī, he “ordered a person well acquainted with Hindūstānī to interpret [his] words to him [sc. Daulat Xān], one after another” (Bābur 1922: 459). Because of his princely education as a boy, Bābur, of course, knew Persian fluently, but his note indicates clearly that this was not the case of Daulat Xān.

⁷⁷ Unfortunately, the data concerning the early Indo-Aryan period are scarce and incomplete, most often deficient, and sometimes simply non-existent, as in case of Sindhi or Panjabi (Masica 1991: 54). The dating of some of the borrowings, however, may be inferred from their phonology (cf. paragraph 3.1.3.).

constituted the first motivation for borrowing – the people's requirement to refer to some unfamiliar objects or concepts, for which they lacked terms in their own speech. Before long a new important factor appeared and greatly influenced the linguistic contact situation between Persian and local Indian dialects. This change came in the form of Sufi hospices, which, since the early 13th c.,⁷⁸ had been increasing in number all over India as a result of the rapid spread and the growing impact of Islamic mysticism. The Sufi idea of practising a classless society and avoiding inequity and intolerance among the disciples attracted many people. In a little while, the dwelling places of *šaix*-s and *pīr*-s became significant centres for Islamization and conversions. Through the Sufi hospices, Persian literary culture and Persian vocabulary were gradually incorporated into Indo-Aryan languages (and in Deccan also into Dravidian languages, e.g. Telugu). As Schimmel observed:

(...) the Indo-Muslim languages (Sindhi, Panjabi, and to a great extent Urdu and Pashto) are in large part the result of the speech and song of mystical leaders, who could not address their simple disciples in high-flown theological Arabic or poetical Persian; in order to express the mysteries of divine love and devotion, they were obliged to use the vernaculars, making them vehicles of the most lofty thought. Then they emerged as languages well able to serve as a literary medium for nonmystical writers. (Schimmel 2007: 33)

Therefore, even though their major works were mostly composed in Persian, the Sufi saints employed the spoken languages in their everyday sermons, teachings and short verse. Obviously, they were gradually introducing various elements of their own

⁷⁸ The first Sufi *xānqāh* in the subcontinent was established by Mu'īnuddīn Čištī around 1200, near Ajmer, though Sufi mystic poets and theologians were reaching India much earlier. Cf. Schimmel 2007, esp. Ch. 8: "Sufism in Indo-Pakistan".

tongues into these vernaculars – vocabulary, whole phrases, or even forms of versification, contributing in this way to the emergence of a language “able to reflect the traditions of western Asia and the Islamic *weltanschauung* [sc. world-view]” (Schimmel 1975: 128). Thus, from the point of view of sociolinguistics, they constituted an extremely important bilingual group, which had an immense influence on the way of living, thinking and speaking among the largest strata of the South Asian population.⁷⁹

1.3.6. As has been previously stated, the greatest blossoming and expansion of Persian literature and culture in South Asia took place during the Mughal times.⁸⁰ Persian, which had already established itself in India as the language of most Muslim authorities and elites, rose, in the times of Akbar, to a position comparable to no other part of the Islamic world, ruled, similarly to the Mughal empire, by Turkic dynasties: neither the Ottomans in Anatolia nor the Uzbeks in Central Asia exhibited a similar eagerness for Persian (Spear 1978: 48). In India, however, certain aspects of Mughal politics ultimately led to the emergence of the Indo-Persian literary culture in its fullest efflorescence. The basic principles of Akbar’s rule: close and friendly relations with Safavid Persia and extraordinarily tolerant domes-

⁷⁹ Nile Green underscores the fact that knowledge was located in persons rather than in books and stresses the great role of oral transmission of knowledge in Sufi religious circles: “Books were not considered independent sources of knowledge, but were appendages to the personal pedagogical relationships through which knowledge was transferred and within which writing served to provide only one dimension of the knowledge being transferred. (...) those in search of knowledge looked for a master rather than a bookshop or library, with libraries in any case generally being private collections accessed through personal affiliation to a teacher.” (Green 2010: 243). I am grateful to Rahul Peter Das for drawing my attention to Green’s article.

⁸⁰ For an overview of Persian under the Mughals see Alam 1998 and Alam 2003: 159-174.

tic policy, the basic guideline of which was “peace with all” (*ṣulḥ-i kull*), resulted in his empire earning a reputation as a place of refuge and a realm of peace (*dār al-amān*). This, in turn, caused many Iranians – especially Sunnis, who did not approve of the Safavid theocratic regime and opposed mandatory conversion to Shiah Islam, which became the official religion of Persia – to flee from religious and political persecution and to seek protection and better fortune in the neighbouring and lenient India.⁸¹ There were many noted scholars, Persian writers, and poets among them, and Akbar, who spared no effort and used every opportunity to invite Iranian literati to India, welcomed them most generously.⁸² All major forms of Persian literature were extensively cultivated under his munificent patronage: literary anthologies (*tazkirah*), consisting of brief selections from the

⁸¹ The Mughal sultanate was not the only destination in the subcontinent chosen by the newcomers. The constant influx of Iranians, Turks and Arabs, who were later named *āfāqī*-s (or “cosmopolitans”) by the locals, to the Bahmani kingdom, dated back to the 14th c. and had a decided effect upon the culture and the future history of the Deccan. It brought many men of learning, literature and arts to South India, and significantly influenced the linguistic situation there, by contributing to, i.a., the formation of Dakkani (cf. Sherwani 1985: 76ff.). This trend continued in the 16th and 17th c. in post-Bahmanī sultanates, especially in Bijapur under the ʿĀdil-šāhī dynasty and Golkonda ruled by the Qutb-šāhīs.

⁸² Muzaffar Alam suggested that this extended invitation for Persian scholars and writers was caused by Akbar’s desire to repay Iranians for the support the Mughals received from them to reconquer Hindustan (Alam 2003: 160). Another reason, however, for promoting Iranians in Mughal service, could have been the attempt to neutralize influences of powerful and ambitious Turkic nobles. Rivalry in the sphere of culture may also be seen as a manifestation of Akbar’s aspiration to equal and surpass the Safavid rulers of Persia. But whatever the underlying motives, they undoubtedly had a political dimension and remain a good example of how extra-linguistic factors can affect the trajectory of linguistic development.

works of various poets, usually with prefatory comments about each poet, which were not only written in Persian, but also used critical vocabulary borrowed from Persian; the recordings of utterances attributed to Sufi *šaix*-s (*malḡūḡāt*); model letters (*inṣā*), commonly used for official and diplomatic purposes and characterized by an elegance of style;⁸³ romances (*dāstān*) in prose or verse, telling the deeds of one or more particular heroes and usually full of war, magic, trickery, and love; glossaries and dictionaries (*farhang*), mono- or multilingual lexicographical works, regarded by many scholars as the greatest legacy of South Asia in the field of linguistic inquiry into Persian.⁸⁴ But more than anything else, poetic art (*šā'irī*) was loved and appreciated by Indo-Persian society; all the main poetic genres were commonly represented and *ġazal* was in the forefront of them all.

1.3.7. Persian had become the most important and, at the same time, the most desirable language for both Muslims and Hindus, regardless of whether they sought careers as men of letters and learning or as functionaries of the Mughal administration. The administrative reforms introduced by Akbar's Hindu minister of finance, Rājā Ṭoḡar Mal, were essential for the development of the latter. Under his instructions, since the 1580s, all the revenue accounts in the Empire were to be kept in Persian. His contemporary, Abū 'l-Faḡl, a chronicler of Akbar's court, underlines the great significance of this regulation:

The most important reform introduced by Ṭoḡar Mal is the change in the language and the character used for the revenue accounts. Formerly they have been kept in Hindī by Hindū *Muḡarris* [sc. clerks]. Ṭoḡar Mal ordered that all government accounts should henceforth be written in Persian. He thus forced his co-religionists

⁸³ The origin and growth of *inṣā* literature in India until the 16th c. has been scrutinized in Zilli 2000.

⁸⁴ An outline study of the Persian lexicographical tradition in South Asia can be found in Maciuszak 2013.

to learn the court language of their rulers (...) The study of Persian therefore became necessary for its pecuniary advantages. (Abū'l-Faẓl 1873: 377)

Indeed, Tōḍar Mal's reform was of great consequence: as a result of its implementation, the members of indigenous Hindu castes who traditionally played the role of record-keepers and dominated departments of accountancy, revenue and draftmanship – Kāyasth and Khatrī in particular – started learning Persian and “their achievements in the language were soon to be extraordinary. (...) [They] joined *madrāsas* in large numbers to acquire excellence in Persian language and literature, which now promised good careers in imperial service” (Alam 1998: 326). Until the 19th c., when the British colonial government replaced Persian in its official function as the medium of state administration by the pre-eminent world imperial language English (1837), the knowledge of Persian had been a prerequisite for applying for any state job (cf. also Gilchrist's remark used as a motto for this chapter).

1.3.8. Another reform, instated by Akbar, and important in terms of language contact, pertained to the system of education.⁸⁵ Traditional Islamized curricula of Muslim schools (*maktab*-s and *madrāsa*-s) were enriched by the introduction of many “secular” subjects, fixed rules, and a canon for the purpose of teaching Persian on both the elementary and the advanced levels. This also attracted many non-Muslim students who undertook education in order to gain skills and knowledge necessary for the job of an accountant or a *munšī* [sc. secretary]. Schools became a very important and powerful factor fostering language interference, since Persian was often taught through the medium of Hindavi,

⁸⁵ Cf. Kahane's observation: “In literate societies, one of the primary MOTIVATIONS for acquiring the prestige language is its identification with education, which transfers to it the values of a class symbol” (Kahane 1986: 495).

and Arabic – through the medium of Persian (Rahman 2013: 370). Pupils, who started their education between the ages of four and half and seven, transformed, during their studies, into an influential group of educated bilinguals. And while, initially, the teachers in charge of these schools came mostly from Iran, in the course of time Hindu masters also started to teach. Their writings contributed to Persian syllabi and thus enhanced the language contact situation.⁸⁶

1.3.9. With the passage of time the motivation for learning and using Persian gradually and inevitably changed, and the factor of prestige became at least as important as that of need. The sublime Indo-Persian culture was generally the culture of the ruling class and the Muslim elites. This social strata, known as *ašrāf*,⁸⁷ or respectable Muslim gentry, who claimed its nobility

⁸⁶ More information on the system of education reformed by Akbar cf. Alam 1998: 326-330. The issue of teaching Persian in South Asia, including pedagogical material and educational institutions, was also recently addressed by Rahman (2013). The exemplary curriculum of Persian during Aurangzeb (r. 1659-1707) times was quoted and discussed in Syed 2012: 289-291.

⁸⁷ Since the earliest times of their presence in the subcontinent, South Asian Muslims have been sharply divided into two categories: those belonging to *ašrāf*, noble-born (A *ašrāf*, pl. of *šarīf* ‘nobles, noblemen, grantees, persons of high extraction, gentlemen, gentlefolk; honourable men’, cf. Platts 56) and those not belonging to *ašrāf*, labelled as *ajlāf*, commoners (A. *ajlāf*, pl. of *jilf*, ‘base, vile, mean, or ignoble people; the rude, the vulgar; the lower orders’, cf. Platts 24). The former have had a superior status derived from their foreign ancestry and have included all the undoubted descendants of foreign Muslims (Arabs, Persians and Afghans, etc.) as well as the converts from the upper caste of Hindus. *Ajlāf* has comprised the various functional castes (such as the weavers, oil-pressers, barbers, tailors, etc.) and the converts of originally humble castes. The social distinction between *ašrāf* and *ajlāf* is a South Asian device and seems to correspond to the division into varnas of twice-born people (*dvija*), and the *śūdra*-s or aborigines, who perform functions of serving the *dvija*-s in the Hindu

and a higher social status based on their foreign (Arabian, Iranian, Turkic, Afghan) or in some cases even northern Indian ancestry, sustained and transmitted from generation to generation the best ideals and traditions of the great Perso-Islamic culture, with its standards of cultivation and behaviour. In the whole of the subcontinent, *āsrāf*'s alleged alien, non-Indian descent went hand in hand with their non-Indian cultural orientation and their identification with the ethos of Persianate aristocratic culture of the Mughal courts. Zafar Ahmad, writing about traditional Bengal Muslim society, draws attention to the fact that

[t]he most critical component of this extraterritorial cultural orientation of Bengali *ashraf* was their utter contempt for the Bengali language and literature, and a corresponding leaning towards Persian and Urdu. (Ahmad 2000: 80-81)

Similar enthusiasm toward Persian culture and language has been observed in every place where the local born aspirants tried to rise above their own status and ascend the social ladder. Abdul Halim Sharar, describing the subtleties of the *āsrāf* culture prevalent in nineteenth century Lucknow, gives a vivid testimony of the highest prestige enjoyed by Persian in the mixed Indo-Muslim society:

(...) throughout India, Persian alone was the road to literary and social advancement. (...) Persian was the mark of distinction not only for the better classes but also for the masses. (...) Apart from Muslims, the leanings of élite Hindus were towards Persian writing and literature, so much so, that distinguished Persian compositions flowed from the pens of Hindu authors. (...) there was as much if not more Persian poetry written in India as there was in Persia itself. (...) Persian forms, constructions and genitives are on the tongues of all, even the uneducated and the women. (...) Even

community. On social stratification of Muslims in South Asia cf. Ahmad 1967; Bhatti 1996; on caste system among South Asian Muslims see Das 2005: 75-78 (there also further literature).

women composed Persian verses and every child could make himself understood in the language. (Sharar 1975: 100-101)

Persian also reached the common man and was, at least to some extent, used and understood outside the elites. Local languages were encrusted with Persian casuals⁸⁸ or occasional quotations, aimed at adding colour and making an utterance more elegant and stylish.⁸⁹ They appeared in both speech and text, regardless of already assimilated loanwords, and were usually introduced by sentences like ‘as it is said in Persian’ or ‘a Persian would say’.⁹⁰ India came under the spell of Persian: “general command over idiomatic Persian was a matter of pride; deficiency in elegant self expression meant cultural failure” (Alam 2003: 167).

1.3.10. Most Indo-Aryan languages and regional cultures show various imprints of Persian influence. They can be seen even today in the forms of Bangla verses, in Marathi vocabulary and grammar, in the script (cf. 2.1.3.), in the lexical and literary content of, i.a., Pashto, Sindhi, Kashmiri, Panjabi, Gujarati and,

⁸⁸ Casuals, called also Gastworts (from German *Gastwörter* ‘guest-words’) are unassimilated borrowings that are not widely used but are simply alien words or phrases taken into the language as they are and retaining their original pronunciation, grammar, spelling and meaning (cf. Hughes 2000: 54).

⁸⁹ It was a common practice not only in India, e.g. on Persian casuals in Syriac, cf. Ciancaglini 2008: 23-24.

⁹⁰ This phenomenon, called by linguists “flagged code-switching” (Poplack and Sankoff 1988: 1178ff), is generally one of the strategies used by bilingual speakers to avoid ungrammatical utterances and to solve the problem of combining elements from two different languages. Thus, usually, flagged switches are explained as a kind of indication that the speaker is aware of making a mistake and is apologizing for it (cf. Romaine 1995: 135). In the case of prestigious languages, however, such as Persian in our case, “mistakes” of this type must be understood as deliberate “deviations”, which hint at the “transformation of the instinctive error into accepted norm” (cf. Kahane 1986: 502).

above all, Hindustani.⁹¹ However, the impact was never one-sided and the Persian language was also effected by local Indian literary and linguistic tradition (some say up to the “hybridization” of Indian Persian, cf. Abidi & Gargesh 2008: 105). Persian was employed by the Hindus to express their traditional religious, mythological and literary themes; numerous translations of Indian classical literary and scholarly texts were produced under royal order (Truschke 2012), partly thanks to specially established departments of translation;⁹² the confluence of Islamic and Indian philosophical thought manifested itself in the writings of Sufi and Bhakti mystics.⁹³ All this resulted in the “Indianization” of Persian, which rested not only on “a large-scale borrowing of words from Indian languages, but also (...) expressions that carry the emotional content and thought processes which are Indian” (Abidi & Gargesh 2008: 109). From the 16th c. onward, a specific Indian style of Persian began to emerge. It was called *sabk-i hindī*⁹⁴ and was cultivated both by Indian and Iranian poets (i.e.

⁹¹ It is worth noting that the use of Persian was not restricted to the regions governed by Muslims (i.e. Northern India and the Deccan), but was common also in those parts of the subcontinent which were controlled by non-Muslims, e.g. Marathas or Sikhs (Robinson 1991: 113).

⁹² It happened that classical Sanskrit texts were translated into Persian through the medium of rough oral Hindustani versions provided by the Indian interpreters (Truschke 2012: 62, 187, *passim*).

⁹³ Bhakti mystical poets, such as Nāmdēv (fl. 14th c. ?), Kabīr (fl. 15th c.) or Guru Nānak (1469–1539), used a special mixed language called Sādhukkāṛī ‘holy men’s jargon’ for their literary works. It was based on vernacular Hindustani and contained elements of the early Braj b^hāṣā, Bhojpurī and Avadhī (Nespital 1994: 308–309; McGregor 1984: 40ff.) as well as a considerable amount of Perso-Arabic vocabulary (on Persian loans in the *Ādi Granth* see Shackle 1978).

⁹⁴ *Sabk-i hindī*, or the ‘Indian style’, is the third term of a classification of Persian literature into three stylistic periods. The other terms, *sabk-i xurāsānī* (initially also called *sabk-i turkistānī*) and *sabk-i ʿirāqī*, refer

those Iranian poets, who spent long periods of their creative life in India). It was characterized by code mixing and code-switching, use of hybrid expressions and semantic shifts in the vocabulary, an influence of Indian poetic moods (*rasa*) and the integration of typically Indian motives – “descriptions of places, people, festivals, seasons, flora, fauna and experiences which are Indian and not Iranian” (Abidi & Gargesh 2008: 111).⁹⁵ It also retained a number of archaic and locally determined features, apparent mostly in phonology (the morphology and syntax of both the styles being practically identical) and, above all, in the phonology of the vowels, where Indian Persian retained the archaic system, corresponding closely to that of Hindustani and distinguishing South Asian Persian from Farsi spoken in Iran (cf. 3.1.3.).⁹⁶ The deviation of *sabk-i hindī* from Iranian Persian, which still remained the benchmark, was immense, and the Iranians from Persia disapproved of this style, even to the point of disdain. Therefore, already during Akbar’s reign, attempts were made to purify the language. They mostly took the shape of various lexicons and dictionaries prepared in such a way as to purge the Persian lexis of all non-Iranian words (Alam 2003: 174). These efforts did not produce the expected results. However,

respectively to the eastern and the western parts of mediaeval Persia. (Bruijn 2012). An in-depth analysis of *sabk-i hindī* as well as the extensive bibliography on the phenomenon may be found in Faruqi 2004.

⁹⁵ Examples of these phenomena are given by Abidi & Gargesh 2008: 109-114 and Faruqi 2004. An exposition of Indian Persian vs. Iranian Persian has been provided by Alam 2003: 174-186.

⁹⁶ As argued by Banārsī Prasād Saksenā: “Indo-Persian literature (...) is certainly not purely Persian. It is futile to accept it to be so. The Persian language had come to stay in India, and it could not for long keep itself aloof from its new and powerful environments. It absorbed Indian ideas and Indian thoughts, and it was used for Indian subjects. Necessarily it developed distinct character” (qtd. after Qureshī 1965: 39).

they contributed greatly to the development of Persian lexicography in South Asia.⁹⁷

1.3.11. The importance of Persian in India began to slowly decline already in the 18th c., following the gradual weakening of the Mughal empire. Due to the rise and development of local vernaculars, mainly Hindustani which was strongly backed up by the British, and because of the growing significance of English, Persian began to lose its privileged and unique position, though it still remained the language of Muslim rule, of the judiciary and of education, as well as one of the most important literary mediums. Macaulay's famous *Minute...* (Macaulay 1835) and the subsequent immediate reforms of Lord William Bentinck (Bentinck 1835), introduced the western model of teaching in the British Raj, in order to provide more educated Indians for service in the British bureaucracy, and ultimately led to the replacement of Persian by the English language in both the educational and the court systems. However, Persian was not completely supplanted. As before, it was often used as a tool of intercultural communication in the whole subcontinent – Spear mentions, for example, that the German missionary, Christian Schwartz (*fl.* 1750-1798), communicated in Persian when he went on a mission to the raja of Tanjore in the extreme South of India (Spear 1978: 49). Until the 1930s, it was also chosen, in general, as one of the degree subjects by Hindu students (*ibid.*), despite the predominant position of the English-medium education.

⁹⁷ The first “dictionary of pure Persian words” was *Farhang-i Jahāngīrī* (“Dictionary in honour of Jahangir”), compiled by a native of Shiraz, Mīr Jamāl-al-Dīn Ḥosayn b. Fakr-al-Dīn Ḥasan Enjū. The work was ordered by Akbar but completed only in 1608, after the king's death. It is believed to be one of the earliest serious attempts to discuss the grammatical structure of Persian and to classify its lexicon (Maciuszak 2013: 349; cf. also Bayevsky 1999).

1.3.12. Being one of the main members of the Iranian linguistic family, which comprises the nearest – both geographically and genetically – relatives of the Indo-Aryan languages, Persian shows a considerable amount of linguistic convergence with many IA languages of the Indian subcontinent. This closeness and concurrence of linguistic features made linguistic interference possible: for centuries of its presence in South Asia, Persian held a profound and enduring sway over Indian vernaculars (particularly in northern and western parts of the subcontinent), which manifested in the introduction of Persian phonemes, morphemes, lexical items and, although rarely, also syntactic constructions into these languages. These strictly linguistic conditions were supported by a range of social and cultural factors which strongly furthered the language interference.⁹⁸ Persian remained the language of administration in South Asia for nearly a millennium. It enjoyed enormous prestige, both among its native and non-native speakers, and – favoured by certain political and social constellations – it played a crucial role as a cross-regional medium of culture, education and literary activity.⁹⁹ The legacy of the language contact between Indian languages

⁹⁸ Among the non-structural factors which influence the phenomenon of interference Weinreich enumerates, i.e., attitudes toward each language (whether idiosyncratic or stereotyped) and toward the culture of each language community; size of bilingual group and its socio-cultural homogeneity or differentiation; the relations of social or political nature between the sub-groups existing within this bilingual group etc. (Weinreich 1953: 3-4). These extra-linguistic aspects of language contact between Persian and Indo-Aryan languages have been discussed in the above chapter.

⁹⁹ Various reasons for the supra-regional use of languages range from “dominance by sheer power of conquest, over a combination of political and cultural or technological/commercial predominance, to purely cultural pre-eminence” (Hock 1991: 473). As has been demonstrated, in the case of Persian in South Asia all these motives were equally relevant and important.

and Persian remains extensive and vivid in South Asia, even though Persian is no longer spoken in the subcontinent. It shows itself in the classical forms of poetry (like *ġazal*, *rubāʿī* or *marṣiyah*), cultivated with great eagerness also by contemporary writers, in the vibrant tradition of performative literary forms (*qavvālī*), or in the heavily Perso-Arabicized, popular film songs. But, above all, Persian still provides the Indo-Aryan languages with their most substantial set of non-Indian elements.



Fig. 4. A trilingual (Kannada, Hindustani and English) guidepost at the Gol Gumbaz in Bijapur, with the indication written in Kannada, Nagari and Roman scripts. The sentence reads in Kannada: *golagummaṭakke hōguva dāri*, and in Hindustani: *gol'gummaṭ jānekā rāstā*. The way in which the Persian name of the monument has been noted in Nagari, viz. *golgummaṭ*, clearly corresponds to its form in Kannada (cf. P *gol gumbaz* / H *gol gumbad* 'rose dome'). (Photo by the author, 2012; I am grateful to Prof. Adapa Satyanarayana for his help in transliterating the Kannada version)

2. Scripts and Orthography

(...) nearly every people has a separate alphabet
(Abū'l-Faẓl 1873: 105)

Hindústání, which had no alphabet of its own, was written sometimes in Persian, sometimes in Devanágari characters (...)
(Monier Williams 1859: 258)

2.1. Writing Systems in Persian and Hindustani

2.1.1. As a result of the tremendous Islamic expansion and conquest carried on since the 7th c. CE, the Arabic alphabet, in which the Islamic holy book, the *Qurʾān*, was written, reached a vast area of the Middle East, North and Central Africa and South and Southeast Asia; then it was adapted to many Semitic as well as non-Semitic languages, among them Turkic and Iranian.

2.1.2. Persian¹⁰⁰ is written in the so-called Perso-Arabic alphabet, based on the Arabic script and read from right to left. It consists of 32 graphemes;¹⁰¹ four of them, namely: *pe* (پ), *če* (چ), *že* (ژ) and *gāf* (گ) were added by the Persians to the Ara-

¹⁰⁰ Persian (Fārsī) belongs to the southwestern group of Iranian languages and is characterized by a wide number of linguistic varieties (ethnolects) within its three major variants: Farsi – used in Iran, Dari – used in Afghanistan and Tajiki – spoken in Central Asia (Tajikistan). The major differences have occurred between the spoken forms of different regions. In the present book, if not specified otherwise, I use the term Persian in its broader sense.

¹⁰¹ In some sources the number 33 is given, depending on whether *hamzah* is or is not counted as a letter (cf. Hashabeiky 2005: 63).

bic alphabet. The majority of the letters of the alphabet have four forms: initial (joined on the left), medial (joined on both sides), final (joined on the right) and isolated (or final/initial unjoined); the appearance of a letter changes depending on its position in the word. Seven letters: *alif* (ا), *dāl* (د), *zāl* (ذ), *re* (ر), *ze* (ز), *že* (ژ), *vāv* (و) cannot join the following letter and, therefore, they do not have distinctive initial¹⁰² or medial forms but the isolated and the final forms are used instead. Short vowels /æ/, /o/ (/ʊ/) and /ɛ/ (/ɪ/)¹⁰³ are not written and even though there is a means of marking them,¹⁰⁴ it is used only occasionally and mostly in the school textbooks. Short vowels in final position, contrary to Arabic, are always displayed and indicated by *vāv* (و) = /o/ and by *he* (ه) = /æ/ or /ɛ/. A few letters occur only in words

¹⁰² The term 'initial' is used not only to denote an initial position of a letter in a word but also a medial, when it follows one of the letters which cannot join the following letter.

¹⁰³ The vowels occurring in numerous Perso-Arabic words borrowed by Hindustani depict an early New-Persian stage of development of the Persian vowel system (10th-15th c.), hence the vocalization of these loanwords differs significantly from the modern Persian forms of the same words. The older forms of short vowels (/ʊ/ and /ɪ/) are specified in parentheses. These forms were current in Persian before the 16th c. and in many cases survived unchanged in words borrowed by Hindustani, while in modern Persian they evolved into /o/ and /ɛ/ respectively (more see in paragraph 3.1.3.; cf. also Pisowicz 1985: 12ff.; Kuczkiewicz-Fraś 2003b: 68-69; Kłagisz 2010: 29).

¹⁰⁴ The signs (called *ḥarakāt*), used to denote short vowels and written above or below the letter to which they refer, are respectively: َ *fathah* (= /æ/), ِ *kasrah* (= /ɪ/ ~ /ɛ/) and ُ *zammah* (= /ʊ/ ~ /o/); there exist three more orthographic signs: ْ *sukūn* or *jazm* (written over the final consonant of a shut syllable, shows that a consonant is not vocalized), ̣ *tašdīd* (used to indicate that a consonant is doubled), and ٓ *tanvīn* (or nunation, an adverbial ending pronounced /æn/, written over the final *alif* of a noun) – but they may also be omitted and do not always appear in texts (cf. Lambton 2000: xix; Kłagisz 2010: 41).

of Arabic origin (their number varies from 6 to 8 according to different scholars (cf. Hashabeiky 2005: 63); these are: *he* (ح), *ṣād* (ص), *zād* (ض), *lā* (ط), *zā* (ظ), *‘ain* (ع), *ṣe* (ث) and *qāf* (ق) (the two latter being disputable).

2.1.3. The process of dispersal and accommodation of Arabic script may serve as a perfect illustration of Diringer’s opinion that the alphabet follows religion (Diringer 1948). With the spread of Islam, the Perso-Arabic writing system which developed on Persian ground, together with a bulk of Arabic and Persian vocabulary, has been adopted – following various necessary modifications – by many other languages adjacent to Persian and remaining under its cultural influence, e.g.: Pashto, Kashmiri, Saraiki, Urdu, Sindhi, and many others. Some languages, like Panjabi or Gujarati, use this script alongside their own writing systems (more on IA languages using Perso-Arabic script cf. Salomon 2003: 99-100; Masica 1991: 151-153).

2.1.4. The writing system of a Semitic language, which Arabic is, had to become appropriated for the needs of Indo-European tongues, since Arabic did not possess certain sounds found in other languages, and it did possess certain nuances of sounds not required for others.¹⁰⁵ Hence, in Perso-Arabic script, we are dealing with a number of orthographic simplifications and modifications. The earliest additions and subtractions appeared already in Persian, which was the first Indo-European language to adopt Arabic script for its purposes.¹⁰⁶ Some of them were of a purely visual nature, for example, writing the final *kāf* (ك) with-

¹⁰⁵ One of the major problems of Semitic scripts used for Indo-European languages is that the latter are much richer in vowels, which Semitic scripts (at least Arabic and Hebrew) tend mostly to ignore.

¹⁰⁶ The oldest preserved and dated Persian text written in the Perso-Arabic script was written down by Asadī Tūsī in 1055, but there are scholars’ opinions that New Persian must have been written in the Arabic script much earlier than this date (cf. Hashabeiky 2005: 76 ff).

out a flourish (in Arabic it would be ك). Other changes had a more significant and enduring effect on the evolution of New Persian lexis. All these changes were incorporated into the languages which later on adopted the Perso-Arabic alphabet. For this reason, to understand the processes and phenomena observed in other languages using – partly or entirely – the Perso-Arabic writing system, it is necessary to possess at least basic knowledge concerning these modifications on the Persian ground.

2.1.5. Major adaptations of the Arabic characters in Persian are as follows:¹⁰⁷

(i) *hamzah* (ء) – a glottal stop (mostly in the Arabic loanwords), in Arabic it usually takes a ‘seat’ (*kursī*),¹⁰⁸ which might be *alif* (أ), *vāv* (و) or dotless *yā* (ي). In Persian an initial *hamzah* is represented by a mere *alif*, e.g. A اثر (*A²-ṣ-r*) ≥ P اثر (*A-ṣ-r*) /æsaer/ ‘a footprint, sign, mark; result, consequence’ (Stein. 16); in the middle of a word *hamzah* based on *alif* or *vāv* is often

¹⁰⁷ Apart from these changes, most Arabic loanwords are written in Persian, and in other languages which use Perso-Arabic script, exactly as they are written in Arabic (although the Kurdish dialect Sorānī is an exception). There are two probable reasons for this: firstly, the sacredness of Arabic as a vehicle of the *Qurʾān* was a decisive factor upholding the traditional forms and defending them against any alterations; secondly – Arabic elements were, at the early stages, introduced into Persian mainly by bilingual scholars who had no need to conform them to the vernacular usage (Perry [1992]; Perry 2002: 229; cf. also Perry 2012: 77-81). Additionally, the strong tradition of teaching a reading knowledge of the *Qurʾān* (which according to the overriding principle ought to be learned in Arabic, not in translation) has been a factor reinforcing the adherence to the Arabic standards of spelling (Shackle & Snell 1990: 48). Thanks to their traditional Arabic spelling these loanwords are easily recognizable in most non-Arabic texts.

¹⁰⁸ Except for the situation when it follows a long syllable, like e.g. in Arabic إن شاء الله (*A²-n-ṣ-A^{-c}-ll-h*) /ɪnfəʔ allah/ ‘God willing’.

omitted, e.g. A رأس ($r-A^2-s$) $\geq$ P راس ($r-A-s$)¹⁰⁹ /raʔs/ ‘(...) the head’ (Stein. 562); A رؤيا ($r-V^2-Y-A$) $\geq$ P روىا ($r-V-Y-A$) /ruʔja/ ‘a vision, dream’ (Stein. 598); *hamzah* based on dotless *yā* is always marked, e.g. A مطمئن ($m-t-m-Y^2-n$) $\geq$ P مطمئن ($m-t-m-Y^2-n$) /mutmæʔin:/ ‘quiet, secure (...)’ (Stein. 1262). In the most common Arabic loanwords one can observe a tendency to write *hamzah* above dotless *yā* instead of placing it above *alif* or *vāv*, as in the original Arabic forms, e.g.: A مسألة ($m-s-A^2-l-i$) $\geq$ P مسئله ($m-s-Y^2-l-h$) /mæʔsʔælæ/ ‘a problem’ (Hayyim II: 898);¹¹⁰ more examples – see Kłagisz 2010: 45–47;

(ii) *tāʔ marbūṭah* – the Arabic feminine ending, marked by a unitary graph (ة). In Persian, it is written as two distinct graphic and lexical endings:¹¹¹ $-(a)t$ (ت), e.g. A قيمة ($q-Y-m-i$) $\geq$ P قیمت ($q-Y-m-t$) /qimæʔt/ ‘price, value, worth (...)’ (Stein. 999) and $-(a)h$ (ه), e.g. A قصة ($q-ṣ-i$) $\geq$ P قصه ($q-ṣ-h$) /qis:æ/ ‘a thing, affair (...)’; history; a tale, fable (...)’ (Stein. 974), the latter pronounced /æ/ or /ɛ/ in New-Persian;¹¹²

¹⁰⁹ Steingass usually gives both the forms (traditional Arabic and modified Persian), indicating at the same time the modified forms as current in Persian. This is also the case in the Hayyim’s dictionary, though sometimes exclusively Arabic spelling is used (even in sample sentences). However, this tendency towards simplification is not confirmed in modern Persian dictionaries (e.g. Šadrī Afšār et al. 1994), where we can still find both the forms (esp. in the words where *hamzah* is based on *alif*).

¹¹⁰ Such notations do not occur in Steingass’s dictionary, cf. “مسألة *masʔalat*, *masʔala*, A question, query (...)” (Stein. 1266).

¹¹¹ There are cases where the same word is written both with ت and ه, but with different significations, e.g. آلة ($\tilde{A}-l-i$) /alæʔt/ is written in this way when it’s general meaning is ‘an instrument, apparatus’, but آلت ($\tilde{A}-l-t$) /alæʔt/ when it signifies ‘penis’ (cf. Phillot 1919: 20).

¹¹² This dichotomous form class comprises about 1500 Arabic loans in modern Persian, which is almost one-fifth of the Arabic vocabulary in Persian (Perry 2002: 230), including about forty pairs of doublets (items lexicalized in Persian with both endings, showing semantic, lexical, or

(iii) the final *yā* (ی) is written without two dots underneath (in Arabic it would be ی). This apparently only graphical simplification has in fact further consequences, leading to the unification of two Arabic characters: final dotted *-yā* and final dotless *-yā* called *ʿalif maqṣūrah* and sometimes written with a small *alif* sign (*alif-e kūṭāh*) above (آ). As a result, since the two cannot be differentiated, in Persian Arabic words ending with *ʿalif maqṣūrah* (which is pronounced as long [a:]) are usually either written with final *alif*, following the pronunciation, e.g. A معنی (*m^c-n-ā*) ≥ P معنا (*m^c-n-A*) [ma:ʔna:] ‘sense, meaning’ (Stein. 1276), or – in a few cases – are pronounced with long [i:] in pausa, following the spelling (cf.: “A معنی *maʿnā*, in P. also *maʿnī*, Sense, meaning...”, *ibid.*). Cf. also 2.3.9. (vii), 2.3.9. (x), 2.3.11. (viii).

2.1.6. The Arabic alphabet came into use in Persia in the 10th-11th c. Since that time many scholars and writers have pointed out its deficiencies even in respect to the Arabic language. They argued that the shortcomings of the consonantal script, designed originally for Semitic languages, caused many failings in the Persian orthography resulting in misspelling as well as misreading. Hashabeiky (2005: 66-67) enumerates some frequently discussed shortcomings of the Persian writing system which affect not only Persian, but also other languages written in the Perso-Arabic script, including Hindustani/Urdu: (i) the usage of several characters for one and the same phoneme; (ii) the usage of one character for displaying several phonemes; (iii) omission of short vowels; (iv) the presence of two different types of letters, namely those that cannot join the following letter and those that can be bound to the left; (v) the presence of Arabic words and phrases which maintain their original Arabic spelling;

stylistic distinctions in the use of the two forms). More about this group of words and their influence on the evolution of New-Persian lexis cf. Perry 1995; Perry 2012: 79-80.

(vi) the presence of diacritic marks which are not obligatory and can be skipped; (vii) some omissions and changes in pronunciation, e.g. *vāv* (و) which is not pronounced in some words after *xe* (خ) or the consonant cluster *-nb-* (نب) which is phonemically realised as /mb/. All these problems will be discussed in greater detail later in the book, in relation to Hindustani/Urdu.

2.1.7. There are two main forms of script in which Persian can be written: *nasx*, which is the style developed in the 9th c. and still most commonly used in Arabic, Persian, Pashto, Sindhi and some other languages for printing of mass-circulated books, newspapers etc., and *nasta'liq* ($\leq$ *nasx-e ta'liq* 'hanging *nasx*') – a style predominant in Persian calligraphy, employed for fine printing, correspondence, bureaucratic purposes etc. Created in the 14th c. and widely used in the Persianate world since then, *nasta'liq* has become a preferred style of writing in Urdu, Kashmiri, Panjabi, Balochi, Brahui and other languages of contemporary Afghanistan, Pakistan and India; it is also used alongside *nasx* in Pashto (Phillot 1919: 35-36; Mace 2003: 19; Morton 2012: 148).

بر اساس گزارش‌های دریافتی، نیروهای دولتی در سوریه حملات مخالفان مسلح در شهرهای دمشق و حلب را دفع کرده اند. تلویزیون دولتی سوریه تصاویری از خرابی‌ها و اجساد آنهایی را که "تروریست" توصیف کرده، به نمایش گذاشته است.

Fig. 5. A sample of modern Persian text in *nasx*.

بر اساس گزارش‌های دریافتی، نیروهای دولتی در سوریه حملات مخالفان مسلح در شهرهای دمشق و حلب را دفع کرده اند. تلویزیون دولتی سوریه تصاویری از خرابی‌ها و اجساد آنهایی را "تروریست" توصیف کرده، به نمایش گذاشته است.

Fig. 6. The same Persian text written in *nasta'liq*.

2.1.8. One of the important factors that contributed to the spread and popularization of Perso-Arabic script was its aesthetic value. Combining “stenographic terseness” with “extreme ornamentation” (cf. Rypka 1968: 75), the Perso-Arabic script was soon appreciated by artists, who “raised fine penmanship to one of the highest ranks in the field of Iranian art” (ibid.). Furthermore, writers, especially poets, sometimes deliberately altered the form of letters to suit the ideas expressed in their works.

2.1.9. In Hindustani/Urdu, as in other Indo-Aryan languages which are written – exclusively or not – in the Perso-Arabic script, the alphabet follows the order and system of Persian, with the addition (mainly by various kinds of diacritic modification) of several extra letters which are necessary to represent characteristic IA phonemes not occurring in Persian or Arabic (see *infra* 2.2.). However, the orthography of the Persian and Arabic loanwords is conservative, retaining and representing the phonetic distinctions that are no longer audible in their pronunciation.¹¹³

2.1.10. While Hindustani/Urdu uses the modified Perso-Arabic consonantal alphabet,¹¹⁴ Hindustani/Hindi is written in Nagari, which is one of the most important South Asian alphasyllabaries¹¹⁵, used to record, i.a., Sanskrit, Hindi, Marathi and Nepali. This fact distinguishes Nagari from most other South

¹¹³ Many of these distinctions in pronunciation of Arabic words and phrases were already abandoned in Persian (even though they keep their Arabic spelling), and this practice is continued in the languages into which these Arabic loanwords were incorporated through the mediation of Persian, cf. Lambton 2000: 181-182; Mace 2003: 181 ff.

¹¹⁴ This standing assignment, propping the traditional ideology proclaiming that the choice of script is crucial in differentiating between Urdu and Hindi, has been successfully challenged by several recent studies (e.g.: Abbi et al. 2004; Shahabuddin 2006; Ahmad 2011).

¹¹⁵ Or syllabic alphabet (Montaut 2004: 16).

Asian writing systems, usually linked with and limited to a single language.¹¹⁶

2.1.11. Nagari script used in Hindustani/Hindi is almost identical (with a few minor variations/additions) to that, which is used for Sanskrit and which serves as the prototype for other languages using this script. It is written from left to right and consists of thirty three¹¹⁷ basic graphemes (H *akṣar*), built of CV sequences,¹¹⁸ and eleven independent characters to denote syllabic vowel forms (cf. Tables of Adequacy, p. 116-120). Apart from these main characters there are also ten diacritic marks (*mā-trā*) used to notate intra-syllabic vowel forms, which are combined with consonant characters: का (*kā*), कि (*ki*), की (*kī*), कु (*ku*),

¹¹⁶ In India, since the country's independence, Nagari has enjoyed to some extent the special status of the 'national script', being connected to the 'national language' (Hindi), cf. Salomon 2003: 75; Saran 1969: 121-122. However, in the 19th and early 20th c. it had to fight for its position of the regional script (i.e. used for Hindi) with several other writing systems. Nagari's most serious opponent was Kaithī (or Kāyasthī, 'script of the Kāyasth cast of writers and clerics'), widely used for writing legal, administrative, and private records in parts of North India, primarily in the former North-Western Provinces, Avadh and Bihar. Nagari won the battle very narrowly and Kaithī lost, mainly due to the political decisions. The conflict between Nagari and Kaithī has been scrutinized e.g. by King (1994).

¹¹⁷ This number does not include graphemes modified by adding a dot beneath, created to mark specific Hindi retroflex flaps *ṛa* (ॠ) and *ṛha* (ॡ) (which are actually allophones of retroflex plosives *ḍa* (ጅ) and *ḍha* (ጆ) respectively) or to represent borrowed Perso-Arabic phonemes, alien to the IA sound system (these are discussed in paragraph 2.3.4).

¹¹⁸ In Nagari, a fully represented underlying consonant is assumed to have the inherent vowel. If there is no other vocalizing diacritic mark, this role is played by the neutral vowel ('schwa' /ə/). It is not marked in any way and its presence should be guessed by a reader. The inherent neutral vowel is not normally vocalized at the end of a word. Cf. Pandey 2007: 146.

कू (*kū*), के (*ke*), कै (*kai*), को (*ko*), कौ (*kau*), कृ (*kr*)¹¹⁹ and a few other diacritics to denote such phenomena as homorganic nasals (*anusvāra*) and nasalisation (*ˆ candrabindu*), absence of an inherent vowel (*˘ virāma*) or voiceless aspiration (*: visarga*).¹²⁰ The consonant clusters are indicated either by means of *virāma* or through the creation of consonant ligatures of two or more characters.¹²¹ This rule, however, is applied very inconsequently in writing and there are many examples of consonant clusters which are commonly pronounced but not marked in the script.¹²²

¹¹⁹ In Hindustani, the character ऋ (*r*), together with its diacritical version (ॠ), is vocalic only in terms of the script (in Nagari it has separate syllabic and intrasyllabic forms), but it is actually pronounced as a CV sequence of an alveolar tapped [r] and following short [ɪ], and appears exclusively in borrowed Sanskrit words (*tatsama*). Cf. McGregor 1999: xxv; Shapiro 1989: 10.

¹²⁰ Note that *visarga*, which in Sanskrit was supposedly unvoiced, becomes voiced in Hindustani (Srivastava 1969: 925).

¹²¹ According to the general rule, half the sign of the first consonant precedes the full sign of the second consonant or, in case of three-consonant clusters, the first two consonants are specified for half their letters, and the third is fully specified, though for some clusters a brand new grapheme has been formed.

¹²² The massive phonological changes that occurred during the development of New Indo-Aryan resulted, i.e., in a widespread process of cluster simplification. Thus there are only a few indigenous consonant clusters in Hindustani, while the majority of initial and final clusters occurs in borrowed Sanskrit, Perso-Arabic or English lexical items. On the other hand, new consonant clusters are formed as a result of the complex and greatly irregular phonological process of “schwa deletion” (Pandey 2007: 147), in words of more than one syllable and in consequence of the effects of interaction between such vowel dropping and affixation. Generally, they are not represented in the Nagari script as ligatured consonants since they are not syllable internal consonant clusters (more cf. Kachru 2006: 21-29). See also remarks on notation of consonant clusters in Perso-Arabic loanwords in the Nagari script in 2.3.5. and 2.3.6.

2.1.12. The Nagari script (as almost all indigenous Indic scripts that evolved from the ancient Brāhmī script)¹²³ was devised to record Old Indo-Aryan phonemes, and since the sound system within Hindustani/Hindi is generally that of the OIA, Nagari characters correspond to this system strictly and allow a speaker to pronounce even an unknown word in the accurate way.¹²⁴ Furthermore, thanks to its greater stability, Nagari also allows for clearer representation of the part of Hindustani lexis which has been borrowed from other languages – be it Persian, Arabic or English.¹²⁵ This quality remains the most serious difference between Nagari and Perso-Arabic script – in the case of the latter the correct pronunciation of a previously unknown word always remains uncertain – one has to know the language already to be able to read it properly in the Perso-Arabic script (cf. 2.3.20-2.3.21).

¹²³ On the historical emergence of the Nagari script see e.g. Montaut 2004: 15-16; Salomon 2003: 72-75.

¹²⁴ Cf. the opinion of Monier Williams: “Deva-nāgarī (...) although deficient in two important symbols (represented in Roman by *z* and *f*), is on the whole the most perfect and symmetrical of all known alphabets. (...) too perfect for the practical purposes of this utilitarian age” (Monier Williams 1859: 261).

¹²⁵ On notation of Perso-Arabic loanwords in Nagari cf. Chapter 2.3.

1.2. Adaptation of the Persian Writing System in Hindustani/Urdu

2.2.1. As has been already mentioned, the Perso-Arabic script needed a number of necessary modifications in order to adjust to the Hindustani language system.¹²⁶ Some of them were dictated by phonological differences, others – by peculiar grammatical rules. The most essential amendment was to create a method of representing sounds and phenomena specific to the Indo-Aryan tongue but alien to Persian and the more so to Arabic – like the retroflex and aspirated consonants, some elements of the vowel system, nasalization of vowels etc. As a result, a number of modified graphemes appeared in the Perso-Arabic alphabet adapted for the needs of Hindustani.

2.2.2. Retroflex consonants. To mark Hindustani retroflex consonants [ʈ], [ɖ] and [ʂ], three separate characters were added to the Perso-Arabic script's inventory, created by the addition of four diacritic dots or, alternatively, by the sign ^ٲ (the Arabic letter *ṭā'* in superscript) written above a letter. These added characters are: ^ٲ / ^ٲ (*ṭ*), ^ٲ / ^ٲ (*ḍ*), and ^ٲ / ^ٲ (*ṣ*). The variant with ^ٲ is definitely prevalent.

2.2.3. Aspirated consonants. The aspirated consonants are indicated by writing the medial form (^ٲ)¹²⁷ of the consonant *he* (so-called *čʰoṭī he*) after the corresponding unaspirated consonant (both in the initial/medial and in the final position). Even though it influences the utterance of the syllable in which it occurs, this *he* of aspiration does not count metrically as a letter (Pritchett & Khaliq 1987). Since this form of *he* is used in Hindustani/Urdu

¹²⁶ In today's Hindustani/Urdu, the *nasta'liq* style is almost exclusively used for both ordinary writing and for publications (cf. Naim 1999 I: 53).

¹²⁷ The so-called "butterfly- or spectacle-shaped" *h* (P *he-ye do čašm* (Adle 1982: 206); H *do-čašmī he*).

exclusively to denote aspiration, the characters which have been created by its means function as fixed graphemes in the writing system.¹²⁸ The table below presents all the graphemes coined for the aspirates and their appearance in different positions within the words.

ASPIRATE

س (b ^h)	INI./MED.	سبا (s-b ^h -a) /səb ^h a/ 'assembly'
	FIN.	جیبہ (j-Y-b ^h) /dʒib ^h / 'tongue'
پ (p ^h)	INI./MED.	پھیلنا (p ^h -Y-l-n-A) /p ^h aɪlna/ 'expand'
	FIN.	--
ت (t ^h)	INI./MED.	سُھرا (s-t ^h -r-A) /su ^h ra/ 'neat, clean'
	FIN.	کے ساتھ (k-e s-A-t ^h) /ke sa ^h / 'along with'
ٹ (ʈ ^h)	INI./MED.	لاٹھی (l-A-ʈ ^h -Y) /la ^h ʈi/ 'stick, club'
	FIN.	بیٹھ جانا (b-Y-ʈ ^h -j-A-n-A) /baɪ ^h dʒana/ 'sit down'

¹²⁸ In some grammar books (for instance Forbes 1846; Platts 1909; Barannikov & Barannikov 1956; Naim 1999) the number of letters added to express sounds peculiar to Hindustani is restricted only to those three, which mark retroflex *ʈ*, *ɖ*, and *ʀ*. Graphemes representing aspirated consonants are not recognized by their authors as separate and independent units of alphabet. The reason for this practice is probably those authors' opinion (although expressed directly in none of these books) that aspirated consonants are clusters of unaspirated consonants plus /h/. Indeed, from a purely graphical point of view, in the Perso-Arabic script these are ligatured consonants. The question whether they are or are not clusters also in terms of phonology remains a subject of inconclusive discussion (cf. e.g. Narang & Becker 1971: 651-652 or Shapiro 1989: 15-17). However, if for no other reason, then by analogy to Nagari script, where characters denoting aspirated consonants are inherent and independent items of the alphasyllabary, these fourteen graphemes should also be included in the Perso-Arabic (or Hindustani/Urdu) alphabet.

ج (j ^h)	INI./MED.	منجھلا (m-n-j ^h -Y-l-A) /məndʒ ^h ela/ 'delay'
	FIN.	بوجھ (b-V-j ^h) /bodʒ ^h / 'load, burden'
چ (č ^h)	INI./MED.	اچھا (A-čč ^h -A) /aʃʃ ^h a/ 'good'
	FIN.	باچھ (b-A-č ^h) /baʃ ^h / 'corner of the mouth'
د (d ^h)	INI./MED.	بندھنا (b-n-d ^h -n-A) /bənd ^h na/ 'be tied'
	FIN.	دودھ (d-V-d ^h) /dud ^h / 'milk'
ڈ (ḍ ^h)	INI./MED.	ڈھنگ (ḍ ^h -n-g) /d ^h əŋg/ 'manner'
	FIN.	گڈھ (g-ḍ ^h) /gəḍ ^h / 'fort' (= g-r ^h) /gəṛ ^h /
ڑ (ṛ ^h)	INI./MED.	آڑھت (Ā-r ^h -t) /aṛ ^h ət/ 'commission'
	FIN.	پاڑھ (b-A-r ^h) /baṛ ^h / 'volley'
ک (k ^h)	INI./MED.	کھی (m-kk ^h -Y) /mæk ^h hi/ 'fly'
	FIN.	سکھ (s-k ^h) /suk ^h / 'comfort'
گ (g ^h)	INI./MED.	گھر (n-g ^h -r-A) /niḡhəra/ 'homeless'
	FIN.	باگھ (b-A-g ^h) /baḡ ^h / 'tiger'

2.2.4. The eleven graphemes presented above correspond strictly to eleven letters of the Nagari script denoting Hindustani aspirated sounds. In the Perso-Arabic script, three more consonants, while followed by /h/, may be thought of and marked as aspirates, namely: nasals /m^h/ and /n^h/, and a lateral alveolar /l^h/.¹²⁹ When aspirated none of them occur at the end of a word,

¹²⁹ In some descriptions of the Hindi consonantal inventory, these three consonants in their aspirated forms are also enumerated as allophones of /m/, /n/ and /l/ respectively, represented in writing by the conjunct consonants म्हा (mha), न्हा (nha) and ल्हा (lha) (e.g. Tivārī 1969: 58). For Ohala (1999: 101), who mentions, beside these three, also /r^h/ (represented in writing as र्ह), these are clusters of liquids or nasals and /h/. Ejaz writes about total number of fifteen aspirated consonants (Ejaz 2004: 140).

they can only be found medially. However, their spelling might be confusing – in texts they appear (rarely) represented by a unitary grapheme, built according to the same rule as in the case of the ‘regular’ aspirates enumerated in the table above, or as a sequence of two independent letters. This ambiguity in writing is caused by the fact that /mh/, /nh/ and /lh/ not always are considered as single aspirates. When they result from vowel-dropping or affixation they are rather thought of as clusters of two consonants, where the first component goes with the preceding syllable and the latter – with the following (see also Naim 1999 I: 7 and remarks in fn. 122 on p. 66). Cf. the following examples:

د (l^h)

دلھن (d-l^h-n) /dul^hən/ vs. دلبن (d-l-h-n) /dul^hən/ ‘bride’

چولھا (č-V-l^h-A) /ʃul^ha/ vs. چولھا (č-V-l-h-A) /ʃul^ha/ ‘stove’

م (m^h)

کھار (k-m^h-A-r) /kum^har/ vs. کھار (k-m-h-A-r) /kum^har/ ‘potter’

تمھارا (t-m^h-A-r-A) /tum^hara/ vs. تمھارا (t-m-h-A-r-A) /tum^hara/ ‘your’

ن (n^h)

چینھنا (č-Y-n^h-n-A) /ʃin^hna/ vs. چینھنا (č-Y-n-h-n-A) /ʃin^hna/ ‘recognize’

انھیں (A-n^h-Y-ñ) /un^hẽ/ vs. انھیں (A-n-h-Y-ñ) /un^hẽ/ ‘them (Object)’¹³⁰

¹³⁰

It is worth to notice here that inconsistency might be observed, though very rarely, also in Nagari notation (but not in phonetic realisation), where /mh/, /nh/ and /lh/ are represented either by ligatures indicating the consonant cluster, or as a sequence of two independent letters (and then the lack of vocalization between them, often being a result of vowel dropping or morphological processes, is not marked), e.g.: दुलहन (dul^hhan) ‘bride’ (McGr. 505) but दूल्हा (dūlhā) ‘bridegroom’ (ibid. 508). However, despite this fact, Nagari notation of words of this type is much more stable than their spelling in Perso-Arabic script; cf.: चूल्हा (cūlhā) ‘stove’ (ibid. 328); कुहा

2.2.5. When *ĉhoī* *he* is to be pronounced as a separate letter, the form **ر** is used in the middle of the word, usually jointly with the diacritic mark of inverted comma (,) written beneath a letter. Cf.: بهار (*b-h-A-r*) /bəhar/ ‘spring’ vs. بهاری (*b^h-A-r-Y*) /b^hari/ ‘heavy’. In the final position the usual final form of the letter *he* is used, cf.: فواكه (*f-V-A-k-h*) /fəvakih/ ‘fruits’ vs. مکه (*m-k^h*) /muk^h/ ‘face’. In *nasta‘līq* style the grapheme *he* has a number of variants when joining different neighbouring letters, e.g.: هـ (*h-A*), هـ (*h-ī*), كهـ (*k-h-A*), دهـ (*q-h-d*), هـ (*Y-h*), هـ.¹³¹ (*b-h-h*).

2.2.6. Nasals and nasalization. In the Perso-Arabic system of writing adapted for Hindustani, the letter *nūn* (ن) stands for four nasal consonants of Indic origin: dental /n/, retroflex /ɳ/, palatal /ɲ/, and velar /ŋ/, which in Nagari are represented by four distinctive characters: न (*n*) ण (*ṇ*) ञ (*ṇ̥*) and ङ (*ṅ*) respectively, and for *anusvāra* (ँ) which in Nagari stands for postvocalic homorganic nasals (cf. 2.3.3.(v)). *Nūn* is also used to indicate nasalization of vowels (then it is called *nūn-e gunnah*).¹³²

(*kumhār*) ‘potter’ (ibid. 206); तुम्हारा (*tumhārā*) ‘your’ (ibid. 457); चीन्हना (*cīnh’nā*) ‘to recognize’ (ibid. 321); उन्हें (*unhem*) ‘them (Object)’ (ibid. 321).

¹³¹ This variant appears either in roots or imperative forms of the verbs in which root ends in /h/, e.g.: कह (*k-h-h*) /kəh/ ≤ كهـ (*k-h-n-A*) /kəhna/ ‘to say’; به (*b-h-h*) /bəh/ ≤ بهـ (*b-h-n-A*) /bəhna/ ‘to flow’ or in certain other words to indicate a final /h/ after a front long vowel, e.g.: بهـ (*b-Y-h-h*) /beh/ ‘perforation, hole’ (Platts 212); تشبیه (*t-š-b-Y-h-h*) /təʃbih/ ‘similitude, comparison...’ (ibid. 324). Note that some examples given by Platts are erratic, e.g.: کُنْه (*k-n-h-h*) /kunh/ ‘substance, essence...’ (ibid. 856), where the preceding vowel is neither front nor long, cf. کُنْ (*k-n-h*) /kunh/ ‘essence (of)...’ (STCD 500).

¹³² Platts uses the label ‘indistinct’ (*xafā*) for this form of *nūn*, which describes the variable character of this phenomenon very well (cf. Platts 1909: 7). *Nūn-e gunnah* is – like the *he* of aspiration – also invisible metrically and does not count as a letter, although it affects the pronunciation of the syllable in which it occurs (Pritchett & Khaliq 1987).

2.2.7. Depending on its position in the word or syllable, *nūn* may be realized as a consonant, as nasalization, or as both – the rules are not strict,¹³³ and depend also on individual choice (*vide infra* 3.2.7.). But regardless of whether it means a consonant or nasalization, it is always notated by a *due* (i.e. initial, middle or final) form of the letter *nūn* (ن),¹³⁴ with one exception: *nūn-e ġunnah*, while representing the final nasalized long vowel, has its independent variation written without the diacritical dot (ن).¹³⁵ Medially, however, *nūn-e ġunnah* is written as regular *nūn* (ن) and one has to judge from the context whether it represents nasalization or the consonant. Cf. the following examples:

بندکی (*b-n-d-k-Y*) /bundki/ ‘tiny drop’
 cf. بُندکی (*bumd’kī*)
 بُندکی (*bund’kī*)

¹³³ Cf.: “The contrast between a homorganic nasal stop and a nasalized vowel can be identified, but is unstable, particularly with long vowels” (Schmidt 2003: 309).

¹³⁴ In Roman transliteration, *nūn-e ġunnah* is usually depicted by *ñ* or by the placement of a tilde (˜) over the vowel sign (e.g. *meñ* / *mē* ‘in’; *jā’ūngā* / *jā’ūgā* ‘I will go’).

¹³⁵ *Nūn-e ġunnah* is occasionally written as a standard *nūn* with the dot also in the final position. For example, in his dictionary, Platts does not differentiate between *nūn* and *nūn-e ġunnah* in Perso-Arabic script, even though he indicates nasalization in both the Nagari and Roman transliteration of words, cf. “یہاں *yahān*, (...) adv. In this place, here...” (Platts 1252). This is strange because in his grammar which precedes the dictionary by fourteen years, nasal *n*, when final, is printed without the dot (ن) and in the Roman character it is indicated, in all positions, by *n* with the dot (*ñ*). Inconsistencies in notation of *nūn-e ġunnah* occurring finally are not rare, differences are observed especially between the colloquial and literary use, cf.: “جان (*jān*, lit. also جان *jān*) life, soul...” (STCD 223); “بَدَان (*bād-bān*, poet. also بَادَان *bād-bān*) sail” (STCD 92); “تَسْكِين (*taskīn*’ (lit. also تَسْكِين *tas’kīn*) consolation, comfort...” (STCD 190). It is worth to note the shift of accentuation in the last two examples.

دنتیلا (<i>d-n-t-Y-l-A</i>) /dəntila/ 'toothed'	cf. दँतीला (<i>damtīlā</i>) दन्तीला (<i>dantīlā</i>)
ٹھنڈی (<i>tʰ-n-ḍ-Y</i>) /tʰəṇḍi/ 'cold (f.)'	cf. ठंडी, ठण्डी (<i>tʰaṇḍī</i>)
گھنٹا (<i>gʰ-n-ṭ-A</i>) /gʰəṇṭa/ 'hour'	cf. घंटा, घण्टा (<i>gʰaṇṭā</i>)
پنجرا (<i>p-n-j-r-A</i>) /piɲdʒra/ 'cage'	cf. पिंजरा (<i>piñj'rā</i>)
کھینچنا (<i>kʰ-Y-n-č-n-A</i>) /kʰiɲtʃna/ 'to pull'	cf. खींचना (<i>kʰimc'nā</i>)
مانگنا (<i>m-A-n-g-n-A</i>) /māgna, maɲgna/ 'to request'	cf. माँगना (<i>mām̃g'nā</i>)
کلتکی (<i>k-l-n-k-Y</i>) /kələŋki/ 'blemished'	cf. कलंकी (<i>kalankī</i>)
هنس (<i>h-n-s</i>) /həns/ 'goose'	cf. हंस (<i>hams</i>)
ہاں (<i>h-A-n</i>) /hā/ 'yes'	cf. हाँ (<i>hām̃</i>)
دھواں (<i>dʰ-V-A-n</i>) /dʰuā/ 'smoke'	cf. धूआँ (<i>dʰūām̃</i>), धुआँ (<i>dʰuām̃</i>)
میزیں (<i>m-z-Y-n</i>) /mezē/ 'tables (pl. of mez)'	cf. मेज़ें (<i>mezem̃</i>)
کروں (<i>k-r-V-n</i>) /kərũ/ 'I would do (subj.)'	cf. करूँ (<i>karũm̃</i>)
دوں ¹³⁶ (<i>d-V-n-g-A</i>) /dũga, duŋga/ 'I shall give (fut.)'	cf. दूँगा (<i>dũm̃gā</i>)

¹³⁶ In Hindustani/Urdu script (contrary to Nagari) morphological units such as suffixes, some grammatical endings or postpositions are often written separately and do not join the preceding word/stem (though they are pronounced jointly). Cf. بت پرستی (*b-t p-r-s-t-Y*) /butpərəsti/ 'idolatry' (STCD 99) vs. बुतपरस्ती (*but'paras'tī*) 't.s.' (McGr. 741); کھوگا (*k-h-V-n g-A*) /kəhũga/ (the joint spelling کھوگے (*k-h-V-n-g-A*) also might occur sporadically) 'he will tell (3 pers. sg. fut.)' vs. कहूँगा (*kahũm̃gā*) 't.s.'; آپ (*A-p k-A*) /apka/ 'your' vs. आपका (*āp'kā*) 't.s.' Cf. also further remarks on spacing (2.2.27).

2.2.8. In Persian, *nūn* can stand for bilabial nasal /m/ when it is immediately followed by ب (*b*) or پ (*p*). Perso-Arabic loan-words of this type generally keep to this rule also in Hindustani, though sometimes in colloquial or corrupted forms the letter *mīm* can occur instead of *nūn*, e.g.: تيمبو (*t-m-b-V*) || تيمبو (*t-n-b-V*) /tambu/ 'tent, canopy' (Platts 336, 338); تمباکو (*t-m-b-A-k-V*) || تنباکو (*t-n-b-A-k-V*) /tambaku/ 'tobacco' (ibid.). Similar alternation may occur also in the case of genuinely Indic words when noted in Perso-Arabic script, but here the rule works the other way round, i.e. the cluster *-nb-* appears in words traditionally written in Nagari with *-mb-*, e.g.: چھپت (*ĉ-n-p-t*) || چھپت (*ĉ-m-p-t*) 'hidden, invisible', cf. चम्पत, चंपत (*campat*) 't.s.' (Platts 441, 443); لنبا (*l-n-b-A*) || لمبا (*l-m-b-A*) 'long, tall' (Platts 963, 964; STCD 550), cf. लंबा (*lambā*) 't.s.'; کنبل (*k-n-b-l*) || کمبل (*k-m-b-l*) 'blanket' (Platts 847, 851), کمبل (*k-m-b-l*) || कम्बल (*k-mm-l*) 't.s.' (STCD 496), cf. कंबल (*kambal*), कम्मल (*kammal*) 't.s.' (McGr. 154, 170); کنبه (*k-n-b-h*), کنبا (*k-n-b-A*) || کمبه (*k-m-b-h*), کمبا (*k-m-b-A*) 'family' (Platts 847, 851), کنبه (*k-n-b-h*) 't.s.' (STCD 498), کنبا (*k-n-b-A*) 't.s.' (Shak. 1373), cf. कुनबा (*kun'bā* ≤ *kutumba-*) 't.s.', कुटुंब (*kuṭumb*) 't.s.' (McGr. 205, 203), कुटुम्ब (*kuṭumb*) 't.s.' (Awasthi & Awasthi 1981: 449; Sharma 2006: 263).¹³⁷

2.2.9. As one can conclude from the examples given above, the notation (and pronunciation) of the cluster /mb/ is characterized in both the writing systems by a lack of consistency (cf. 3.2.7.(iii)).

2.2.10. Notation of vowels. The symmetrical ten-vowel system of Hindustani is more extensive than those of Arabic or Persian,¹³⁸ but generally the method of marking Hindustani vowels

¹³⁷ McGr (205) and the MHK (I, 552) record the form कुनबा (*kun'bā*) as well. In this case the alternate pronunciation [kun.ba:] is also observed, which according to Narang and Becker, might be an effect of metanalysis (cf. Narang & Becker 1971: 658, fn. 12).

¹³⁸ Vide *infra* 3.1.2.

and diphthongs in Perso-Arabic script follows the same general rules, that apply in Persian: short vowels are not indicated or, provided it is necessary, the diacritical signs of *fathah* (in Persian called also *zabar*), *kasrah* (P *zer*) and *zammah* (P *pīš*) are used; to denote long vowels and diphthongs the obligatory consonantal signs *alif*, *ye* and *vāv* are employed in accordance with the convention described below.

2.2.11. Long vowel /a/ after a consonant in the middle and final position is represented by *alif* (ا), e.g.: با (b-A-r) /bar/ 'time, turn'; جا (j-A-n-A) /džana/ 'to go', and in the initial position by *alif* with the sign of prolongation called *madd* (آ),¹³⁹ e.g.: آدمی (Ā-d-m-Y) /admi/ 'a man'.¹⁴⁰

2.2.12. Long vowels /i/, /e/ and the diphthong /ai/ after consonants are represented by the letter *ye* in its two allographic variants: *čʰoṛī ye* (ی), which stands for long /i/ and *barī ye* (ے) indicating /e/ or /ai/.¹⁴¹ Both the variants of *ye* have the same initial and medial forms, cf.:

¹³⁹ In a few words borrowed from Arabic, *alif* with *madd* appears to represent also medial long /a/, e.g.: قرآن (q-r-Ā-n) /qurʾan/ 'the Qurʾān' (Barz & Yadav 1993: 285), which is conditioned metrically (Pritchett & Khaliq 1987).

¹⁴⁰ *Alif* in the initial position invariably shows the presence of a vowel or a diphthong. Short vowels may be marked by the diacritical signs placed above or beneath *alif*, e.g.: اندر (A-n-d-r) /əndər/ 'within'; انصاف (A-n-ṣ-A-f) /insaf/ 'justice'; آتار (A-t-r-n-A) /utərna/ 'to get down', but this is a rare practice, used mostly to differentiate heterophones, e.g.: اس (A-s) /is/ 'this' vs. اُس (A-s) /us/ 'that'.

¹⁴¹ *Barī ye* (also known by its Persian name *yā-e maʿkūsī* or 'reversed *yā*') is an innovation devised in Hindustani/Urdu script mainly because it was necessary to clearly represent the vowel /e/ in the final position and differentiate it from the long /i/, since in Hindustani they both perform the grammatical functions of various inflectional endings. In Persian, however, *ye* also has two variants, though they do not differ in script and the distinction between them is only phonemic: *yā-majhūl* ('unknown' *yā*) – when it

<i>čʰoḡī ye</i> :	ايجاب (ijāb) 'assent'	فيل (fīl) 'elephant'
<i>baṛī ye</i> :	ایک (ek) 'one'; پیش (peš) 'front';	ایک (aibak) 'slave' پیک (paik) 'messenger'

The difference between *čʰoḡī* and *baṛī ye* in these examples, as well as the correct phonemic realization, must be judged from the context and requires prior knowledge of a word (or at least some experience of the language). Also, the lack of clear orthographical differences results in a large number of homographs, e.g.: (i)  (*p-Y-r*) /pir/ 'Monday'; (ii)  (*p-Y-r*) /pir/ 'spiritual guide, saint'; (iii)  (*p-Y-r*) /pair/ 'foot'; (iv)  (*p-Y-r*) /pir/ =  (*p-Y-r*) /piṭ/ 'pain'; or (i)  (*s-Y-n-A*) /sina, saina/ 'Sinai'; (ii)  (*s-Y-n-A*) /sena/ 'to hatch (eggs)'; (iii)  (*s-Y-n-A*) /sena/ 'army, force'; (iv)  (*s-Y-n-A*) /sina/ 'to sew, stitch'. **Bibl Jag**

2.2.13. However, when occurring at the end of a word or syllable, these two allographs of *ye* have visibly distinct forms, cf.:

<i>čʰoḡī ye</i>	<i>baṛī ye</i>
 (<i>b-čč-Y</i>) /bəḡf:i/ f. sg. 'girl'	 (<i>b-čč-e</i>) /bəḡf:e/ pl. 'children'
 (<i>s-Y</i>) /si/ adj. f. 'like'	 (<i>s-e</i>) /se/ ppn. 'by means of'
 (<i>s-m-jʰ-Y</i>) /səmdʒʰi/	 (<i>s-m-jʰ-e</i>) /səmdʒʰe/
'(she) understood'	'(they) understood'

Yet this distinction, even though requisite for grammatical reasons, makes the matter easier only to some extent, because of numerous heterophones existing within both the groups, cf.: (i)  (*p-Y*) /pi/

represents the sound /e/, which is alien to the Arabic vowel system and *yā-ma'rūf* ('known' *yā*), standing for long /i/, also known in Arabic (cf. Phillot 1919: 21).

m. ‘(woman’s) beloved’ (= پیا /pija/); (ii) پی (p-Y) /pi/ f. sg. perf. ‘(she) drank’; (iii) پی (p-Y) /pəj/ m. ‘anything to drink; milk, water’ (= S पय *paya*); (iv) پی (p-Y) /pai/ f. ‘fault, blemish’ (= H पै *pai*) vs. (i) पे (p-e) /pe, pai/ ppn. & conj. ‘on, upon’; (ii) पे (p-e) /pai/ m. ‘foot, footstep’.

2.2.14. Long vowels /u/, /o/ and the diphthong /au/ occurring after consonants are represented by the letter *vāv*. Here again the choice depends on the specific word, cf.:

سوراخ (s-V-r-A-x) /surax/ ‘hole’	بُ (b-V) /bu/ ‘smell’
فیروزہ (f-Y-r-V-z-h) /firoza/ ‘turquoise’	چالو (č-A-l-V) /tʃalo/ ‘go!’
دوڑنا (d-V-r-n-A) /dauṛna/ ‘to run’	نَو (n-V) /naʊ/ ‘nine’ ¹⁴²

2.2.15. Apart from long /a/, all other long vowels and diphthongs, when in the initial position, are represented, respectively, by the letters *ye* or *vāv*, which follow the initial *alif*, e.g.: ایزاد (A-Y-z-A-d) /izad/ ‘addition (of)’; ایلچی (A-Y-l-č-Y) /elʃi/ ‘envoy’; ایمان (A-Y-m-A-n) /aiman/ ‘right-hand’; اوپر (A-V-p-r) /upər/ ‘up; above’; اوکھنا (A-V-r^h-n-A) /oṛ^hna/ ‘to cover (body)’; اولاد (A-V-l-A-d) /aulad/ ‘children, offspring’.

2.2.16. In some Arabic loanwords the initial vowel or diphthong is indicated by the letter *‘ain* (ع) (which otherwise, i.e. in a non-initial position, is pronounced as a glottal stop¹⁴³), e.g.: عبادت

¹⁴² Platts gives a number of heterophones of the form نُ (n-V): [naʊ], [nəʊ], [no], [nə.ʊ] (cf. Platts 1157). But there are also homographs: نُ /naʊ/ ‘a boat’ and نُ /naʊ/ ‘new, fresh’ (with its heterophonic variant [naʊ]).

¹⁴³ Actually, *‘ain* is pronounced as a glottal stop only in the highly Arabicized, official style of Hindustani/Urdu (cf. Schmidt 2003: 310). Most speakers, however, commonly pronounce it as a vowel. Note that in Arabic *‘ain* represents a consonant (viz. a voiced pharyngeal fricative /ʕ/) and is never pronounced as a vowel, but such an utterance has remained unadopted by Persian, and all the more so by Hindustani (Phillot 1919: 16).

(^c-b-A-d-t) /ibadət/ 'worship'; عَرْت (^c-V-r-t) /aurət/ 'woman'; اَم (^c-A-m) /am/ 'common'.¹⁴⁴

2.2.17. Arabic elements, such as *'alif maqṣūrah* and *tā' marbūṭah* occurring in some borrowed Arabic words which maintain the traditional spelling, follow the same rules of adapting and simplifying in Hindustani/Urdu script as they did in Persian.¹⁴⁵ Cf. *supra*, 2.1.5.(ii) and (iii).

2.2.18. *Hamzah*, however, in Hindustani/Urdu script has two functions: (i) it occurs in traditionally written Arabic loans, including those which went through the process of simplification in Persian;¹⁴⁶ (ii) in all words, regardless of their origin, it indicates vowel clusters, i.e. places where two vowels occur without an intervening consonant. In this function, *hamzah* is always placed between the two vowel signs but points towards the second vowel and, except when it is combined with *vāv* (و), requires a 'seat' (*kursī*), which has the shape of the *b*-series letters. *Hamzah* itself has two graphical variants: (َ) – which seems to be more common, and (؀) – which looks like a slightly rotated initial form of the letter *ain* (or like its upper part) in superscript.

2.2.19. The possible vowel combinations with *hamzah* may comprise of:

(i) two long vowels, e.g.: بھائی (*b^h-A-Y²-Y*) /b^ha.i/ 'brother'; کھیں (*k^h-e-Y²-Y-ñ*) /k^he.ĩ/ 'rowed' (perf. pl. f. of *k^henā*); سوے (*s-V-²-e*) /so.e/ 'slept' (perf. pl. m. of *sonā*);

¹⁴⁴ For more examples see paragraphs 2.3.8.-2.3.17.

¹⁴⁵ Although they are preserved in writing by Hindustani/Urdu typographers to indicate careful spelling of some Arabic nouns, usually of central religious status (Shackle & Snell 1990: 49).

¹⁴⁶ In Arabic loanwords in Persian *hamzah* is retained after a consonant, but generally, in not very formal speech, it is realized as a prolongation of the vowel before a consonant, and as a glide between the vowels (Perry 2002: 230).

(ii) long vowel following a short vowel, e.g.: كُنِيَ ($k-Y^2-Y$) /kə.i/ 'many, several'; رِاسَت ($r-Y^2-A-s-t$) /rɪ.asət/ 'State'; لِي ($l-Y^2-e$) /lɪ.e/ 'because of';

(iii) short vowel following a long vowel, e.g.: تَاب ($t-A-Y^2-b$) /ta.ɪb/ 'penitetnt'; ذَاك ($z-A-Y^2-q-h$) /za.ɪqa/ 'taste'.¹⁴⁷

In its combinations with *vāv*, *hamzah* must take a *kursī* only when it is preceded by a letter that can join on the left (a connector), e.g.: شَرَبْ ($p-Y^2-V$) /pɪ.o/ 'drink!' (imperat.); جَوْن ($j-Y^2-V-ñ$) /dʒɪ.ũ/ | جَيُون ($j-Y-Y^2-V-ñ$) /dʒɪ.ũ/ 'I may live' (subj.); otherwise it is posted directly over the independent form of a grapheme *vāv*, e.g.: گَاوَدِي ($g-A-V^2-d-Y$) /ga.odi/ 'fool'; پِلَا ($p-l-A-V^2$) /pula.o/ 'pilaf'; رُوْف ($r-V^2-f$) /rə.uf/ 'forgiving'. If, in a sequence of vowels, *vāv* precedes the long /a/ (*alif*), the sign of *hamzah* is usually omitted or, very rarely, it occurs over *vāv*, cf.: مَوَا ($m-V^2-A$) /mu.a/ 'dead' (perf. of *mūnā*); سَوَال ($s-V^2-A-l$) /so.al/ | سَوَال ($s-V-A-l$) /səval/ 'question'; فَوَاد ($f-V-A-d$) /fo.ad/ 'heart'; مَوَازَن ($m-V-A-z-n-h$) /mo.azəna/ 'comparison'.¹⁴⁸

¹⁴⁷ To avoid a hiatus, *hamzah* in speech is often realized as a glide and also in writing its sign is sometimes replaced by a semivowel *y*, e.g.: سَرَبِي ($s-V-Y-e$) /soje/, لِي ($l-Y-e$) /lije/, رِاسَت ($r-Y-A-s-t$) /rɪjasət/, شَرَبِي ($p-Y-V$) /pijo/. Naim observes that "most people write these words the way they pronounce them" (Naim 1999 I: 38) and in contemporary texts this is the noticeably increasing tendency. Such shifts in spelling can lead, however, to the formation of homographs, as in the case of the postposition لِي ($l-V^2-e$) /lɪ.e/ 'because of', which when written with a glide is identical to the perfect participle form of the verb *lenā*: لِي ($l-Y-e$) /lije/, though in careful writing the common practice is to write them differently (ibid.). Cf. also 2.3.18.

¹⁴⁸ In colloquial speech this vowel cluster with *hamzah* is usually replaced by a sequence /va/, as in سَوَال ($s-V-A-l$) /səval/ 'question' (*vide supra*). Literary pronunciation, however, requires the utterance of a *hamzah*, even if it is not written.

2.2.20. The complete Hindustani/Urdu alphabet, with all the additional signs (including the aspirated consonants), consists of 53 items:

ا (A) آ (Ā) ب (b) بھ (b^h) پ (p) پھ (p^h) ت (t) تھ (t^h) ث (ṭ)
 ڈ (ḍ) ڈھ (ḍ^h) ج (j) جھ (j^h) چ (č) چھ (č^h) ح (ḥ) خ (x) د (d)
 دھ (d^h) ذ (ḏ) ذھ (ḏ^h) ز (z) زھ (z^h) ر (r) رھ (r^h) س (s) ش (š)
 ص (ṣ) ض (ḍ) ط (ṭ) طھ (ṭ^h) ع (ʿ) غ (ġ) ف (f) ق (q) ک (k) کھ (k^h) گ (g) گھ (g^h) ل (l) لھ (l^h) م (m) مھ (m^h)
 ن (n) نھ (n^h) و (V) ہ (h) ی (Y) ے (e) ء (ʾ)

2.2.21. Specific difficulties and inconsistencies. Variations in spelling have always been typical for Hindustani/Urdu, and the orthography of old texts in particular is characterized by instability.¹⁴⁹ Inconsistencies concern, i.a., the use of *kāf* and *gāf*, *yā-majhūl* and *yā-maʿrūf*, graphemes representing aspirates and retroflex sounds or the use of the dot for the letter *nūn*. As Beg observes, the way the words are spelled does not obey hard and fast rules, which, obviously, makes it difficult to determine the exact pronunciation of certain words (Beg 1988: 6). The probable reason for these inconsistencies might have been the parallel, two-fold process of phonological nativization of the Perso-Arabic borrowings in Hindustani, which – depending on various factors – has

¹⁴⁹ Cf. Saran 1969: 67-68: “During middle ages, when education was a luxury, scribes indulged into a practice of dispensing with not only vowel marks but the diacritical dots too which were so essential to distinguish the different consonantal characters. Hence the words ‘pitna’ and ‘bunna’ were written in the same way and were distinguished only through the context in which they occurred. In 19th century, with the introduction of litho printing and spread of education, this pedantic luxury was lessened and consonantal marks were clearly given in the written texts.” Cf. also Ahmad 2008: 1165.

taken the form of either adoption or adaptation.¹⁵⁰ Both the routines described by Hock, in the case of Hindustani/Urdu, can also affect the graphical notation of loanwords; at the same time, the way a loanword is spelled in the Perso-Arabic script may indicate, and even to some extent may influence, the type of nativization or its scale (cf. 3.2.2.).¹⁵¹ Moreover, the known shortcomings of the Perso-Arabic script – like the usage of several characters for one and the same phoneme, the usage of one character for displaying several phonemes, and the omission of short vowels – contribute significantly to the complexity of the problem.

2.2.22. A metaplasm, or a change in the spelling of a word made by the addition, omission, transposition (a change in the order of letters) or substitution of particular graphemes, occurs – in its various forms – in Hindustani/Urdu words, including the borrowed Perso-Arabic lexical items. These changes often reflect phonological alterations that have occurred in (mostly oral) loanwords; sound changes of this kind are discussed in detail in

¹⁵⁰ Hans H. Hock, who discusses the question of nativization (mainly at the phonological and lexical levels) at length in his writings (e.g. Hock 1991; Hock & Joseph 1996) explains the difference between adoption and adaptation as follows: “In the various nativization processes one can recognize two recurrent routines. One of these consists in the unmodified adoption of foreign words and their morphology and/or phonology. In such cases, nativization results only slowly and almost grudgingly. The other approach, with different sub-routines, can be more properly called nativization, in that it attempts to integrate borrowed words, their morphology, and their phonology into the structure of the borrowing language. We can refer to this approach as adaptation.” (Hock 1991: 408). Both the approaches, quite apparent in the Hindustani context, are discussed more extensively in the following chapters.

¹⁵¹ This phenomenon is quite common in the modern literate languages, where nativization often takes place through spelling (Hock 1991: 391; also *vide infra* 3.2.2.). In the case of Hindustani/Hindi, represented by a totally different script (Nagari), for obvious reasons we can speak exclusively about adaptation.

paragraphs 3.3.-3.6. Below, only a few examples are given to illustrate various graphical divergences which can be found in Perso-Arabic loanwords in Hindustani/Urdu.¹⁵²

(i) addition is a commonly observable process in Hindustani, though rarely are the changes reflected in the Perso-Arabic spelling of the loanwords, since most of them consist in adding an epenthetic short vowel; the spelling of loans in Perso-Arabic script may be influenced by the addition of a long vowel, e.g.: P باغچه (*b-A-ġ-ċ-h*) /baɣʃæ/ 'a little garden, a garden' ≥ H بَیْجِه (*b-ġ-Y-ċ-h*) /bəɣiʃa/ || بَغِي (*b-ġ-ċ-h*) /bəɣiʃa/ 'grove, orchard, kitchen garden'; P دبز (*d-b-z*) /dæbz/ 'thick, coarse' ≥ H دَبَز (*d-b-Y-z*) /dəbiz/ 't.s.'; there are also a few other examples, e.g.: P مست (*m-s-t*) /mæst/ 'drunk, intoxicated' ≥ H اَلْمَسْت (*A-l-m-s-t*) /əlmæst/ 't.s.'; P دارانی (*d-A-r-A-Y'-Y*) /dara'ni/ 'a kind of silk stuff or shawl' ≥ H دَرِیائی (*d-r-Y-A-Y'-Y*) /darja.i/ 't.s.';

(ii) omission is, to great extent, connected with the phenomenon of schwa-dropping and also includes frequent skipping of other short vowels which are not marked in the script; therefore, omission is manifested in Perso-Arabic script only in certain cases, e.g.: P سنگتره (*s-n-g-t-r-h*) /sængtæɾæ/ 'a kind of orange' ≥ H سَنَتَرَه (*s-n-t-r-h*) /səntərə/¹⁵³ || سَنَتَرَه (*s-n-g-t-r-h*) /sængtərə/

¹⁵² Sometimes – especially in the language of poetry – changes in word spelling (followed by changes in pronunciation) result from the author's intentional intervention in order to maintain a particular meter maintenance. In such cases the spelling is meant to inform the reader of the intended scansion. A popular maneuver of this kind is the optional use of *tašdid*, e.g.: شکر (*š-k-r*) /ʃəkər/ || شکر (*š-kk-r*) /ʃək:ər/ 'sugar' or the spontaneous lengthening or shortening of vowels, e.g.: آخِر (*Ā-x-Y-r*) /axir/ || آخِر (*Ā-x-r*) /axir/ || آخِر (*A-x-Y-r*) /əxir/ 'finally'. However, most of such alterations concerning Persian and Arabic words appeared already in Persian and have been adopted for the sake of Hindustani/Urdu metrics which is based largely on Persian poetic theory (more cf. Pritchett & Khaliq 1987).

¹⁵³ It is worth noting that the STCD (p. 386) indicates the contracted form as the more common (similarly McGr. 965, 963).

'tangerine'; P غرغرة (*ġ-r-ġ-r-ī*) /*ɣæɾɣæɾæ*/ 'gargling' ≥ H غراره (*ġ-r-A-r-h*) /*ɣæɾara*/ 't.s.'; P شبهة (*š-b-h-ī*) /*ʃubhæ*/ 'doubt' ≥ H شبه (*š-b-h*) /*ʃubh*/¹⁵⁴ 't.s.';

(iii) transposition (a change in the order of sounds/letters) is a common phenomenon but, because it concerns mainly short vowels, it is not noticeable in Perso-Arabic script;¹⁵⁵ here are rare examples of loanwords in which the order of graphemes has been visibly changed comparing to their Persian source words: P زیادة (*z-Y-A-d-ī*) /*zɪjadæ*/ 'augmentation, increase' ≥ H زیاده (*z-Y-A-d-h*) /*zɪjada*/ | اِزَادَة (*A-Y-z-A-d*) /*izad*/ 'addition (of)'; P لباب (*l-b-A-b*) /*lɒbab*/ '(pl. of *lubb*) quintessence' ≥ H لبوب (*l-b-V-b*) /*lɒbub*/ 'essences';

(iv) substitution is one of the most frequent orthographical changes occurring in Hindustani/Urdu, mainly reflecting sound changes,¹⁵⁶ e.g.: P پایچه (*p-A-Y-č-h*) /*pajtʃæ*/ 'breeches, drawers, or trousers' ≥ H پائچہ (*p-A-Y²-n-č-h*)¹⁵⁷ /*pa.ɪŋʃa*/ 'one of the legs of trousers'; P افیون (*A-f-Y-V-n*) /*æfjun*/ 'opium' ≥ H افيون (*A-f-Y-V-n*) /*æfjun*/ | انیم (*A-f-Y-m*) /*əfɪm*/¹⁵⁸ 't.s.'; P قصاب (*k-šš-A-b*) /*kæːsab*/ 'a butcher' ≥ H قصاب (*k-š-A-b*) /*kəsab*/ | قصابی (*k-š-A-Y²-Y*) /*kəsa.i*/ |

¹⁵⁴ This loanword is colloquially pronounced /*ʃubæ*/, analogically to nouns which end with the 'silent' *h*, even though the final *h* in this word is not a remnant of *tāʾ marbūʿah*.

¹⁵⁵ Contrary, of course, to Nagari, in which all the changes, including those which concern short vowels, are clearly visible.

¹⁵⁶ The most numerous groups of changes in this category are: replacement of *ʿain* by a vowel; replacement of a (short) vowel by another (short) vowel (see paragraph 3.3.).

¹⁵⁷ Cf. an older form quoted by Platts, which was written identically as its Persian source word, but pronounced with a hiatus: پائچہ (*p-A-Y²-č-h*) /*pa.ɛʃa*/, though still without nasalization, which is a later development (cf. also nasalized form *pāyamca* in McGr. 623).

¹⁵⁸ The adapted form (*afīm*) is indicated in SCTD as the prevalent one; this is confirmed also by derivatives such as *afīm-cī* or *afīmī* 'the opium-eater' (vs. Persian source words *afyūn-cī*, *afyūnī* 't.s.').

قَسَالِ (*k-s-A-Y'-Y*) /kəsa.i/ 't.s.';¹⁵⁹ P بطک (*b-t-k*) /bætæk/ 'a duckling' ≥ H بَطْخ (*b-t-x*) /bætəx/ 't.s.';¹⁶⁰ but often also involving graphemes of a similar shape,¹⁶¹ like د (*d*) | ذ (*z*); ک (*k*) | گ (*g*); ج (*ċ*) | ج (*j*), س (*s*) | ش (*š*), e.g.: P استاد (*A-s-t-A-d*) /ustad/ 'teacher, instructor' ≥ H استاد (*A-s-t-A-d*) /ustad/ | استاز (*A-s-t-A-z*) /ustaz/¹⁶² استاز (*A-s-t-A-z*) /ustaz/¹⁶³ 't.s.'; P کبگ (*k-b-g*) /kæbg/ 'a partridge' ≥ H کبک (*k-b-k*) /kæbk/ 't.s.'; P خرج (*x-r-j*) /xærj/ 'expense, cost' ≥ H خرچ (*x-r-ċ*) /xærġ/ 't.s.'; P کچکول (*k-ċ-k-V-l*) /kæŋkul/ 'a beggar's cup or bowl; a wallet' ≥ H کفکول (*k-š-k-V-l*) /kæŋkol/ | کجول (*k-j-k-V-l*) /kædʒkol/ | کچول (*k-ċ-k-V-l*) /kæŋkol/ 't.s.'; P طاسه (*t-A-s-h*) /tasæ/ 'a kind of drum' ≥ H طاسه (*t-A-s-h*) /tasa/ | تاش (*t-A-š-h*) /tafa/ 't.s.'; replacing the liquid consonant -y- (which functions in many Persian words as a glide placed between two vowels to avoid a hiatus) by a *hamzah* is another very common change of this type,

¹⁵⁹ Actually, there are three changes in this loanword: 1. degemination of -šš- (which is a phonological change, but noticeable also in script since the *taššid* is omitted over -š-); 2. possible replacement of *šād* by *sin* (-š- > -s-); 3. loss of the final consonant -b and addition of the vowel -ī, resulting in the occurrence of hiatus noted by a *hamzah*; this change has been probably caused by mixing two Persian words, viz. /qas:ab/ 'a butcher' and /qas:abi/ 'the trade of a butcher'; the intervocalic -b- in the latter has been weakened and subsequently lost, giving way to the hiatus.

¹⁶⁰ This is the only form of this loanword given by the STCD, but cf. *bat(t)ak*, *bat(t)akh* in McGr. 702.

¹⁶¹ This fact may suggest that in such cases the orthographical variances appeared not to depict previous phonological changes but as a result of careless notation of Persian words in Hindustani texts, especially in uncareful hand-writing (cf. Ahmad 2008: 1165).

¹⁶² This form, labelled in the STCD (p. 50) as a 'pedantic' variant of the prevailing form /ustad/, is confirmed also by the plural استاذ (*A-s-A-t-z-h*) /əsatɪza/ 'teachers' (ibid.), which is sometimes used instead of the more frequent, regular Hindustani plural (in this case the same as the singular form, i.e. /ustad/ 'teachers').

¹⁶³ This form can be found in Maddāh 86 as an alternative for /ustad/.

e.g.: P فرمایش (*f-r-m-A-Y-š*) /*færmajɪʃ*/ 'order, will, commands' ≥ H فرمایش (*f-r-m-A-Y'-š*) /*færmajɪʃ*/ 't.s.'; P آینده (*Ā-Y-n-d-h*) /*ajændæ*/ 'coming; next; future' ≥ H آینده (*Ā-Y'-n-d-h*) /*a.inda*/ 't.s.'; P بجای (*b-j-A-Y*) /*bædʒaj*/ 'in lieu' ≥ H بجا (*b-j-A-Y'-e*) /*bædʒa.e*/ 't.s.'

2.2.23. Looking at the examples given above, one can notice that many Persian words nativized in the Hindustani context display alternate spelling, retaining – apart from the orthographically adopted variants – also the exact forms they have in the source language (Persian, or even Arabic) and attesting them to be introduced to Hindustani also – if not entirely – through the medium of written texts.¹⁶⁴ A good example might be the compound loanword P فار غخطی (*f-A-r-g-x-t-Y*) /*fariɣxæt:i*/ 'a written release', which in Hindustani/Urdu is always spelled precisely in the same way, even though its common pronunciation differs significantly from the Persian source word – cf. H فار غخطی (*f-A-r-g-x-t-Y*) /*farxət:i*/ 'written acquaintance; deed of divorce' (STCD 448).¹⁶⁵

2.2.24. Hindustani has a large number of homophones. This phenomenon results from two facts: firstly – lexically it is a mixed language and its vocabulary has been derived freely from different sources (cf. Lehmann 1992: 133-135); secondly – in the Perso-Arabic script, it maintains the traditional spelling of words, even though the pronunciation of originally distinct sounds has been, in many cases, unified and differs vastly from the original.

¹⁶⁴ Hindustani, used for oral and written purposes, was the everyday North Indian language during the centuries of the Muslim rule, and especially during the rule of the Mughal Empire, but the official and administrative language was Persian. Its enormous influence on (especially written) Hindustani can also be seen in the conservative way the loanwords are spelled in the modified Perso-Arabic script.

¹⁶⁵ The variation in utterance is illustrated, however, in the Nagari notation of this loanword, cf.: फ़ारिग़-ख़ती (*fārig'hkhatī*), फ़ारिख़ती (*fārikh(a)tī*) फ़ारिख़ती (*fār'khātī*) (McGr. 681).

The latter reason results in the fact that Hindustani homophones, when written in Perso-Arabic script, are simultaneously heterographs.¹⁶⁶ Cf. for instance the following example of heterographical homophony:

- (i) اسیر (*A-s-Y-r*) /əsir/, m. 'prisoner, captive' (STCD 55)
- (ii) اشر ¹ (*A-š-Y-r*) /əsir/, adj. 'high, lofty' (STCD 34)
- (iii) اثر ² (*A-š-Y-r*) /əsir/, m. 'ether' (STCD 34)
- (iv) عسر (*ʿ-s-Y-r*) /əsir/, adj. 'difficult, poor' (STCD 434)
- (v) عصر (*ʿ-š-Y-r*) /əsir/, m. 'extract; juice pressed from grapes' (Maddāh 29)
- (vi) غش (*ʿ-š-Y-r*) /əsir/, m. 'dust' (Maddāh 29)

2.2.25. Another complication is caused by the lack of clear orthographical differences between words, which results in large number of homographs. E.g. a word بن (*b-n*) may be read as: (i) /bən/ (H) 'wood, forest'; (ii) /bun/ (P) 'root'; (iii) /bɪn/¹ (A) 'son (of)'; (iv) /bɪn/² (H) 'without, except'. Cf. also *supra* 2.2.12. and 2.2.13.

2.2.26. The Hindustani/Urdu set of numerals is based on those used in Arabic and Persian, but some numerals are unique in their shape:

	1	2	3	4	5	6	7	8	9
Arabic/Persian:	١	٢	٣	٤	٥	٦	٧	٨	٩
Hindustani/Urdu:	١	٢	٣	٤(٤)	٥(٥)	٦	٧(٧)	٨	٩

2.2.27. Last but not least, there is the problem of dividing a written text into semantically or grammatically distinct units. A natural quality of the Perso-Arabic script is its cursiveness – which means that the subsequent graphemes in the script join

¹⁶⁶ On the notation of homophones in Nagari script see fn. 219 on p. 113.

together into units (called ligatures) to form words. The Hindustani/Urdu script, and especially the *nasta'liq* style of writing, which is even more convoluted than *naxs*, does not have the concept of space to separate words.¹⁶⁷ Consequently, the readers are expected to parse the ligatures into words as they read along the text – but to do it successfully, they must know the language at least to some extent – otherwise, marking out the sensible units of speech borders on the impossible. What further complicates the whole thing is that the Perso-Arabic system of writing, while very well suited to indicate the process of root-modification characteristic for Semitic languages, does not cope well with long polysyllabic words resulting from the preference for word-formation through compounds which prevail in the Indo-European linguistic environment (Shackle & Snell 1990: 61). Thus, in Hindustani/Urdu orthography, a common practice is to break such complex forms (compounds as well as derivatives and inflected words) up into their constituent elements. The problem is that there are no established rules concerning which elements should be written separately and which should not.¹⁶⁸ Lack of consistency in this regard and flexibility in connecting or disconnecting individual components of words may cause big trouble for the reader.¹⁶⁹

¹⁶⁷ Contrary to most writing systems (including Latin) which, as a rule, concentrate mainly on representing words as graphic units (Bodrogligeti 1971: 22).

¹⁶⁸ E.g.: رہے گا (*r-h-e g-A*) /rəhe-ga/ vs. رہے گا (*r-h-Y-g-A*) /rəhega/ 'he will stay' (fut.); مل کر (*m-l k-r*) /mil-kər/ vs. مل کر (*m-l-k-r*) /milkər/ 'having met' (conjunctive part.); ان سے (*A-n s-e*) /un-se/ vs. ان سے (*A-n-s-e*) /unse/ 'with him/her' (pron. + ppn.); مکان دار (*m-k-A-n d-A-r*) /məkan-dar/ vs. مکان دار (*m-k-A-n-d-A-r*) /məkandar/ 'householder' (n. + suf.); بے دل (*b-e d-l*) /be dīl/ vs. بیدل (*b-Y-d-l*) /bedīl/ 'sad' (pref. + n.).

¹⁶⁹ Nearly a hundred years ago A. Yusuf Ali, in his short article, formulated a few rules concerning the conjunctive and disjunctive spelling of

2.3. Notation of Borrowed Words in Nagari

2.3.1. There exist a number of rules that have allowed to establish the notation of the borrowed words of Perso-Arabic origin in Nagari, but in many a case this notation does not make it possible to reconstruct the exact spelling of a given word in the original script. This is caused by the nature of the process of adopting these foreign items by Hindustani/Hindi – which was at first based on oral contact between the users of different tongues¹⁷⁰ and was primarily phonetic in its character. Perso-Arabic vocabulary began to seep into texts written in Nagari later on.¹⁷¹ The principally oral character of these loanwords in Hindustani/Hindi is testified also by the manner in which the Perso-Arabic loanwords are represented in Nagari – which has mostly (though not entirely – *vide infra*) pictured how a word is pronounced, not how it is spelled in the original script. Therefore, it would be difficult to address the question of writing adaptation

complex words in Hindustani/Urdu (cf. Ali 1920). Since then little has changed in this respect.

¹⁷⁰ On the contrary, both the process of assimilating Arabic lexis by Persian, and even the way in which Persian and Arabic elements were absorbed by Hindustani/Urdu, were to a much greater extent literary and were additionally supported by the use of the same writing system. Cf. e.g.: “[T]he emergence of New Persian, via partial relexification of *Dari* by Arabic, was essentially a literary process in which Arabic vocabulary was initially used in the writings of bilingual scholars and poets from 9th century on...” (Perry 2012: 81), and “this Persianization [of Hindustani – AKF] was carried to an extreme by the pliant Hindū Kāyasths and Khattrīs employed in the Muṣul administration and acquainted with Persian” (Grierson 1918: 51).

¹⁷¹ Kuczkiewicz-Fraś 2003b: 55-57. The earliest known literary text written in Nagari, which contains Persian and Arabic words is probably *Pr̥t̥vīrāj-rasau*, a heroic poem by Cand Baṛdāī, dated as 13th-14th c. (cf. Chatterji 1960: 195-196). McGregor remarked that an early form of Braj b^hāṣā is dominant in the language of the poem (McGregor 1984: 11).

of borrowings without at least some references to phonological and phonetic issues. Consideration of these matters, however, has been kept to a minimum in this chapter and relevant remarks appear only when necessary in order to understand the discussed changes and adaptations in spelling. The analysis of phonological and phonetic changes in the loanwords, with references to the material presented in this chapter, is given in 3.2.-3.5.

2.3.2. On the other hand, the extraordinarily strong correlation between the elements of the Nagari writing system and Hindustani speech sounds (Shapiro 1989: 20) has made it possible to notate, in a much more stable and exact way, the nuances of pronunciation of the previously adapted borrowed words, which often demonstrate inconsistencies or ambiguities resulting from the shortcomings of the Perso-Arabic script (e.g. notation of vowels).

2.3.3. Even though Persian and Hindustani/Hindi employ different writing systems, the phonological inventories of these closely related languages are largely congruent (Bahri 1960: 57; Tivārī 1969: 59). This was undoubtedly a favourable factor which fostered the process of adapting Perso-Arabic loanwords in terms of both the phonetic and graphical aspects. Most of the genuinely Persian consonants which are pronounced like their Hindustani counterparts can be easily represented by the Nagari consonantal signs. This applies to the following seventeen Persian phonemes:

/p/ پ → प	/b/ ب → व	/m/ م → म
/tʃ/ چ → च	/dʒ/ ج → ज	/ʃ/ ش → श
/k/ ک → क	/g/ گ → ग	/h/ ه, ح → ह
/r/ ر → र	/l/ ل → ल	/v/ و → व
/j/ ی → य	/n/ ن → न	/d/ د → द
/t/ ط, ت → त	/s/ ث, ص, س → स	

But this seemingly simple and clear system of correspondences conceals a number of details which require a closer look and additional explanation:

(i) the Persian phoneme /t/ (voiceless, slightly aspirated dental plosive) is represented by two heterographs in Perso-Arabic script, one of them (ط – *tā*) occurs rarely and only in Arabic words.¹⁷² In Nagari this graphical difference has been discarded and one sign, त (*ta*), stands for both Persian variants. Therefore it is not possible to infer the original spelling of a loanword in such cases;

(ii) the Persian phoneme /h/ (voiceless glottal fricative) is also represented by two graphical variants in Perso-Arabic script, one of which (ح – *hāṭī* *hoṭṭī*, in Hindustani also known as *barī he* ‘big *h*’¹⁷³) is restricted entirely to words of Arabic origin. In Nagari, both heterographs are represented by the same grapheme ह (*ha*);¹⁷⁴

(iii) the Persian phoneme /s/ (voiceless dento-alveolar fricative) has three heterographs in Perso-Arabic script. Two of them, namely ث (*ṣā*) and ص (*ṣād*), occur only in Arabic words,¹⁷⁵ sym-

¹⁷² It may happen, though rarely, that graphemes representing typically Arabic sounds occur also in genuinely Persian words as an effect of incidental Arabicization (Kłagisz 2010: 22).

¹⁷³ Contrary to ه – *hāṭī* *havvaz*, called *choṭī he* (‘small’ *h*) in Hindustani.

¹⁷⁴ Note that in Persian, in certain words, *hāṭī* *havvaz* (ه) in final position is not pronounced when preceded by the short vowel /æ/ which is not marked in writing. In such a case it is called the ‘silent’ *h* (Lambton 2000: xviii) and realized as [ɛ] (= like /e/ in English words ‘pet’, ‘met’). In Nagari the grapheme representing *h* is omitted in loanwords of this type and they are usually written accordingly to how they are pronounced: with long *ā* in pausa, e.g.: *khānā* (*khānā*) ‘house’ ≤ خانه (*x-A-n-h*) [xa:nɛ], and, in a few cases, with -e in pausa, e.g.: *cunāmce* ‘hence’ ≤ چنانچه (*č-n-A-n-č-h*) [ʃʊnā:ʃɛ] or with -i in pausa, e.g.: *ki* conj. ‘that’ ≤ که (*k-h*) [kʰɛ].

¹⁷⁵ With one exception – the grapheme *ṣād* appears in the spelling of the indigenous Persian numeral صد (*ṣ-d*) /sæd/ ‘hundred’ (Kłagisz 2010: 22).

bolizing genuinely Arabic sounds, the pronunciation of which has been unified with the prevalent sound /s/ indicated by the grapheme س (*sīn*). The differences between the graphical variants are not retained in Nagari, where one sign, स (*sa*), stands for all of them;

(iv) the Persian phoneme /ʃ/ (voiceless postalveolar fricative), which is uttered somewhere between the Hindustani voiceless postalveolar fricative /ʃ/ and the voiceless retroflex fricative /ʂ/ (indicated by Nagari signs श and ष respectively), is always represented by the grapheme श (*śa*) in Nagari (the sign ष (*ṣa*) being reserved exclusively for Sanskrit loanwords);

(v) the Persian phoneme /n/ (alveolar nasal) can be pronounced in Hindustani as a homorganic nasal, depending on the following consonant, and, in such cases, it is represented in Nagari by *anusvāra*,¹⁷⁶ e.g.: शतरंज (*śat'rañj*) [ʃətrəɳ ʈɔɳ] 'chess' ≤ P شترنج (*ṣ-t-r-n-j*) 't.s.'; फ़रंगी (*farāṅgī*) [fəraŋgi:] m., f. & adj. 'Frankish; European' ≤ P فرنگی (*f-r-n-g-Y*) [fræŋgi:] ~ [firiŋgi:]¹⁷⁷ 'French, Italian; an European Christian'; मुंतज़िर (*muntazir*) [montəzɪr] 'expecting' ≤ P [≤ A] منتظر (*m-n-t-z-r*) 't.s.';

(vi) the Persian phoneme /n/, in a few Persian loanwords in Hindustani, is pronounced as a nasalization of a preceding vowel,

¹⁷⁶ Theoretically these consonant clusters could be spelled according to standard rules concerning cluster notation in Nagari. Then, instead of शतरंज we would write शतरञ्ज, and फ़रंगी would be written फ़रङ्गी, the difference being purely graphical. Such cluster-type notation, even though absolutely correct, is used only sporadically for merely economic reasons – the adhibition of *anusvāra* enables to simplify the notation of a word and is therefore exceedingly common (cf. McGregor 1999: xxx; Montaut 2004: 20).

¹⁷⁷ The Persian alveolar nasal /n/ has a velar allophone [ŋ], which is positionally conditioned – it occurs before velar and uvular consonants: [k], [g], [x], [ɣ] and [ŋ], in words like *rang* 'colour'; both the allophones are represented by the same grapheme नून.

and is represented in Nagari by *candrabindu*,¹⁷⁸ e.g.: चूँकि (*cūmki*) [tʃū:ki] ‘since, because’ ≤ P چونکہ (*č-V-n-k-h*) [tʃu:ŋkiɛ]¹⁷⁹ ‘for now; seeing that’; हालाँकि (*hālāmki*) [ha:lā:ki] ‘although; however’ ≤ P حال آن کہ (*h-A-l Ā-n k-h*) [ha:la:ŋkiɛ] ‘the fact is that’; कँगूरा (*kaṁgūrā*) [kōgu:ra:] ‘cornice; parapet’ ≤ P کنگرہ (*k-n-g-r-h*) [kungurɛ] ‘t.s.’;

(vii) the Persian grapheme ن (*n*), in some Persian and Arabic loanwords, when followed immediately by ب (*b*) or پ (*p*), stands for the homorganic bilabial nasal /m/.¹⁸⁰ In these cases, in Nagari, it is always represented by *anusvāra* or (less commonly) by the letter म (*ma*), e.g.: अंबार, अम्बार (*ambār*) ‘heap, pile’ ≤ P انبار (*A-n-b-A-r*) [æmba:r]; दुंबाल, दुम्बाल (*dum'bāl*)¹⁸¹ ‘tail’ ≤ P دنباله (*d-n-b-A-*

¹⁷⁸ Note that in printed Hindustani/Hindi, the distinction between the nasalization of a vowel marked by *candrabindu* and the combination of a nasal consonant with an oral consonant have been largely neutralized in recent years and both are often represented by *anusvāra* (Montaut 2004: 18), though Kachru’s observation that “nasalization symbol / · / is no longer used in Hindi printing” (Kachru 2006: 39) is too categorical and may be contradicted by many recently printed publications. On the significant functional differences between *anusvāra* and *candrabindu* see Srivastava 1969: 915-916.

¹⁷⁹ The phonetic transcription of Persian source words is given here beside the usual transliteration to depict the adaptation of the Persian phoneme /n/ to the Hindustani sound system. The way in which the borrowed words are written mirrors the phonetic changes and only their understanding makes the system of Nagari notation fully comprehensible. The vocalization of Persian words indicates the early New Persian stage, not the contemporary pronunciation (this issue will be discussed in paragraph 3.1.3).

¹⁸⁰ Actually, I didn’t find a single loanword with the cluster *-np-*.

¹⁸¹ Steingass notates in Persian both the forms: دمبال (*d-m-b-A-l*) /dumbal/ ‘a tail; a man’s back’ (Stein. 535), and دنباله (*d-n-b-A-l-h*) /dumbale/ ‘the tail...’ (with many derivatives). In Hindustani/Urdu the form with *nūn* is prevalent, cf. دنبال (*d-n-b-A-l*), دنباله (*d-n-b-A-l-h*) (Platts 526, 527); دنباله (*d-n-b-A-l-h*) (STCD 313). In Nagari notation, if nasalization is not marked by *anusvāra*, the consonants *m* and *b* are written sepa-

l-h) [dumba:lɛ]; सुंबुल, सुम्बुल (*sumbul*) ‘hyacinth’ ≤ P [≤ A] سنبل
(s-n-b-l) [sumbul]; मिवर, मिबर (*mimbar*) ‘pulpit’ ≤ P [≤ A] منبر
(m-n-b-r) [mimbær];

(viii) the Persian graphemes و (*v*) and ی (*y*) may also be represented in Nagari by a number of various vowel signs (including diacritical), when they stand for a vowel or a diphthong (*vide infra*).

2.3.4. Perso-Arabic phonemes nonexistent in the IA phonological system are usually marked by a dot (called *nuqṭa*¹⁸²) added beneath (sometimes slightly shifted leftward) certain Nagari characters.¹⁸³ These are:

/q/ ق → क / क (<i>ka</i>)	/ɣ/ غ → ग / ग (<i>ga</i>)
/x/ خ → ख / ख (<i>kha</i>)	/z/ ز → ज / ज (<i>za</i>) or ज (<i>ja</i>)
/f/ ف → फ / फ (<i>fa</i>)	/z/ ظ, ض, ز, ذ → ज / ज (<i>za</i>)

Once again, this scheme of correlation between the graphemes and phonemes requires a few explanatory remarks:

rately (not as a consonant cluster), which indicates the possible morphological border between them. But this is probably again the effect of incorrectly breaking this word into segments understandable for the speakers, i.e. *dum* ‘tail’ and *bāl* ‘hair’.

¹⁸² From A نقطة (*n-q-ṭ-i*) ‘a dot; a diacritical point’ (Stein. 1420).

¹⁸³ The fact that Nagari notation of these phonemes consists merely in a slight modification of already existing graphemes has always caused a tendency towards uttering them in the same way as the phonemes represented by the primary, dotless forms of the signs, i.e. /k/ instead of /q/, /j/ instead of /z/ etc. – especially among those speakers who know only Nagari and among illiterate non-Muslims. This notice, however, is based on the author’s casual observations and is not confirmed by any more systematic research, though one can find similar remarks by other authors, e.g.: “Sounds characteristic for Persian, Arabic and English loans are pronounced [properly] mainly by [the speakers coming from] the learned urban social strata” (Varmā 1933: 98; cf. also Bahri 1960: 58).

(i) the Perso-Arabic phonemes /q/ and /ɣ/, marked by separate graphical signs (*qāf* (ق) and *ḡayn* (غ) respectively), in Persian merged into one phoneme with three positionally conditioned allophones: voiceless uvular plosive [q], voiced uvular plosive [g] and voiced uvular fricative [ɣ]. As an effect of this phonemic merger¹⁸⁴ a number of homophones appeared, differing in script by only one letter and having different meanings, but essentially the same pronunciation. In Hindustani, however, these borrowed phonemes do not reveal any trend towards merging with each other and are pronounced differently.¹⁸⁵ This phonemic situation is reflected in script – in Nagari the graphical differentiation between the two original phonemes has been preserved and they are marked by dotted क (*ka*) and dotted ग (*ga*) respectively. Cf.: बाकी (*bākī*) [ba:qi:] ~ [ba:ki:] ‘remaining’ ≤ P [≤ A] باقی (*b-A-q-Y*) /baqi/ ‘t.s.’; बागी (*bāgī*) [ba:ɣi:] ~ [ba:gi:] ‘rebellious’ ≤ P [≤ A] باغی (*b-A-ḡ-Y*) /baɣi/ ‘t.s.’;

(ii) the Persian phoneme /ʒ/, very rare in Persian, appears only in two loanwords adapted by Hindustani: अज्दहा (*aj’dahā*)¹⁸⁶ ‘a large snake’ ≤ P اژدها (*A-ž-d-h-A*) /æʒdæha/ ‘a dragon’ and पुरज़ा (*pur’zā*) ‘part, piece’ ≤ P پرژه (*p-r-ž-h*) /purʒæ/, پرزه (*p-r-z-h*) /purzæ/ ‘t.s.’;

¹⁸⁴ Hoenigswald 1960: 77ff. Cf. also Pisowicz 1985: 42ff.

¹⁸⁵ Shapiro 1989: 20; Shukla 2000: 37. More information on phonetic adaptation of these phonemes can be found in Chapter 3.2.

¹⁸⁶ Maddāh (p. 10) gives a form written in Nagari with the dotted letter ज and the ligature: अज्दहा (*azdahā*) but his notation can hardly be acknowledged as standard since his intention was rather to give the most exact and the most correlated (one is tempted to use a label ‘ideal’) transliteration of words in both Perso-Arabic and Nagari scripts. The form अज्दहा (*aj’dahā*) is confirmed, i.a., by McGr. 17 and the MHK I 49. Varmā, in his table of correlation between Nagari and Perso-Arabic characters, assigns Persian ج (ž) to the Nagari sign झ (*jʰ*), but does not support this correspondence with any example (Varmā 1933: 197).

(iii) the Persian phoneme /z/ (voiced dento-alveolar fricative) is represented by four heterographs in Perso-Arabic script. Two of them, viz. ض (*zād*) and ظ (*zā*) occur only in Arabic words, symbolizing genuinely Arabic sounds; another typically Arabic grapheme ذ (*zā*), which in Arabic represents the voiced dental fricative /ð/, appears, though rarely, also in originally Persian words. The pronunciation of all of them has been unified already in the Persian language with the prevalent sound /z/ indicated by the grapheme ز (*zā*). The differences between the graphical variants are not retained in Nagari, where one sign, ज़ (*za*), stands for all of them;

(iv) The use of the dots, even though quite common, is never obligatory and depends on the meticulousness of the writer. The dots are often omitted (or their use is inconsistent even within the same text) and the reader has to render the correct pronunciation of the given sound on the basis of his/her prior knowledge of the language (Shapiro 1989: 20).¹⁸⁷ The lack of a

¹⁸⁷ This growing tendency is clearly visible after even a cursory scan of descriptive texts and dictionaries of Hindustani/Hindi. For example, in agreement with the above-quoted remark, Shapiro, in his handbook, shows a great inconsistency in marking phonemes of Perso-Arabic origin, both in Nagari notation and in transliteration (e.g. on p. 91 the word ख़रिदे is transliterated *kharīde* in one place and *kharīde* in another, on p. 106 we find dotted ज़्यादा and dotless ज्यादा – both transliterated as *zyādā*, etc.). While older sources, like Maddāh (1959), MHK (1962) or Besk. (1972) notate the borrowed sounds with great precision, more recent studies do not pay much attention to this issue. In Awasthi & Awasthi (1981), Perso-Arabic sounds are still indicated, though many discrepancies also occur, as e.g. तकलीफ़ (*tak'lif*) written with and without a dot in the same entry ('trouble') or rather consistent avoidance of dotting the letter क (*k*), as in बाकी (*bākī*), करीब (*karīb*), किला (*kilā*) and many others (*passim*). In Sharma (2006), Perso-Arabic phonemes are rarely marked at all and when they are, it is done with a great lack of care (inconsistencies occur even within the same entry). Such heterogeneity in notation confirms that 'alien' sounds are in fact no

dot may also bring confusion in not uncommon cases when a dot is the only visible differentiator that allows to distinguish between words of different meaning, and often of different etymology (cf. Baig et al. 2012), e.g.:

फलक (<i>phalak</i>), m. 1. 'board; bench'; 2. 'layer, base'; 3. 'flat surface; blade', etc. (≤ S)	vs. फ़लक (<i>falak</i>), m. 'heaven; the sky' (≤ P [≤ A])
वाग (<i>bāg</i>), f. 1. 'rein'; 2. 'bridle', etc.	vs. वाग (<i>bāg</i>), m. 'large garden, park', etc. (≤ P)
वाजी ¹ (<i>bājī</i>), m. 'a horse' वाजी ² (<i>bājī</i>), f. 'elder sister' (≤ T)	vs. बाजी (<i>bāzī</i>), f. 1. 'game, play', etc. (≤ P)
ख़ान (<i>khān</i>), f. 1. 'a main, a quarry; source'; 2. 'a store; treasury', etc.	vs. ख़ान (<i>khān</i>), m. (≤ P) 1. 'lord, prince', etc.;
चाक ¹ (<i>cāk</i>), m. 1. 'a wheel; potter's wheel; grindstone; pulley', etc. चाक ² (<i>cāk</i>), m. & adj. 1. m. 'a tear; a slit (...)' ; 2. adj. 'torn' (≤ P) चाक ³ (<i>cāk</i>), m. 'chalk' (≤ E)	vs. चाक (<i>cāk</i>), adj. 1. 'trim'; 2. 'active, vigorous' (≤ P [≤ T])

2.3.5. Conjunct (ligatured) consonants used in the Nagari script to indicate pronunciation of consonant clusters are commonly employed also in Perso-Arabic loanwords.¹⁸⁸ All of the Nagari consonantal signs described above can function as first members of a consonant cluster (though some of them only in geminates) and, therefore, they can be represented by half of the character (cf. fn. 121 on p. 66). The rules regarding when to write and when not to write consonantal ligatures are not quite clear.

longer perceived as alien and have been almost fully assimilated by the Hindustani sound system (*vide infra* 3.2.).

¹⁸⁸ McGregor 1999: xxviii ff.; Shapiro 1989: 23 ff.

Generally, conjunct consonants are not written if the cluster¹⁸⁹ has emerged from the dropping of the inherent vowel /ə/ or as an effect of affixation, i.e. across morpheme boundaries (Masica 1991: 149). However, in the case of Perso-Arabic loanwords this principle is observed in a rather arbitrary manner and the statement that practically every sequence of consonants can be represented in both ways – either by a ligature or by separately written graphemes – would not be an exaggeration. The following example, chosen at random and in no respect exceptional, depicts and confirms the lack of clear rules regarding not only the notation of consonant clusters, but also the way in which the adapted ‘alien’ sounds are exhibited in the Nagari script:

P مغز (*m-g-z*) /mæɣz/ ‘brain’ in Nagari may be written as:¹⁹⁰

- (i) मगज (*mag'z*) – McGr. 780
- (ii) मगज़ (*magaz*) – Bahri 1989: 490
- (iii) मगज़ (*mag(a)z*) – Besk. II 324; Sharma 2006: 84; Awasthi & Awasthi 1981: 140
- (iv) मगज (*magaz*) – Caturvedi 1970: 567
- (v) मगज़ (*magz*) – Maddāh 465
- (vi) मगज [अ° मगज़] (*mag(a)j* [*a. magz*]) – MHK IV 780; SHŚ 770
- (vii) मगज (*magj*) /magja/ [mægʤ(ə)] – SHABDKOSH^{TM191}

¹⁸⁹ Here the term ‘cluster’ is understood in its broad sense, as a sequence of consonants before, after, or between vowels (Matthews 1997: 57).

¹⁹⁰ Transliterations in examples (i), (ii), (iv) and (vii) follow the original sources and show the presence or the absence of a consonant cluster; sources cited in examples (iii) and (vi) present the reader with neither transliteration nor any guidance on pronunciation and, therefore, it is not possible to deduce the proper utterance of a word.

¹⁹¹ An online translator for English and several Indian languages, that provides notation in original script, transliteration, simplified phonemic transcription and an audio sample reproducing the phonetic realization of a

2.3.6. In McGregor's material (which has been the main basis for the selection of Perso-Arabic loans in Hindustani presented in Kuczkiwicz-Fraś 2008) there are 450 Perso-Arabic loanwords (excluding derivatives which were not counted) containing ligatured consonants. Apart from ग (*ga*), every consonant occurs at least once as a first member of the consonantal sequence. Graphemes ह (*ha*) and य (*ya*) occur in this position only once each, cf.: बह (*bahr*) 'the sea' (McGr. 717) and तैयूर (*tagaiyyur*) 'becoming changed' (McGr. 434). Graphemes च (*ca*), द (*da*), ज (*ja*), न (*na*), य (*ya*), and व (*va*) appear exclusively in ligatures representing geminates, and the grapheme प (*pa*) – solely at the beginning of a few words where it precedes the semivowel *y*. Cf. respectively: बच्चा (*baccā*) 'child' (McGr. 698); मुकद्दर (*mukaddar*) 'fate' (McGr. 818); ताज्जुब (*tājjub*) 'surprise' (McGr. 447); सुन्नत (*sunnat*) 'rite' (McGr. 1027); तैयूर (*tagaiyyur*) *supra*; अव्वल (*avval*) 'first' (McGr. 65); प्याज़ (*pyāz*) 'onion' (McGr. 651). The most common consonantal conjunctions are: स्त (*sta*), श्त (*śta*) and the geminate ल्ल (*lla*). The table below presents qualitative and quantitative data gathered on the basis of McGregor's dictionary.

SIGN	POSITION OF A CLUSTER IN A WORD			GEMINATE	TOTAL
	FINAL	MEDIAL	INITIAL		
व (<i>b</i>)	-व (<i>-br</i>) ⁽³⁾ -व (<i>-br</i>) ⁽³⁾ -व (<i>-bz</i>)	-व (<i>-br</i>) ⁽²⁾ -व (<i>-bj</i>) -व (<i>-br</i>) -व (<i>-bs</i>)	व (<i>br</i> -)	-व (<i>-bb</i>) ⁽¹¹⁾	24
च (<i>c</i>)	--	--	--	-च (<i>-cc</i>) ⁽³⁾	3
द (<i>d</i>)	--	--	--	-द (<i>-dd</i>) ⁽¹⁸⁾	18
फ (<i>f</i>)	-फ (<i>-fi</i>) -फ (<i>-fz</i>)	-फ (<i>-fi</i>) ⁽⁷⁾ -फ (<i>-fl</i>)	--	-फ (<i>-ff</i>) ⁽³⁾	13

ह (h)	-ह (-hr)	--	--	--	1
ज (j)	--	--	--	-ज्ज- (-jj-) ⁽¹¹⁾	11
क (k)	-क्ल (-kl) ⁽²⁾ -क्स (-ks) ⁽²⁾ -क (-kr)	-क्र- (-kr-) -कश- (-ks-)	--	-क्क- (-kk-) ⁽¹¹⁾	18
ख (kh)	-खल (-khl) ⁽⁸⁾ -ख्सा (-khs) ⁽²⁾ -खस (-khs)	-खल- (-khl-) ⁽¹⁰⁾ -खव- (-khv-) ⁽²⁾ -ख्सा- (-khs-)	खव- (khv-) ⁽²⁾	--	26
क (k)	-क्फ (-kf) -क्त (-kt)	--	--	-क्क- (-kk-) ⁽¹⁴⁾	16
ल (l)	-ल्द (-ld) ⁽²⁾ -ल्क (-lk) ⁽²⁾ -ल्ब (-lb) -ल्फ (-lf) -ल्म (-lm)	-ल्ल- (-ll-) ⁽³⁾ -ल्फ- (-lf-) -ल्क- (-lk-) -ल्म- (-lm-)	--	-ल्ल- (-ll-) ⁽³²⁾	45
म (m)	-म्ल (-ml)	-म्ल- (-ml-) ⁽²⁾ -मद- (-md-) -म्र- (-mr-)	म्य- (my-)	-म्म- (-mm-) ⁽¹⁴⁾	20
न (n)	--	--	--	-न्न- (-nn-) ⁽¹³⁾	13
प (p)	--	--	प्य- (py-) ⁽³⁾	--	3
र (r)	-र्द (-rd) ⁽⁸⁾ -र्फ (-rf) ⁽⁵⁾ -र्ग (-rg) ⁽⁴⁾ -र्श (-rs) ⁽⁴⁾ -र्ज (-rz) ⁽⁴⁾ -र्ग (-rg) ⁽³⁾ -र्क (-rk) ⁽²⁾ -र्म (-rm) ⁽²⁾ -र्व (-rb) -र्व (-rc) -र्ज (-rj) -र्स (-rs) -र्त (-rt) -र्ख (-rkx)	-र्द- (-rd-) ⁽⁴⁾ -र्ज- (-rj-) ⁽²⁾ -र्ज- (-rz-) ⁽²⁾ -र्क- (-rk-) -र्ख- (-rkx-)	--	-र्- (-rr-) ⁽²²⁾	70

स (s)	-स्त (-st) ⁽²⁰⁾ -स्म (-sm) ⁽⁴⁾ -स्व (-sb) -सद (-sd) -सल (-sl) -सन (-sn)	-स्त- (-st-) ⁽³⁹⁾ -स्फ- (-sf-) ⁽²⁾ -स्म- (-sm-) ⁽²⁾ -स्क- (-sk-) -सल- (-sl-) -सन- (-sn-) -स्प- (-sp-) -स्र- (-skh-) -स्व- (-sv-)	स्य- (sy-) ⁽²⁾	-स्स- (-ss-) ⁽¹¹⁾	90
श (ś)	-श्त (-śt) ⁽¹⁰⁾ -श्क (-śk) ⁽⁶⁾ -श्द (-śd) -श्च (-śr)	-श्त- (-śt-) ⁽¹⁹⁾ -श्म (-śm-) ⁽³⁾ -श्व (-śv) ⁽²⁾	--	-श्श- (-śś-)	43
त (t)	-त्फ (-tf) -त्ल (-tl) -त्म (-tm) -त्र (-tr)	--	--	-त्त- (-tt-) ⁽¹⁰⁾	14
व (v)	--	--	--	-व्व- (-vv-) ⁽¹⁰⁾	1
य (y)	--	--	--	-य्य- (-yy-)	1
ज़ (z)	-ज़ब (-zb) -ज़म (-zm) -ज़र (-zr)	-ज़ब- (-zb-)	ज़्य- (-zy)	-ज़ज़- (-zz-) ⁽⁶⁾	11
					450

Table 1. Occurrence of conjunct consonants in Perso-Arabic loanwords in the *Oxford Hindi-English Dictionary* by R. S. McGregor

2.3.7. Since Perso-Arabic script does not possess separate graphemes to denote vowels, Nagari notation of vowels in Perso-Arabic loanwords is based, perforce, on how they were/are pronounced rather than how they are written in the source languages. Persian vowels and diphthongs,¹⁹² compatible to a large

¹⁹² Since in Arabic vocabulary adopted by Persian all Arabic vowels (short: /æ/, /i/, /u/ and long: /a/, /i/, /u/) and both diphthongs (/ai/, /au/) were equated in utterance with those of Persian (Perry 2002), we can safely

degree with the Hindustani sound system, are easily represented in Nagari by means of both independent characters denoting syllabic vowel forms and diacritic marks, which are combined with consonant characters and indicate intra-syllabic vowel forms.

2.3.8. Short *a* /ə/ in Hindustani can occur at the beginning and in the middle of a word.¹⁹³ In Nagari it is represented by the independent character अ (*a*) always in the initial position (i and ii) and, occasionally, in the middle of a word when it occurs as a second member of a vowel cluster, preceded by the inherent vowel /ə/ (iii). In Perso-Arabic loanwords it stands for Persian short /æ/ signalled by: *alif* (i), *‘ain* (ii and iii), diacritical *fathah* (iv) or not marked at all (v).

- (i) अहल (*ah'l*) /əhl/ ‘people’ ≤ P [≤ A] اهل (*A-h-l*) /æhl/ ‘t.s.’
- (ii) अहद (*ah'd*) /əhd/ ‘contract; reign’ ≤ P [≤ A] عهد (*‘-h-d*) /æhd/ ‘t.s.’
- (iii) तअज्जुब (*taajjub*)¹⁹⁴ /tə.ədʒ:ub/ ‘surprise’ ≤ P [≤ A] تعجب (*t-‘-j-b*)¹⁹⁵ /tæʔædʒ:ub/ ‘t.s.’
- (iv) कदम (*kad'm*) /qədm/ ‘foot’ ≤ P [≤ A] قَدَم (*q-d-m*) /qædæm/ ‘t.s.’
- (v) दफ्तर (*daftar*) /dæftər/ ‘office’ ≤ P [≤ A] دفتر (*d-f-t-r*) /dæftər/ ‘t.s.’

speak of a Persian (not Perso-Arabic) vowel inventory that appeared in Hindustani, so henceforth this label will be used in this context for all loanwords, including those which are genuinely Arabic.

¹⁹³ Historically, short /a/, realized phonetically as [ā] or [ə], can also occur: (i) at the end of some Persian loanwords and (ii) in monosyllables but it is represented only in the Perso-Arabic script, corresponding in orthography to the ‘silent’ *h* (e.g. (i) پَرَوَانَه (*p-r-v-A-n-h*) [pərva:nā] ‘moth’; (ii) نَه (*n-h*) [nə] ~ [nā] ‘no’) or to *‘ain* (e.g. شمع (*š-m-‘*) [ʃəmʔə] ‘candle’) respectively. In Nagari these vowels are marked as long (cf. 2.3.9.). More see Kelkar 1968: 47-48.

¹⁹⁴ Generally, *‘ain* is transcribed in Nagari as long *ā* (*vide infra*), and also, in the case of this word the more common form is ताज्जुब (*tājjub*) with long /a/ instead of a vowel cluster, though both forms may occur in texts.

¹⁹⁵ Note the omission of *tašdīd* over *j* to indicate a doubled consonant.

2.3.9. Long *ā* /a/ occurs in all three positions in a word. In Nagari it is represented by the independent grapheme आ (*ā*) at the beginning of a word (i, ii and viii), and also in a few sporadic cases when it occurs medially (vii) or finally (xi) as a second member of a vowel cluster. In the middle position (iii-vi) – also when it is the first member of a cluster (iv) – and at the end of a word (vii-xi) it is represented by the diacritical sign ɪ placed after a consonant. In the loanwords it mostly stands for Persian long /a/, and sometimes for Persian /æ/, which in the Perso-Arabic script may be marked by: *alif-madd* (i and viii), *alif* (iii and viii), *‘ain* (v and xi), *‘alif maqṣūrah* (vii and x), *hāṭ havvaz*¹⁹⁶ (ix), juxtaposition of *‘ain* and *alif* (ii, vi and vii), or juxtaposition of *‘ain* and *hāṭ havvaz* (xii):

- (i) आल (*āl*) /a/ ‘family’ ≤ P [≤ A] آل (*Ā-l*) /a/ ‘t.s.’
- (ii) आदत (*ādat*) /adat/ ‘habit’ ≤ P [≤ A] عادت (*‘-A-d-t*) /adæt/ ‘t.s.’
- (iii) बाग (*bāg*) /bag/ ‘garden’ ≤ P باغ (*b-A-g*) /baɣ/ ‘t.s.’
- (iv) गुंजाईश (*guñjāiś*) /gũdʒa.ɪʃ/ ‘space’ ≤ P گنجایش (*g-n-j-A-y-š*) /gundʒajɪʃ/ ‘t.s.’
- (v) लाल (*lāl*) /lal/ ‘ruby’ ≤ P لعل (*l-‘-l*) /lal/ ‘t.s.’
- (vi) माफ़ (*māf*) /maf/ ‘forgiven’ ≤ P [≤ A] معاف (*m-‘-A-f*) /muʔaf/ ‘t.s.’
- (vii) तआला (*taālā*) /tə.ala/ ‘most high (God)’ ≤ P [≤ A] تعالیٰ (*t-‘-A-l-ā*) /tæʔala/ ‘t.s.’
- (viii) आशना (*āsʾnā*) /afna/ ‘acquaintance’ ≤ P آشنا (*Ā-š-n-A*) /afna/ ‘t.s.’
- (ix) परदा (*par’dā*) /pərda/ ‘curtain’ ≤ P پرده (*p-r-d-h*) /pærdæ/ ‘t.s.’
- (x) मूसा (*mūsā*) /musā/ ‘Moses’ ≤ P [≤ A] موسى (*m-V-s-ā*) /musā/ ‘t.s.’¹⁹⁷
- (xi) ज़िला (*zilā*) /zila/ ‘district’ ≤ P [≤ A] ضلع (*z-l-‘*) /zila/ ‘t.s.’
- (xii) मतवूआ (*mat’būā*) /mætbu.a/ ‘printed’ ≤ P [≤ A] مطبوعة (*m-t-b-V-‘-ī*) /mætbuʔæ/ ‘t.s.’

¹⁹⁶ See also remarks on the ‘silent’ *h* in fn. 174 on p. 92 and in 3.3.1.

¹⁹⁷ *‘alif maqṣūrah*, which occurs in a few Perso-Arabic loanwords (mostly proper names), is usually marked in Nagari by the diacritic sign used to denote long /a/, and sometimes long /i/ as well (*vide infra* – *mānī*). See also remarks on adaptation of *‘alif maqṣūrah* in 2.1.5. (iii).

2.3.10. Short *i* /i/ in Hindustani appears mostly at the beginning and in the middle of a word, and finally only in the monosyllabic conjunction *ki* and its combinations (vi)¹⁹⁸. In Nagari, in the initial position (i and ii) and medially, as a second member of a vowel cluster (iii) it is represented by the independent character इ (i) and in the middle of a word by the diacritical sign ि posted before a consonant (iv and v). In Perso-Arabic loanwords it corresponds to Persian short *i* /i/ which may be marked by: *alif* (i), *‘ain* (ii and iv), diacritical *kasrah* (v), *hāʾi havvaz* (vii) or may not be indicated at all (iii and vi):

- (i) इस्म (*ism*) /ism/ ‘name’ ≤ P [≤ A] اسم (*A-s-m*) /ism/ ‘t.s.’
- (ii) इल्म (*ilm*) /ilm/ ‘knowledge’ ≤ P [≤ A] علم (*‘-l-m*) /ilm/ ‘t.s.’
- (iii) नुमाइश (*numāiś*) /numa.iʃ/ ‘show’ ≤ P نمایش (*n-m-A-y-š*) /numaj.iʃ/ ‘t.s.’¹⁹⁹
- (iv) साइत (*sāit*) /sa.it/ ‘hour’ ≤ P [≤ A] ساعت (*s-A-‘-t*) /saʔæt/ ‘t.s.’
- (v) जिक्र (*zikr*) /zikr/ ‘mention’ ≤ P [≤ A] ذکر (*z-k-r*) /zikr/ ‘t.s.’
- (vi) करिश्मा (*kariśmā*) /kəriʃma/ ‘wink’ ≤ P [≤ T] کرشمه (*k-r-š-m-h*) /kɪrɪʃmæ/²⁰⁰ ‘t.s.’
- (vii) कि (*ki*) /ki/ conj. ‘that’ ≤ P که (*k-h*) /kʰæ/ ‘t.s.’

¹⁹⁸ It can appear finally in some Sanskrit loanwords of more than one syllable (as लिपि (*lipi*) ‘script’, cf. Kelkar 1968: 46), but these cases are of no consequence to our considerations. The Persian enclitic ‘of’ (*izāfar*), sometimes realized phonetically as short *i* /i/ (e.g. कोहिनूर *koh-i nūr*), mentioned by Kelkar (ibid.), never occurs in absolute pausa, being followed by the second member of the juncture, and therefore cannot be included here.

¹⁹⁹ In Persian pronunciation of similar words the glide -y- often occurs to avoid a hiatus and to facilitate the utterance of two juxtaposed vowels. It is depicted in Persian notation of these words (cf. Stein. 1425; Hayyim II 1111). However, in Hindustani, this glide is not marked, in either Nagari (*vide supra*), or in the Perso-Arabic notation, cf. نمایش (*n-m-A-Aʾ-š*) /numa.iʃ/ ‘show’ (STCD 649). It may be a confirmation of the oral character of this and similar loanwords.

²⁰⁰ On qualitative vowel changes see Chapter 3.3.

2.3.11. Long $\bar{i}$ /i/ can occur in initial, medial and final position in a word. In Nagari it is symbolized by the independent grapheme ई ($\bar{i}$) at the beginning of a word (i and ii), and when it functions as a second member of a vowel cluster both medially (iii) and in the final position (iv). After a consonant, in the middle position (v) and at the end of a word (vi-viii) it is represented by the diacritical sign $\bar{i}$. In the loanwords it stands for Persian long $\bar{i}$ /i/ indicated in the Perso-Arabic writing system by: a juxtaposition of *alif* and the grapheme $y\bar{a}$ (i), a juxtaposition of ^cain and the grapheme $y\bar{a}$ (ii and iv), the grapheme $y\bar{a}$ (iii-v and vii), ^cain (vi) or, rarely, $^3alif\ maq\bar{s}urah$ (viii):

- (i) ईमान ($\bar{i}m\bar{a}n$) /iman/ 'belief' $\leq P [\leq A]$ ایمان ($A-m-A-n$) /iman/ 't.s.'
- (ii) ईद ($\bar{i}d$) /id/ 'festival' $\leq P [\leq A]$ عید (^c-Y-d) /id/ 't.s.'
- (iii) रईस ($ra\bar{i}s$) /rə.is/ 'nobleman' $\leq P [\leq A]$ رئیس ($r-Y^2-Y-s$) /ræ̃is/ 't.s.'
- (iv) बिदाई ($bid\bar{a}i$) /bida.i/ 'dismissal' $\leq P$ وداعی ($V-d-A-^c-Y$) /vɪda'i/ 'valedictory'
- (v) तीर ($t\bar{i}r$) /tir/ 'arrow' $\leq P$ تیر ($t-Y-r$) /tir/ 't.s.'
- (vi) जमीयत ($jam\bar{i}yat$) /dʒəmiyat/ 'association' $\leq P [\leq A]$ جمعیت ($j-m-^c-Y-y-t$) /dʒæm̃iʃat/ 't.s.'
- (vii) बाज़ी ($b\bar{a}z\bar{i}$) /bazi/ 'game' $\leq P [\leq A]$ بازی ($b-A-z-Y$) /bazi/ 't.s.'
- (viii) मानी ($m\bar{a}n\bar{i}$) /mani/ 'meaning(s)' $\leq P [\leq A]$ معنی ($m-^c-n-ā$) /mani/ 't.s.'²⁰¹

2.3.12. Short u /u/ in Hindustani, with a few exceptions,²⁰² may stand only at the beginning and in the middle of a word. In the initial position (i and ii) and medially (rarely), as a second member of a vowel cluster (iii), it is represented in Nagari by the

²⁰¹ Note that *'alif maqṣūrah* is not always written with a small *alif* sign (*alif-e kūṭāh*) above (ئ) and in such cases telling it apart from Persian ending *-ī* may be problematic. In Persian, the *'alif maqṣūrah* is therefore often written as *alif* (Phillot 1919: 20).

²⁰² Similarly to short *i* (cf. fn. 198), short *u* /u/ may appear finally in some Sanskrit loanwords of more than one syllable, e.g.: बिंदु (*bindu*) 'point' (cf. Kelkar 1968: 46), which are not important for this analysis.

independent character उ (*u*); in the middle of a word it is marked by the diacritical sign ُ placed under the consonant (iv), apart from the specific ligatured combination with the grapheme representing *r* (v). In Perso-Arabic loanwords it represents Persian short *u* /*u*/ which may be indicated by: *alif* (i and iii), *‘ain* (ii), diacritical *ẓammah* (iv) or is not marked at all (v):

- (i) उलफ़त (*ul’fat*) /*ulʃat*/ ‘affection’ ≤ P [≤ A] الفت (*A-l-f-t*) /*ulʃæt*/ ‘t.s.’
- (ii) उम्र (*umr*) /*umr*/ ‘life; age’ ≤ P [≤ A] عمر (*‘-m-r*) /*umr*/ ‘t.s.’
- (iii) नाउम्मीद (*nāummīd*) /*na.ʊm:ɪd*/ ‘desperate’ ≤ P نامید (*n-A-A-m-Y-d*) /*na.ʊm(:)ɪd*/ ‘t.s.’
- (iv) चावुक (*cābuk*) /*tʃabuk*/ ‘whip’ ≤ P چابک (*č-A-b-k*) /*tʃabuk*/ ‘t.s.’
- (v) रुख (*rukḥ*) /*ruχ*/ ‘face’ ≤ P رخ (*r-x*) /*ruχ*/ ‘t.s.’

2.3.13. Long *ū* /*u*/ can occur in all three word positions, though the initial position is rare. In Nagari it is symbolized by the independent grapheme ऊ (*ū*) at the beginning of a word (i and ii),²⁰³ and when standing as the second member of a vowel cluster, both medially (iii and iv) and in the final position.²⁰⁴ In the middle position (v and vi) and at the end of a word (vi and vii) it is represented by the diacritical sign ू placed beneath the adequate consonant, with the exception of the ligatured combination with the grapheme representing *r* (vi). In Perso-Arabic loanwords it corresponds to Persian long *ū* /*u*/ indicated sporadically by the juxtaposition of *alif* and the grapheme *vāv* (i), very rarely by the

²⁰³ In McGregor’s dictionary there is actually not a single Perso-Arabic loanword beginning with *ū*. Platts quotes a few examples (mainly from Arabic) but with no Nagari notation. They can be found also in Maddāh, where the Nagari form is given as well. The examples (i) and (ii) cited above are based solely on Maddāh’s dictionary (p. 87).

²⁰⁴ Such clusters appear very rarely and only medially in Perso-Arabic loanwords. However, they are common in the subjunctive forms (1. pers. sg.) of the regular verbs formed from the borrowed nominal parts of speech (Kuczkiewicz-Fraś 2003b: 82, 125-129), e.g.: शर्माऊँ (*śarmāūm*) /*ʃərma.ū*/ ‘I may be ashamed’ (≤ P *šarm* ‘bashfulness’). In all these cases *ū* is nasalized.

juxtaposition of ^ʿ*ain* and the grapheme *vāv* (ii and iv), or by the grapheme *vāv* (v-vii):

- (i) ऊँकियः (*ūḱiya:*)²⁰⁵ /uqijæt/ ‘ounce’ ≤ P [≤ A] اوقية (*A-V-q-Y-ī*) /uqijæt/ ‘t.s.’
- (ii) ऊद (*ūd*) /ud/ ‘wood, timber’ ≤ P [≤ A] عود (^ʿ*-V-d*) /ud/ ‘t.s.’
- (iii) ताऊस (*tāūs*) /ta.us/ ‘peacock’ ≤ P [≤ A] طأوس (*t-A^ʿ-V-s*) /taʷus/ ‘t.s.’
- (iv) ताऊन (*tāūn*) /ta.un/ ‘plague’ ≤ P [≤ A] طاعون (*t-A^ʿ-V-n*) /taʷun/ ‘t.s.’
- (v) कबूतर (*kabūtar*) /kəbutər/ ‘pigeon’ ≤ P كبوتر (*k-b-V-t-r*) /kəbutər/ ‘t.s.’
- (vi) रूबरू (*rūbārū*) /rubaru/ ‘face to face’ ≤ P روبارو (*r-V-b-A-r-A*) /rubaru/ ‘t.s.’
- (vii) पहलू (*pah’lū*) /pəhlu/ ‘side’ ≤ P پہلو (*p-h-l-V*) /pəhlu/ ‘t.s.’

2.3.14. The vowel *e* /e/ can occur in Hindustani in all three word positions. In Nagari it is represented by the independent grapheme ए (*e*) at the beginning of a word (i, ii and iii), and medially (very rarely) when it is the second member of a vowel cluster (iv). In the middle position (v-vii) and at the end of a word (viii) it is represented by the diacritical sign ̣ written above the consonantal sign. In Perso-Arabic loanwords it stands for the Persian vowel /e/ (and sometimes also for Persian vowels /i/ or /æ/), which may be indicated by: *alif* (i), ^ʿ*ain* (ii and vi), the juxtaposition of *alif* and ^ʿ*ain* (iii), the grapheme *yā* (v), *hāʾī havvaz* (iv and xi), diacritical *kasrah* (vii) or may not be indicated at all (viii):

- (i) एलची (*el’cī*) /elʃi/ ‘envoy’ ≤ P [≤ T] الچی (*A-l-č-Y*) /ilʃi/ ‘t.s.’
- (ii) एवजी (*evazī*) /evəzi/ ‘substitute’ ≤ P [≤ A] عوضی (^ʿ*-V-ž-Y*) /ivæzi/ ‘t.s.’

²⁰⁵ In Nagari notation Maddāh uses a colon to mark Arabic *tāʾ mar-būṭah*. Also Steingass and Hayyim maintain the original Arabic notation of *tāʾ marbūṭah*. Note that the latter presents only the form with the initial consonant – وقية (*V-q-Y-ī*) /væqijə/ ‘an ounce’ (Hayyim II 1175).

- (iii) एतकाद (*et'kād*) /etqad/ 'confidence' ≤ P [≤ A] اعتقاد (*A-^c-t-q-A-d*) /i²⁰⁶tiqad/ 't.s.'
- (iv) किराएदार (*kirāēdār*) /kīra.edar/ 'tenant' ≤ P کرایہ دار (*k-r-A-y-h-d-A-r*) /kīrajædar/ 't.s.'²⁰⁷
- (v) फ़रेब (*fareb*) /færeb/ 'fraud' ≤ P فریب (*f-r-Y-b*) /firib/ 't.s.'
- (vi) इस्तेमाल (*istemāl*) /īstema/ 'use' ≤ P [≤ A] استعمال (*A-s-t-^c-m-A-l*) /īstri²⁰⁸mal/ 't.s.'
- (vii) सेहत (*sehat*) /sehət/ 'health' ≤ P [≤ A] صحت (*ṣ-h-t*)²⁰⁸ /sih:æt/ 't.s.'
- (viii) रेहन (*reh'n*) /rehn/ 'pledge' ≤ P رهن (*r-h-n*) /ræhn/ 't.s.'
- (ix) चुनाँचे (*cunānce*) /ʃunāŋʃe/ 'therefore' ≤ P چنانچه (*č-n-A-n-č-h*) /ʃunantʃæ/ 't.s.'

2.3.15. The vowel *o* /o/ in Hindustani can take initial, medial and final position, but occurs relatively often only in the middle of a word. In Nagari, the independent grapheme ओ (*o*) represents this vowel at the beginning of a word²⁰⁹ (i and ii). In the medial (iii-iv) and final (v) position it is marked by the diacritical sign े written after a consonant (the latter occurrence being extremely rare).²¹⁰ It depicts the Persian vowel /o/ (and in some cases also the Persian vowel /u/), represented most often by the grapheme

²⁰⁶ The vocalization of this and a number of similar loanwords may suggest that they appeared in Hindustani relatively late (after the 15th c.), i.e. after the modern Persian vocalization became well established (Pisowicz 1985: 95). Cf. also: *e^cteghad* (Hayyim I 108), *e^ctiqād* (STCD 60).

²⁰⁷ Cf. alternative spelling in Hindustani/Urdu: کرایے دار (*k-r-A-^e-d-A-r*) /kīra.edar/ vs. کرایہ دار (*k-r-A-y-h-d-A-r*) /kīrajædar/ (STCD 485).

²⁰⁸ Note the omission of *tašdīd* over *h* to indicate a doubled consonant.

²⁰⁹ Similarly to long *ū* /u/, the vowel /o/ in Perso-Arabic loanwords also occurs extremely rarely at the beginning of a word. McGregor quotes only one such example (i), Maddāh (p. 91), on the other hand, gives a few heterophonic variants of words otherwise commonly written and pronounced with *u* /u/, like /ostad/ ~ /ustad/ 'teacher' (ii).

²¹⁰ The vowel /o/ occurs in Persian in the final position only in rare cases and at the beginning of a word solely in a few loanwords (Kłagisz 2010: 33, 36).

vāv (iii and v), *‘ain* (i and iv), and occasionally by the juxtaposition of *alif* and the grapheme *vāv* (ii):

- (i) ओहदा (*oh'dā*) /ohda/ 'duty' ≤ P [≤ A] عهدہ (*‘-h-d-h*) /uhdæ/ 't.s.'
- (ii) ओस्ताद (*ostād*) /ostad/ 'teacher' ≤ P اوستاد (*A-V-s-t-A-d*) /ustad/ 't.s.'²¹¹
- (iii) बोसा (*bosā*) /bosa/ 'kiss' ≤ P بوسه (*b-V-s-h*) /bosæ/ 't.s.'
- (iv) रोब (*rob*) /rob/ 'fear' ≤ P [≤ A] رعب (*r-‘-b*) /ruʔb/ 't.s.'
- (v) सरो (*saro*) /səro/ 'cypress' ≤ P [≤ A] سرو (*s-r-v*) /særv/ 't.s.'
- (vi) बानो (*bāno*) /bano/ 'lady' ≤ P بانو (*b-A-n-V*) /bano/ ~ /banu/ 't.s.'

2.3.16. The diphthong *ai* /aɪ/ can occur in Hindustani in all three word positions. In Nagari it is represented by the independent grapheme ऐ (*ai*) at the beginning of a word (i and ii). In the medial (iii-iv) and final (v) position it is represented by the diacritical sign ^ˆ written above the consonantal grapheme. In Perso-Arabic loanwords it corresponds to the Persian diphthong /aɪ/ (and in some cases to a cluster of the vowels /a/ and /i/), which is marked by: the juxtaposition of *alif* and the grapheme *yā* (i), the juxtaposition of *‘ain* and the grapheme *yā* (ii and iv), *‘ain* (v)²¹² or by the grapheme *yā* (iii and vi):

- (i) ऐवान (*aivān*) /aivan/ 'palace' ≤ P ايوان (*A-Y-V-A-n*) /aivan/ 't.s.'
- (ii) ऐब (*aib*) /aɪb/ 'fault' ≤ P [≤ A] عيب (*‘-Y-b*) /aɪb/ 't.s.'
- (iii) कैद (*kaid*) /qaid/ 'imprisonment' ≤ P [≤ A] قيد (*q-Y-d*) /qaid/ 't.s.'

²¹¹ Although the usual spelling of the Persian word استاد (*A-s-t-A-d*) /ustad/ contains a single *alif* at the beginning (pronounced [ʊ]), Steingass gives also the form with long /u/, written as in (ii) (Stein. 120). The same variant form is cited by Maddāh (p. 91) as an equivalent of the Nagari form ओस्ताद (*ostād*), contrary to the pair استاد – उस्ताद (*ustād*).

²¹² The Persian grapheme indicating *‘ain* corresponds only graphically to the sign representing the diphthong *ai* in Nagari. Phonetically, as in the Persian word مستعد (*m-s-t-‘-d*) /mustæʔid:/, there is a juxtaposition of *‘ain* (which can be realized either as a glottal stop or – more frequently – as the vowel /æ/) and short /ɪ/ which is not marked in the script.

- (iv) तैनात (*taināt*) /tainat/ 'appointed' ≤ P تعينات (*t-^c-Y-n-A-t*) /tæ[?]inat/ 't.s.'
- (v) मुस्तैद (*mus'taid*) /mustaid/ 'prompt' ≤ P [≤ A] مستعد (*m-s-t-^c-d*)²¹³ /mustæ[?]id:/ 't.s.'
- (vi) मै (*mai*) /mai/ 'wine' ≤ P می (*mai*) /mai/ 't.s.'

2.3.17. The diphthong *au* /au/ in Hindustani can stand at the beginning, in the middle or the end of a word, but in Perso-Arabic loanwords it occurs finally only in one example and as a variant of final *-av*.²¹⁴ At the beginning of a word (i and ii) it is represented in Nagari by the independent grapheme औ (*au*), while medially (iii-v) and finally (vi) it is marked by the diacritical sign ौ written after a consonant. In Perso-Arabic loanwords it depicts the Persian diphthong /au/ (and, occasionally, a cluster of the vowels /a/ and /u/), represented by: the juxtaposition of *alif* and the grapheme *vāv* (i), the juxtaposition of *ain* and the grapheme *vāv* (ii and v), or simply by the grapheme *vāv* (iii, vi and vii):

- (i) औरंग (*aurang*) /aurəŋg/ 'throne' ≤ P اورنگ (*A-V-r-n-g*) /auræŋg/ 't.s.'
- (ii) औरत (*aurat*) /aurət/ 'woman' ≤ P [≤ A] عورة (*ᶜ-V-r-ī*) /auræt/ 't.s.'
- (iii) मुस्तौफी (*mustaufi*) /mustaufi/ 'auditor' ≤ P [≤ A] مستوفی (*m-s-t-V-f-Y*) /mustaufi/ 't.s.'
- (iv) मौज़ा (*rauza*) /rauza/ 'mausoleum' ≤ P روضة (*r-V-z-ī*) /rauzaet/ 't.s.'
- (v) मौज़ (*nauz*) /nauz/ 'God, preserve us!' ≤ P نعوذ بالله (*n-ᶜ-V-z-b-A-l-l-h*) /næ[?]uzubillah/ 'we seek refuge in God'
- (vi) नौ- (*nav-*) /nav/ 'new' ≤ P نو (*n-V*) [næu] 't.s.'

²¹³ Note the omission of *tašdīd* over *d* to indicate doubled consonant.

²¹⁴ The Persian heterophonic adjective نو (*n-V*) /nav/ ~ /næv/ ~ /nu/ ~ /no/ 'new, fresh' in Nagari may be spelled either with *-av* at the end: नव (*nav*), when it functions as an independent word, or with the diphthong *-au*: नौ (*nav*), when it constitutes a part of compound, e.g.: नौबहार (*nav-bahār*) 'early spring' (McGr. 582); नौ-बनौ (*nav-ba-nav*) 'ever and again' (ibid.); नौजवान (*nav-javān*) 'youthful, young' (ibid.). Regardless of the difference in spelling, in both the cases the pronunciation remains diphthongal.

2.3.18. The glottal stop, marked in Perso-Arabic script by the letter ‘ain or by *hamzah*, which has full phonemic status in Arabic, is practically absent in Hindustani phonology (Shackle & Snell 1990: 48). In Nagari notation of the Perso-Arabic loanwords ‘ain is represented by various vowel signs (cf. *supra*, paragraphs 2.3.8.-2.3.17), *hamzah*, however, is either omitted entirely, e.g.: साईस (*sāīs*) ‘groom’ ≤ P [≤ A] سائس (*s-A-Yʷ-s*) /saʔīs/ ‘t.s.’; इब्तिदा (*ibtidā*) ‘beginning’ ≤ P [≤ A] ابتداء (*A-b-t-d-A-ʔ*) /ibtidaʔ/ ‘t.s.’, or is replaced by a glide -y-, e.g.: दायम (*dāyʷm*) ‘permanent’ ≤ P [≤ A] دائم (*d-A-Yʷ-m*) /daʔim/ ~ /dajm/ ‘t.s.’; फायदा (*fāyʷdā*) ‘advantage’ ≤ P [≤ A] فائدة (*f-A-Yʷ-d-ī*) /faʔidæ/ ~ /fajdæ/ ‘t.s.’

2.3.19. Although, as can be observed in the above paragraphs, the Nagari spelling of Perso-Arabic borrowings is characterized by significant assimilation of foreign phonemes to its own norms, some traditional notations are still preserved, e.g. ‘distorted’ (Platts 1909: 8) or ‘silent’ *vāv* in words like ख्वाव (*khvāb*) ‘dream’ or दरखास्त (*dar’khvāst*) ‘desire’.²¹⁵ The labialization of the velar fricative *x* (≥ H *kh*) in the utterance of these words was discontinued in Persian already in 13th-14th c. (Pisowicz 1985: 122), and also in Hindustani, regardless of the spelling, they are more often (though not always!) pronounced as [xa:b] and [dərxa:st] respectively (instead of the antiquated pronunciation [xʷa:b] and [dərʷa:st]).²¹⁶

²¹⁵ Such spelling is also retained in Perso-Arabic script, both in Persian and in Hindustani, cf.: خواب (*x-V-A-b*), در خواست (*d-r-x-V-A-s-t*). In Nagari (contrary to Perso-Arabic script), the alternative notation without -v- is not only possible, but sometimes also more common, cf. दरख्वस्त (*dar’khvāst*) | दरखास्त (*dar’khāst*) (McGr. 480) or खुश (*khūs*) ‘pleased’ (McGr. 242), which is the only possible form even though ≤ P خوش (*x-V-š*).

²¹⁶ Cf. Fergusson’s notations (1773): *khusamet* ‘flattery’ ≤ P خوشامد (*x-V-š-A-m-d*) /xʷʊʃamæd/ ‘t.s.’ (Stein. 485) or *khav* ‘sleep’ ≤ P خواب (*x-V-A-b*) /xʷab/ ‘t.s.’ (Stein. 478).

2.3.20. In comparison to the Perso-Arabic script, Nagari has a few important advantages for easier and clearer representation of the loanwords. One of the most important among them is the Nagari system of vowel notation, which allows to depict the vocalization of words with greater precision and – above all – in a definitely less complicated but at the same time more standardized way.²¹⁷ This concerns mainly distinguishing between and marking short vowels, which is particularly important from the point of view of those users, who are not very well acquainted with the language.²¹⁸ The Nagari system also desists from some other archaic features which were not continued in Hindustani pronunciation and, therefore, representing them in the script would be futile. It tends, on the one hand, to make the notation simple, and on the other hand – to graphically represent the as-

²¹⁷ Note that even if there are several forms of the same word in Persian (usually differing in the way they are vocalized), as a rule only one of them is represented in Nagari notation, e.g.: P گیس (*g-Y-s*) /gis/ ~ /ges/, گیسو (*g-Y-s-V*) /gesu/ ~ /gisu/ ‘a ringlet of hair, a forelock, a sidelock, curl’ (Stein. 1109) ≥ H गेसू *gesū* ‘t.s.’ (McGr. 275); P فروخت (*f-r-V-x-t*) /firuxt/ ~ /fəroxt/ ~ /furoxt/ ‘sale, selling’ (Stein. 923) ≥ H फरोख्त (*farokht*) ‘t.s.’ (McGr. 678); P (≤ A) زعم (*z-^c-m*) /zæ’m/ ~ /zi’m/ ~ /zu’m/ ‘pretending, assuming’ (Stein. 617) ≥ H ज़ेम (*zom*) ‘assertion; arrogance’ (McGr. 383). The opposite situation may also happen, though very seldom, when a Persian word of established pronunciation displays different spelling in Nagari, e.g.: P (≤ A) قَهَقَهه ا قَهَقَهه (*q-h-q-h-i / q-h-q-h-i*) /qæhqæhæ/ ‘laughing loudly’ (Stein. 956, 996) ≥ H कहकहा (*kah’kahā*), कहकह (*kah’kah*), कहाका (*kā-hākā*) ‘t.s.’ (McGr. 183, 184).

²¹⁸ Greater clearness in vowel notation often helps to prevent such disastrous mistakes, as for instance misreading a word زِکْر (*z-k-r*), which might be vocalized either /zɪkr/, meaning ‘mention (...) repeated invocation of God’s name’ (STCD 332) or /zəkər/, meaning ‘penis’ (ibid.). The same words in the Nagari notation will not be read incorrectly even by not a very well qualified speaker, cf.: ज़िक्र (*zīkr*) and ज़कार (*zakar*) respectively (MHK II 320, 364).

simulated elements in a way which is possibly closest to its own norms.²¹⁹ This includes, for instance, eliminating the depiction of such obsolete phonological phenomena as the glottal stop²²⁰ or the use of several characters for one and the same phoneme (irrespective of the fact that it has been merged from historically different phonemes²²¹). Generally, a more adequate orthography characteristic of the Perso-Arabic loanwords written in Nagari allows for a deeper insight into their actual phonemic structure than in the case of the same loanwords spelled with the use of Perso-Arabic characters.

2.3.21. The greater stabilization of Nagari results also in more consequent notation of these grammatical occurrences, which in the Perso-Arabic script may be signalled by – often omitted in writing – diacritics,²²² namely: (i) lack of vocalisation (marked in the Perso-Arabic script by *sukūn*), in Nagari often indicated by conjunct consonants cf. 2.3.6.; (ii) gemination

²¹⁹ Note that such simplifications may lead to the creation of graphically unified forms for originally heterographic homophones, e.g.: P ($\leq$ A) دفع (*d-f-ʿ*) /dæf/ ‘repelling, averting’ (Stein. 529) and P ($\leq$ A) دفعة (*d-f-ʿ-ī*) /dæfʰe/ ‘a moment’ (Stein. 529) in Nagari are represented by homographic forms: दफ़ा (*dafā*)¹ f. ‘time, occasion; moment’ (McGr. 476) and दफ़ा (*dafā*)² m. inv. ‘repelling, warding off’ (ibid.); similarly P ($\leq$ A) احتراز (*A-h-t-r-A-z*) /ɪhtiraz/ ~ /ehteraz/ ‘guarding against’ (Stein. 20) and P ($\leq$ A) اعتراض (*A-ʿ-t-r-A-z*) /ɪʔtiraz/ ~ /eʔteraz/ ‘opposition’ (Stein. 73) in Nagari are spelled in the same way: एतराज़ (*eʔrāz*)¹ m. ‘guarding against’ (McGr. 142) and एतराज़ (*eʔrāz*)² m. ‘opposition; objection’ (ibid.).

²²⁰ E.g. سوال (*savāl*) /səvāl/ ‘question’ instead of سأل (*s-V-A-l*) /so.al/, مألوم (*mālūm*) /malum/ ‘known’ instead of معلوم (*m-ʿ-l-V-m*) /məʔlum/ or ماف (*māf*) /maf/ ‘forgiven’ instead of ماف (*m-ʿ-A-f*) /muʔaf/; more examples cf. *supra*, 2.3.8.-2.3.17.

²²¹ E.g. जहर (*zahʾr*) ‘poison’, ज़ात (*zāt*) ‘nature’, ज़ब्त (*zabt*) ‘control’, instead of زهر (*z-h-r*) /zæhr/, ذات (*z-A-t*) /ðat/ and ضبط (*z-b-t*) /dʕæbt/ respectively.

²²² Cf. 2.1.2., fn. 104.

(marked in the Perso-Arabic script by *tašdīd*), in Nagari always indicated by the graphical doubling of a letter (ligature), cf. *ibid*; (iii) adverbialisation of substantives (marked in the Perso-Arabic script by *tanvīn*), in Nagari always indicated by adding a separate ending -न (-*an*) at the end of a word; cf. इरादतन (*irādatan*) ‘on purpose’ vs. اِرادَت (A-r-A-d-t-A) /*iradeta*/ ‘t.s.’.²²³ Ignoring these diacritic marks in texts²²⁴ may hinder the correct reading and cause erroneous pronunciation. However, the Nagari notation, in which these grammatical occurrences are in most cases noted clearly by the use of extra letters or ligatures, enables more correct utterance of the Perso-Arabic loanwords even in the case of a novice reader.²²⁵

²²³ This very example of *tanvīn*-omission has been found in the online Urdu-English Dictionary (<http://www.urduenglishdictionary.org/English-To-Urdu-Translation/on%20purpose/Page-1.htm>), but adverbial forms of this kind are frequently found in texts. An interesting fact is that both Maddāh and the STCD give the highly Arabicized form of this adverb, written with *tanvīn* over the *tāʾ marbūʿah* (not *alif*): اِرادَت (A-r-A-d-ī-), but neither Steingass nor Hayyim corroborates the adverb in this form.

²²⁴ Of these three diacritical marks *tašdīd* is used “fairly regularly” (Naim 1999 I: 37), but cf. also: 2.3.8. (iii), 2.3.14. (vii); 2.3.16. (v). The STCD for example marks *tašdīd* only in sporadic cases.

²²⁵ There have been, however, quite opposite opinions on this matter. De Tassy, for instance, in his annual report from 1874 quotes the judgement of Lieutenant-Colonel J. Chambers, Professor of Hindustani at the Oxford University, who expresses his definite preference for Persian script as far as the notation of Hindustani is concerned: “Le caractère dévanagari est, cela va sans dire, admirablement adapté au sanscrit et au pur (*thenth*) hindi; mais il est très-défectueux quand on l’emploie pour exprimer des mots étrangers dont bien des lettres sont intraduisibles dans cet alphabet. En résumé, je pense que l’alphabet persan est préférable pour l’usage général dans l’Inde...” (de Tassy 1875: 4). [In the English translation by S. Kamal Abdali: “The Devanagari script is, needless to say, admirably suited to Sanskrit and the pure (*thainth*) Hindi; but it is quite deficient when used to express foreign words whose letters do not correspond to the ones in this

2.4. Tables of Adequacy for the Nagari and Modified Persian Scripts

2.4.1. To synthesize all that has been written thus far about the correlation between Nagari and Perso-Arabic graphical signs used to represent the sounds of Hindustani, with special reference to the borrowed Perso-Arabic vocabulary, the data discussed and illustrated in the above paragraphs have been presented here in tabularized form, which successfully presents the mutual correspondence of the characters from both the writing systems in a simple schematic method.

2.4.2. Table 2a summarizes the information discussed in paragraphs 2.3.8.-2.3.17. and depicts which vowel signs of the Nagari script can replace individual Persian characters (in addition to the juxtapositions of these characters) used to mark vowels in the Persian source words of the Perso-Arabic loanwords in Hindustani.²²⁶ The position of Persian vowel signs in a word has been indicated by adding a hyphen. The reverse information, showing which graphemes of the Persian script can stand for individual Nagari vowels, has been presented in Table 2b.

script. In essence, I think that the Persian script is better for general use in India" (de Tassy 2011: 143)].

²²⁶ Since notation of vowels in Hindustani/Urdu follows, to a great extent, the original Persian spelling and all possible deviations are sparse, irregular and difficult to systematize, it has not been included here. Relevant information can be found in paragraphs 2.2.10-2.2.16.

PERSIAN SIGNS INDICATING VOWELS		NAGARI VOWEL GRAPHEMES	
-ا (alif)		अ (a), इ (i), उ (u), ए (e)	
-آ (alif-madd) / ل- (alif)		आ (ā)	
ع (ʿain)	-ع	अ (a), इ (i), उ (u), ए (e), ओ (o)	
	-عـ	अ (a), आ (ā), ओ (o), ऐ (ai)	
	عـ	आ (ā)	
ی (yā)	-ی	ई (ī), ए (e), ऐ (ai)	
	یـ	ई (ī), ऐ (ai)	
-و, و (vāv)		ऊ (ū), ओ (o), औ (au)	
ا (alif) +	ی (yā) ⇒ -ای	ई (ī), ए (e), ऐ (ai)	
	و (vāv) ⇒ -او	ऊ (ū), ओ (o), औ (au)	
	ع (ʿain) ⇒ -اع	ए (e)	
ع (ʿain) +	ا (alif) ⇒ -عا / -عا-		आ (ā)
	ی (yā)	-عی	ई (ī), ऐ (ai)
		-عی-	ऐ (ai)
		عی	ई (ī)
	و (vāv)	-عو	औ (au)
		-عو-	ऊ (ū), औ (au)
ی- / ی (ʿalif maqṣūrah)		आ (ā), ई (ī)	
ه- / هـ (hāʾi havvaz)		आ (ā), इ (i), ए (e)	

Table 2a. Persian signs indicating vowels²²⁷ and corresponding Nagari graphemes

²²⁷ One must not forget that in Persian neither ‘ain nor alif in the initial position are proper vowel signs. These graphemes only serve as a signal that there is an unmarked (but pronounced) vowel following them.

NAGARI VOWEL GRAPHEMES	PERSIAN SIGNS INDICATING VOWELS
अ (a)	-ا, -ع, -ا
आ (ā)	-آ, -ا, -ا, -عا, -ها, -ا, ع, ه, ی, ی
इ (i)	-ا, -ع, -ه
ई (ī)	-ای, -عی, -ی, -ا, -ی, -ی, ی, ی
उ (u)	-ا, -ع
ऊ (ū)	-و, -و, -عو, -او
ए (e)	-ا, -ع, -ا, -ی, -ه,
ओ (o)	-و, -و, -او, -ا, -ع
ऐ (ai)	-ی, -ی, -ا, -عی, -ای, -ی
औ (au)	-و, -عو, -و, -او

Table 2b. Nagari vowel graphemes and corresponding Persian signs indicating vowels

2.4.3. The correlation between Persian consonant graphemes and their Hindustani counterparts (both in modified Perso-Arabic script and in Nagari), including the phonemic value of individual consonants has been demonstrated in Table 3. The synthesized data depicted in Table 3 have been discussed in detail in paragraphs 2.2.2.-2.2.8. and 2.3.3.-2.3.4.

PERSIAN CONSONANT SIGNS	MODIFIED PERSIAN SCRIPT USED FOR HINDUSTANI	NAGARI CONSONANT GRAPHEMES
ب (b) /b/	پ (b) /b/	व (ba) /bə/
	بھ (b ^h) /b ^h /	भ (b ^h a) /b ^h ə/
پ (p) /p/	پ (p) /p/	प (pa) /pə/
	پھ (p ^h) /p ^h /	फ (p ^h a) /p ^h ə/
ت (t) /t/	ت (t) /t/	त (ta) /tə/
	تھ (t ^h) /t ^h /	थ (t ^h a) /t ^h ə/
	ت (t) /t/	ट (ta) /tə/
	تھ (t ^h) /t ^h /	ठ (t ^h a) /t ^h ə/
ث (s) /s/	ث (s) /s/	स (sa) /sə/
ج (j) /dʒ/	ج (j) /dʒ/	ज (ja) /dʒə/
	جھ (j ^h) /dʒ ^h /	झ (j ^h a) /dʒ ^h ə/
چ (č) /tʃ/	چ (č) /tʃ/	च (ca) /tʃə/
	چھ (č ^h) /tʃ ^h /	छ (c ^h a) /tʃ ^h ə/
ح (h) /h/	ح (h) /h/	ह (ha) /hə/
خ (x) /x/	خ (x) /x/	ख (k ^h a) /xə/
د (d) /d/	د (d) /d/	द (da) /də/
	دھ (d ^h) /d ^h /	ध (d ^h a) /d ^h ə/
	د (d) /d/	ड (da) /də/
	دھ (d ^h) /d ^h /	ढ (d ^h a) /d ^h ə/
ذ (z) /z/	ذ (z) /z/	ज़ (za) /zə/
ر (r) /r/	ر (r) /r/	र (ra) /rə/

	ر (r) /r/	ड़ (ra) /ɽa/
	ڑ (r ^h) /ɽ ^h /	ढ़ (r ^h a) /ɽ ^h a/
ز (z) /z/	ز (z) /z/	ज़ (za) /zə/
ژ (ž) /ʒ/	ژ (ž) /z/ /j/	ज़ (za) /zə/ /jə/
س (s) /s/	س (s) /s/	स (sa) /sə/
ش (š) /ʃ/	ش (š) /ʃ/	श (śa) /ʃə/
ص (ṣ) /s/	ص (ṣ) /s/	स (sa) /sə/
ض (ẓ) /z/	ض (ẓ) /z/	ज़ (za) /zə/
ط (ṭ) /t/	ط (ṭ) /t/	ट (ta) /tə/
ظ (ẓ) /z/	ظ (ẓ) /z/	ज़ (za) /zə/
ع (ʿ) ^ø /ʔ/ _V	ع (ʿ) ^{/ʔ/} _{V²²⁸}	V ²²⁹
غ (ğ) /ɣ/ ²³⁰	غ (ğ) /ɣ/ /g/ /c/	ग (ga) /ɣə/ /gə/
ف (f) /f/	ف (f) /f/ /p ^h /	फ़ (fa) /fə/ /p ^h a/
ق (q) /q/ ²²¹	ق (q) /q/ /k/	क (ka) /qə/ /kə/
ك (k) /k/	ك (k) /k/	क (ka) /kə/
	ك (k ^h) /k ^h /	ख (k ^h a) /k ^h a/
گ (g) /g/	گ (g) /g/	ग (ga) /gə/

²²⁸ Cf. paragraph 2.2.16., fn. 143.

²²⁹ Cf. *supra*, Table 2a.

²³⁰ Cf. paragraph 2.3.4.(i).

	گ (g ^h) /g ^h /	घ (g ^h a) /g ^h ə/
ل (l) /l/	ل (l) /l/	ल (la) /lə/
	ل ^h (l ^h) /l ^h /	ल्ह / लह (l ^h a) /l ^h ə/
م (m) /m/	م (m) /m/	म (ma) /mə/
	म ^h (m ^h) /m ^h /	म्ह / मह (m ^h a) /m ^h ə/
ن (n) /n/ /m/	ن (n) /n/ /m/ ²³¹	न (na) /nə/ /mə/
		(n) /n/
		(ñ) /ɲ/
		(ṇ) /ɳ/
	ن (ṇ) /n/	ن̣ /Ṇ/ ²³²
		(n) /n/
		(ñ) /ɲ/
		(ṇ) /ɳ/
	ن̣ (n ^h) /n ^h /	न्ह / नह (n ^h a) /n ^h ə/
و (v) /v/	و (v) /v/	व (va) /və/
ه (h) /h/	ه (h) /h/	ह (ha) /hə/
ی (y) /j/	ی (y) /j/	य (ya) /jə/
ء (ʔ) /ʔ/	ء (ʔ) /ʔ/	- - -
	ی- (-y-) /j/	-य- (-y-) /j/ ²³³

Table 3. Persian and Hindustani consonant graphemes

²³¹ Cf. paragraph 2.3.3.(vii).²³² Cf. paragraph 2.3.3.(vi).²³³ Cf. paragraph 2.3.18.

3. Phonology and Phonetics

*The scholar is perhaps more readily misled than others,
for knowing the principal things that he ought to hear,
he easily persuades himself that he does hear them.*

(Bailey 1922: 539)

3.1. Phonological Systems of Persian and Hindustani: Similarities and Differences

3.1.1. Scholars of the Persian language are generally of the unanimous opinion that Persian²³⁴ in its modern (i.e. post-14th c.) form has six simple vowel phonemes/monophthongs (Lambton 2000: xi; Lazard 1992: 7; Pisowicz 1985: 10; Rohany Rahbar 2012: 133; Windfuhr 1997: 677), with a range of allophones (which are not discussed here):

- /i/ – front, high, long, approx. to the vowel in the English word ‘beat’, represented by *ﻯ* in the Arabic script
- /ε/ – front, mid, short, approx. to the vowel in the English word ‘bed’, not represented by a letter in the Arabic script

²³⁴ It must be remembered, especially while dealing with phonological issues, that the term ‘Persian’ does not pertain to one strictly defined linguistic entity, since there are, and there have been, dialectal differences within the Persian of Iran itself (e.g. northern Tehrani dialect vs. southern Shirazi variant) as well as “dialectal differences which – while common to Iranian Persian, separate it from Xorāsani-Tajiki and Afghani Persian, and from the learned Persian in India” (Windfuhr 1979: 131).

-
-
- /æ/ – front, low, short, approx. to the vowel in the English word ‘bad’, not represented by a letter in the Arabic script
 - /ɑ/²³⁵ – back, low, long, approx. to the vowel in the English word ‘park’, represented by ا in the Arabic script
 - /o/ – back, mid, short, approx. to the vowel in the English word ‘go’ (but slightly more rounded), not represented by a letter in the Arabic script
 - /u/ – back, high, long, approx. to the vowel in the English word ‘soon’, represented by و in the Arabic script

The underlying vowel inventory of Persian displays the distinction of length, contrasting the long vowels /i/, /u/, /ɑ/ with the short vowels /ε/, /o/, /æ/ respectively.²³⁶ From the point of view of contemporary Persian phonology, the opposition long-short is a redundant feature, questioned on the basis of experimental data

²³⁵ For practical reasons, in the phonological notation the Persian back /ɑ/ will be marked /a/ in this study. Such a denotation is more adequate for South Asian Persian, in which this vowel was never pronounced as far back as in Iranian Persian (a more frontal pronunciation has also been retained in the vocalisation of Persian loanwords in Hindustani). Cf. Phillot’s description: “In many cases the *alif* is pronounced broad; thus باد *bād* “wind” is pronounced nearly like *bāwd*, the *alif* being given the sound of *a* in the English word *fail*. Some Afghans too have this broad pronunciation, but the Indians always pronounce the *alif* like *a* in the English word *father*. In Persia, however, if this last pronunciation be used, the word will be taken to mean بعد *ba’d* “after”, and even an educated Persian will be puzzled by it” (Phillot 1919: 11).

²³⁶ Phillot, whose Persian grammar was based on the Indian Persian, claimed: “The original Persian vowel system was that of the Sanskrit, with the semi-vowel *rī* peculiar to the latter excluded” (Phillot 1919: 22). Beyond doubt this statement would be a far-reaching simplification in relation to the modern vocalic system of the literary Persian from Iran, but it is justified in relation to the Persian language in its Indian variant.

(Pisowicz 1985: 12) and neutralized in contemporary Iranian Persian in most environments (Windfuhr 1997: 679). Thus also the primordial dimensions of contrasts in the Persian vowel system are widely believed to be qualitative rather than quantitative, with quantity playing a secondary, non-contrastive role (thorough critical discussion of this question has been recently provided in Rohany Rahbar 2012). But there are arguments which justify the traditional division of Persian vowels into long and short, the first of them being of an orthographic nature, since such a division is reflected in the way Persian vowels are represented in Arabic script: genetically long vowels /a/, /i/ and /u/ are marked as long by the Arabic characters *alif*, *yā* and *vā* respectively, while the short ones are represented only sporadically by the relevant diacritics (q.v. 2.1.2. and fn. 104 p. 58). In prosody, as well, long vowels are counted as long and short – as short (cf. Elwell-Sutton 1976: 84-85). But the evidence provided by diachronic phonology is the most important, confirming that from classical to modern New Persian, historically short high vowels *ī and *ū lowered and transformed into /ɛ/ and /o/, and historically long vowels *ē and *ō raised and transformed into /i/ and /u/ respectively (e.g. Pisowicz 1985; Windfuhr 1997). The following schema presents these phonetic changes in a clearer and more detailed way:²³⁷

²³⁷ When one tries to define the sources of sounds in Persian loanwords in Hindustani, the knowledge of historical phonological changes that had occurred in Persian is essential for avoiding errors of the kind that appeared, for example, in Beg's interpretation. Here the New Persian forms, representing the later (modern) vocalic system of Persian have been indicated as sources of several Hindustani (Urdu) forms, which had been borrowed from Persian either before the described sound changes or irrespective of them and which retain the earlier (classical) Persian vocalization. Cf. Beg's anachronical examples: H/U *dil* < P *del* 'heart'; H/U *jigar* < P *jegar* 'liver' (Beg 1988: 35); H/U *betāb* < P *bītāb* 'impatient'; H/U *tez* < P *tīz*

modern short vowels	originated from	earlier short vowels
/ɛ/	←	*ĩ
/o/	←	*ũ
/æ/	←	*ă
modern long vowels	originated from	earlier long vowels
/i/	←	*ī, *ē
/u/	←	*ū, *ō
/a/	←	*ā

3.1.2. Hindustani, in contrast to Persian, natively possesses a symmetrical²³⁸ ten-vowel system, which is a continuation of the Old Indo-Aryan system.²³⁹ The vowels: /ə/, /ɪ/, /ʊ/ are always short in length, while the vowels: /a/, /i/, /u/, /e/, /o/, and the diphthongs /aɪ/ and /aʊ/ are always regarded as long. When we take into consideration the close vowels, what were thought to have been primarily quantitative distinctions (or distinctions of vowel length, i.e. [i] ~ [i:] and [u] ~ [u:]) in OIA, have become qualitative distinctions (or rather length accompanied by quality, i.e. [ɪ] ~ [i:] and [ʊ] ~ [u:]) in Hindustani, as in other New Indo-

‘sharp’ (ibid.: 36); H/U *pur* < P *por* ‘full’; H/U *xuṣbū* < P *xoṣbū* ‘fragrance’ (ibid.: 39); H/U *roz* < P *rūz* ‘day’; H/U *afsos* < P *afsūs* ‘regret’ (ibid.: 40) etc.

²³⁸ Viz. all front vowels have corresponding back vowels, all oral vowels have corresponding nasal vowels and all long vowels (except two monophthongized diphthongs) have corresponding short vowels (cf. Shukla 2000: 60).

²³⁹ Masica 1991: 110. Note that this number also comprises the diphthongs /aɪ/ and /aʊ/, which in most cases are monophthongized in pronunciation to [æ] and [ɔ], with various dialectal or acrolectal (i.e. closest to a standard language) realizations (cf. McGregor 1999: xvi; Montaut 2004: 21; Shukla 2000: 59). However, since analysing the phonetic differences and nuances of utterance within Hindustani itself is not the subject of this book, they will be conventionally treated as diphthongs.

-Aryan languages (Ohala 1983: 10; Masica 1991: 111). This rule generally pertains also to Persian loanwords in Hindustani. However, for the sake of clarity and typographical convenience, this study relies on the traditional quantitative contrast.

3.1.3. For the purpose of analyzing Persian loanwords in Hindustani, the diachronic symmetry of the classical Persian vocalic system, marked by a length distinction, is much more important than the modern inventory of Persian vowels and mutual dependencies between them in their synchronic dimension. The reason is simple: a great part of Persian loanwords in Hindustani manifest the Early New Persian stage of vocalisation, antecedent to the changes described and distinguishing the borrowed forms from their modern Iranian Persian (Farsi) counterparts.²⁴⁰ For centuries these loanwords have been functioning as the preservers of the old vocalisation.²⁴¹ Their phonological interpretation is often greatly facilitated and more verifiable due to the fact that they were also written down in Nagari, which allowed for noting vocalic elements more accurately. The minimal pairs,

²⁴⁰ Cf. "The *majhūl* sounds *o* and *e* are still preserved in the Persian spoken by Afghans and Indians, but they are now unknown in Persia: in modern Persian an 'ant' is called *mūr*, and there is nothing in pronunciation to distinguish the word for 'lion' from *shūr* 'milk'" (Phillot 1919: 22). On the same page Phillot cites the opinion of his predecessor, Friedrich Rosen, who stated that: "The Persian of India may be looked upon as a petrification of the old classical language. It has also preserved the *majhūl* vowels *e* and *o* for *ī* and *ū*, and many other differences of pronunciation. The Persian-speaking Indians, whose studies are mostly confined to the classics and poetic exercises, have followed none of the developments of the modern language" (Rosen 1898: XII-XIII). According to Phillot, the Persian of India is decidedly far purer, both in idiom and pronunciation, than the language of Persia, but "sounds pedantic, and is almost unintelligible to ordinary Persians" (ibid.). Also cf. Shackle & Snell 1990: 60.

²⁴¹ A similar phenomenon of retaining the old vocalic system can be observed in Persian spoken in Afghanistan (cf. Henderson 1975).

although few in number, confirm that the contrasted long vowels: /e/ : /i/ and /o/ : /u/ constitute separate phonemes in words borrowed from Persian:

/e/ : /i/

شیر (*šer*), شیر (š-Y-r) /šer/ 'lion' : شیر (šīr), شیر (š-Y-r) /šīr/ 'milk'

زیر (*zer*), زیر (z-Y-r) /zer/ 'under' : زیر (zīr)²⁴², زیر (z-Y-r) /zir/ *mus.* 'treble'

/o/ : /u/

گول (*gol*), گول (g-V-l) /gol/ 'crowd' : گول (*gūl*)²⁴³ گول (g-V-l) /gul/ 'demon'

رو (*ro*)²⁴⁴, رو (r-V) /ro/ suf. 'growing' : رو (*rū*), رو (r-V) /ru/ 'face'

زور (*zor*), زور (z-V-r) /zor/ 'strength' : زور (*zūr*)²⁴⁵, زور (z-V-r) /zur/ 'untruth'

گو (*go*), گو (g-V) /go/ conj. 'though'²⁴⁶ : گو (*gū*), گو (g-V) /gu/ 'ball'²⁴⁷

The vowel changes that occurred between classical and modern New Persian coincided with the massive influx of Arabic vocabulary which thus became subject not only to the usual adaptation processes but also to the Persian "vowel revolution" (within the limits set by Arabic phonemic inventory, of course). The traces of these changes, which retain a certain degree of regularity, can also be noticed in Perso-Arabic loanwords in Hindustani. They might be generalized as follows:

(i) the changes (1) *ē → /i/ and (2) *ō → /u/ affected genuinely Persian words since Classical Arabic, which generally has

²⁴² MHK II 371; Maddāh 252; Platts 620.

²⁴³ Maddāh 202; Platts 773.

²⁴⁴ Present stem of P رونیدن (*r-V-Y'-Y-d-n*) /ro.idan/, 'to grow, increase', functioning (in comp.) as act. part.; Platts 602; STCD 347.

²⁴⁵ Maddāh 257; Platts 619.

²⁴⁶ Also present stem of P گفتن (*g-f-t-n*) /guftan/ 'to say', functioning in H as part. adj. & s.m. 'saying'; MHK II 132; Platts 921; STCD 529.

²⁴⁷ Maddāh 202; STCD 529; Platts (921) gives also a variant form /go/.

a three-vowel inventory (in their short and long variants), and does not possess the vowels /e/ and /o/ (Watson 2007: 21-22);

(ii) the changes (3a) *ī → /ɛ/ and (4) *ū → /o/ have widely affected New-Persian words of both Persian and Arabic origin; in Hindustani these loanwords retain their original (Arabic or Persian) vocalisation with the exception of a few genuinely Arabic words, in which short /i/ (unmarked) was followed by *‘ain* (3b).

The table below shows examples of Persian loanwords of different etymologies, common in Hindustani, in which the Persian old vocalic system, antecedent to the changes, has been preserved (except 3b) and can be easily observed.

Hindustani	Classical Persian (Steingass)	Modern Persian (Hayyim)
------------	----------------------------------	----------------------------

(1) *ē → /i/

پیش (<i>peš</i>) پیش (<i>p-Y-š</i>) /pef/	پیش (<i>p-Y-š</i>) /pef/	پیش (<i>p-Y-š</i>) /piʃ/ ‘forward’ (P)
گسّ (<i>gesū</i>) گسّ (<i>g-Y-s-V</i>) /gesu/	گیسو (<i>g-Y-s-V</i>) /gesu/, /gisu/	گیسو (<i>g-Y-s-V</i>) /gisu/ ‘lock of hair’ (P)
مخّ (<i>mekḥ</i>) مخّ (<i>m-Y-x</i>) /mex/	میخ (<i>m-Y-x</i>) /mex/	میخ (<i>m-Y-x</i>) /mix/ ‘nail’ (P)

(2) *ū, *ō → /u/

گوشت (<i>gošt</i>) گوشت (<i>g-V-š-t</i>) /goʃt/	گوشت (<i>g-V-š-t</i>) /goʃt/	گوشت (<i>g-V-š-t</i>) /guʃt/ ‘meat’ (P)
پوشاک (<i>pošāk</i>) پوشاک (<i>p-V-š-A-k</i>) /poʃak/	پوشاک (<i>p-V-š-A-k</i>) /poʃak/	پوشاک (<i>p-V-š-A-k</i>) /puʃak/ ‘clothing’ (P)

मोम (<i>mom</i>) موم (<i>m-V-m</i>) /mom/	موم (<i>m-V-m</i>) /mom/, /mum/	موم (<i>m-V-m</i>) /mum/ 'wax' (P)
--	--------------------------------------	---

(3a) *ī → /ε/²⁴⁸

तिलिस (<i>tilism</i>) ظلم (<i>t-l-s-m</i>) /tilism/	ظلم (<i>t-l-s-m</i>) /tilism/	ظلم (<i>t-l-s-m</i>) /telesm/ 'talisman' (≤ A ≤ G)
मिज़ाज (<i>mizāj</i>) مزاج (<i>m-z-A-j</i>) /mɪzadz/	مزاج (<i>m-z-A-j</i>) /mɪzadz/	مزاج (<i>m-z-A-j</i>) /mezadz/ 'temperament' (≤ A)
फ़िरनी (<i>fir'nī</i>) فرنی (<i>f-r-n-Y</i>) /firmi/	فرنی (<i>f-r-n-Y</i>) /firmi/	فرنی (<i>f-r-n-Y</i>) /ferni/ 'kind of dish' (P)

(3b) *ī + ^c → /ε/

पेदा (<i>medā</i>) مده (<i>m-^c-d-h</i>) /meda/	معدة (<i>m-^c-d-i</i>) /mɪ'dæ/	معدة (<i>m-^c-d-h</i>) /mɛ'dɛ/ 'stomach' (≤ A)
फ़ैल (<i>fel</i>) فعل (<i>f-^c-l</i>) /fel/	فعل (<i>f-^c-l</i>) /fi'l/	فعل (<i>f-^c-l</i>) /fɛ'l/ 'action' (≤ A)
नेमत (<i>nemat</i>) نعت (<i>n-^c-m-t</i>) /nemət/	نعمت (<i>n-^c-m-t</i>) /nɪ'mæt/	نعمت (<i>n-^c-m-t</i>) /nɛ'mæt/ 'affluence' (≤ A)

(4) *ū → /o/

बुजुर्ग (<i>buzurg</i>) بزرگ (<i>b-z-r-g</i>) /buzurg/	بزرگ (<i>b-z-r-g</i>) /buzurg/	بزرگ (<i>b-z-r-g</i>) /bozorg/ 'great' (P)
--	-------------------------------------	---

²⁴⁸ Among the loanwords there is (at least) one exception which did not retain the old vocalization: H सेहत (*sehat*) / سَهِت (*ṣ-ḥ-t*) /seh:ət/, cf. Stein.: صَهِت (*ṣ-ḥ-t*) /sih:ət/; Hayyim: صَهِت (*ṣ-ḥ-t*) /seh:æt/ 'health'.

गुबार (<i>gubār</i>) غبار (<i>ġ-b-A-r</i>) /gubar/	غبار (<i>ġ-b-A-r</i>) /ɣubar/	غبار (<i>ġ-b-A-r</i>) /ɣobar/ 'dust' (≤ A)
बुकचा (<i>buk'cā</i>) بُقچہ (<i>b-q-ċ-h</i>) /buqʃa/	بُقچہ (<i>b-q-ċ-h</i>) /buqʃæ/	بُقچہ (<i>b-q-ċ-h</i>) /boqʃɛ/ 'bundle' (≤ A ≤ T)

Table 4. Retaining of classical Persian vocalic system in Perso-Arabic loanwords in Hindustani

3.1.4. The question of diphthongs in Persian is complicated and specialists do not agree on their number (from six to one, and occasionally even none) or classification (phonemic or phonetic).²⁴⁹ But as far as the subject of this study is considered, most of the discussion concerning Persian diphthongs can be skipped. From the point of view of Persian loanwords in Hindustani, the knowledge of diachronic change of two vocalic sequences, /ai/ → /ej/ and /au/ → /ou/, once again remains the most important and necessary factor. However, as pointed out by Windfuhr (1979: 137), this change occurred only in post-classical Persian and, additionally, did not occur in several modern Persian dialects. Such was also, at least to some extent, the case of the Persian spoken in South Asia. The confirmation of this fact can be found in Phillot's observations on how two Persian diphthongs, which he distinguished, were pronounced in the middle of words:

By combining *faṭha* with *ی* and with *و*, the diphthongs *ai* or *ay*, and *au* or *aw*, are formed, and pronounced like *ai* in *aisle* (rare in Persian) or *ey* in *they*, and *ou* in *stout*; Ex. میل *mail* or *mayl* 'inclination' صوم *ṣaum* or *ṣawm* 'fasting'.

²⁴⁹ Phonemic or 'real' diphthongs are thought to be tautosyllabic and to occupy the syllable nucleus, while phonetic diphthongs are those whose glide element functions as an onset or coda and is not a part of the nucleus (Rohany Rahbar 2012: 119).

In modern Persian the *ay* is most often pronounced like *ey* in *they*; Ex. خلی is pronounced both *kḥailī* and *kḥeylī* (also *kḥailé*, etc., etc.) (...); also the *au* sound, characteristic of the Arabic, Afghan, or Indian accent, tends towards the sound of *o* in *hose*. (Phillot 1919: 22)

It seems clear on the basis of Phillot's description that in this case the discussion (e.g. Pisowicz 1985; Rohany Rahbar 2012; Windfuhr 1979) concerning the phonemic or non-phonemic status of these vocalic sequences, i.e. whether they are acting as one vowel with two parts which are not separable or as a sequence of a vowel followed by a consonant, would be of secondary importance (it is, however, crucial for what might be regarded diphthongs in Iranian Persian). What matters more from the point of view of this study is the question of these diphthongs' potential monophthongal utterance, which would put them closer to the Hindustani vocalic system. An important factor, which critically influenced the frequency of diphthongs, was their monophthongization – one of the basic and most regular changes in the Middle period of both the Indo-Aryan and Iranian linguistic families. In consequence of this change the old diphthongs /aɪ/ and /aʊ/ in Middle Persian and Middle Indo-Aryan became long /e/ and /o/ respectively.²⁵⁰ As a result of the monophthongization, the number of genuinely Persian loanwords with diphthongs is relatively small in Hindustani,²⁵¹

²⁵⁰ E.g.: /aɪ/ > /e/: S /taila/ 'oil' > Pali /tela/ > H /tel/; OP /daiva/ 'demon' > MP /dev/ > P /div/; /aʊ/ > /o/: S /gaura/ 'white, yellowish' > Pali /gora/ > H /gora/; OP /raucah/ 'day' > MP /roz/ > P /ruz/.

²⁵¹ Genuinely Persian diphthongs appear in following loanwords: /aɪvan/ 'palace'; /mai/ 'wine'; /naɪfʃa/ 'tube of a hookah'; /naɪrəŋg/ 'deception'; /pai/ 'foot' (and its derivatives); /pada/ 'born' (and its derivatives); /paɪgam/ 'message' (and its derivatives); /paɪkər/ 'face, form'; /paɪma.ɪf/ 'measuring'; /paɪmana/ 'cup'; /paɪvənd/ 'patch'; /paɪvəsta/ 'joined (to)'; /fajda/ 'possessed'; /aʊrəŋg/ 'throne'; /fʃaugan/ 'polo stick'; /gəuhər/ ~ /dʒəuhər/ 'jewel'; /naʊkər/ 'servant'; /naʊsadər/ ~ /naʊfədər/ 'sal-ammoniac'; /saʊgənd/ 'oath'; /fʃəuhər/ 'husband'; /taʊsən/ 'a young unbroken horse'.

diphthongs occur mainly in words of Arabic origin which were incorporated by Persian only during its transition into the New Persian stage, and therefore were not subject to the Middle Persian changes (monophthongization). The examples presented in the table below show how diphthongs are realized in Persian loan-words in Hindustani in contrast to their modern Iranian (Farsi) equivalents:

Hindustani	Classical Persian (Steingass)	Modern Persian (Hayyim)
خیمہ (<i>khaimā</i>) خیمہ (<i>x-Y-m-h</i>) /xaima/ ~ /xæma/	خیمہ (<i>x-Y-m-h</i>) /xaimæ/	خیمہ (<i>x-Y-m-h</i>) /xejmε/ 'tent' (≤ A)
فہملا (<i>fais'lā</i>) فیصلہ (<i>f-Y-s-l-h</i>) /faisla/ ~ /faisla/	فیصلہ (<i>f-Y-s-l-h</i>) /faisælæ/	فیصلہ (<i>f-Y-s-l-h</i>) /fejsæle/ 'decision' (≤ A)
پیدائش (<i>paidāiš</i>) پیدائش (<i>p-Y-d-A-Y'-š</i>) /paida.ɪʃ/	پیدائش (<i>p-Y-d-A-Y'-š</i>) /paida'ɪʃ/	پیدائش (<i>p-Y-d-A-Y-š</i>) /pejdajæʃ/ 'birth'
سُہیل (<i>suhel</i>) سہیل (<i>s-h-Y-l</i>) /suhæl/ ~ /suhail/	سہیل (<i>s-h-Y-l</i>) /suhail/	سہیل (<i>s-h-Y-l</i>) /suhejl/ 'the star Canopus' (≤ A)
رینک (<i>raunak</i>) رونق (<i>r-V-n-q</i>) /raunæq/ ~ /rɔnæq/	رونق (<i>r-V-n-q</i>) /raunæq/	رونق (<i>r-V-n-q</i>) /rounæq/ 'radiance' (≤ A)
موجود (<i>maujūd</i>) موجود (<i>m-V-j-V-d</i>) /moɖʒud/	موجود (<i>m-V-j-V-d</i>) /maujud/	موجود (<i>m-V-j-V-d</i>) /moujud/ 'existing' (≤ A)

शौहर (<i>śauhar</i>) شوهر (<i>š-V-h-r</i>) /ʃauhær/ ~ /ʃɔhær/	شوهر (<i>š-V-h-r</i>) /ʃauhær/	شوهر (<i>š-V-h-r</i>) /ʃovhær/ 'husband'
कौम (<i>kaum</i>) قوم (<i>q-V-m</i>) /qaum/	قوم (<i>q-V-m</i>) /qaum/	قوم (<i>q-V-m</i>) /qovm/ 'people, nation' (≤ A)

Table 5. Realization of diphthongs in Perso-Arabic loanwords in Hindustani

3.1.5. As for the Persian consonants, their number depends on which phones are categorized as allophones of the regular consonantal phonemes, and mostly varies between 23 (Windfuhr 1997: 667) and 24 (Pisowicz 1985: 18; Lazard 1992: 7). The table below is a listing of Persian contrasting segmental consonantal phonemes.

	bilabial	labio-dental	dento-alveolar	alveo-palatal	velar	uvular	glottal
stops							
voiceless	/p/		/t/		/k/	[q] ⁽ⁱ⁾	/ʔ/
voiced	/b/		/d/		/g/	[ɢ] ⁽ⁱ⁾	
affricates							
voiceless				/tʃ/			
voiced				/dʒ/			
fricatives							
voiceless		/f/	/s/	/ʃ/		/x/	/h/
voiced		/v/	/z/	/ʒ/	[ŋ] ⁽ⁱⁱ⁾	[ʁ] ⁽ⁱ⁾	
liquids			/r/ /l/				
nasals	/m/		/n/				
semivowels	[v] ⁽ⁱⁱⁱ⁾			/j/			

Table 6. Schema of the Persian consonantal system

Notes:

(i) As mentioned already in paragraph 2.3.4.(i), in Persian the phonemes /q/ and /ɣ/ merged into one phoneme with three positionally conditioned allophones: the voiceless uvular plosive [q], the voiced uvular plosive [ɢ] and the voiced uvular fricative [ɣ].

(ii) The phone [ŋ] is the positionally conditioned allophone of the phoneme /n/ and appears exclusively before /k/, /g/, and /q/. Due to its similarity to the analogous genuine Hindustani phone [ŋ], it is a natural and easily articulated sound for the speakers of Hindustani.

(iii) The status of the phone [v] in the Persian consonantal system is very complicated and unclear: it may be regarded either as an allophone of the spirant /v/, as an allophone of the vowel /u/, or as a quite separate phoneme (cf. discussion in Pisowicz 1985: 24-27). Its distribution is restricted to the single position after the vowel /o/ which approximates it functionally to the second element of the genuine Hindustani diphthong /au/. In Persian loanwords in Hindustani the segment [o] + [v] is always realized diphthongically, with a tendency to monophthongization (*vide supra*, fn. 239 and examples in Table 5).

3.1.6. The table on page 137 presents the total range of contrasting segmental phonemes as well as some important allophones (in square brackets) that constitute the Hindustani consonantal system. The chart comprises both the genuinely Hindustani phonemes (i.e. those Indic consonants which developed from the OIA through the MIA stage) and the loan phonemes, including those, which were adopted from Persian/Arabic, English (e.g. the voiced alveo-palatal fricative ʒ, like in the English word *pleasure* [pleʒər]) or Sanskrit (e.g. the voiceless postalveolar (retroflex) sibilant /ʂ/, like in *viṣay* [vɪʂəj] 'subject, matter'). Of course, since this is the maximum inventory, one must remember that hardly any speaker of Hindustani will use it in its entirety. Some of the phonemes listed in the table occur solely in the most

puristic registers of the language and in the most formal contexts.²⁵² Furthermore, the two flaps: /ɾ/ and /ɽ^h/ have marginal phonemic status and are often considered the allophones of the retroflex voiced stops (respectively /d/ and /d^h/) and replaced by them both in speech and in writing.²⁵³ The alveo-palatal [ɲ] and the retroflex [ɳ] are positionally conditioned allophones of the phoneme /n/. The borrowed phonemes /q/ and /ɣ/ ~ /g/ and their mutual interchangeability will be discussed below (3.2.5.). The Arabic glottal stop /ʔ/ (and also /ʕ/) which in Persian is often either dropped completely or replaced by the glide /j/ (Pisowicz 1985: 47 ff.), occurs only sporadically in Hindustani in the extreme formal Arabicized speech (e.g.: [bæʔd], [ʃaʔr] instead of the more colloquial [ba:d] 'after' or [ʃa:jr] 'poet'). The labio-dental voiced approximant /v/ is often pronounced bilabially – either as a bilabial semivowel or as the bilabial voiced fricative [β]. In Persianized style of speech it can also be realized as the labio-dental voiced fricative [v] (the way it is pronounced in Persian). Moreover, /v/ and /b/ are neutralized by some speakers (e.g. बीबी (*bīvī*), بی بی (*b-Y-V-Y*) [bi:vi:] ~ [bi:bi:] 'wife'; नवाब (*navāb*), نواب (*n-V-A-b*) [nəv:a:b] ~ [nəba:b] 'a governor'). In Hindustani the voiced alveolar /r/ is realized either as a trill [r] or as a tongue tap [ɾ], articulated anywhere from the alveolar ridge to the back of the front teeth.

²⁵² For example, the phonemic contrast between the postalveolar (retroflex) and alveo-palatal sibilants (/ʃ/ ~ /s/) which operates only on the highly Sanskritized level of Hindustani/Hindi (in words like विशेष (*viśeṣ*) 'special' or संलेप (*śleṣ*) 'coniunction') or the hyper correct Arabicized utterance of the glottal stop /ʔ/ (e.g. pronouncing the word مافق (*m-V-A-f-q*) 'suitable' with *hamzah*: [muʔa:fi:q] instead of a glide: [muva:fi:q]).

²⁵³ On the phonemic or non-phonemic status of the retroflex flaps see e.g. Kelkar 1968: 39-42.

	bilabial	labio-dental	dental	alveolar	post-alveolar	alveo-palatal	velar	uvular	glottal
stops									
voiceless non aspirated	/p/		/t/		/tʃ/		/k/	/q/	/ʔ/
voiced non aspirated	/b/		/d/		/dʒ/		/g/	[ɢ]	
voiceless aspirated	/pʰ/		/tʰ/		/tʃʰ/		/kʰ/		
voiced aspirated	/bʰ/		/dʰ/		/dʒʰ/		/gʰ/		
affricates									
voiceless non aspirated						/tʃ/			
voiced non aspirated						/dʒ/			
voiceless aspirated						/tʃʰ/			
voiced aspirated						/dʒʰ/			
fricatives									
voiceless		/f/		/s/	/ʃ/	/ʃ/		/x/	/h/
voiced	[β]	[v]		/z/		/ʒ/		/ɣ/	
lateral			/l/						
trill				/r/					
flaps									
non aspirated					/ɾ/				
aspirated					/ɾʰ/				
nasals	/m/		/n/		/ɳ/	[ɲ]	[ŋ]		
approximants		/ʋ/				/j/			

Table 7. Schema of maximum Hindustani phonemic (consonantal) inventory

3.1.7. To sum up this chapter on the similarities and differences between Persian and Hindustani phonological systems, a general observation can be made, namely, that in comparison with the previous stages, both Hindustani and Persian seem to tend towards maximum simplification. Since the break of the Indo-Iranian community sometime at the beginning of the second millenium BC, these languages evolved separately through the consecutive Old, Middle and New stages, yet their independent development shows many congruous and comparable tendencies. Diachronical analysis of both the systems reveals, among others, haplology and simplification of consonantal groups²⁵⁴ and diphthongs (monophthongization),²⁵⁵ decaying of phones of compound articulation (in Persian these are labiovelars²⁵⁶ and aspirates,²⁵⁷ in Hindustani – syllabic sonants,²⁵⁸ in both languages geminates are becoming simplified²⁵⁹); weakening and, in some cases, loss of intervocalic consonants (usually preceded by stages

²⁵⁴ E.g.: MIA (Pr.) /lǝllǝkkǝ/ > H /lǝlkar/ 'challenge'; MIA (Pr.) /buddhǝtti/ > H /bʊrʰauti/ 'old age'; MP /dadistan/ > NP /dastan/ 'story'; OP /xʃajǝθijǝ/ > NP /ʃah/ 'shah'.

²⁵⁵ *Vide supra* fn. 250.

²⁵⁶ E.g.: Av /xwǝp-/ > NP /xab/ 'sleep'; Av /xwǝr-/ > NP /xur-/ ~ /xor-/ 'eat'. In New Persian there is a trace of them in spelling but they are not pronounced; they still exist partly in Tajiki and Dari. Cf. also Pisowicz 1985: 121-123.

²⁵⁷ As a matter of fact, in Iranian languages aspirates have never existed – they had already disappeared when the Indo-Iranian group collapsed; compare OIA /bʰratrǝ/ and OP /bratǝr/ 'brother'.

²⁵⁸ E.g.: S पृच्छक (*prccʰaka*) /prɪʃʃʰəkǝ/ > H पूछता (*pūcʰtā*) /puʃʃʰtǝ/ 'asking'.

²⁵⁹ E.g.: MIA (Pali) /kǝmmǝ/ > H /kam/ 'work'; MIA (Pr.) /nissǝkki/ > H /nisǝk/ 'weak, powerless'. In Persian, both native and Arabic geminates are pronounced stronger/longer, not twice, e.g.: پَر (p-rr) [par:] 'feather'; بَچَه (b-čč-h) [bæʃʃ:ɛ] 'child'.

of vocalization and spirantization).²⁶⁰ Suprasegmental changes in accentuation from moving and tonal accents, still existing in the Middle Indian and Middle Iranian periods, into steady and dynamic ones²⁶¹ have led to the reduction of vowels in the final and unstressed syllables (and, as a result, the shortening of the words' length).²⁶² These phonological changes had an immense effect on the whole system by influencing morphological and syntactic modifications in both the languages. The decay of final sounds was one of the reasons for declining inflexions (expecially in Persian) and the appearance of analytic constructions in the place of the inflexional endings. Subsequently, it induced the development of grammatical words²⁶³ (prepositions, postpositions, articles) and periphrases (analytic forms or verbal compounds acting as periphrastic predicates²⁶⁴). As a result of these changes the whole language structure transformed into analytic, with some inclinations to agglutination. The decline of inflexions has also

²⁶⁰ E.g.: S /gata/ > MIA (Pr.) /gədə/ > H /gəja/ 'gone'; Av. /pāda-/ > MP /paj/ > NP /pa(j)/ 'leg'.

²⁶¹ In New Persian stress is steady, generally oxytonical (with some exceptions) and has a delimitative role (e.g. Lambton 2000: *passim*; Mace 2003: 31-34); in Hindustani accent is not settled yet and leans towards long syllables, irrespective of their position in the word; its phonemic status remains a matter of controversy (see e.g. discussion in Nair 2001).

²⁶² E.g.: MIA (Pr.) /rət.ti/ > H /rət/ 'night'; MIA (Pr.) /me-hə/ > H /megh/ 'cloud'; Av. /æhurə mæzdah/ > OP /'auræ.mæzdah/ > MP /'o:hrmæzd/ > NP /'ormozd/ 'Ahura Mazda (the name of god)'; OP /'mærtijæ/ > NP /'mærd/ 'man'

²⁶³ Since these words have lost their lexical value and perform solely grammatical functions, they are also called empty or auxiliary words.

²⁶⁴ They are replacing simple verbal predicates (when synonymous) or – if there is no simple predicate – they are the only way for expressing any definite contents. They are formed by the operator indicating the grammatical category (tense, aspect, mood etc.) and an abstract noun conveying the actual predicate meaning (Kuczkiewicz-Fraś 1997: 48).

become the reason for the systematization and reduction of paradigms (they were originally differentiated notably, because of different groups of lexemes) or the decaying of rection and congruence (it is connected with the decaying of gender, declension and dual numbers, replaced by periphrastic constructions, postpositions or constructions with *izāfat*). Some of these phenomena, important from the point of view of this study, will be discussed later in the book.

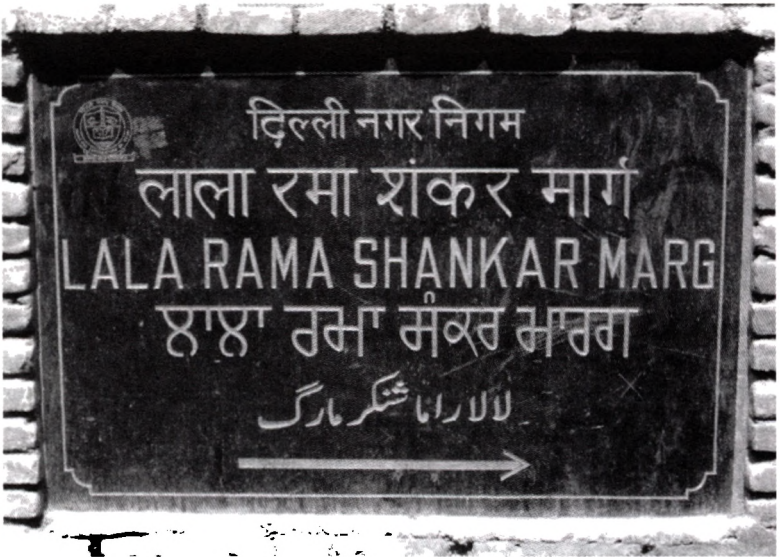


Fig. 7. “Street allography”. A multigraphic plaque with the name of a street in New Delhi, written down in four scripts: Nagari, Roman, Gurmukhi and Perso-Arabic. (Photo by the author, 2012)

3.2. Phonological Adaptation of the Perso-Arabic Loanwords

The borrowing of foreign words always entails their phonetic modification. (...) Frequently we have phonetic compromises.

(Sapir 1921: 210)

3.2.1. Before entering into an analysis of how Perso-Arabic loanwords have been adapted to the Hindustani system on the level of phonology, two relevant observations should be made. The first is connected with the ongoing general discussion on the nature of the input to the loanword phonology which can be boiled down to the simplified question whether it is phonetic or phonological.²⁶⁵ The phonetic approximation theory assumes that it is perceptual (or acoustic) similarity, unfaithful by definition (since speakers of the borrowing language cannot properly recognize foreign sounds and structures), which is responsible for mapping forms from the source language onto the borrowing language. In other words, the foreign segment is substituted by the native one that most closely approximates the input in articulatory and/or acoustic properties. Furthermore, the phonetic approximation can be either intentional (importation of an alien sound) or unintentional (misinterpretation of an alien sound).²⁶⁶ Phonological approximation theory, in turn, relies on the assumption that a foreign segment from the source language is replaced by a segment from the borrowing language that is the closest phonologically (i.e. in terms of features) even if there is a phonetically closer sound in the inventory of the borrowing language. Obviously, the favourable condition for the latter is the

²⁶⁵ E.g.: LaCharité & Paradis 2005 (with extensive references); Heffernan 2005; Chang 2012.

²⁶⁶ Apart from prevailing cases of analogy, non-phonological adaptations include also cases of orthography influence (cf. e.g. 3.5.4.(vi)).

bilingual competence of the speaker. It may be guessed that both the adaptation strategies can be traced in Perso-Arabic loanwords in Hindustani. Moreover, it may be the sociolinguistic aspect of the language contact that determines which of the strategies has been employed. An explicit and detailed analysis of the adaptation of Perso-Arabic loanwords in Hindustani in terms of its phonetic and/or phonological character would require a separate and comprehensive study – due to the vastness and diversity of the research material – backed by solid verification with Hindustani native speakers representing different sociolinguistic strata. Unfortunately, the aim and the scope of this book leave no room for a more detailed investigation of this issue.

3.2.2. The second remark concerns the high level of irregularity in sound change occurring in Perso-Arabic loanwords in Hindustani. Although throughout the process of adopting Persian and Arabic lexical items to its own phonological system, Hindustani has worked out a sophisticated system-based substitution scheme (*vide infra*), in practice the phonetic realization of many borrowed words differs from their underlying representation. Sometimes these variations occur in a particular context and are therefore predictable, but in most of the cases they show random behaviour, additionally influenced by the language register, the style, the speaker, etc., so that they are hard to predict or generalize into phonological rules.²⁶⁷ This sociolinguistically conditioned diversity is further complicated by such phenomena as etymological nativization (for example, when literate speakers deliberately Arabicize the pronunciation of etymologically Arabic loanwords,²⁶⁸

²⁶⁷ For many Hindustani native speakers changing the language register, including the phonetic level, in accordance with the interlocutors/hearers with whom they communicate, is a common linguistic competence.

²⁶⁸ Cf. Bailey's observation "Thus a Mawlavī will import what he thinks are correct (Arabic) sounds into Urdu words. He will discourse on

even though formally these loanwords have been previously adjusted to the Persian and then to the Hindustani phonological systems) or nativization through spelling (cf. Hock 1991: 391ff) – which, in the case of Hindustani, a quintessential example of synchronic digraphia, is an important factor influencing sound change. Furthermore, it is hard to formulate universally applicable rules since both Persian and Hindustani exhibit a strong tendency to fluctuations in the use of vowels.²⁶⁹ An important – perhaps even the most important – reason for this phenomenon is again Perso-Arabic orthography. As noticed by Saran, since vowel marks have been neglected in written texts for centuries, the result is that even educated people do not agree among themselves regarding the vowel contents in a great number of Persian and Arabic words and often it is impossible to determine a steady pattern of their vocalization (Saran 1969: 68).²⁷⁰

the hamza, on 'ain, will assure the unfortunate student that words written *ḥukm*, *fikr*, *ṣubḥ*, are monosyllables, that *jadd o jihad* should be pronounced *jidd o jahad*, that *fi'l* is different from *fel*, that the first syllable of *mahdi* is not the same as that of *mahfi'l*" (Bailey 1922: 540). Thomason & Kaufman (1988: 78) notice exactly the opposite process: "(...) uneducated Moslems who are native speakers of Urdu nativize the pronunciation of the numerous Arabic loanwords borrowed into Urdu, and have no syntactic features borrowed from Arabic either".

²⁶⁹ Examples of fluctuation of vowels in Persian, both on a lexical and on a morphological level, see Pisowicz 1985: 14ff. Windfuhr remarks: "The unstable short vowels *a*, *e*, *o* show certain degree of assimilation in height, rounding, and length. There is certainly a continuum between social register and dialect register" (Windfuhr 1997: 679).

²⁷⁰ The following examples may illustrate the scale of vowel variation, especially when considered diachronically: H /færox/ ~ /firox/ 'sale' ≤ Class. P /færox/ ~ /firux/ ~ /furox/ (modern P /forux/) 't.s.'; H /māhar/ ~ /muhar/ ~ /mohar/ 'camel's nose string' ≤ Class. P [≤ A] /mīhar/ (modern P /mæhar/) 't.s.'.

3.2.3. Taking into consideration all that was said above and also bearing in mind the character of the linguistic material analyzed in this study (i.e. loanwords gathered from and confirmed by written sources), the following paragraphs will focus on the range and nature of phonological adaptation to which Persian and Arabic lexical borrowings have been subject in Hindustani. The description of changes attempts to mirror the rules on the one hand and deviations from the rules (alternative phonetic realization) on the other. It must also be remembered that Hindustani, as defined in this book (cf. 1.2.14) is characterized by a great diversity in pronunciation (this pertains especially to the realization of vowels). Therefore, it was necessary to define a compromised common ground which would be presumably most representative for phonetically non-homogenous Hindustani. Thus the pronunciation of the so-called Modern Standard Hindi (as given in McGr.) and Modern Standard Urdu (as given in the STCD) was adopted as such a basis for the presentation of phonological adaptation and sound changes that affect Perso-Arabic loanwords.²⁷¹ Since Hindustani's twofold orthography remains closely related to its phonological level, usually reflecting sound changes and even prompting some of them, both the original (Nagari and Perso-Arabic) notation of loanwords and the Roman transcription are given beside the phonological record.

3.2.4. A number of genuinely Arabic consonants are alien to Indo-European languages (even though they have been preserved in the script) and in the process of lexical borrowing they were assimilated to the closest phonemes of the recipient language. Since an overwhelming part of the loanwords from Arabic was incorporated into Hindustani via Persian, their assimilation to the Indo-European (or, more specifically, Indo-Iranian) phonological

²⁷¹ The reference point was predominantly the way in which Hindustani is spoken in Delhi, as well as in other large urban centers of northern South Asia, like Lahore or Lucknow.

structure firstly took place in Persian, and only then, if it was necessary, they were further adapted to meet specific Hindustani phonological constraints.²⁷² The fundamental consonantal sound changes in the Arabic loanwords in Persian (and, accordingly, in Hindustani) include the following assimilations:²⁷³ (i) three voiceless fricatives: dental /θ/ and alveolar /s/, /sʰ/ are realized as [s]; (ii) four voiced fricatives: dental /ð/ and alveolar /z/, /zʰ/, /dʒ/ are realized as [z]; (iii) voiceless dental plosives /t/ and /tʰ/ are realized as [t]; (iv) two voiceless fricatives: pharyngeal /ħ/ and laryngeal /h/ are realized as the voiceless aspirate [h]; (v) voiceless uvular stop /q/ and voiced uvular fricative /ɣ/ are pronounced as a voiced velar stop or affricate;²⁷⁴ (vi) Arabic labiovelar approximant /w/ is realised in Persian as labio-dental voiced fricative [v]²⁷⁵ (cf. Perry 2002: 229-230; Lambton 2000: xviv-xviii).

3.2.5. Phonological adaptation and realization of Persian consonantal sounds in Hindustani has already been partly discussed in the previous chapter, chiefly in paragraphs 2.3.3. and 2.3.4. devoted to orthographical adjustments. The high congruence of Persian and Hindustani phonological inventories has already been mentioned as well. Indeed, the assimilation of Persian loanwords into Hindustani has been eased by phonological similarities.²⁷⁶ Both Persian and Hindustani are generally characterized by a five-way distinction in place of articulation but only

²⁷² A short synchronical analysis of phonological assimilation of Arabic lexical elements in Hindustani (Urdu) can be found in Danecki 1971.

²⁷³ On the graphical realization of these phonemes cf. paragraphs 2.3.3.(i-iii) and 2.3.4.(iii).

²⁷⁴ *Vide supra* 2.3.4.(i).

²⁷⁵ However, in many other languages, including Hindustani, it becomes a bilabial or semi-vowel, *vide infra* 3.2.6.(v).

²⁷⁶ Congruence is observable also at the level of word phonology – Hindustani, like other modern Indo-Aryan languages, shares the Persian preference for simple patterns of the CVCVC type.

four of them are common for both systems (labial, dental, palatal and velar). In Hindustani the fifth group consists of retroflex consonants (defined also as prepalatal or postalveolar, and in older, mainly indological literature – as cerebral), which are present in all Indian languages but are alien to Persian.²⁷⁷ For Hindustani a brand new (sixth) place of articulation appeared together with Persian uvular sounds, which have been adapted to the Indic system as new, supplemental consonants: voiceless /q/ and voiced /ɣ/ ~ /g/.²⁷⁸ In Hindustani the realization of /q/ and /ɣ/ is transient and depends fundamentally on the language register which is connected with the speakers' education – mainly those who learnt High Urdu or studied Persian or Arabic will be able to recognize and articulate them properly. On the other hand, well-educated Urdu speakers are much more careful and conservative in their pronunciation of these sounds than an average native-speaker of Persian. The existing minimal pairs may argue for /q/ and /ɣ/ (with the allophonic variant [g]) being recognized as two separate phonemes:

²⁷⁷ Retroflex consonants, however, resembling Indian cerebrals, are common in Pashto (Henderson 1983: 596).

²⁷⁸ Windfuhr finds characterisation of /q/ to be the most problematic issue of the phonetic system of Iranian Persian and enumerates its possible allophonic realizations: "1) In prevocalic position, *q* tends to be a voiced stop, often followed by a fricative release [this is controversial – AKF]. 2) In final position, it is a devoiced stop, or fricative, both with equal frequency. 3) In clusters, voicing is neutralized. (...) In addition, in intervocalic position *q* is a voiced fricative, which supports its systematic pairing with *x*." (Windfuhr 1997: 682; cf. also Pisowicz 1985: 42–47). Phillot, on the basis of Indian Persian, observed that "ق (...) sometimes interchanges with غ", but at the same time seemed to be quite critical of these two sounds merging, seeing the source of this phenomenon in the usage of Persian Turks ("a Turk of Persia will often spell the same word with either letter indifferently"). He also warned against confusing /q/ with /k/: "It must not be pronounced like *k*, a common English mistake" (Phillot 1919: 17).

कदर (<i>kad'r</i>)		गदर (<i>gad'r</i>)	
قد (<i>q-d-r</i>)	/qədr/ 'value'	غدر (<i>ġ-d-r</i>)	/ɢədr/ 'revolt'
बाकी (<i>bākī</i>)		बागी (<i>bāgī</i>)	
باقی (<i>b-A-q-Y</i>)	/baqi/ 'remaining'	باغی (<i>b-A-ġ-Y</i>)	/baɣi/ 'of garden'
करीब (<i>karīb</i>)		गरीब (<i>garīb</i>)	
قريب (<i>q-r-Y-b</i>)	/qərib/ 'near'	غریب (<i>ġ-r-Y-b</i>)	/ɢərib/ 'poor'

In casual Hindustani, however, and among non-educated speakers, the Perso-Arabicized pronunciation of /q/ and /ɣ/ is relatively uncommon, but not entirely absent.²⁷⁹ Cf. the following examples:

(i) depending on the speaker, /q/ may be (and, actually, very often is) realized as the velar voiceless plosive [k]:

क़फ़स (*kafas*), قفص (*q-f-s*) [qəfəs] ~ [kəfəs] 'cage'
 मुक़ाम (*muḳām*), مقام (*m-q-A-m*) [muqa:m] ~ [muka:m] 'staying'
 बर्राक (*barrāk*), برق (*b-r-A-q*) [bər:a:q] ~ [bər:a:k] 'shining'

(ii) before the front close vowels /i/ and /ɪ/ the articulation of /q/ tends to be more frontal even in Hindustani/Urdu, e.g.:

किस्म (*kism*), قسم (*q-s-m*) [kɪsm] 'kind',²⁸⁰
 बाकी (*bākī*), باقی (*b-A-q-Y*) [ba:ki:] 'remaining'

²⁷⁹ Therefore, Bahri's critical claim, that the exact pronunciation of these sounds is "used for accuracy mainly by a few hypersensitive persons" (Bahri 1960: 58) seems to be too far-reaching.

²⁸⁰ In initial position it may still be realized as [q], simultaneously influencing the following vowel by pushing its articulation slightly back, e.g.: [qybla:] 'the direction of Mecca', [qysmæt] 'fate', [qys:ə] 'story'.

(iii) before a voiceless stop /q/ may be realized as the velar voiceless fricative [x],²⁸¹ e.g.:

वक्फा (*vakfā*), وقف (V-q-f-h) [vəqfə] ~ [vəxfə] 'pause'

बुकचा (*buk'cā*), بقی (b-q-ċ-h) [buqʃə] ~ [buxʃə] 'bundle'

वक्त (*vakṭ*), وقت (v-q-t) [vəqt] ~ [vəxt] 'time'

(iv) before a voiced stop /q/ may be realized as the velar voiced fricative [ɣ],²⁸² e.g.:

तक्दीर (*tak'dīr*), تقدیر (t-q-d-Y-r) [təqdi:r] ~ [təɣdi:r] 'destiny'

मक्बरा (*mak'barā*), مقبره (m-q-b-r-h) [məqbərə] ~ [məɣbəərə] 'tomb'

(v) depending on the speaker, /ɣ/ may be realized in every position as: the uvular voiced plosive [g] (especially initially and intervocalically), the velar voiced stop [g] (especially before /i/) or the velar voiced fricative [ɣ]:

गालिब (*gālib*), غالب (ġ-A-l-b) [ga:lɪb] ~ [ga:lɪb] ~ [ɣa:lɪb] 'triumphant'

आगा (*āgā*), آغا (Ā-ġ-A) [a:ga:] ~ [a:ga:] ~ [a:ɣa:] 'lord, chief'

बगैर (*bagair*), بغیر (b-ġ-Y-r) [bəgaɪr] ~ [bəgaɪr] ~ [bəɣaɪr] 'without'

3.2.6. In the consonantal inventory of Hindustani, fricatives are less well represented than in Persian. It is the influence of Persian which principally accounts for the presence of /f/, /z/, /ʃ/ and /x/ in the Hindustani consonantal system, although, of course, /f/ and /z/ also appear in the loanwords from English and /ʃ/ – in the loanwords from Sanskrit (*tatsama*). In everyday speech (especially in Hindustani/Hindi) these fricatives are often subject to the process of fortition (defricativization) and substi-

²⁸¹ Cf. Kelkar 1968: 32; Schmidt 2003: 310.

²⁸² Ibid.

tuted with phonetically similar intrinsic consonants.²⁸³ However, their actual phonetic value is not fixed and depends largely on such factors as language register or social environment of the speaker.²⁸⁴ The most widespread and most regular changes in the fricatives are discussed below:

(i) voiceless labio-dental fricative /f/ is quite often replaced in speech by the aspirated bilabial plosive /p^h/, e.g.:

फ़िक्र (fīkr), فکر (f-k-r) [fīkr] ~ [p^hīkr] ‘thought, reflection’

काफ़ी (kāfī), کافی (k-A-f-Y) [ka:fi:] ~ [ka:p^hi:] ‘enough, sufficient’

साफ़ (sāf), صاف (ṣ-A-f) [sa:f] ~ [sa:p^h] ‘clear, clean’²⁸⁵

It is worth noting that in the case of many Hindustani speakers the reverse process of fricativization occurred and the alien phoneme /f/ nearly completely replaced /p^h/ in native Indic words,²⁸⁶ e.g.:

²⁸³ Cf. Bahri 1960: 57-59; Ohala 1983: 2; Shukla 2000: 51; Tivārī 1969: 67-68. Also *vide supra* 2.3.4.(iv) and fn. 175.

²⁸⁴ This high volatility of fricatives may be confusing even for specialists. For example, Kachru who does not use Nagari in her study, states: “The borrowed sounds q, x, r, f, z are transcribed as k, k^h, g, p^h, j respectively. In some places, however, the free variation in pronunciation has been indicated to signal the occurrence of these sounds in educated speech” (2006: 21) but it is difficult to figure out a rule according to which the loanwords are represented in her system (e.g. on p. 25 alternative forms k^h/xvab, j/zyada are given, while on p. 60 we find only “Persianized” transcription: zyada-se-zyada – as, in fact, happens in most cases, cf. qarīb (p. 60), hazar (p. 58), etc.).

²⁸⁵ Such assimilation of /f/ to /p^h/ is observed most commonly at the onset of a word and much less frequently at the end of a word.

²⁸⁶ Cf. Shapiro 2003: 260. I do not know about any detailed study devoted to this phenomenon but on the basis of personal observations I would risk an opinion that the lenition of this type (i.e. replacing /p^h/ with /f/) is

फिर (*p^hir*), فیر (*p^h-r*) [p^hir] ~ [fir] 'anew, again'

सफल (*sap^hal*), سफल (*s-p^h-l*) [səp^həl] ~ [səfəl] 'fruitful'

सौंफ (*saump^h*), سونف (*s-V-n-p^h-h*) [saũp^h] ~ [saũf] 'fennel'²⁸⁷

(ii) the voiced alveolar or postdental fricative /z/ tends to be replaced by the voiced alveo-palatal affricate /ɖʒ/ and this propensity is very strong, especially in Hindustani/Hindi, e.g.:

ज़ोर (*zor*), زور (*z-V-r*) [zor] ~ [ɖʒor] 'strength, power'

मज़ाक (*mazāk*), مزاح (*m-z-A-q*) [məza:q] ~ [məɖʒa:q] ~ [məɖʒa:k] 'taste'

मग़ज़ (*mag'z*), مغز (*m-g-z*) [məɣz] ~ [məɣɖʒ(ə)] 'brain'

ज़ुल्म (*zulm*), ظلم (*z-l-m*) [zulm] ~ [ɖʒulm] 'wrong-doing'

हैज़ा (*haizā*), حيه (*h-Y-z-h*) [haizə] ~ [haɖʒa:] 'cholera'

(iii) the voiceless alveo-palatal fricative /ʃ/, as a result of the unification of sibilants that occurred in the Middle Indo-Aryan (cf. Bubenik 1996: 9, 34) has disappeared from the indigenous words, having been replaced by alveolar or postdental /s/. Therefore, in Hindustani, the phoneme /ʃ/ appears only in borrowings from Persian (or English, or Sanskrit) and by many speakers may be substituted with /s/, e.g.:

शायद (*śāyad*), شاید (*ś-A-y-d*) [ʃa:jəd] ~ [sa:jəd] 'perhaps'

शादी (*śādī*), شادی (*ś-A-d-Y*) [ʃa:di:] ~ [sa:di:] 'wedding'

more common than the opposed process, especially among the better educated urban dwellers, who have relatively frequent contact with English. As Ohala points out, what was previously a result of hypercorrection, now is recognized as a non-standard pronunciation (Ohala 1983: 4, fn. 4).

²⁸⁷ Following the pronunciation, this word, both in Perso-Arabic and Nagari scripts, is sometimes written with -f instead of -p^h, cf.: सौंफ (*saumf*), सौंफ (*sorimf*), سونف (*s-V-n-f*) 'anise seed' (Platts 702); سونف (*s-V-n-f*) 't.s.' (STCD 392); सौंफ (*saumf*) 'aniseed, fennel' (Bahri 1989: 677).

पेशानी (*pareśānī*), پشانی (*p-r-Y-ṣ-A-n-Y*) [pəreʃa:ni:] ~ [pəresa:ni:] 'disorder'
 दुश्मान (*duś'man*), دشمن (*d-ṣ-m-n*) [duʃmən] ~ [dusmən] 'enemy'
 किशमिश (*kiś'miś*), کشش (*k-ṣ-m-ṣ*) [kiʃmiʃ] ~ [kismis] 'raisin'²⁸⁸

(iv) the voiceless uvular²⁸⁹ fricative /x/, quite common in Persian and Arabic loanwords, shows a strong tendency towards merging with the aspirated plosive /k^h/ of genuine Indic origin and, more often than not, is substituted by it in Hindustani, e.g.:

ख़त्म (*khatm*), ختم (*x-t-m*) [xətm] ~ [k^hət(ə)m] 'end'
 नाखुन (*nākhun*), ناخن (*n-A-x-n*) [na:xun] ~ [na:k^hʊn] 'nail'
 फ़ारख़ (*farākh*), فراخ (*f-r-A-x*) [fəra:x] ~ [fəra:k^h] 'open, generous'

The labialization of the velar fricative /x^w/, which was discontinued in Persian already in the 13th-14th c., is sometimes retained in Hindustani,²⁹⁰ e.g.:

ख़ाव (*khvāb*), خواب (*x-V-A-b*) [xvə:b] ~ [xə:b] 'dream'
 ख़ाहिश (*khvāhiś*), خواہش (*x-V-A-h-ṣ*) [xvə:hiʃ] ~ [xə:hiʃ] 'wish, desire'

(v) the labio-dental voiced fricative /v/ in Perso-Arabic loanwords in Hindustani is commonly pronounced either like a

²⁸⁸ Although replacing /ʃ/ with /s/ in speech is a widespread phenomenon, especially among people belonging to lower castes and rural residents, it never occurs in script and even in simplified Roman notation (like that which is used by the Internet surfers) /ʃ/ is always transcribed as *sh*, not *s*, e.g.: *sha(a)di*, *dushman*, *pareshani*.

²⁸⁹ In various sources this consonant is labelled as velar (cf. McGregor 1984: xxii; Shapiro 2003: 260), but it is rather retracted and with a uvular scrape. Schmidt classifies this fricative as dorso-velar (Schmidt 2003: 308).

²⁹⁰ Cf. also 2.3.19.

labio-dental or bilabial approximant [v]²⁹¹ or like the bilabial fricative [β] e.g.:

वतन (*vat'n*), ڀٽ (*V-t-n*) [vət(ə)n] ~ [βət(ə)n] 'country'

जवाब (*javāb*), جواب (*j-v-A-b*) [ʤəva:b] ~ [ʤəβa:b] 'answer'

In Hindustani, the labiodental fricative /v/ occurs at the end of a word only sporadically, mainly in *tatsama* words borrowed from Sanskrit and it is commonly replaced by the vowel /o/ or the diphthong /au/ ~ /ɔ/ in colloquial speech. In a few Perso-Arabic loanwords ending with /v/ an analogous process is observed, e.g.:

सरो (*saro*), سرو (*s-r-V*) [səro] ~ [səɔ] 'cypress'

नगाव (*nar'gāv*), نرگاز (*n-r-g-A-V*) [nərga:ɔ] ~ [nərgɔ] 'bull'

दवाव (*davādav*), دواو (*d-V-A-d-v*) [dəvadau] ~ [dəvado] 'running'

3.2.7. The Persian dento-alveolar nasal /n/ that occurs in the Perso-Arabic loanwords adjusted to the phonological system of Hindustani as follows:

(i) it remains unchanged when occurring: 1. at the beginning of a word or syllable; 2. at the end of a word or syllable when preceded by a short vowel; 3. in some Persian and Arabic loanwords if it is not immediately followed by the bilabial stops /b/ or /p/,²⁹² e.g.:

²⁹¹ Phillot noticed that the bilabial realization of /v/, resembling the approximant in the English words "we" or "went", was current "amongst Arabs, Afghans and Indians", while for Persians it was "a consonant like v" (Phillot 1919: 18; cf. also Platts 1909: 8). Probably not without significance in this case is also the fact that in the classical period the modern Persian labio-dental fricative /v/ was a semivowel and did not belong to the fricatives (Pisowicz 1985: 118). In contemporary Persian, such old pronunciation is observed, for example, in Dari (cf. Henderson 1975: 653).

²⁹² There is an observable tendency also in this group of loanwords to shift the pronunciation of *nūn* in consonant clusters towards nasalization (or

नेक (*nek*), نیک (*n-Y-k*) [nek] 'good'

तन (*tan*), تن (*t-n*) [tən] 'body'

बंदूक (*bandūk*), بندوق (*b-n-d-V-q*) [bənduːq] 'rifle'

(ii) depending on the following consonant, /n/ can be realized as one of its positionally conditioned allophones: 1. it becomes velarized to [ŋ] before /k/, /g/ and /q/; 2. it is palatalized to [ɲ] before /tʃ/ and /dʒ/, e.g.:

नेक (*nek*), نیک (*n-Y-k*) [nek] 'good'

तंगी (*taṅgī*), تنگی (*t-n-g-Y*) [təŋgiː] 'narrowness'

इनकलाव (*in'kalāb*), انقلاب (*A-n-q-l-A-b*) [ɪŋqəlaːb] ~ [ɪŋkɪlaːb] 'revolution'

गुचा (*guñcā*), غنچه (*ġ-n-ċ-h*) [ɣuŋtʃa] 'a bud; blossom'

अंजाम (*añjām*), انجام (*A-n-j-A-m*) [əɳdʒaːm] 'end'

(iii) before bilabial stops /p/ and /b/ the nasal /n/ is realized phonologically as [m], e.g.:

तानपूर (*tān'pūrā*), تنپور (*t-n-b-V-r-h*) [taːmpuːraː] ~ [taːmbuːra] 'kind of lute'

गुंबद (*gumbad*), گنبد (*g-n-b-d*) [gumbəd] 'dome'

(iv) before the labio-dental fricative /f/ the nasal /n/ is realized as the labiodental nasal [ɱ], e.g.:

सिफ़ (*simf*), صنف (*ʃ-n-f*) [simf] 'kind, sort'

सौफ़ (*saumf*), سونف (*s-V-n-f*) [saumf] 'anise seed'

(v) in the middle and at the end of the word, the nasal /n/ can be elided and substituted by the nasalization of the vowel when:

pronouncing it both ways); this tendency is confirmed by the way in which the newer lexical sources transliterate this character, e.g.: چند (≤ P: Platts: *ĉand*; STCD: *ĉaṇd* 'several'); منجمد (≤ A: Platts: *munjamid*; STCD: *munjamid* 'congealed'); غنچه (≤ P: Shak. *guñcāh*; STCD: *guñcāh* 'bud').

1. it is quiescent and preceded immediately by a long vowel; 2. it is quiescent and immediately followed by a quiescent letter, e.g.:

आसमान (*ās'mān*), آسمان (*A-s-m-A-n*) [asman]²⁹³ ~ [asmā:] 'heaven'

मुंशी (*muṃśī*)²⁹⁴, منشی (*m-n-š-Y*) [munʃi] ~ [mūʃi] 'clerk, writer'

خ xē ख़े	ح barī hē वड़ी हे	چ cē चे	ج jīm जोम	ث ṣē से	ث iē टे	ت tē ते	پ pē पे	ب bē बे	ا alif अलिफ़
ص ṣuād मुआद	ش ṣīn शीन	س sīn सीन	ژ žē झे	ز zē जे	ڑ rē डे	ر rē रे	ڑ zāl ज़ाल	ڈ dāl डाल	د dal दाल
ل lām लाम	گ gāf गाफ़	ک kāf काफ़	ق qāf काफ़	ف fē फे	غ gān गैन	ع ān ऐन	ظ zōē जोए	ط tōē तोए	ض zuād जुआद
		ے barī yē वड़ी ये	ی chōtī yē छोटी ये	ء hamzah हमज़ा	ہ dō čaśmī hē दो चश्मी हे	ہ chōtī hē छोटी हे	و wāō वाओ	ن nūn नून	م mīm मीम
۰ ṣīfār सिफ़ार	۱ ēk एक	۲ dō दो	۳ tīn तीन	۴ čār चार	۵ pāñc पांच	۶ chē छे	۷ sāt सात	۸ āth आठ	۹ nō नौ

Fig. 8. "Learn Urdu through Hindi" – a table presenting *nasta'liq* script with names of the characters transliterated in Nagari and Roman scripts – popular teaching material on various "Teach-yourself-Urdu" webpages. (Source: Goldsztajn/Wikipedia)

²⁹³ Alternative spelling with *nūn-e gunnah*: آسمان (*A-s-m-A-n*). Cf. 2.2.7.

²⁹⁴ Alternative spelling मुंशी (*muṃśī*).

3.3. Quantitative and Qualitative Change of Vowels in the Perso-Arabic Loanwords

I shall feel amply rewarded if I succeed in inducing others to feel and enjoy in some measure the chaotic beauty of this phonologic system

(Kelkar 1968: 85)

3.3.1. Quantitative changes are relatively rare apart from the regular and morphophonemically motivated lengthening of the final vowel in Persian loanwords (mostly nouns and adjectives) which end with the 'silent' *h* (*vide supra* 2.3.3., fn. 174). In these cases the short vowel /æ/²⁹⁵ which preceded the 'silent' *h* and was not marked in writing, became elongated in Hindustani and started to be realized as [a:]²⁹⁶ e.g.: H जमाना (*zamānā*), زمانه (*z-m-A-n-h*) /zəmana/²⁹⁷ m. 'time' ≤ Class. P زمانه (*z-m-A-n-h*) /zæmanæ/ (modern P /zæmanɛ/) 't.s.'; H रास्ता (*rāstā*), راه (*r-A-s-t-h*) /rasta/ m. 'road' ≤ Class. P راسته (*r-A-s-t-h*) /rastæ/ (modern P /rastɛ/)

²⁹⁵ In modern Iranian Persian the vowel preceding the 'silent' *h* is realized as [ɛ], however, in Middle and early New Persian it was pronounced as /a/ (cf. Perry 2007: 979) and this old pronunciation has been retained, for example, in Dari, where it is continued as [æ] (Henderson 1975: 651) and in Persian spoken in India (cf. Phillot 1919: 19). In the following examples, if there is any difference in terms of form, then both Classical (quoted after Stein.) and modern Persian (quoted after Hayyim) forms are given because the differences in their phonological realization help to clearly indicate a source form and, at the same time, determine the chronology of the analyzed loanwords.

²⁹⁶ This is a result of the morphologization of loans (cf. Beskrovnyĭ 1960a: 129-131) which may also be called inflectional hybridization (Kuczkiewicz-Fraś 2003b: 76).

²⁹⁷ The final /-a/ may be realized phonetically in a slightly different manner, varying from full long [a:] to the neutral vowel [ə]. Since the difference is not phonemic, for reasons of convenience it is always recorded as /-a/ in the phonological notation.

‘t.s.’;²⁹⁸ H मासूका (*māsūkā*), معسوقه (*m-^c-š-ū-q-h*) /mæsuqa/ f. ‘mistress’ ≤ Class. P [≤ A] معسوقه (*m-^c-š-ū-q-h*) /mæsuqæ/ (modern P /mæsuqe/) ‘t.s.’; H ताजा (*tāzā*), تازه (*t-A-z-h*) /taza/ adj. ‘fresh’ ≤ Class. P تازه (*t-A-z-h*) /tazæ/ (modern P /tazε/) ‘t.s.’. A similar change (lengthening of the final vowel) takes place when the ‘silent’ *h* is preceded by *‘ain* (which in Persian is either realized as a glottal stop²⁹⁹ or dropped entirely in this position), e.g.: H दफ़ा (*dafā*), دفة (*d-f-^c-h*) /dəfa/ f. ‘time, occasion’ ≤ Class. P [≤ A] دفعة (*d-f-^c-i*) /dæfʔæ(t)/ (modern P دفعة (*d-f-^c-h*) /dæfʔε/) ‘t.s.’; H बज़रिया (*bazāriā*), بزریه (*b-z-r-Y-^c-h*) [bəza:ɾiːˈaː] adv. ‘by means (of)’ ≤ Class. P [≤ A] بزریعه (*b-z-r-Y-^c-h*) [bəza:ɾiːˈæ] (modern P [bəza:ɾiːˈε]) ‘t.s.’. It is worth noting that in the latter example the lengthening of the final (stressed) /a/ is accompanied by the shortening of the preceding vowel /i/ > /ɪ/ which is clearly visible in the Nagari notation of this word.

3.3.2. Other quantitative changes should rather be regarded as alternations, since they always occur in alternative forms of loanwords, parallel to forms retaining the original vocalization – even though the latter, being more conservative and puristic, sometimes are definitely less common, especially in colloquial speech. Cf. the following examples:

(i) P /a/ > H /ə/, e.g.: H बागीचा (*bagīcā*) ~ बागीचा (*bāgīcā*), باغچه (*b-A-ġ-Y-ċ-h*) /bəyitʃa/ ~ /bəyitʃa/ ‘small garden’ ≤ Class. P باغچه (*b-A-ġ-ċ-h*) /bəyitʃæ/ (modern P /bəyitʃε/) ‘t.s.’;³⁰⁰ H तassa (*tāsā*) ~ तassa (*taśā*), تاسا (*t-A-s-h*) /taśa/ ~ /təsā/ ~ /tasa/ ‘a kind of drum’ ≤ P

²⁹⁸ This word is one of many examples of semantic narrowing, or restriction of meaning in a lexical item (Crystal 2008: 320) which is a common feature of Perso-Arabic loanwords in Hindustani. *Vide infra* 5.1.2.

²⁹⁹ I.e. a catch in the breath. As Mace observes, in modern Persian it is “a very weak sound, barely audible in the speech of most Iranians” (Mace 2003: 28)

³⁰⁰ The alternative form باغچه (*b-A-ġ-ċ-h*) /bəyitʃa/, identical to the Persian source word, occurs very rarely in Hindustani/Urdu (cf. STCD 95).

[≤ A] طاسه (*t-A-s-h*) /tasæ/ (modern P /tase/) ‘t.s.’; H बलिश्त (*bālišṭ*) ~ वलिश्त (*baliṣṭ*), باشت (*b-A-l-ṣ-t*) ~ بلشت (*b-l-ṣ-t*) /bəliʃt/ ~ /bəliʃt/ ‘a span of the hand’ ≤ P بالشت (*b-A-l-ṣ-t*) /bəliʃt/ ‘t.s.’;³⁰¹ H कलवूत (*kal'būt*) ~ कालवूत (*kāl'būt*), کالوت (*k-A-l-b-V-t*) /kəlbud/ ~ /kalbut/ ‘mould, frame’ ≤ P کالبد (*k-A-l-b-d*) /kalbud/ ~ /kalbæd/ ‘t.s.’;³⁰²

(ii) P /u/ > H /u/, e.g.: H कलवूत (*kal'būt*) ~ कालवूत (*kāl'būt*), کالوت (*k-A-l-b-V-t*) /kəlbud/ ~ /kalbut/ ‘mould, frame’ ≤ P کالبد (*k-A-l-b-d*) /kalbud/ ~ /kalbæd/ ‘t.s.’; H कँगुरा (*kaṁgūrā*), کنگره (*k-n-g-r-h*) ~ کنگوره (*k-n-g-V-r-h*) /kəṁgura/ ~ /kəṁgura/ ~ /kuṁgura/ ~ /kəṁgura/ ‘parapet’ ≤ Class. P کنگره (*k-n-g-r-h*) /kuṁgura/ (modern P /kongərə/) ‘t.s.’;

(iii) P /i/ > H /i/, e.g.: H मिल्कियत (*milkiyat*), ملکیت (*m-l-k-Y-t*) /milkijæt/ ~ /milkij:æt/ ‘property’ ≤ Class. P [≤ A] ملکیت (*m-l-k-Y-t*) /milkijæt/ (modern P /molkij:æt/ ~ /melkij:æt/) ‘t.s.’;

(iv) P /ɪ/ > H /i/, e.g.: H मिरज़ा (*mir'zā*) ~ मीरज़ा (*mīr'zā*), مرزا (*m-r-z-A*) /mirza/ ~ /mirza/ ‘Persian and Mughal title: prince; Mr.’ ≤ P مرزا (*m-r-z-A*) /mirza/ ‘t.s.’;³⁰³ H शिकरा (*śik'rā*) ~ सीकरा (*sīk'rā*), شکره (*ṣ-k-r-h*) /ʃikra/ ~ /sikra/ ‘hawk’ ≤ Class. P شکره (*ṣ-k-r-h*) /ʃekæra/ (modern P /ʃekære/) ‘t.s.’.

3.3.3. Monophthongization, or the reduction of a diphthong to a single sound occurs in several cases:

(i) P /aɪ/ > H long mid front [e:], e.g.: H खैमा (*khēmā*) ~ खैमा (*khaimā*), خیمه (*x-Y-m-h*) /xema/ ~ /xaima/ ‘tent’ ≤ Class. P خیمه

³⁰¹ Probably this change in the length of the vowel /i/ was caused by mixing this word with a similar Persian form بلیست (*b-l-s-t*) /bīlist/ ‘a span’, in which both the vowels /i/ are short and not marked in writing.

³⁰² The alternative form کالبد (*k-A-l-b-d*) /kalbud/ ~ /kalbæd/, identical to the Persian source word, occurs only in Hindustani/Urdu in the meaning ‘(human) body’ (cf. STCD 475).

³⁰³ According to Steingass, this word comes from میرزا (*m-Y-r-z-A*) ‘a prince; son of a great lord, noble cavalier, knight; a gentleman; a writer, scribe’ (cf. Stein. 1361) ≤ A امیر (*A-m-Y-r*) ‘an emperor, prince, lord, governor, chief, leader [etc.]’ + زه (*z-h*) ‘offspring of men (or of cattle)’.

(*x-Y-m-h*) /xaimæ/ (modern P /xejmɛ/) 't.s.'; H सुहेल (*suheḷ*) ~ सुहैल (*suhail*), سہیل (*s-h-Y-l*) /suhel/ ~ /suhail/ ~ /sohail/ 'the star Canopus' ≤ P سہیل (*s-h-Y-l*) /suhail/ (modern P /sohejl/) 't.s.'; H शेख (*śekh*), شَیخ (*ṣ-Y-x*) /ʃex/ ~ /ʃaix/ 'a venerable old man' ≤ P [≤ A] شیخ (*ṣ-Y-x*) /ʃaix/ (modern P /ʃɛjx/) 't.s.';³⁰⁴

(ii) P /au/ > H long mid back [o:]: e.g.: H गोणा (*gogā*), غَوَا (*ḡ-V-ḡ-A*) /goga/ ~ /gauga/ 'noise' ≤ P [≤ A] غَوَاء (*ḡ-V-ḡ-A-ʔ*) /qauqæ/ (modern P /qovqa/) 't.s.'; H रोणन (*roḡan*), رَوْنَن (*r-V-ḡ-n*) /rogæn/ ~ /raugæn/ 'grease' ≤ P روغن (*r-V-ḡ-n*) /rauqæn/ (modern P /rovqan/) 't.s.'; H सनोबर (*sanobar*), صَنَبَر (*ṣ-n-V-b-r*) /sənobær/ ~ /sənaubær/ 'conifer' ≤ P صنوبر (*ṣ-n-V-b-r*) /sənaubær/ (modern P /sənaubær/) 't.s.'.

3.3.4. The reverse process of diphthongization happens rarely and can affect both short and long vowels:

(i) P /ɪ/ > H /aɪ/, e.g.: H एलान (*elān*) ~ ऐलान (*ailān*), اعلان (*A-^c-l-A-n*) /elan/ ~ /ailan/ ~ /eʔlan/ 'announcement' ≤ Class. P [≤ A] اعلان (*A-^c-l-A-n*) /ɪʔlan/ (modern P /eʔlan/) 't.s.';

(ii) P /e/ > H /aɪ/, e.g.: H सफ़ेदा (*safedā*) سفیدہ (*s-f-Y-d-h*) /səfedæ/ ~ /sufaida/ 'white lead' ≤ Class. P سفیدہ (*s-f-Y-d-h*) /səfedæ/ 't.s.' (modern P /sæfide/ 'albumen');

(iii) P /u/ > H /aʊ/, e.g.: H फौलाद (*faulād*) فُولَاد (*f-V-l-A-d*) /faulad/ 'steel' ≤ P فولاد (*f-V-l-A-d*) /fulad/ 't.s.'.

3.3.5. Qualitative changes. In Hindustani, the Persian basic vowel /æ/ is substituted with its natural vowel /ə/; since they both differ only slightly in features and perform a similar function in their respective phonological systems, the replacement of /æ/ by /ə/ is not regarded here as an actual sound change.

(i) P /æ/ > H /ɪ/; this change occurs in several dozen loanwords, sometimes as an alternative for forms retaining traditional

³⁰⁴ At least in one example the vowel cluster /a.i/ has been analogically reduced to /e/: H हमेल (*hamel*) ~ हमायल (*hamāyal*), حَمَل (*h-m-A-Y^c-l*) /hæmel/ ~ /həmajəl/ ~ /həma.ɪl/ 'a necklace of flowers, or of rupees' ≤ P [≤ A] حَمَل (*h-m-A-Y^c-l*) /həmaʔɪl/ (modern P /həmaʔel/) 't.s.'

Persian vocalization, e.g.: H अखलाक (*akh'lāk*) ~ इखलाक (*ikh'lāk*), اخلاق (*A-x-l-A-q*) /əɣlaq/ ~ /ɪɣlaq/ 'manners, ethics' ≤ Class. P [≤ A] اخلاق (*A-x-l-A-q*) /æɣlaq/ (modern P /ɛɣlaq/) 't.s.'; H बाशिंदा (*bāšindā*), باشنده (*b-A-š-n-d-h*) /baʃinda/ 'resident' ≤ Class. P باشنده (*b-A-š-n-d-h*) /baʃændæ/ (modern P /baʃændɛ/) 't.s.'; H हिमाकत (*himākāt*), حماقت (*h-m-A-q-t*) /himaqət/ ~ /həmaqət/ 'stupidity' ≤ Class. P [≤ A] حماقت (*h-m-A-q-t*) /həmaqət/ (modern P /həmaqət/) 't.s.'; H जिरह (*jirah*), جز (*j-r-h*) /dʒirəh/ ~ /dʒirh/ ~ /dʒərh/ 'cross-questioning; wound' ≤ P [≤ A] جز (*j-r-h*) /dʒərh/ 't.s.'; H किसब (*kisab*), کسب (*k-s-b*) /kɪsəb/ ~ /kəsɪb/ 'gain' ≤ P [≤ A] کسب (*k-s-b*) /kəsɪb/ 't.s.'; H किला (*kilā*), قلعه (*q-l-ʿ-h*) /kɪla/ ~ /qəʔla/³⁰⁵ 'fort' ≤ Class. P [≤ A] قلعه (*q-l-ʿ-ī*) /qəʔlæ/ (modern P /qəʔlɛ/) 't.s.'; H परिंदा (*parindā*) ~ परिंद (*parind*), پرند (*p-r-n-d-h*) /pərinda/ ~ /pərɪnd/ 'winged; bird' ≤ Class. P پرند (*p-r-n-d-h*) /pærændæ/ (modern P /pærændɛ/) 't.s.'; H शिकवा (*śikvā*), شکوه (*ś-k-V-h*) /ʃɪkva/ 'complaint' ≤ Class. P [≤ A] شکوه (*ś-k-V-ī*) /ʃækvæ/³⁰⁶ (modern P /ʃækvɛ/) 't.s.';

(ii) P /a/ > H /ɪ/; this change appears in only one example and is one of the few essential modifications concerning this loanword: H हावन (*hāvan*) ~ हिमाम (*himām*), هاون (*h-A-V-n*) /havən/ ~ /himam/ 'mortar' ≤ P هاون (*h-A-V-n*) /havæn/ 't.s.'; cf. also: H इमाम-दस्ता (*imām-dastā*), امام-دست (*A-m-A-m d-s-t-h*) ~ امام-دست (*h-m-A-m d-s-t-h*) ~ امام-دست (*h-A-V-n d-s-t-h*) /ɪmamdəsta/ ~ /himamdəsta/ ~ /havəndəsta/ 'mortar and pestle' ≤ Class. P هاون-دسته (*h-A-v-n d-s-t-h*) /havæn dastæ/ 't.s.';

³⁰⁵ In Hindustani such pronunciation is rare and regarded as 'pedantic' (cf. STCD 468), used solely in highly Persianized speech. Note that in Persian the word قلعه /qɪʔlɛ/ has a different meaning, namely, 'the half' (Stein. 985).

³⁰⁶ Steingass also gives the nativized Persian expression گله شکوه (*g-l-h ś-k-V-h*) of the same meaning ('complaint') but different vocalization: /gɪlɛ ʃɪkɪvɛ/ (Stein. 1096).

(iii) P /æ/ > H /u/, e.g.: H बुज़ (*bugz*), بۇز (*b-ġ-z*) /buʒz/ 'malice' ≤ P [≤ A] بغز (*b-ġ-z*) /bæʒz/ 't.s.'; H जहन्नम (*jahannum*), جہنم (*j-h-n-m*) /dʒəhən:ʊm/ 'hell' ≤ P [≤ A] جهنم (*j-h-n-m*) /dʒæhæn:æm/ 't.s.'; H बरक (*barrāk*) ~ बुरक (*burrāk*), برآق (*b-r-A-q*) /bər:aq/ ~ /bur:aq/ 'flashing' ≤ P [≤ A] براق (*b-r-A-q*) /bær:aq/ 't.s.';³⁰⁷ H मुहब्बत (*muhabbat*), محبت (*m-h-b-t*) /muhəb:ət/ 'love' ≤ P [≤ A] محبت (*m-h-b-t*) /mæhæb:æt/ 't.s.';

(iv) P /æ/ > H /e/;³⁰⁸ this change may happen in two situations. The first is phonemically conditioned and occurs in proximity to the tautosyllabic /h/, e.g.: H रहन (*reh'n*) ~ रहन (*rah'n*), رهن (*r-h-n*) /rehn/ ~ /rəhn/ ~ /rəhn/ 'pledge, mortgage' ≤ P [≤ A] رهن (*r-h-n*) /rəhn/ 't.s.'; H तहसील (*tah'sīl*), تحصیل (*t-h-ṣ-Y-l*) /təhsil/ ~ /tehsil/ 'a revenue district' ≤ P [≤ A] تحصیل (*t-h-ṣ-Y-l*) /təhsil/ 't.s.'. In many cases the pronunciation of the vowel preceding or following /h/ is further lowered and tends toward [æ] or even the slightly diphthongized realization [ɛɪ],³⁰⁹ e.g.: H कहत (*kah't*), قہت (*q-h-t*) [qəht] ~ [qɛht] 'famine' ≤ P [≤ A] قہت (*q-h-t*) /qæht/ 't.s.'; H सहन (*sah'n*), سہن (*ṣ-h-n*) [səhn] ~ [sehn] ~ [seihn] 'inner courtyard' ≤ P [≤ A] صحن (*ṣ-h-n*) /sæhn/ 't.s.'. The second situation results from a morphophonemical change typical for Hindustani which affects the final vowel in the first member of some compounded loanwords, e.g.: H फ़ाक़ेमस्त (*fāk'emast*), فاقہ مست (*f-A-q-h m-s-t*)

³⁰⁷ In Persian (and Arabic) براق (*b-r-A-q*) /bur:aq/ is the name of the celebrated animal, smaller than a mule and bigger than an ass, upon which Muhammad is said to have gone from Jerusalem to heaven (Stein. 168).

³⁰⁸ In Hindustani, the mid central vowel [ə] in proximity to the tautosyllabic /h/ has a front allophone [ɛ], e.g.: कहना (*kahnā*) 'to say' and बहन (*bah'n*) 'sister' are often pronounced [kehna:] and [bəhn] respectively. It must be noted, however, that this alternation is restricted only to the phonemic level and except for a few cases (like *rehn* quoted above), never noted in script (this rule applies both to the genuinely Hindustani words and to the loanwords).

³⁰⁹ Such pronunciation can be observed, for example, among educated urban dwellers in Delhi or Lucknow.

/faqemæst/ ~ /faqamæst/³¹⁰ ‘cheerful though starving’ ≤ Class. P فاقه مست (*f-A-q-h m-s-t*) /faqæmæst/ (modern P /faqemæst/³¹¹) ‘t.s.’; H *ijāredār* (اجاره دار, *A-j-A-r-h d-A-r*) /ɪdʒaredar/ ~ /ɪdʒaradar/ ‘lease-holder’ ≤ Class. P *ajarah dār* (*A-j-A-r-h d-A-r*) /ɪdʒarædar/ (modern P /ɪdʒaredar/) ‘t.s.’; H *išārebāzī* (ایشاره بازی, *A-š-A-r-h b-A-z-Y*) /ɛʃarebazi/ ~ /ɪʃarabazi/ ‘conducting a mute dialogue (of gestures, glances)’ ≤ Class. P *ishārah* (*A-š-A-r-h*) /ɪʃaræ/ ‘pointing...’ + *bāzī* (*b-A-z-Y* /bazi/ ‘play...’; H *zāy’kedār* (زایقده دار, *z-A-Y’-q-h d-A-r*) /za.ɪkedar/ ~ /za.ɪkædar/ ‘savoury’ ≤ Class. P *zānqeh dār* (زانه دار, *z-A-Y’-q-h d-A-r*) /za’ɪkædar/ (modern P /za’ɪkedar/) ‘t.s.’;

(v) P /ɪ/ > H /ə/; this change is very common but it usually appears as an alternative for forms retaining traditional Persian vocalization, e.g.: H *ahātā* (احاطه, *A-h-A-t-h*) /əhata/ ~ /ehata/ ‘enclosure’ ≤ Class. P [≤ A] *ahāṭa* (*A-h-A-t-ī*) /ɪhætæ/ (modern P /æhætæ/) ‘t.s.’; H *ak’sīr* (اکسیر, *A-k-s-Y-r*) /əksir/ ~ /ɪksir/ ‘elixir’ ≤ P [≤ A] *aksir* (*A-k-s-Y-r*) /ɪksir/ ‘t.s.’; H *bahišt* (بهشت, *b-h-š-t*) /bəhiʃt/ ~ /bɪhiʃt/ ~ /behɪʃt/³¹² ‘paradise’ ≤ Class. P *bēšt* (*b-h-š-t*) /bɪhiʃt/ (modern P /bəheʃt/) ‘t.s.’; H *cahal* (چهل, *č-h-l*) /ʃəhəl/ ~ /ʃehəl/ ‘forty’ ≤ Class. P *čehl* (*č-h-l*) /ʃɪhɪl/ ~ /ʃɪhæl/ (modern P /ʃehɪl/) ‘t.s.’; H

³¹⁰ Such pronunciation is characteristic of the Persianized, more conservative register of Hindustani/Urdu.

³¹¹ Note that phonemically both the forms: H /faqemæst/ and modern P /faqemæst/ are very similar to each other, yet this is only secondary similarity (this remark concerns all the examples of this type). The vowel /e/ in Hindustani results from the modification of the inflectional morpheme /a/ while in Persian it is a genuine vowel developed in Early New Persian from the word-final /æ/.

³¹² In this case the alternation /ɪ/ ~ /e/ is justifiable in the light of diachronical change /ɪ/ > /e/ in Persian, cf.: early NP /bɪhiʃt/ (Stein. 211) vs. modern NP /bəheʃt/ (Hayyim I 295). There are many more examples of this type among Perso-Arabic loanwords in Hindustani.

चिकन (*cikan*) چکن (*ĉ-k-n*) /ʃikən/³¹³ ‘embroidered muslin’ ≤ Class. P چکن (*ĉ-k-n*) /ʃækɪn/ ~ /ʃikɪn/ ~ /ʃikin/ ‘t.s.’ (modern P /ʃækən/); H درخت (*darakh̄t*), درخت (*d-r-x-t*) /dəræxt/ ‘tree’ ≤ Class. P درخت (*d-r-x-t*) /dɪræxt/ (modern P /dəræxt/ ‘t.s.’; H गिरह (*girah*), گره (*g-r-h*) /gɪrəh/ ~ /gɪreh/ ‘knot’ ≤ Class. P گره (*g-r-h*) /gɪrɪh/ (modern P /gɪrɛh/) ‘t.s.’; H हादसा (*hād’sā*), حادثه (*h-A-d-ṣ-h*) /had(ə)sa/ ~ /hadɪsa/ ‘accident’ ≤ Class. P [≤ A] حادثه (*h-A-d-ṣ-ī*) /hadɪsæ/ (modern P /hadɛsɛ/) ‘t.s.’; H इंतज़ार (*intazār*) ~ इंतिज़ार (*intizār*), انتظار (*A-n-t-ṣ-A-r*) /ɪntəzar/ ~ /ɪntizar/ ‘waiting’ ≤ Class. P [≤ A] انتظار (*A-n-t-ṣ-A-r*) /ɪntizar/ (modern P /ɛntɛzar/) ‘t.s.’; H खज़ाना (*k̄hazānā*), خزانه (*x-z-A-n-h*) /ɣəzana/ ~ /ɣɪzana/ ‘treasury’ ≤ Class. P [≤ A] خزانه (*x-z-A-n-h*) /ɣɪzanæ/ (modern P /ɣɛzanɛ/) ‘t.s.’; H नायब (*nāyab*), نائب (*n-A-Y²-b*) /najəb/ ~ /na.ɪb/ ‘deputy’ ≤ Class. P [≤ A] نائب (*n-A-Y²-b*) /naʔɪb/ (modern P نائب (*n-A-Y-b*) /najæb/ ‘t.s.’.³¹⁴ The phonological change /ɪ/ > /ə/ can also occur in words with ^cain, in Persian commonly pronounced as a vowel, e.g.: H अस्मत् (*asmat*), عِصْمَت (*ʿ-ṣ-m-t*) /əsmət/ ~ /ɪsmət/ ‘chastity’ ≤ Class. P [≤ A] عِصْمَة (*ʿ-ṣ-m-ī*) /ɪsmæt/ (modern P /ɛsmæt/) ‘t.s.’; H इश्ताल (*istāāl*), اشتعال (*A-ṣ-t-ʿ-A-l*) /ɪʃtə.al/ ~ /ɪʃtɛ.al/ ‘inflaming’ ≤ Class. P [≤ A] اشتعال (*A-ṣ-t-ʿ-A-l*) /ɪʃtɪʔal/ (modern P /ɛʃtɛ.al/ ~ /ɛʃtɛʔal/) ‘t.s.’; H वाकई (*vākai*), واقعی (*v-A-q-ʿ-Y*) /vaqə.i/ ~ /vaqe.i/ ‘real; really’ ≤ Class. P [≤ A] واقعی (*v-A-q-ʿ-Y*) /vaqɪʔɪ/ (modern P /vaqeʔɛ/) ‘t.s.’;

³¹³ Even though this word has three differently vocalized versions in Persian (including the alternation short /ɪ/ ~ long /i/), it occurs in only one form in Hindustani, with additional change of sound /ɪ/ > /ə/.

³¹⁴ Note that the phonetic realizations of this word in Hindustani (in most cases) and in modern Persian are almost identical, with the glide /j/ which is pronounced instead of the glottal stop and, in both the cases, may also be marked in script. However, the Hindustani/Urdu orthography, contrary to the modern Persian notation, retains the original Arabic spelling with *hamzah*.

(vi) P /i/ > H /e/, e.g.: H बेहतर (*beh'tar*), بَہتر (*b-h-t-r*) /behtər/ 'better' ≤ Class. P بہتر (*b-h-t-r*) /bihtər/ (modern P /behtər/) 't.s.'; H एहसान (*eh'sān*), احسان (*A-h-s-A-n*) /ehsan/ ~ /ihsan/ 'favour, kindness' ≤ Class. P [≤ A] احسان (*A-h-s-A-n*) /ihsan/ (modern P /ehsan/) 't.s.'; H एहतमाम (*eh'tamām*), اہتمام (*A-h-t-m-A-m*) /ehtāmam/ ~ /ehtimam/ 'care' ≤ Class. P [≤ A] اہتمام (*A-h-t-m-A-m*) /ihtimam/ (modern P /ehtāmam/) 't.s.'; H मेहमान (*meh'mān*), مہمان (*m-h-m-A-n*) /mehmən/ 'guest' ≤ P مہمان (*m-h-m-A-n*) /mihman/ (modern P /mehman/) 't.s.';³¹⁵ H मेहनत (*meh'nat*), محنت (*m-h-n-t*) /mehnət/ 'hard work' ≤ Class. P [≤ A] محنت (*m-h-n-t*) /mihnæt/ (modern P /mehnæt/) 't.s.'; H मेहराब (*meh'rab*), محراب (*m-h-r-A-b*) /mehrāb/ 'an arch' ≤ Class. P [≤ A] محراب (*m-h-r-A-b*) /mihrab/ (modern P /mehrāb/) 't.s.'; H सेहत (*sehat*), صحت (*ṣ-h-t*) /seh(:)ət/ 'health' ≤ Class. P [≤ A] صحت (*ṣ-h-t*) /sih:ət/ (modern P /seh:æt/) 't.s.'. The phonological change /i/ > /e/ can also occur in words with *ain*, in Persian commonly pronounced as a vowel, e.g.: H जाया (*jāyā*) ~ जाए (*zāe*),³¹⁶ ضائع (*z-A-Y^o-^c*) /za.e/ 'lost' ≤ Class. P [≤ A] ضائع (*z-A-Y^o-^c*) /za.iʔ/ ~ /zaʔiʔ/ (modern P /za.ɛʔ/) 't.s.'; H मजारा (*majārā*), مزارع (*m-z-A-r-^c*) /mədʒara/ ~ /muzare/ 'peasant' ≤ Class. P [≤ A] مزارع (*m-z-A-r-^c*) /muzariʔ/ (modern P /mozareʔ/ ~ /zareʔ/) 't.s.'; H शेर (*śer*), شعر (*ṣ-^c-r*) /fer/ ~ /fɛʔr/ 'poetry' ≤ Class. P [≤ A] شعر (*ṣ-^c-r*) /fiʔr/ (modern P /fɛʔr/) 't.s.'; H एतबार (*et'bār*), اعتبار (*A-^c-t-b-A-r*) /etbar/ ~ /eʔtibar/ 'confidence' ≤ Class. P [≤ A] اعتبار (*A-^c-t-b-A-r*) /iʔtibar/ (modern P /ɛʔtebar/) 't.s.'; H इस्तेमाल (*istemāl*), استعمال (*A-s-t-^c-m-A-l*) /istemal/ ~ /isteʔmal/ 'use, practice' ≤ Class.

³¹⁵ Steingass gives another form of this word, with long /i/, namely: میهمان (*m-Y-h-m-A-n*) /mihman/ 'a guest' (Stein. 1365), but it is not confirmed by Hayyim in this sense (cf. میهمان (*m-Y-h-m-A-n*) /mihman/ 'etc., and so forth', Hayyim II 1046).

³¹⁶ Maddāh 240. जाया /dʒaja/ is also the regular perf. part. of the verb *jānā* 'to go' (although used only in comp.); so far-reaching assimilation of this loanword was probably due to the semantic proximity of both the words (H /dʒaja/ 'gone' ≈ P /za.iʔ/ 'lost; perished').

P [$\leq$ A] استعمال (*A-s-t-^c-m-A-l*) /ɪstɪʔmal/ (modern P /ɛsteʔmal/) 't.s.'. The Persian enclitic 'of' (*izāfat*), realized phonetically as short *i* /ɪ/ in Class. Persian, is sometimes pronounced as /e/ in Hindustani,³¹⁷ e.g.: H तपेदिक (*tapedik*), تپ دق (*t-p d-q*) /təpedɪq/ 'pulmonary tuberculosis' $\leq$ Class. P تپ دق (*t-p d-q*) /təpɪdɪqq/ (modern P /tæbɛdɛy/) 't.s.'; H कोह-नूर (*koh-nūr*), کوه نور (*k-V-h n-V-r*) /kohnur/ $\sim$ /kohenur/ 'the Koh-i-Noor diamond' $\leq$ Class. P کوه نور (*k-V-h n-V-r*) /kohɪnur/ (modern P /kuhenur/) 't.s.';

(vii) P /i/ > H /e/, e.g.: H चेचक (*cecak*), چچک (*č-Y-č-k*) /tʃɛtʃək/ 'small-pox' $\leq$ P چچک (*č-Y-č-k*) /tʃɪtʃək/ 't.s.'; H दिलेर (*diler*), دلیر (*d-l-Y-r*) /dɪlɪr/ $\sim$ /dɛlɪr/ 'bold, brave' $\leq$ Class. P دلیر (*d-l-Y-r*) /dɪlɪr/ (modern P /delɪr/) 't.s.'; H हमेशा (*hamešā*), ہمیشہ (*h-m-Y-š-h*) /həmeʃa/ 'always' $\leq$ Class. P همیشه (*h-m-Y-š-h*) /həmiʃæ/ (modern P /həmiʃɛ/) 't.s.'; H ग़रेबान (*garebān*), گریان (*g-r-Y-b-A-n*) /gəreban/ $\sim$ /gireban/ 'collar' $\leq$ Class. P گریان (*g-r-Y-b-A-n*) /gɪriban/ (modern P /gɛriban/) 't.s.'; H मेवा (*mevā*), میوه (*m-Y-V-h*) /meva/ 'fruit' $\leq$ Class. P میوه (*m-Y-V-h*) /mivæ/ $\sim$ /maivæ/ (modern P /mivɛ/) 't.s.'; H नेस्त (*nest*), نیست (*n-Y-s-t*) /nest/ 'non-existent' $\leq$ P نیست (*n-Y-s-t*) /nist/ 'it is not'; H ریش (*rešā*), ریشہ (*r-Y-š-h*) /reʃa/ 'fibre' $\leq$ Class. P ریشہ (*r-Y-š-h*) /rɪʃæ/ (modern P /rɪʃɛ/) 't.s.'; H ریشمی (*reš'mī*), ریشمی (*r-Y-š-m-Y*) /reʃ(ə)mi/ 'silken' $\leq$ Class. P ریشمی (*r-Y-š-m-Y*) /rɪʃmi/ (modern P /abriʃomi/) 't.s.'; H सरेस (*sares*), سریش (*s-r-Y-š*) /səres/ $\sim$ /səreʃ/ 'glue' $\leq$ Class. P سریش (*s-r-Y-š*) /sɪrɪʃ/ (modern P /serɪʃ/) 't.s.'; H तवेल (*tavelā*), طویله (*t-v-Y-l-h*) /təvela/ 'stable, stall' $\leq$ Class. P طویله (*t-v-Y-l-h*) /təvilæ/ (modern P /tavilɛ/) 't.s.'. The change /i/ > /e/ also occurs in words with *ʔalif maqṣūrah*, e.g.: H माने (*māne*), معنی (*m-^c-n-Y*) /mane/ $\sim$ /mani/ $\sim$ /maʔna/³¹⁸ 'meaning' $\leq$ Class. P [$\leq$ A] معنی (*m-^c-n-Y*) /maʔni/ (modern P /maʔni/ $\sim$ /maʔna/) 't.s.';

³¹⁷ More on the *izāfat* construction *vide infra* 4.5.5.ff.

³¹⁸ Such utterance (compatible with the Arabic pronunciation of *ʔalif maqṣūrah*) is characteristic of a Persianized, more conservative register of Hindustani/Urdu and regarded as 'pedantic' (cf. STCD 597).

(viii) P /ɪ/ > H /ʊ/, e.g.: H चुनिंदा (*cunindā*),³¹⁹ چنیده (*č-n-Y-d-h*) /tʃʊnɪdā/ ~ /tʃʊnɪdā/ 'chosen; gathered' ≤ P چنیده (*č-n-Y-d-h*) /tʃɪnɪdæ/³²⁰ 't.s.'; H शगूफ़ा (*śagūfā*), شگوفه (*š-g-V-f-h*) /ʃəgufā/ ~ /ʃigufā/ ~ /ʃugufā/ 'bud, blossom' ≤ Class. P شگوفه (*š-g-V-f-h*) /ʃigufæ/ (modern P /ʃokufæ/) 't.s.';

(ix) P /ʊ/ > H /ə/, e.g.: H चुगल (*cugal*), چغل (*č-ğ-l*) /tʃʊyəl/ ~ /tʃʊyɪ/³²¹ 'informer' ≤ Class. P چغل (*č-ğ-l*) /tʃʊyul/ (modern P /tʃʊyol/) 't.s.'; H बदसलूकी (*bad'salūkī*), بدسلوکی (*b-d-s-l-V-k-Y*) /bædsəluki/ 'misbehaviour' ≤ Class. P بدسلوکی (*b-d-s-l-V-k-Y*) /bædsuluki/ (modern P /bædsoluki/) 't.s.'; H دروغ (*darog*), دروغ (*d-r-V-ğ*) /dəroɣ/ 'falsehood' ≤ Class. P دروغ (*d-r-V-ğ*) /duroɣ/ (modern P /doruɣ/) 't.s.'; H फ़तूर (*fatūr*) ~ फ़ितूर (*fitūr*) ~ फ़ुतूर (*futūr*), فتور (*f-t-V-r*) /fətur/ ~ /fitur/ ~ /futur/ 'defect' ≤ Class. P [≤ A] فتور (*f-t-V-r*) /futur/ (modern P /fotur/) 't.s.'; H جمہور (*jam'hūr*), جمهور (*j-m-h-V-r*) /dʒəmhur/ ~ /dʒumhur/ 'populace' ≤ Class. P [≤ A] جمهور (*j-m-h-V-r*) /dʒumhur/ (modern P /dʒomhur/) 't.s.'; H گواہ (*gavāh*), گواه (*g-V-A-h*) /gəvāh/ 'witness' ≤ Class. P گواه (*g-V-A-h*) /gʊvāh/ (modern P /govāh/) 't.s.'; H کراول (*karāval*), قراول (*q-r-A-v-l*) /qəraʊl/ 'sentinel' ≤ Class. P [≤ T] قراول (*q-r-A-v-l*) /qəraʊl/ (modern P /ɣəraʊl/) 't.s.'; H مزدور (*maz'dūr*), مزدور (*m-z-d-V-r*) /mæzdur/ ~ /muzdur/ 'labourer' ≤ Class. P مزدور (*m-z-d-V-r*) /muzdur/ ~ /muzdvær/ (modern P /mozdur/) 't.s.'; H وصول (*vasūl*), وصول (*V-š-V-l*) /vəsul/ ~ /vusul/ 'acquired' ≤ Class. P [≤ A] وصول (*V-š-V-l*) /vusul/ (modern P /vosul/) 't.s.'. The phonological change /ʊ/ > /ə/ can also occur in words with *‘ain*, in Persian commonly pronounced as a vowel, e.g.: H अजूबा (*ajūbā*), آجوبه

³¹⁹ The spontaneous nasalization that occurs in the first syllable of this loanword is accompanied by the shortening of the long vowel /i/ in the following syllable.

³²⁰ This is the perf. part. of P *čīnīdan* 'to gather; to admonish; to drop'.

³²¹ As a result of the process of schwa deletion.

(^c-j-V-b-h) /əɖʒuba/³²² 'wonderful; miracle' ≤ Class. P [≤ A] عجوبه
 (^c-j-V-b-h) /ʊɖʒubæ/ 't.s.'; H शऊर (*śāūr*), شعور (*š-^c-V-r*) /ʃə.ur/ ~
 /ʃo.ur/ 'knowledge' ≤ Class. P [≤ A] شعور (*š-^c-V-r*) /ʃuʔur/ (mod-
 ern P /ʃoʔur/) 't.s.'; H तवाज़ा (*tavāzā*), تواضع (*t-V-A-z-^c*) /təvaza/ ~
 /təvazoʔ/ 'hospitality' ≤ Class. P [≤ A] تواضع (*t-V-A-z-^c*) /təvazuʔ/
 (modern P /təvazoʔ/) 't.s.';

(x) P /u/ > H /a/, e.g.: H मातवर (*māt'bar*) ~ मोतवर (*mot'bar*), معتبر
 (*m-^c-t-b-r*) /matbər/ ~ /mot(ə)bər/ 'reliable' ≤ Class. P [≤ A] معتبر
 (*m-^c-t-b-r*) /muʔtəbər/ (modern P /moʔtəbər/) 't.s.';

(xi) P /u/ > H /i/, e.g.: H फ़तूर (*fatūr*) ~ फ़ितूर (*fitūr*) ~ फ़ुतूर (*fu-
tūr*), فتور (*f-t-V-r*) /fətur/ ~ /fitur/ ~ /futur/ 'defect' ≤ Class. P [≤ A]
 فتور (*f-t-V-r*) /futur/ (modern P /fotur/) 't.s.'; H गिरोह (*giroh*), گروه
 (*g-r-V-h*) /giroh/ ~ /guroh/ 'group, troop' ≤ Class. P گروه (*g-r-V-h*)
 /guroh/ (modern P /goruh/) 't.s.';

(xii) P /u/ > H /o/, e.g.: H मोहलत (*moh'lat*), مهلت (*m-h-l-t*)
 /mohlət/ ~ /muhlət/ 'delay' ≤ Class. P [≤ A] مهلة (*m-h-l-ī*)
 /muhlæt/ (modern P /mohlæt/) 't.s.'; H मोतवर (*mot'bar*) ~ मातवर
 (*māt'bar*), معتبر (*m-^c-t-b-r*) /motbər/ ~ /matbər/ ~ /muʔtəbər/ 'reli-
 able' ≤ Class. P [≤ A] معتبر (*m-^c-t-b-r*) /muʔtəbər/ (modern P
 /moʔtəbər/) 't.s.'; H आहदा (*oh'dā*), عهده (*^c-h-d-h*) /ohda/ 'office'
 ≤ Class. P [≤ A] عهدة (*^c-h-d-ī*) /ʊhdæ/ (modern P /ohdɛ/) 't.s.'; H रोब
 (*rob*) ~ रोआव (*roāb*), رعب (*r-^c-b*) /rob/ ~ /ro.ab/ ~ /roʔob/ 'fear'
 ≤ Class. P [≤ A] رعب (*r-^c-b*) /rub/ ~ /ruʔub/ (modern P /roʔb/) 't.s.';
 H सोहबत (*soh'bat*), صحب (*š-h-b-t*) /sohbət/ 'companionship'
 ≤ Class. P [≤ A] صحبة (*š-h-b-ī*) /suhbæt/ (modern P /sohbæt/) 't.s.';
 H तोहफा (*toh'fā*), تحفة (*t-h-f-h*) /tohfə/ 'gift' ≤ Class. P [≤ A]
 تحفة (*t-h-f-ī*) /tuhfæt/ ~ /tuhæfæt/ ~ /tuhfæ/ (modern P /tohfɛ/) 't.s.';

³²² In higher register of Hindustani/Urdu this form is regarded as a colloquial and corrupted variant of the original Arabic العجوبة (*A-^c-j-V-b-h*) /oʔɖʒuba/, which greatly resembles the form used in modern Persian (cf. Hayyim I 108: /oʔɖʒubɛ/).

(xiii) P /u/ > H /o/, e.g.: H अंदोह (*andoh*), اندوه (*A-n-d-V-h*) /ændoh/ ‘grief’ ≤ P اندوه (*A-n-d-V-h*) /ænduh/ ‘t.s.’; H लोबिया (*lobiyā*), لوبيا (*l-V-b-Y-A*) /lobija/ ‘pulse’ ≤ Class. P لوبيا (*l-V-b-Y-A*) ~ لوبيه (*l-V-b-Y-h*) /lubija/ ~ /lubijæ/ (modern P /lubja/) ‘t.s.’; H नीलोफर (*nilofar*) ~ नीलोपर (*nilopar*), نیلوفر (*n-Y-l-V-f-r*) /nilofær/ ~ /nilopær/ ‘water-lily’ ≤ P نیلوفر (*n-Y-l-V-f-r*) /nilufær/³²³ ‘t.s.’; H शोशा (*śośā*), شوشه (*š-V-š-h*) /ʃoʃa/ ‘mixing’ ≤ Class. P شوشه (*š-V-š-h*) /ʃuʃæ/ (modern P /ʃuʃæ/) ‘ingot’; H रोदा (*rodā*), روده (*r-V-d-h*) /roda/ ‘gut’ ≤ Class. P روده (*r-V-d-h*) /rudæ/ (modern P /rudæ/) ‘t.s.’;

(xiv) P /u/ > H /e/, e.g.: H गुलेला (*gulelā*), غلله (*ġ-l-Y-l-h*) ~ غلوله (*ġ-l-V-l-h*) /gulela/ ~ /ġələla/ ‘pellet’ ≤ P غلوله (*ġ-l-V-l-h*) /yulula/ ‘t.s.’.



Fig. 9. “Street allography”. A multigraphic advertising banner in Hyderabad. The name of the hotel has been written down in four scripts: Roman, Telugu, Perso-Arabic and Nagari. (Photo by the author, 2012)

³²³ Both in Classical and in modern Persian this word has several variations, but they differ only in consonants while the vocalization remains immutable, cf.: نیلوبرگ (*n-Y-l-V-b-r-g*) /nilubærg/, نیلوپر (*n-Y-l-V-p-r*) /nilupær/, نیلوپل (*n-Y-l-V-p-l*) /nilupæl/, نیلوفر (*n-Y-l-V-f-r*) /nilufær/, نیلوفل (*n-Y-l-V-f-l*) /nilufæl/ (Stein. 1444; Hayyim II 1135).

3.4. Changes in Distribution of Vowel Phonemes in the Borrowed Words

3.4.1. Changes in the distribution of vowels in Perso-Arabic loanwords in Hindustani are common and numerous. Two adversative processes are chiefly responsible for this situation: the addition of a euphonic vowel in the middle position and the loss of the neutral vowel /ə/.

3.4.2. In a number of borrowed glosses, mostly between a stop and another consonant, a euphonic vowel appears in order for the stops to be pronounced with an audible release. The exact value of this vowel is vague,³²⁴ it is most often realized as /ə/ or /ɪ/ and in a few cases as /ʊ/. Structurally, it very likely belongs to the short vowels.³²⁵ The epenthetic /ə/ appears mostly in one-syllable nouns (often of Arabic origin) with a consonant cluster in the coda position (CVCC).³²⁶ The euphonic /ɪ/, however, chiefly breaks apart the cluster of a consonant and the following approximant /j/.³²⁷

³²⁴ They are basically nonsegmental “pseudo-vowels” (Hock 1991: 125) but in many a case their segmentalization results in fully segmental vowels which is reflected also in script (Nagari); cf. examples in 3.4.2.(i-iii).

³²⁵ Long vowels may also occur, though sporadically, as a result of anaptyxis, e.g.: H दबीज़ (*dabīz*), دبیز (*d-b-Y-z*) /dəbiz/ ‘thick’ ≤ P دبز (*d-b-z*) /dæbz/ ‘t.s.’.

³²⁶ Hindustani phonotactics does not forbid a two-consonantal coda, but the genuine Hindustani tendency towards a noun pattern of the CVCVC type is probably an important reason for anaptyxis. On the other hand, the opposing schwa deletion process remains a far more frequent phenomenon in Hindustani than the occurrence of an apertic vowel. (On deletion vs. insertion of schwa see Ohala 1983: 145-147.)

³²⁷ Epenthesis may be strengthened in Hindustani for orthographical reasons. In Perso-Arabic notation it would be the omission of the *sukūn* (the diacritic sign used to indicate lack of vocalization). In such a situation the consonants in a cluster might be separated by a vowel. A similar effect can

(i) examples of the epenthetic /ə/: H अमर (*amar*), امر (*A-m-r*) /əməɾ/ ~ /əmr/ 'order, command' ≤ P امر (*A-m-r*) /əmr/ 't.s.'; H ववर (*babar*), ببر (*b-b-r*) ~ بابر (*b-A-b-r*) /bəbər/ ~ /babər/³²⁸ 'lion' ≤ P ببر (*b-b-r*) /bəbr/ 't.s.'; H फ़तह (*fatah*), فتح (*f-t-h*) /fətəh/³²⁹ ~ /fəth/ 'victory' ≤ P [≤ A] فتح (*f-t-h*) /fæth/ 't.s.'; H ग़बन (*gaban*), غبن (*ġ-b-n*) /gəbən/ ~ /yəbən/ 'misappropriation' ≤ P [≤ A] غبن (*ġ-b-n*) /yəbn/ 'deceit, fraud'; H हलक (*halak*), طلق (*h-l-q*) /hələq/ ~ /həlq/ 'throat' ≤ P [≤ A] حلق (*h-l-q*) /hælq/ 't.s.'; H हतक (*hatak*), تحك (*h-t-k*) /hətək/ ~ /hətk/ 'disrespect' ≤ P [≤ A] هتك (*h-t-k*) /hætk/ 't.s.'; H इतरदान (*itar'dān*),³³⁰ عطران (*ʿ-t-r-d-A-n*) /itərdan/ ~ /itrdan/ 'container for perfume' ≤ P [≤ A] عطران (*ʿ-t-r-d-A-n*) /itrdan/ 't.s.'; H जिरह (*jirah*), جرح (*j-r-h*) /dʒirəh/³³¹ ~ /dʒirh/ ~ /dʒərh/ 'wound' ≤ P [≤ A] جرح (*j-r-h*) /dʒærh/ 't.s.'; H कवक (*kabak*), كبك (*k-b-k*) /kəbək/ ~ /kəbk/ 'partridge' ≤ P کبگ (*k-b-g*) /kæbg/³³² 't.s.'; H किसव (*kisab*),³³³ کسب (*k-s-b*) /kɪsəb/ ~ /kəsəb/ 'gain; trade' ≤ P [≤ A] کسب (*k-s-b*) /kəsəb/ 't.s.'; H ख़लक (*khalak*), خلق (*x-l-q*) /xələq/ ~ /xəlq/ 'creation' ≤ P [≤ A] خلق (*x-l-q*) /xælq/ 't.s.'; H सतह (*satah*), سطح (*s-t-h*) /sətəh/ ~ /səth/ 'surface' ≤ P [≤ A] سطح (*s-t-h*)

be observed in the case of loanwords in Nagari notation since both the *virāma* diacritic sign and the ligatures are often desisted in script.

³²⁸ Anaptyxis with accompanying lengthening of the preceding vowel ([ə] → [a]).

³²⁹ In this case the euphonic /ə/ occurred probably not only to ease the pronunciation but also to distinguish the consonant cluster /th/ from the aspirated /tʰ/.

³³⁰ But cf. Nagari notation of the word इत्र (*itr*) /itr/ 'perfume' where the ligature /tr/ clearly indicates the original lack of vocalization between the two consonants (= P [≤ A] عطر (*ʿ-t-r*) /itr/ 't.s.').

³³¹ Anaptyxis accompanied by the qualitative change of the preceding vowel (cf. 3.3.5.(i)).

³³² In Hayyim's dictionary both کبگ (*k-b-g*) /kæbg/ and کبک (*k-b-k*) /kæbk/ are confirmed (cf. Hayyim II 610).

³³³ Anaptyxis accompanied by the qualitative change of the preceding vowel (cf. 3.3.5.(i)).

/sæth/ 't.s.';³³⁴ H सुबह (*subah*), سُبْح (š-b-h) /subəh/ ~ /suboh/ ~ /subh/ 'morning' ≤ P [≤ A] صَبَح (š-b-h) /subh/ 't.s.'; H वजह (*vajah*), وَجْه (V-j-h) /vəɖʒəh/ ~ /vəɖʒh/³³⁵ 'reason' ≤ P [≤ A] وَجْه (V-j-h) /vəɖʒh/ 't.s.'; H ज़बह (*zabah*), زَبْح (z-b-h) /zəbəh/ ~ /zəbh/ 'slaughtering (an animal for food)' ≤ P [≤ A] نَبَح (z-b-h) /zæbh/ 't.s.';

(ii) examples of the epenthetic /ɪ/: H आशियाना (*āśiyānā*), آشیانه (*Ā-š-Y-A-n-h*) /aʃijana/ 'nest' ≤ P آشیانه (*Ā-š-Y-A-n-h*) /aʃjana/ 't.s.'; H बखिया (*bakhīyā*), بَخِي (b-x-Y-h) /bəxija/ 'stitching' ≤ P بَخِي (b-x-Y-h) /bəxja/ 't.s.'; H दुनिया (*duniyā*), دُنْيَا (d-n-Y-A) /dunija/ 'the world' ≤ P دُنْيَا (d-n-Y-A) /dunja/ 't.s.'; H होशियार (*hośiyār*), هُوشِيَار (h-V-š-Y-A-r) /hoʃijar/ ~ /hoʃjar/ ~ /hoʃʃjar/ 'clever' ≤ P هُوشِيَار (h-V-š-Y-A-r) /hoʃjar/ 't.s.'; H हुलिया (*huliyā*), طِيَة (h-l-Y-A) /hulija/ 'appearance' ≤ P [≤ A] طِيَة (h-l-Y-A) /hulja/ 't.s.'; H दरियाई (*dariyāī*), دَرِيَائِي (d-r-Y-A-Y^c-Y) /dərija.i/³³⁷ ~ /dərja.i/ 'satin' ≤ Class. P دارانی (d-A-r-A-Y^c-Y) /dara?i/ 'a red silken stuff'; H बागीचा (*bāgīcā*), باغِيچَا (b-A-ğ-Y-č-h)³³⁸ /baɣiʃʃa/ ~ /baɣiʃʃa/³³⁹ 'garden' ≤ P باغچه (b-A-ğ-č-h) /baɣʃʃæ/ 't.s.'; H किलिक (*kilik*), کَلَب (k-l-k) /kılık/ 'reed, pen' ≤ P کَلَب (k-l-k) /kilk/ 't.s.';

³³⁴ Here again the euphonic /ə/ plays a double function – it makes the pronunciation easier and prevents the confusion of taking the consonant cluster /th/ for the aspirated consonant /tʰ/.

³³⁵ Gilchrist quotes this word as *wujih* (cf. OL 115).

³³⁶ In this alternate form (STCD 681) the vowel has been shortened ([o] → [u]) and the alternation is clearly visible in the orthography.

³³⁷ In this example epenthesis of the vowel /ɪ/ is combined with epenthesis of the consonant (glide) /j/ and, additionally, the preceding vowel is shortened ([a] → [ə]).

³³⁸ The Persian form بَغِيچَة (b-A-ğ-č-h) is also quoted in the STCD, but qualified as rare (STCD 95).

³³⁹ Note that in this case the euphonic vowel /i/ is long and anaptyxis may be accompanied by the shortening of the preceding vowel [a] → [ə], which can be observed in the Nagari notation of this loanword.

(iii) examples of the epenthetic /ʊ/: H कुर (kuf^r), कुर (kufur), کفر (k-f-r) /kuf^r/ ~ /kufur/ 'paganism' ≤ P کفر (k-f-r) /kuf^r/ 't.s.'; H कुतुब (qutub) قطب (k-t-b) /qutub/ ~ /qutb/ 'the pole star' ≤ P [≤ A] قطب (k-t-b) /qutb/ 't.s.'.

3.4.3. Paragoge (or the addition of a sound at the end of a word) is a very rare phenomenon in Hindustani and may be observed only in a few cases, cf.: H फी सदी (fī sadī), فی صدی (f-Y ṣ-d-Y)³⁴⁰ /fī sadī/ 'per cent' ≤ P فی صد (f-Y ṣ-d) /fī sad/ 't.s.'; H उमराव (um'rāv),³⁴¹ امرأ (A-m-r-A-V) /ʊmra.o/ 'Umrao'³⁴² ≤ Class. P [≤ A] امراء (A-m-r-A-ʔ) /æmæraʔ/ (modern P /omæra/) pl. of /amir/ 'princes, &c.'.

3.4.4. Elision (or loss of sound) occurs in Hindustani mainly in the form of syncope, i.e. the loss of an unstressed vowel from the interior of a word, and is significantly connected with the schwa-dropping phenomenon. In Hindustani, the schwa /ə/, defined as a mid central vowel that appears in unstressed syllables (and is implicit in each consonantal sign of the Nagari script), is "obligatorily deleted" in pronunciation at the end of words (apocope) and in certain other contexts.³⁴³ Ohala (1983) has

³⁴⁰ Traditional (Persian) spelling فی صد (f-Y ṣ-d) is also possible in Hindustani/Urdu; it is less common though.

³⁴¹ Note that even though the consonantal sign व is added in the Nagari notation, it is always pronounced as a vowel /o/ ~ /ɔ/ (cf. similar tendency to vocalization in words like H [≤ S] पुलाव (pulāv), पला (p-l-A-V)³⁴² /pula.o/ 'a dish of fried or boiled rice and meat with spices' or H [≤ P] सरो (saro), सर (s-r-V) /səro/ ~ /səru/ 'cypress').

³⁴² This word, when used in its basic meaning and not as a proper name, differs both in spelling and in pronunciation, cf.: H उमरा (um'rā), امرأ (A-m-r-A) /um(ə)ra/ 'persons of high birth or rank' (= P).

³⁴³ This phenomenon has been termed the 'schwa syncope rule' (cf. Bhatia 1987: 40) or the 'schwa deletion rule' (cf. Ohala 1983: 117-154). Essentially, it is defined as a morphological process (e.g. Singh & Agnihotri 1997) since it is associated with specific morphological operations and is also an integral part of the word formation strategy. But the morphologically

summarized this rule as $VCəCV \rightarrow VC\emptyset CV$ (when a schwa-succeeded consonant is followed by a vowel-succeeded consonant, the schwa inherent in the first consonant is deleted).³⁴⁴ In the words borrowed by Hindustani schwa replaced the short vowel [ā],³⁴⁵ and became subjected to the process of schwa deletion,³⁴⁶ e.g.: P احدى (*A-h-d-Y*) /æhædi/ ‘a kind of Indian military corps’ ($\leq$ A /æhæd/ ‘one, an individual’) $\geq$ H अहदी (*ah’dī*), احدی (*A-h-d-Y*) /əhdi/ ‘one in league, collaborator’; P [$\leq$ A] اهم (*A-h-m*) /æhæm/ ‘important’ $\geq$ H अहम (*ah’m*), فم (*A-h-m*) /əhm/ ‘t.s.’; P [$\leq$ A] الغرض (*A-l-ġ-r-z*) /æɫqæræz/ ‘in short; finally’ $\geq$ H अलगज़

dependent schwa deletion also results in certain phonological changes. Therefore, it can be classified as belonging to both the phonological and morphological domains and should be analyzed on the morphophonemical level (as proposed e.g. by Elizarenkova 1988). Such an approach justifies the allocation of this specific phenomenon to the chapter devoted to sound changes.

³⁴⁴ This systematization, however, is neither exact nor complete and can generate errors, since sometimes it deletes a schwa when it shouldn’t or, at other times, fails to delete it when it should (more see Ohala 1983; Choudhury et al. 2004; Pandey 1990). In Hindustani, correct schwa deletion is critical because, depending on the context, the same letter-sequence is pronounced differently, and failure to delete the appropriate schwas can even change the sense of the word.

³⁴⁵ In some loanwords schwa appears in place of other short vowels as well, cf. 3.3.5.(v) and 3.3.5.(ix) and subsequently, in rare cases, can be deleted, e.g.: P /fʊyʊl/ ‘an informer’ $\geq$ H /fʊyəl/ ~ /fʊyɫ/; P /ʊsturæ/ ‘a razor’ $\geq$ H /ʊstəra/ ~ /ʊstra/.

³⁴⁶ It should be noted that schwa deletion in loanwords is not an unequivocal or fixed phenomenon and many borrowed words may be pronounced both ways, i.e. with or without a schwa in the unstressed syllable, depending on the language register (speakers of Persianized Hindustani/Urdu will retain the Persian vocalization pattern), eg.: हकत / حرکت [hærkət] vs. [hərəkət] ‘motion, movement’ $\leq$ P [$\leq$ A] /hærækæt/; मुबारकी / مبارکی [muba:rki:] vs. [muba:ræki:] ‘blessings, congratulations’ $\leq$ P [$\leq$ A] /mubæræki/; نظریہ / نظریہ [nəzrija:] vs. [nəzərija:] ‘point of view, theory’ $\leq$ P [$\leq$ A] /næzærija/.

(*al'gar'z*), الغرض (*A-l-ġ-r-z*) /əlyərz/ 't.s.'; P [≤ A] عمله (^c-*m-l-h*) /æmælæ/ (modern P /æmæle/) 'workers, staff' ≥ H अमला (*am'lā*), عمل (^c-*m-l-h*) /əmla/ 't.s.'; P آدمي (*Ā-d-m-Y*) /adæmi/ 'a man' ≥ H आदमी (*ād'mī*), آدمي (*Ā-d-m-Y*) /admi/ 't.s.'; P گروی (*g-r-V-Y*) /girævi/ (modern P /gerovi/) 'pledged, pawned' ≥ H गिरवी (*gir'vī*), گردی (*g-r-V-Y*) /girvi/ 't.s.'; P [≤ A] حوصله (*h-V-ṣ-l-h*) /hausælæ/ (modern P /housele/) 'stomach' ≥ H हैमला (*haus'lā*), حوصله (*h-V-ṣ-l-h*) /hausla/ 't.s.'; P قيمتي (*q-Y-m-t-Y*) /kimæti/ 'expensive' ≥ H कीमती (*kīm'tī*), قيمي (*q-Y-m-t-Y*) /qimti/ 't.s.'; P لا جورد (*l-A-j-V-r-d*) /lædʒæværd/ 'lapis lazuli' ≥ H लाजवर्त (*lāj'vart*), لا جورد (*l-A-j-V-r-d*) /ladʒvært/ ~ /ladʒværd/ 't.s.'; P [≤ A] مثلاً (*m-ṣ-l-Ā*) /mæsælæn/ 'for example' ≥ H मसलन (*mas'lan*), مثلاً (*m-ṣ-l-Ā*) /mæslən/ 't.s.'; P [≤ A] معتبر (*m-^c-t-b-r*) /mʊ'tæbær/ (modern P /mo'tæbær/) 'reliable' ≥ H मोतबर (*mot'bar*) ~ मातबर (*māt'bar*), معتبر (*m-^c-t-b-r*) /motbær/ ~ /matbær/ 't.s.'; P شکره (*ṣ-k-r-h*) /ʃikæræ/ (modern P /ʃikære/) 'rapacious birds trained to hunt; vulva' ≥ H शिकरा (*śik'rā*) ~ शीकरा (*śīk'rā*), شکره (*ṣ-k-r-h*) /ʃikra/ ~ /ʃikra/ 'a hawk'; P [≤ A] طبقة (*t-b-q-ī*) /tæbæqæt/ (modern P /tæbæqe/) 'storey, floor' ≥ H तबका (*tab'kā*), طبقه (*t-b-q-h*) /təbqa/ 't.s.'.

3.4.5. Other cases of syncope are also common, especially in non-conservative registers of Hindustani, e.g.: H काबला (*kāb'lā*), قابله (*q-A-b-l-h*) /qabla/ 'a midwife; a bolt' ≤ Class. P قابله (*q-A-b-l-h*) /qabilæ/ (modern P /qabela/) 't.s.'; H अलहदा (*alah'dā*) ~ अलाहदा (*alāh'dā*), ³⁴⁷علا (*ḥ-l-Y-l-ḥ-d-h*) /əlahda/ 'separate' ≤ Class. P [≤ A] ³⁴⁷علا (*ḥ-l-Y-l-ḥ-d-h*) /əlahidæ/ (modern P /alaheda/) 't.s.'; H आरज़ा (*ār'zā*), عارضه (*ḥ-A-r-z-h*) /arza/ 'sickness' ≤ Class. P [≤ A] عارضه (*ḥ-A-r-z-h*) /ærizæ/ (modern P /æreze/) 't.s.'; H एतबार (*et'bār*), اعتبار (*A-^c-t-b-A-r*) /etbar/ 'confidence' ≤ Class. P [≤ A] اعتبار (*A-^c-t-b-A-r*) /i'tibar/ (modern P /ε'tebar/) 't.s.';³⁴⁸ H कुलचा

³⁴⁷ The alternative spelling without dots under the letter yā is possible.

³⁴⁸ Analogical loss of the vowel [ɪ] can be observed in a whole group of similarly structured loanwords, viz.: H एतकाद / اعتقاد / *etqad* / 'confidence'

(*kul'cā*), کچھ (*k-l-č-h*) /*kulʃa*/ 'a small round bread-roll made of flour, milk and butter' ≤ Class. P کلیجہ (*k-l-Y-č-h*) /*kuliʃæ*/ 't.s.';³⁴⁹ H बाकला (*bāk'lā*), باقلا (*b-A-q-l-A*) /*baqla*/ 'kind of bean' ≤ Class. P [≤ A] باقلاء (*b-A-q-l-A-*) /*baqla*?/ (modern P /*baqela*?/) 't.s.'; H जाबता (*zāb'tā*), ضابطہ (*z-A-b-t-h*) /*zabta*/ 'procedure' ≤ Class. P [≤ A] ضابطہ (*z-A-b-t-h*) /*zabitæ*/ (modern P /*zabetæ*/) 't.s.'; H हौलदिला (*haul'dilā*), هول دلا (*h-V-l d-l-A*) /*hauldila*/ 'frightened' ≤ P هول دل (*h-V-l d-l*) /*hauledil*/ 't.s.';³⁵⁰ H खुनकी (*khun'kī*), خنکی (*x-n-k-Y*) /*xunki*/ 'coolness' ≤ Class. P خنکی (*x-n-k-Y*) /*xunuki*/ (modern P /*xonoki*/) 't.s.'; H नाश्ता³⁵¹ (*nāštā*), ناشتا (*n-A-š-t-A*) /*našta*/ 'breakfast' ≤ P ناشتا (*n-A-š-t-A*) /*našta*/ 't.s.'; H वास्ता (*vāstā*), واسطہ (*V-A-s-t-h*) /*vasta*/ 'sthg. intermediate or between' ≤ P [≤ A] واسطہ (*V-A-s-t-ī*) /*vasitæ*/ (modern P /*vasetæ*/) 't.s.'; H ज्यादा (*zyādā*), زیادہ (*z-Y-A-d-h*) /*zjada*/ 'more; very many' ≤ P [≤ A] زیاده (*z-Y-A-d-ī*) /*zijadæ*/ (modern P /*zjadæ*/) 'augmentation, increase'; H मामला (*mām'lā*), معاملہ (*m-^c-A-m-l-h*) /*mamla*³⁵² 'an affair, matter' ≤ Class. P [≤ A] معامله (*m-^c-A-m-l-ī*) /*mo'amælæ(t)*/ (modern P /*mo.amele*/) 't.s.'; H माफ (*māf*), معاف (*m-^c-A-f*) /*maf*³⁵³ 'forgiven' ≤ Class. P [≤ A] معاف (*m-^c-A-f*) /*mo'af*/ (modern P /*mo.af*/) 't.s.';

≤ P /*i'tiqad*/; H एतमाद / اعتماد /*etmad*/ 'reliance, trust' ≤ P [≤ A] /*i'timad*/; H एतराज / اعتراض /*etraz*/ 'opposition' ≤ P [≤ A] /*i'tiraz*/.

³⁴⁹ This is a rare example of a long vowel loss.

³⁵⁰ The lost vowel in this case is short /e/ indicating the Persian enclitic 'of' (*izāfat*), noted in the Persian source word by the diacritical *kasrah*.

³⁵¹ As may be observed, it happens only sporadically (and inconsistently) that the syncope is indicated in the Nagari script by ligatured consonants – as is the case in this and the two following examples.

³⁵² Double syncope – in the first and third syllables. The more traditional pronunciation /*mo.amlā*/ prevails in Persianized Hindustani/Urdu.

³⁵³ The traditional pronunciation /*mo.af*/ prevails in Persianized Hindustani/Urdu.

3.5. Quantitative and Qualitative Change of Consonants in the Perso-Arabic Loanwords

3.5.1. Quantitative change of consonants in Perso-Arabic loanwords in Hindustani relates to the phenomenon of consonant lengthening (gemination) and shortening (degemination). Both in Hindustani and in Persian, consonant length is phonemic³⁵⁴ (it is also distinctive in Arabic). In the pronunciation of elongated consonants a definite prolongation of the blockage of the air flow is often observed. This entails the conclusion that in Hindustani (and in Persian as well) gemination is realized as a single drawn out sound rather than a sequence of two distinct sounds. In Hindustani all consonants can be lengthened except /b^h/, /t̪/, /t̪^h/ and /h/. Geminates occur only monomorphemically (i.e. a morpheme boundary may never be posited in between)³⁵⁵ and medially, always preceded by a non-peripheral vowel /ə/, /ɪ/ or /ʊ/ (Ohala 1999: 101). In some cases geminates are preserved in final position by orthography (e.g. हज्ज (hajj) / ھَجَّ (h-jj) */hədʒ:/ ‘pilgrimage to Mecca’), but in registers other than the most conservative they are usually pronounced as singletons – unless there is a following vowel in the same word or phrase,³⁵⁶ e.g.: /hədʒ/ but /hədʒ: o

³⁵⁴ Cf. some illustrative examples of the geminate contrasts in Hindustani: /bəʃʃ:a/ m. ‘a child’ vs. /bəʃʃa/ part. ‘saved’; /pət:a/ m. ‘leaf’ vs. /pəta/ m. ‘trace; address’; /bən:o/ f. ‘lady, darling’ vs. /bəno/ imperat. ‘make!’; /kəl:a/ 1. ‘shoot, sprout’; 2. [≤ P] ‘upper jaw’ vs. /kəla/ ‘a small part of anything’ and in Persian (cf. Windfuhr 1997: 681): /kor:ɛ/ ‘foal’ vs. /kore/ ‘globe’; /zæn:/ ‘opinion’ vs. /zæn/ ‘a woman, wife’.

³⁵⁵ Hence the difference in pronunciation (as well as in spelling) between a genuine geminate and a cluster of the same consonants which can appear on the morpheme boundary as a result of the schwa deletion. Cf. e.g.: बन्ना (bannā) / ٻَ (b-n-A) /bən:a/ m. ‘bridegroom’ vs. बनना (ban’nā) / ٻَ (b-n-A) /bən.na/ inf. ‘to be made’ (where -/na/ is the infinitive suffix).

³⁵⁶ Similarly in Persian, cf. Mace 2003: 30.

umra/ '«great» and «small» pilgrimage to Mecca'. This rule entails that final geminates, common in Arabic, in loanwords are reduced to meet the constraints of Hindustani phonotactics, e.g.: H अहम (*ahm*), اَهِم (*A-h-m*) /əh(ə)m/ 'important' ≤ Class. P [≤ A] اَهِم (*A-h-m*) /æhæm:/ (modern P /æhæm/) 't.s.'; H अवाम (*avām*), اَوَام (*^c-V-A-m*) /əvām/ 'the common people' ≤ Class. P [≤ A] عوام (*^c-V-A-m*) /ævam:/ (modern P /ævam/) 't.s.'; H बेशक (*beśak*), بَـشَـك (*b-e š-k*) /beʃək/³⁵⁷ 'undoubtedly' ≤ Class. P [≤ A] بی شک (*b-Y š-k*) /bi ʃæk:/ 't.s.'; H बिल्कुल (*bil'kul*), بِـاَکـل (*b-A-l-k-l*) /bilkul/ 'entirely, completely' ≤ Class. P [≤ A] بالکل (*b-A-l-k-l*) /bilkol:/ (modern P /bēkol/) 't.s.'; H दिक् (*dik*), دِـق (*d-q*) /dɪq/ 'troubled, disturbed' ≤ Class. P [≤ A] دِـق (*d-q*) /dɪq:/ (modern P /dēq/) 'teasing, trouble'; H जद् (*jad*), جـد (*j-d*) /dʒəd/ ~ /dʒɪd/ 'effort, exertion' ≤ Class. P [≤ A] جـد (*j-d*) /dʒɪd:/ (modern P /dʒɛd/) ³⁵⁸ 't.s.'; H कट (*kat*), قـط (*q-ṭ*) /qəṭ/ 'cutting: nib (of a pen)' ≤ Class. P [≤ A] قـط (*q-ṭ*) /qæt:/ (modern P /qæt/) 't.s.'; H रब (*rab*), رـب (*r-b*) /rəb/ 'the Lord, God' ≤ Class. P [≤ A] رـب (*r-b*) /ræb:/ (modern P /ræb/) 't.s.'. Erratic degemination may occur also medially, in the intervocalic position, ³⁵⁹ e.g.: H बज़ाज़ (*bazāz*), بـزـز (*b-z-A-z*) /bəzaz/

³⁵⁷ But the geminate is retained in the phrase /beʃək: o fubh/ lit. 'without doubts-and-intricacies' (STCD 135).

³⁵⁸ Hayyim retains gemination and indicates it very clearly with a *tašdīd*, cf.: "جـد (*jedd*) Noun 1. Effort, endeavor; painstaking, exertion. 2. Seriousness." (Hayyim I 536).

³⁵⁹ In Persian, gemination often serves to distinguish words of different meanings, e.g.: /næx:as/ 'a seller of captives' vs. /næxas/ 'a thong sewed betwixt two pieces of leather'; /dæl:al/ 'a broker' vs. /dælal/ 'an amorous glance'; /bəz:az/ 'a dealer' vs. /bəzaz/ 'a shoe-latchet', etc. Since usually only one word from each 'pair' has been borrowed by Hindustani, gemination lost its distinctive character in these particular loanwords. The lack of gemination in pronunciation may also be caused by the omission of *tašdīd*, especially among those speakers who don't know Arabic or Persian but use Perso-Arabic script. Those who read and write Nagari do not make

'cloth-merchant, draper' ≤ P [≤ A] بزاز (*b-z-A-z*) /bæz:az/ 't.s.'; H دلال (*dalāl*), دلال (*d-l-A-l*) /dəlal/ ~ /dəl:al/ 'a broker' ≤ P [≤ A] دلال (*d-l-A-l*) /dəl:al/ 't.s.'; H फ़ारख़ती (*fār'khatī*), فآر ختلی (*f-A-r-ġ-x-t-Y*) /færxəti/³⁶⁰ 'deed of release (from debt or obligation)' ≤ Class. P [≤ A] فار غطی (*f-A-r-ġ-x-t-Y*) /færɣxæt:i/ (modern P /færɣxæt:i/) 't.s.'; H नख़ास (*nakhās*), نخاس (*n-x-A-s*) /nəxas/ ~ /nəx:as/ 'a market (for livestock)' ≤ P [≤ A] نخاس (*n-x-A-s*) /nəx:as/ 'a seller of captives, or cattle taken as plunder'; H सेहत (*sehat*), صحت (*ṣ-h-t*) /sehət/ ~ /seh:ət/ 'health' ≤ Class. P [≤ A] صحة (*ṣ-h-t*) /sih:æt/ (modern P /seh:æt/) 'recovering from sickness, soundness'.

3.5.2. The reverse phenomenon, namely gemination, or elongation of originally short/single consonants, occurs irregularly and not very often, e.g.: H अब्बा (*abbā*), ابا (*A-b-A*) /əb:a/ 'father, daddy' ≤ P [≤ A] اب (*A-b*) /ab/ 't.s.' × ابااء (*A-b-A-'*) /abaʔ/ pl. 'fathers, ancestors'; H बुज़्ज़ा (*buzzā*), بزاز (*b-z-A*) /buz:a/ 'the white or the black ibis' ≤ P بزک (*b-z-k*) /buzæk/ 'name of a black-coloured, long-beaked bird'; H जच्चा (*jaccā*) ~ जच्चा (*zaccā*), जच्चा (*j-č-A*) ~ जच्चा (*z-č-A*) /dʒəʃ:a/ ~ /zəʃ:a/ 'a woman who has recently given birth' ≤ Class. P زجه (*z-j-h*) /zædʒæ/ ~ زجه (*z-č-h*) /zæʃæ/ (modern P /zadz/ ~ /zaʃ/) 't.s.'; H मुरव्वत (*muravvat*), مروء (*m-r-V-h*) /murəv:ət/ ~ /muruv:ət/ 'humanity' ≤ Class. P مروء (*m-r-V-t*) /muruv:ət/ (modern P /morov:ət/) 't.s.'; H रवैया (*ravaiyā*), رويء (*r-V-Y-h*) /rəvəj:a/ 'conduct, behaviour' ≤ Class. P رويء (*r-V-Y-h*) /rəvɨjæ/ (modern P /rəvɨj:ɛ/) 't.s.'; H रुक्का (*rukḱā*), رقه (*r-q-^c-h*) /ruq:a/ 'piece, scrap (as of paper)' ≤ Class. P رفعة (*r-V-Y-t*) /ruqʔæ(t)/ (modern P /roqʔɛ/) 't.s.'.

3.5.3. Changes in manner of articulation (**qualitative**) are commonly observed in the borrowings and the P voiceless uvular

such errors since gemination (or it's lack) is clearly notified in Nagari (cf. 2.3.21.).

³⁶⁰ The Persianized pronunciation /færɣxæt:i/, labelled as 'pedantic' (STCD 448), occurs rarely and only in conservative Hindustani/Urdu. Note that even in this variant of utterance the gemination of /t/ is not retained.

stop /q/ (with its voiced allophones: the uvular stop [g] and the fricative [ɣ]) is probably the most assimilated phoneme,³⁶¹ e.g.: H चिक (*cik*) ~ चिक (*cik*), چق (*č-q*) ~ چک (*č-k*) /ʃiq/ ~ /ʃik/ 'folding bamboo, screen' ≥ P چق (*č-q*) /ʃiq/ 'a Venetian blind'; it remains also a common object of defricativization,³⁶² e.g.: Class. P [≤ A] غيب (*ġ-Y-b*) /ɣaib/ (modern P /ɣejb/) 'invisible, hidden' ≥ H गैब (*gaib*), غيب (*ġ-Y-b*) /carb/ ~ /gaib/ 't.s.'; Class. P مغل (*m-ġ-l*) /muɣul/ (modern P /moɣol/) 'a Mogul, Mongolian' ≥ H मुगल (*mugal*), مغل (*m-ġ-l*) /mugəl/ ~ /mugəl/ 't.s.'.

3.5.4. Some qualitative changes characterized by greater regularity, have already been described in paragraph 3.2.6., whereas those alternations which are observed in the borrowed words more infrequently are enumerated below:

(i) approximation of the Persian voiced labio-dental fricative /v/ by replacing it with the bilabial nasal /m/, e.g.: P دیوک (*d-Y-V-k*) /dɛvæk/ ~ /dɛvuk/ 'weevil; moth' ≥ H दीमक (*dīmak*), دیمک (*d-Y-m-k*) /dimək/ 't.s.'; P هوان (*h-V-A-n*) /havæn/ 'mortar' ≥ H हावन (*hāvan*) ~ हिमाम (*himām*), هاون (*h-A-V-n*) ~ هام (*h-m-A-m*)³⁶³ /havən/ ~ /himam/ 't.s.';

(ii) approximation of the Persian voiced labio-dental fricative /v/ by replacing it with the bilabial stop /b/,³⁶⁴ e.g.: Class. P وا وילה (*V-A V-Y-l-A-h*) /va varlah/ (modern P /vavejla/) 'alas! woe!' ≥ H बावेली (*bāvelī*), واولیه (*V-A V-Y-l-A-h*) /baveli/ ~ /bavela/ 't.s.'; Class. P [≤ A] وداع (*V-d-A-^c*) /vida/ (modern P /veda/) 'adieu, farewell' ≥ H विदा (*bidā*) ~ विदाय (*vidāy*), بد (*b-d-A*) ~ وداع (*V-d-A-^c*) /bida/ ~ /vida(j)/ 'departure'; Class. P [≤ A] جوار (*j-V-A-r*) /dʒivar/

³⁶¹ This assumption finds confirmation in Nagari writing (common use of dotless क, ग vs. क, ग), *vide supra* 2.3.4.

³⁶² This type of change has already been discussed in paragraph 3.2.5.

³⁶³ Especially in the expression هام دست (*h-m-A-m d-s-t-h*) /həmam dəstə/ 'pestle and mortar'.

³⁶⁴ Cf. also 3.2.6.(v).

(modern P /dʒevar/) 'neighbourhood' ≥ H जवार (*jabār*), جوار (*j-V-A-r*) /dʒəbar/ ~ /dʒivar/ 't.s.';

(iii) fricativization of the Persian voiceless velar stop /k/, e.g.: P بکتر (*b-k-t-r*)³⁶⁵ /bæktær/ 'a coat of mail' ≥ H बखतर (*bakh'tar*) ~ बकतर (*bak'tar*), بکتر (*b-k-t-r*) /bæxtær/ ~ /bæktær/ 't.s.'; Class. P زنکه (*z-n-k-h*) /zænækæ/ (modern P /zænækæ/) 'a little woman, a wretch of a woman' ≥ H जनखा (*zan'khā*), زخا (*z-n-x-A*) ~ زخ (*z-n-x-h*) /zænxa/ 'effeminate; eunuch'; P بطک (*b-t-k*) /bætæk/ 'a duckling' ≥ H बतख (*batakḥ*) ~ बतक (*batak*), بٹخ (*b-t-x*) /bætəx/ ~ /bətək/ 't.s.';

(iv) defricativization of the Persian voiceless labio-dental fricative /f/, e.g.:³⁶⁶ P گنده فیروزه (*g-n-d-h f-Y-r-V-z-h*) /gændæ firozæ/ 'olibanum' ≥ H गन्दा बीरोज़ा (*gandā bīrozā*), گنده بیروزه (*g-n-d-h b-Y-r-V-z-h*) /gənda biroza/³⁶⁷ 't.s.';

(v) depalatalization of the Persian alveo-palatal affricate /dʒ/ by replacing it with Hindustani dental stop, e.g.: P جهیز (*j-h-Y-z*) /dʒæhiz/ 'a dowry' ≥ H दहेज़ (*dahez*), دیز (*d-h-Y-z*) /dæhez/ ~ /dihez/ 't.s.';

(vi) in several loanwords from Arabic the phonetic realization of the voiced fricative /z/ originating from the A voiced interdental fricative /ð/ is variable and may strengthen to the voiced dental stop /d/,³⁶⁸ e.g.: P [≤ A] کاغذ (*k-A-ğ-z*) /kaɣæz/ ~ /kaɣız/ 'paper' ≥ H कागज़ (*kāgaz*) ~ कागद (*kāgad*), کاغذ (*k-A-ğ-z*) /kacəz/ ~ /kacəd/ 't.s.'; P [≤ A] گذاره (*g-z-A-r-h*) /guzaræ/ 'passing' ≥ H

³⁶⁵ This is the collateral form of P بگتر (*b-g-t-r*).

³⁶⁶ Fortition of the fricative /f/ in Hindustani is usually connected with it being replaced by the aspirated bilabial plosive /pʰ/ (cf. 3.2.6.(i)).

³⁶⁷ But cf. P فیروزه (*f-Y-r-V-z-h*) /firozæ/ 'a turquoise' ≥ H फ़ीरोज़ा (*fīrozā*), فیروزه (*f-Y-r-V-z-h*) /firoza/ 't.s.'.

³⁶⁸ Note that the Arabic signs for both the consonants are very similar, a dot being the only observable difference (/ð/ = ð̣ ; /d/ = د), hence, in this case, the guess about the impact of the orthography on the pronunciation seems to be fully justified.

गुजारा (*guzārā*) ~ गुदारा (*gudārā*), گزاریه (*g-z-A-r-h*) /*guzara/* ~ /*guzara/* 't.s.'.³⁶⁹

3.5.5. Changes in phonation type:

(i) devoicing, e.g.: P کبگ (*k-b-g*) /*kæbg/* 'a partridge' ≥ H कवक (*kabak*), کبک (*k-b-k*) /*kəb(ə)k/* 't.s.'; P طنبوره (*t-n-b-V-r-h*) /*tæmburæ/* 'a stringed instrument used to accompany a singer' ≥ H तानपूरा (*tān'pūrā*), طنپوره (*t-n-b-V-r-h*) /*tampura/* ~ /*təmbura/* 't.s.'; P [≤ A] احد (*A-ḥ-d*) /*æhæd/* 'unity, one' ≥ H अहत (*ahat*), احد (*A-ḥ-d*) /*əhət/* ~ /*əhəd/* 't.s.'; P خرج (*x-r-j*) /*hærdʒ/* ~ /*hurɖʒ/* 'expense' ≥ H खर्च (*kharc*), خرج (*x-r-ċ*) /*xærʃ/* 't.s.'; P كاليد (*k-A-l-b-d*) /*kalbud/* ~ /*kalbæd/* 'mould, frame' ≥ H कालवूत (*kāl'būt*) ~ कलवूत (*kal'būt*), کالبت (*k-A-l-b-t*) /*kalbut/* ~ /*kəlbut/* ~ /*kəlbət/* 't.s.'; P خانگه (*x-A-n-g-A-h*) /*xangah/* 'a Šūfī residential establishment' ≥ H खानकाह (*khān'kāh*), خانه (*x-A-n-q-A-h*) /*xanqa(h)/* 't.s.'; P لاجورد (*l-A-j-V-r-d*) /*ladʒværd/* 'lapis lazuli' ≥ H लाजवर्त (*lāj'vart*), لاجورد (*l-A-j-V-r-d*) /*ladʒvərt/* ~ /*ladʒværd/* 't.s.'; P نشانچی (*n-š-A-n-j-Y*) /*nɪʃanji/* 'standard-bearer' ≥ H निशानची (*niśān'cī*), نشانی (*n-š-A-n-ċ-Y*) /*nɪʃanʃi/* 't.s.';

(ii) voicing is restricted practically to one irregular change, namely the utterance of the P /s/ ≤ A pharyngealized /sʕ/ as /z/, e.g.: P قميص (*q-m-Y-š*) /*qæmis/* 'a shirt' ≥ H कमीज (*kamīz*), قميص (*q-m-Y-š*) /*qəmis/*³⁷⁰ 't.s.'; P صفير (*š-f-Y-r*) /*sæfir/* 'whistling; singing (a bird)' ≥ H जफ़ीर (*zafīr*) ~ जफ़ील (*zafīl*), صفر (*š-f-Y-r*) ~ زفر (*z-f-Y-r*) /*zəfir/* ~ /*zəfil/* ~ /*səfir/* 't.s.'.

³⁶⁹ Cf. lenition of /d/ which is being replaced by /z/ as a result of the Arabicization of some genuinely Persian words, e.g.: P استاد (*A-s-t-A-d*) ~ استاذ (*A-s-t-A-z*) /*ustad/* ~ /*ustaz/* 'master, teacher' ≥ H उस्ताद (*ustād*), استاد (*A-s-t-A-d*) ~ استاذ (*A-s-t-A-z*) /*ustad/* ~ /*ustaz/* 't.s.'; Class. P گنبد (*g-n-b-d*) ~ گنبد (*g-n-b-z*) /*gumbæd/* ~ /*gumbæz/* (modern P /*gombæd/* ~ /*gombæz/*) 'a cupola, a dome' ≥ H गुब्द (*gumbad*), گنبد (*g-n-b-d*) ~ گنبد (*g-n-b-z*) /*gumbæd/* ~ /*gombæz/* 't.s.'.

³⁷⁰ The original form /*qəmis/* is rare in Hindustani; the even more "Indianized" variant /*kəmidʒ/* is common in colloquial speech.

3.5.6. Unifying of sibilants /ʃ/ and /s/, a phenomenon commonly observed in Hindustani, manifests itself in two ways:

(i) palatalization of the P voiceless alveolar fricative /s/, e.g.: P خرسند (*x-r-s-n-d*) /xursænd/ 'contented, happy' ≥ H खुशंद (*khur-sānd*), خشن (*x-r-š-n-d*) /xurʃænd/ 't.s.'; P [≤ A] طاسه (*t-A-s-h*) /tasæ/ 'a kind of drum' ≥ H तशा (*tāsā*) ~ तशा (*tasā*), طاش (*t-A-š-h*)³⁷¹ /təʃa/ ~ /taʃa/ 't.s.';

(ii) depalatalization of the P voiceless alveo-palatal fricative /ʃ/,³⁷² e.g.: P سریش (*s-r-Y-š*) /sirɪʃ/ 'glue' ≥ H सेस (*sares*), سریس (*s-r-Y-s*) /səres/ 't.s.'; P [≤ A] شاقل (*š-A-q-l*) ~ شاقول (*š-A-q-V-l*) /ʃaql/ ~ /ʃaql/ 'a plummet' ≥ H साहुल (*sāhul*), साफल (*š-A-q-l*) ~ साफल (*š-A-q-l*) /sahul/³⁷³ ~ /saql/ ~ /saql/ 't.s.';³⁷⁴

(iii) many loanwords of this type, however, retain alternate pronunciation, i.e. /ʃ/ and /s/ are uttered (and spelled) interchangeably (especially in Nagari), depending entirely on the speaker's choice, e.g.: H अर्श (*arś*) ~ अरस (*aras*), عرش (*^c-r-š*) /ərʃ/ ~ /ər(ə)s/ 'roof, canopy; heaven' ≤ P [≤ A] عرش (*^c-r-š*) /ærʃ/ 't.s.'; H खुसफेली (*k^hus'p^heli*), خوش نعلی (*x-V-š f^c-l-Y*) /k^huʃp^heli/ ~ /xuʃfeli/

³⁷¹ Note that in both cases the sound change has also influenced the way these loanwords are spelled in Perso-Arabic script – which is a rather unusual feature. The STCD labels such notation as colloquial or misspelled, but at the same time admits its widespread use.

³⁷² Common depalatalization of /ʃ/ and its replacing by /s/ by the speakers of Hindustani may also be a result of the influence of other IA languages/dialects, where this type of change is more regular (like, for example, in Braj b^hāṣā). The phenomenon has long been known, cf. for example Bābur's remark (16th c.): "It has struck me that as a Hindustani pronounces *shīn* as *sīn* (i.e. *sh* as *s*), and as Kashmīr is the one respectable town in these mountains, no other indeed being heard of, Hindustanis might pronounce it Kasmīr" (Bābur 1922: 484).

³⁷³ Here another change is the debuccalization of the intervocalic stop /q/ → /h/.

³⁷⁴ Here again only such examples are shown in which the phonetic change is mirrored in the orthography. Cf. also 3.2.6.(iii).

3.6. Changes in Distribution of Consonant Phonemes in the Borrowed Words

3.6.1. Generally, changes in distribution can be divided into three types: (i) deletion; (ii) addition; (iii) rearrangement of sounds (metathesis).

3.6.2. Deletion involves the loss of one (or more) consonants, irrespective of their position in a word. Sometimes the loss is preceded by the lenition of a consonant such as degemination, fricativization or debuccalization:

(i) consonant elision occurs predominantly in the middle position, at first, probably, as a result of hasty or careless speech, but, ultimately, also to ease pronunciation and make it more economical. As in the case of other sound changes observable in Perso-Arabic loanwords, consonant elision has a collateral character, and its occurrence is mainly dependent on the speaker. Cf. the following examples: P داروغگی (*d-A-r-V-ġ-g-Y*) /daroʁægi/ 'superintendency' ≥ H داروغائی (*dārogāī*), داروغائی (*d-A-r-V-A-Y²-Y*) /daroga²i/ ~ /darogaji/ 't.s.'; P انبار (*A-n-b-A-r*) /æmbar/ 'full to the brim; storehouse' ≥ H انبار (*ambār*) ~ اعمار (*amār*), انبار (*A-n-b-A-r*) ~ امبار (*A-m-A-r*) /æmbar/ ~ /əmar/ 'heap, pile'; Class. P [≤ A] موافق (*m-V-A-f-q*) /muvafiq/ (modern P /movafeq/) 'agreeing, suitable' ≥ H موافق (*muāfīq*) ~ مافق (*māfīq*), موافق (*m-V-A-f-q*) /mu.afiq/ ~ /mafīq/ 't.s.'; P [≤ A] احتراز (*A-ḥ-t-r-A-z*) /ihtīraz/ 'guarding against' ≥ H احتراز (*et'rāz*), احتراز (*A-ḥ-t-r-A-z*) /etraz/ ~ /ehtīraz/ 't.s.'; P [≤ A] غرغرة (*ġ-r-ġ-r-ī*) /ɣærɣæræ(t)/ 'gargling' ≥ H غرغرة (*garārā*),³⁷⁵ غرغرة (*ġ-r-ġ-r-h*) /ɣərara/ ~ /ɣærɣəra/ 't.s.'; P تباہی (*t-b-A-h-Y*) /təbahī/ 'ruin; destruction' ≥ H تباہی (*tabāhī*) ~ تباہی (*tavāī*), تباہی (*t-b-A-h-Y*) /təbahī/ ~ /təva.i/ 't.s.'; Class. P [≤ A] فارغخطی (*f-A-r-ġ-x-t-Y*) /færɣxæt:i/ (modern P /færɣxæt:i/) 'deed

³⁷⁵ This noun should not be confused with *garārā*, m. l. 'loose trousers (women's)'; 2. 'a large sack, tent-bag'.

of release (from debt or obligation)' ≥ H *فارغی* (*fār'khatī*), *فارغی* (*f-A-r-g-x-t-Y*) /*fārxəti*/ 't.s.'; P خوشحال (*x-V-š-h-A-l*) /*xuʃhal*/ 'happy' ≥ H خوشحال (*khus'ḥāl*), خوشحال (*x-V-š-h-A-l*) /*xuʃhal*/ ~ /*xuʃal*/ 't.s.'; P [≤ A] سنگترة (*s-n-g-t-r-h*) /*sængtæra*/ 'a kind of orange' ≥ H संतरा (*santarā*), سنتره (*s-n-t-r-h*) /*səntərə*/ 't.s.'; Class. P [≤ A] مسودة (*m-s-V-d-ī*) /*musaʊvædæ(t)*/ (modern P /*mosavvædæ*/) 'a rough draft' ≥ H مسودا (*masaudā*), مسوده (*m-s-V-d-h*) /*məsauda*/ ~ /*musvæd:a*/ 't.s.';

(ii) a specific kind of elision is the loss of the intervocalic glide /j/, e.g.: Class. P گنجایش (*g-n-j-A-Y-š*) /*gundʒajɪʃ*/ (modern P /*gondʒajɪʃ*/) 'space, room' ≥ H گنجایش (*guñjāis*), گنجایش (*g-n-j-A-Y-š*) /*gundʒa.ɪʃ*/ 't.s.'; Class. P فرمایش (*f-r-m-A-Y-š*) /*færmajɪʃ*/ (modern P /*færmajɪʃ*/) 'order' ≥ H فرمایش (*far'māis*), فرمایش (*f-r-m-A-Y-š*) /*fərma.ɪʃ*/ 't.s.'; Class. P نمایش (*n-m-A-Y-š*) /*numajɪʃ*/ (modern P /*næmajɪʃ*/) 'show (of), display' ≥ H نمایش (*numāis*), نمایش (*n-m-A-Y-š*) /*numa.ɪʃ*/ 't.s.';

(iii) apheresis, or the loss of a consonant from the beginning of a word occurs sporadically and only as an alternative realization, e.g.: P پختنی (*Y-x-n-Y*) /*jæxni*/ 'cooked' ≥ H پختنی (*yakh'ni*) ~ آبخنی (*akh'ni*), پختنی (*Y-x-n-Y*) /*jæxni*/ ~ /*æxni*/ 'soup';

(iv) in some cases consonant loss occurs at the end of a word, e.g.: Class. P [≤ A] هودج (*h-V-d-j*) /*haudædʒ*/ (modern P /*həudædʒ*/) 'a camel-litter' ≥ H هودا (*haudā*), هودج (*h-V-d-h*) /*hauda*/ 't.s.'; P [≤ A] افراط و تفریط (*A-f-r-A-t o t-f-r-Y-t*) /*ɪfrato.tæfrit*/ 'excess, abundance' ≥ H افراط-تفریط (*af'rā-taf'rī*), افراط و تفریط (*A-f-r-A-t o t-f-r-Y-t*) /*əfra.tæfri*/ ~ /*ɪfrato.tæfrit*/ 't.s.';

(v) it happens that the final voiceless aspirate [h] is dropped and the preceding vowel is elongated (analogously to words which end with the 'silent' *h*), even though in these words /h/ is not a remnant of *tāʔ marbūṭah*,³⁷⁶ e.g.: P کرگه (*k-r-g-h*) /*kærgæh*/

³⁷⁶ Note that in contemporary colloquial Persian similar compensatory lengthening is a common phenomenon, and /h/ is often dropped and re-

'a weaver's loom or shop' ≥ H कर्गा (*kar'gā*), کَرگ (k-r-g-h) /*kərga*/'³⁷⁷ 't.s.'; Class. P [≤ A] شِبْهَة (*š-b-h-ī*) /*ʃubhæ(t)*/ (modern P شِبْهَة (*š-b-h-h*) /*ʃobhæ*/) 'doubt' ≥ H सुवहा (*sub'hā*), شِب (š-b-h) /*subha*/ ~ /*ʃubh*/ ~ /*ʃuba*/ 't.s.'. The final Persian /h/ which in Arabic loanwords has superseded the pharyngeal [h] also can be subject to elision, e.g.: P [≤ A] صَحِيح (*š-h-Y-h*) /*sæhih*/ 'complete, perfect' ≥ H सही (*sahī*), صَح (*š-h-Y-h*) /*səhi*/ ~ /*səhih*/ 't.s.'; P [≤ A] تَسْبِيح (*t-s-b-Y-h*) /*tæsbih*/ 'a rosary' ≥ H तसबी (*tas'bī*), تَسْبِج (*t-s-b-Y-h*) /*təsbi(h)*/ 't.s.';

(vi) consonant elision, occurring word-finally or before voiceless stops, is sometimes accompanied by nasalization of a preceding long vowel,³⁷⁸ e.g.: P [≤ A] أَشْيَان (*Ā-š-Y-A-n*) /*ʔʃjan*/ 'a nest' ≥ H आशियाँ (*āśiyām*), أَشْيَان (*Ā-š-Y-A-n*) /*ʔʃjā*/ 't.s.'; P [≤ A] قَقْنَس (*q-q-n-s*) /*qæqnus*/ 'the phoenix (bird)' ≥ H ककनू (*kaḳ'nūm*), قَقْنَس (*q-q-n-s*) /*qæqnū*/ ~ /*qæqnus*/ 't.s.'; P زمین (*z-m-Y-n*) /*zæmin*/ 'earth, ground' ≥ H जमीँ (*zamīm*) ~ जमीन (*zamīn*), زمين (*z-m-Y-n*) ~ زمين (*z-m-Y-n*) /*zəmī*/ ~ /*zəmin*/ 't.s.'.

3.6.3. Addition primarily involves adding the alveo-palatal approximant /j/ as a glide – which is the most common kind of excrescence in Perso-Arabic loanwords. The glide may appear intervocalically as well as between a vowel and a consonant. Its primary function, however, is to ease the pronunciation, not to prevent hiatus, which is permissible in Hindustani, e.g.: H तवायफ़ (*tavāyaf*), طوائف (*t-V-A-Y²-f*) /*təvajəf*/ ~ /*təvajɪf*/ 'a dancing-girl' ≤ P [≤ A] طوائف (*t-V-A-Y²-f*) /*tævaʔɪf*/ 't.s.'; H जाया (*jāyā*), خالَج (*jāyā*), خالَج (*jāyā*)

placed (especially in high-frequency, everyday words) by a lengthening of the previous vowel e.g.: [kuta:] instead of /kutah/ 'short', [tæ:si:] instead of /tashi:/, [be:tær] instead of /behtær/, etc.

³⁷⁷ The alternative form with metathesis *vide infra* 3.6.4.

³⁷⁸ Instances of nasalized short vowels and of nasalized long vowels before voiced stops, resulting from a deleted nasal consonant, are allophonic (Masica 1991: 117-118), e.g.: P طنز (*t-n-z*) /*tænz*/ 'bantering' ≥ H तंज (*tamz*), طَنَز (*t-n-z*) /*tāz*/ ~ /*tənz*/ 't.s.'.

(*z-A-Y²-^c*) /dʒaja/ ~ /zaje/ 'wasted, lost' ≤ P [≤ A] ضائع (*z-A-Y²-^c*) /zaʔiʔ/ 't.s.'; H पायक (*pāyak*), پیک (*p-Y-k*) /pajək/ ~ /paɪk/ 'footman; messenger' ≤ Class. P پیک (*p-Y-k*) /paɪk/ (modern P /pejk/) 't.s.'; H हमायल (*hamāyal*), حمائل (*h-m-A-Y²-l*) /həmajəl/ 'a necklace of flowers, or of rupees' ≤ Class. P [≤ A] حمائل (*h-m-A-Y²-l*) /həmaʔil/ (modern P /həmaʔel/ ~ /həmajel/) 't.s.'; H ज़ायकेदार (*zāy'kedār*), ذائقه دار (*z-A-Y²-q-h d-A-r*) /zajqedar/ ~ /zaʔiqadar/ 'tasty' ≤ Class. P ذائقه دار (*z-A-Y²-q-h d-A-r*) /zaʔiqedar/ (modern P /za.ɛqedar/) 't.s.'; H दरियाई (*dariyāi*), دارائی (*d-A-r-A-Y²-Y*) /dərija.i/ ~ /daraʔi/ 'satin' ≤ Class. P دارائی (*p-Y-k*) /daraʔi/ (modern P دارایی (*d-A-r-A-Y-Y*) /daraji/) 'a kind of silk stuff'. The labio-dental approximant /v/ may also appear in similar function, although not as frequently as /j/, e.g.: H जौलाने (*jaulānī*), جاولانی (*j-V-l-A-n-Y*) /dʒaulani/ ~ /dʒəvəlani/ 'quickness, alertness' ≤ Class. P [≤ A] جاولانی (*j-V-l-A-n-Y*) /dʒaulani/ (modern P /za.ɛqedar/) 't.s.'. Other cases of excrescence are sparse, e.g.: H अलमस्त (*al'mast*), المست (*A-l-m-s-t*) /əlməst/ 'drunk' ≤ P مست (*m-s-t*) /mæst/ 't.s.'; H चुनिंदा (*cunindā*), چنده (*ĉ-n-n-d-h*) /tʃuninda/ ~ /tʃunīda³⁷⁹ 'gathered' ≤ P چنده (*ĉ-n-Y-d-h*) /tʃinida/ 't.s.'; H जंजाल (*jañjāl*), جنجال (*j-n-j-A-l*) /dʒəndʒal/ 'a large musket' ≤ P جزائل (*j-z-A-Y²-l*) /dʒæzaʔil/ 't.s.'; H तंदूर (*tandūr*), تندور (*t-n-d-V-r*) /təndur/ 'oven' ≤ P تنور (*t-n-V-r*) /tæn:ur/ 't.s.'.

3.6.4. There are few examples of strictly consonantal, adjacent metathesis, e.g.: H खुदा (*khurdā*), خرده (*x-r-d-h*) /xurda/ ~ खुदरा (*khud'rā*) /xudra/ 'small (in size or quantity)' ≤ P خرده (*x-r-d-h*) /xurda/ 't.s.'; H खुदराफरोश (*khud'rāfaroš*) /xudrafəroʃ/ ~ خرده فروش (*x-r-d-h f-r-V-š*) /xurdafəroʃ/ 'a pedlar' ≤ P خرده فروش (*x-r-d-h f-r-V-š*) /xurdafəroʃ/ 't.s.'; H ओहदा (*oh'dā*), عهدہ (*^c-h-d-h*) /ohda/ ~ ओधा (*odhā*) /odhə³⁸⁰ 'an office, duty; position' ≤ Class. P [≤ A]

³⁷⁹ Nasalization of the short vowel in this case is allophonic (see the previous footnote).

³⁸⁰ Here one can observe aspiration through metathesis. Another such example is given in Beg (1988: 77): H/U /əndhə/ 'grief' ≤ P /andoh/ 't.s.'.

عهدہ (°-h-d-ī) /ʰuhdæ(t)/ (modern P عهدہ (°-h-d-h) /ohdɛ/) ‘t.s.’; H पनहा (pan’hā) /pənha/ ~ پنهنا (p-h-n-A) /pəhna/ ‘breadth’ ≤ P پهنا (p-h-n-A) /pəhna/ ‘t.s.’. Other instances of metathesis involve both consonants and vowels, e.g.: H अहवाल (ah’vāl), احوال (A-ḥ-V-A-l) /əhval/ ~ हवाल (havāl) /həval/ ‘state, condition’ ≤ P [≤ A] احوال (A-ḥ-V-A-l) /æhval/ ‘t.s.’; H करणा (kar’gā), کړنه (k-r-g-h) /kərga(h)/ ~ करघा (kar’ghā), کړگه (k-r-gʰ-A) /kərgʰa/³⁸¹ ‘a weaver’s loom or shop’ ≤ P کرگه (k-r-g-h) /kərgah/ ‘t.s.’; H मातहत (māt’hat) /mathət/ ~ ماتحت (m-A-t-ḥ-t) /matəht/ ‘what is under’ ≤ P ماتحت (m-A-t-ḥ-t) /matæht/ ‘the lower part’; H ईजाद (īzād), ایزاد (A-Y-z-A-d) /izad/ ‘increase; addition’ ~ ज्यादा (zyādā), زیاده (z-Y-A-d-h) /zjada/ ‘more; much’ ≤ P زیاده (z-Y-A-d-ī) /zijadæ(t)/ (modern P زیاده (z-Y-A-d-h) /zijadɛ/) ‘t.s.’

³⁸¹ This form is an excellent example of nativization through spelling.

न था कुछ तो खुदा था, कुछ न होता तो खुदा होता
डुबोया मुझको होने ने, न होता मैं तो क्या होता

نہ تھا کچھ تو خدا تھا، کچھ نہ ہوتا تو خدا ہوتا
ڈوبوایا مجھ کو ہونے نے، نہ ہوتا میں تو کیا ہوتا

When nothing was, then God was there,
Had nothing been God would have been;
My being has defeated me
Had I not been, what would have been

Fig. 11. A plaque with a digraphic representation of Ghalib's couplet in Perso-Arabic and Nagari scripts. The verse reads in Nagari: *na t̥hā kuc^h to k^hudā t̥hā, kuc^h na hotā to k^hudā hotā / dūboyā muj^hko hone ne, na hotā mai^{m̐} to kyā hotā*, and in Perso-Arabic script: *n-h t̥h-A k-c^h t-V x-d-A t̥h-A, k-c^h n-h h-V-t-A t-V x-d-A h-V-t-A / d-b-V-Y-A m-j^h k-V h-V-n-e n-e, n-h h-V-t-A m-Y-ñ t-V k-Y-A h-V-t-A*. Mazār-e Gālib (Ghalib's Tomb), New Delhi. (Photo by the author, 2012)

4. Morphosyntax

Indeed, it stands very much to reason that the transfer of morphemes is facilitated between highly congruent structures.

(Weinreich 1953: 33)

4.1. General Morphosyntactic Characteristics of Persian and Hindustani: Similarities and Differences

4.1.1. Due to centuries-long language contact between Persian and Hindustani (which, actually, may be rightly assessed as fruitful co-existence), Persian has had a significant influence on Hindustani not only in the spheres of lexis or phonology, but also at the morphosyntactic level. It is an unquestionable fact that a considerable degree of similarity between the morphological and syntactic systems of both the languages has remained a key force in this process.

4.1.2. In Persian, and in Hindustani, there exist eight parts of speech (nouns, adjectives, pronouns, adverbs, verbs, numerals,³⁸² conjunctions, and interjections) and the so-called grammatical (auxiliary) words: prepositions (mainly in Persian) and postpositions (mainly in Hindustani). In contrast to Persian which has entirely lost inflection of nominal parts of speech³⁸³, in Hindustani there exists a moderate case-system (direct/nominative,

³⁸² Traditionally classified also as adjectives.

³⁸³ Old Persian (600-300 BC) was still a typical inflected language, similar in terms of its system to its contemporaries – Sanskrit and Classical Greek. By the late Middle Persian period (approx. 7th CE), the language had undergone a profound reduction of inflectional morphology and had become close to the analytical structure characteristic of New Persian (cf. Perry 2007: 975).

oblique and vocative) and the distinction of gender (masculine and feminine). In Persian, however, the declining inflexions have been entirely replaced by analytic constructions. Both Hindustani and Persian have the category of number (singular and plural).

4.1.3. Over the centuries, a number of important changes in syntax can also be observed. Previously used, rather free word orders (cf. Staal 1967; Kent 1950: 96), typical for languages with extended inflexions, have become settled and grammatical ones. Both Hindustani and New Persian have a relatively consistent word order, subject-object-verb (SOV), with the verb placed finally in the sentence. Nominal sentences, very common in the archaic period,³⁸⁴ do not exist in contemporary Hindustani, or in Persian,³⁸⁵ as they have been replaced by verbal sentences, which are most commonly formed by combination with the copula (in this function the Hindustani verb /hona/ and the Persian verb /budæn/ – both meaning ‘to be’ – are found). There exists, however, the difference within the construction of the noun phrase, in which the Hindustani order modifier-noun-postposition contrasts with the Persian order preposition-noun-modifier, with elements typically linked by the *izāfat*.

³⁸⁴ In Sanskrit both nominal sentences (including nominal compositions, very popular in this function) and verbal sentences were equally common. For example, the verbal sentence: भक्तिर्ज्ञानाय कल्पते (*bʰaktir jñānāya kalpate*) ‘love is conducive to knowledge’ could be successfully replaced with the nominal sentence: भक्तैर्ज्ञानकल्पन (*bʰakter jñānakalpanā*) / ज्ञानाय भक्तिकल्पना (*jñānāya bʰaktikalpanā*) or with the composition: ज्ञानभक्तिकल्पना (*jñānabʰaktikalpanā*), cf. Staal 1995: 81-83.

³⁸⁵ Except for nominal sentences borrowed from Arabic, which favours this kind of structure, e.g.: P [≤ A] لا حول ولا قوة الا بالله (*l-A ḥ-V-l V l-A q-V-ī A-l-A b-A-ll-h*) /lahaulæ væ la qov:ætæ æl:a bīl:ahæ/ ‘there is neither might nor strength but in God’ and the corrupted interjection in Hindustani: लाहौल विला कूबत (*lāhaul bilā kūbat*), لا حول ولا قوة (*l-A ḥ-V-l V l-A q-V-ī*) /lahaul vɛla quvæt/ ~ /lahaul bīla quvæt/ ‘there is no strength nor power but in God: how horrible! how disgusting!’.

4.1.4. Another common characteristic of both Hindustani and Persian is their proclivity towards derivation and composition. Both the languages inherited extensive and firmly established models of word-formation from their Old Indo-Aryan and Old Iranian ancestors. Moreover, Hindustani remains in a specially favourable position being able to supplement its own richly productive patterns with a number of new elements borrowed from Persian and Arabic (and, during the last decades, also from English). Generally, Hindustani and Persian systems of derivation and composition, both nominal and lexical, are congruous in many respects and comprise similar categories, such as deverbal nominals, nominals created by derivational prefixes and suffixes,³⁸⁶ nominal compounds (copulative, possessive/attributive and determinative), denominal verbs,³⁸⁷ causative verbs and verbal compounds.³⁸⁸ A more detailed theoretical description of these features would exceed the requirements of the present study but they will be referred to in the course of further discussion, devoted to adaptation of Perso-Arabic loanwords at the morphosyntactic level, that will follow in this chapter.

³⁸⁶ Both in Hindustani and in Persian suffixes are the principal means of nominal derivation. The special group within this class of affixes are deverbal (or verb stem) suffixes, very productive in Persian but also quite frequently adopted in a number by Hindustani, *vide infra* 4.2.14.(iii). They occupy the intermediate position between suffixes and second members of compounds.

³⁸⁷ Created by suffixing an infinitive suffix (-/ɪdæn/ in Persian, -/na/ in Hindustani) to a nominal.

³⁸⁸ This category includes two fundamentally distinct classes of verbs: those formed of a verb stem (or a derivative) and a verb, and those composed of a nominal (or adverb/preposition) and a verb. In Hindustani the former are categorized as compound verbs, the latter – as conjunct verbs (cf. Burton-Page 1957).

4.2. Morphosyntactic Classification of the Perso-Arabic Borrowings in Hindustani

4.2.1. The vast majority of Perso-Arabic loanwords in Hindustani are nominal parts of speech, viz. nouns and adjectives. Within this capacious and categorically highly differentiated group one can distinguish several subcategories:

(i) simple nouns (i.e. not derived from other parts of speech), which constitute probably the most numerous group of Perso-Arabic loanwords, e.g.: H शहर (*śah'r*), شهر (*š-h-r*) /ʃəh(ə)r/ 'city' ≤ P شهر (*š-h-r*) /ʃəhr/ 't.s.'; H किताब (*kitāb*), کتاب (*k-t-A-b*) /kɪtab/ 'book' ≤ P کتاب (*k-t-A-b*) /kɪtab/ 't.s.'; H दिल (*dil*), دل (*d-l*) /dɪl/ 'heart' ≤ P دل (*d-l*) /dɪl/ 't.s.'; H मौत (*maut*), موت (*m-V-t*) /maut/ 'death' ≤ P موت (*m-V-t*) /maut/ 't.s.';

(ii) simple adjectives, e.g.: H तर (*tar*), تر (*t-r*) /tər/ 'wet, moist' ≤ P تر (*t-r*) /tær/ 't.s.'; H गलत (*galat*), غلط (*ġ-l-ṭ*) /gələt/ 'incorrect, untrue' ≤ P غلط (*ġ-l-ṭ*) /qələt/ 't.s.'; H खाली (*khālī*), خالی (*x-A-l-Y*) /xali/ 'empty, vacant' ≤ P خالی (*x-A-l-Y*) /xali/ 't.s.';

(iii) Persian participles (verbal adjectives), frequently acting as both adjectives and nouns, e.g.: H दीदा (*dīdā*), دیده (*d-Y-d-h*) /dida/ 1. adj. 'seen'; 2. m. 'eye' ≤ P دیده (*d-Y-d-h*) /didæ/ 't.s.',³⁸⁹ H परिंदा (*parindā*), پرندہ (*p-r-n-d-h*) /pərɪnda/ 1. adj. 'winged; flying'; 2. m. 'bird' ≤ P پرندہ (*p-r-n-d-h*) /pærənda/ 't.s.',³⁹⁰

(iv) Persian verbal nouns (formed from verbal roots and sometimes from infinitives), e.g.: H गिराफ्त (*girāft*), گرفت (*g-r-f-t*) /gɪræft/ f. 'seizing' ≤ P گرفت (*g-r-f-t*) /gɪrɪft/ 't.s.',³⁹¹ H रिवाज (*rivāj*), رواج (*r-V-A-j*) /rɪvadž/ m. 'being in demand' ≤ P رواج (*r-V-A-j*) /rævadž/ 't.s.',³⁹²

³⁸⁹ Perf. part. of P دیدن (*d-Y-d-n*) 'to see, look, &c.'.

³⁹⁰ Act. part. of P پریدن (*p-r-Y-d-n*) 'to fly in the air, to flutter; to throb'.

³⁹¹ 3. sg. praeteritum from P گرفتن (*g-r-f-t-n*) 'to take, seize, &c.'.

³⁹² Verbal (infinitive) noun of A رواج (*r-V-j*) 'to have a ready sale, &c.'.

(v) a number of of derivative forms created by means of affixes, e.g.: H रोशनी (*roś'nī*), روشنی (*r-V-š-n-Y*) /roʃni/ f. 'light, brightness' ≤ P روشنی (*r-V-š-n-Y*) /rouʃæni/ 't.s.' ≤ P روشن (*r-V-š-n*) /roʃæn/ adj. 'light, splendid'; H लाइलाज (*lāilāj*), لاعلاج (*l-A-^c-l-A-j*) /la.ɪladʒ/ adj. 'incurable' ≤ P [≤ A] علاج (*l-A-^c-l-A-j*) /la'ɪladʒ/ 't.s.' ≤ A علاج (*^c-l-A-j*) /ɪladʒ/ 'remedy, cure';

(vi) compound nouns and adjectives, e.g.: H आल-औलाद (*āl-aulād*), آل اولاد (*Ā-l A-V-l-A-d*) /al.aulad/ 'children, descendants' ≤ P آل (*Ā-l*) /al/ 'offspring...' + P اولاد (*A-V-l-A-d*) /aulad/ 'sons...'; H मैखाना (*maikhānā*), میخانه (*m-Y-x-A-n-h*) /maixana/ 'tavern; liquor shop' ≤ P میخانه (*m-Y-x-A-n-h*) /maixanæ/ 't.s.' ≤ P می (*m-Y*) /mai/ 'wine' + P خانه (*x-A-n-h*) /xanæ/ 'house'; H मेहमानवाज़ (*meh'mān'navāz*), مہمان نواز (*m-h-m-A-n n-V-A-z*) /mehman:əvaz/ 'hospitable' ≤ P مهمان نواز (*m-h-m-A-n n-V-A-z*) /mihman.nævaz/ 't.s.' ≤ P مهمان (*m-h-m-A-n*) /mihman/ 'guest' + P نواز (*n-V-A-z*) /nævaz/ 'comfort, satisfaction &c.'.

4.2.2. Adjectives (including verbal adjectives) are often subject to substantivization and, similarly as it happens in Persian (Mace 2003: 47), also in Hindustani most of the borrowed adjectives can be used as nouns if only the meaning permits such usage, e.g.: H जवान (*javān*), جوان (*j-V-A-n*) /dʒəvan/ 1. adj. 'young &c.'; 2. m. 'a youth &c.' ≤ P جوان (*j-V-A-n*) /dʒəvan/ 'young; a young man'; H बहादुर (*bahādur*), بہادر (*b-h-A-d-r*) /bəhadur/ 1. adj. 'brave &c.'; 2. m. 'hero' ≤ P بہادر (*b-h-A-d-r*) /bəhadur/ 't.s.'. There is also a group of loanwords which are classified as both nouns and adjectives, e.g.: H बदनसल (*bad'nas'l*), بد نسل (*b-d-n-s-l*) /bədɲəs/ adj. & m. 'of bad breed; a bastard' ≤ P بد نسل (*b-d-n-s-l*) /bədɲəs/ 't.s.'; H गैरमुल्की (*gair'mulki*), غیر ملکی (*ġ-Y-r-m-l-k-Y*) /ɣairmulki/ adj. & m. 'foreign; a foreigner' ≤ P غیر ملکی (*ġ-Y-r-m-l-k-Y*) /qairmulki/ 't.s.'. The phenomenon of adjectivization of nouns, although rarer, can also be observed, e.g.: H बूत (*but*), بت (*b-t*) /but/ 1. m. 'idol, image &c.'; 2. adj. 'stupefied, stunned, dumb-founded' ≤ P بت (*b-t*) /but/ 'an idol &c.'; H चाक (*cāk*), چاک

(*ĉ-A-k*) /ʃak/ 1. m. 'rent, slit, cut, fissure'; 2. adj. 'torn, lacerated' ≤ P چاک (*ĉ-A-k*) /ʃak/ 'a fissure, rupture &c.'.

4.2.3. Both nouns and adjectives borrowed by Hindustani can be marked (i.e. they have suffixes that change to show gender and number) and unmarked. Regarding such categories as gender, number and case, they are subject to adaptation processes which are discussed in paragraph 4.3.

4.2.4. Within the analyzed lexical material, the group of Persian and Arabic adverbs consists of approximately 180 words. These are both original (underived) adverbs and other parts of speech (mostly adjectives) which have become adverbialized. According to their semantic value, these adverbs may be categorized into several groups:

(i) adverbs of manner (the most numerous category, comprising over one third of all the borrowed adverbs), e.g.: *आहिस्ता* (*āhistā*), آهسته (*Ā-h-s-t-h*) /ahista/ 'slowly, softly' ≤ P آهسته (*Ā-h-s-t-h*) /ahistæ/ 't.s.';³⁹³ *बदिल* (*badiḷ*), بدل (*b-d-l*) /bədil/ 'heartily, cordially' ≤ P بدل (*b-d-l*) /bədil/ 't.s.';³⁹⁴ *फ़र-फ़र* (*far-far*) ~ *फर-फर* (*pʰar-pʰar*), *فر-فر* (*f-r-f-r*) /fərfər/ ~ /pʰərpʰər/ 'quickly; fluently' ≤ P *فر-فر* (*f-r-f-r*) /fərfər/ 'reading or writing expeditiously'; *दिलेराना* (*dilerānā*), دلیرانه (*d-l-Y-r-A-n-h*) /dilerana/ 'boldly, bravely' ≤ P *دلیرانه* (*d-l-Y-r-A-n-h*) /dīliranæ/ 't.s.'; *जबरन* (*jabʾran*), جبراً (*j-b-r-Ā*) /dʒəbrən/ 'forcibly; by compulsion' ≤ P [≤ A] *جبراً* (*j-b-r-Ā*) /dʒəbræn/ 't.s.'; *जल्दी* (*jaldī*), جلدی (*j-l-d-Y*) /dʒəldi/ 'quickly' ≤ P *جلدی* (*j-l-d-Y*) /dʒəldi/ 'speed, haste, quickness'; *मारफ़त* (*mārʾfar*), معرفت (*m-ʿ-r-f-t*) /marfat/ 'knowing, by means (of), through' ≤ P [≤ A] معرفت (*m-ʿ-r-f-t*) /mæʾrifæt/ 't.s.'; *तफ़रीहान* (*tafʾrīhan*), تفریحا

³⁹³ Also possible are: /ahiste/, /ahista-ahista/, /ahiste-ahiste/ and the elided form /aste/.

³⁹⁴ There is a whole group of adverbs created by means of the Persian preposition /bæ/ which in Hindustani functions as a prefix joining nouns and meaning 'with, along with', e.g.: /bæxubi/ 'well, excellently', /bæzor/ 'by force', /bæfauq/ 'with pleasure', etc.

(*t-f-r-Y-h-Ā*) /təfrihən/ 'by way of amusement' ≤ P [≤ A] تفریحاً
(*t-f-r-Y-h-Ā*) /təfrihən/ 't.s.';

(ii) adverbs of place, e.g.: अंदर (*andar*), اندر (*A-n-d-r*) /əndər/
'within, inside' ≤ P اندر (*A-n-d-r*) /ændər/ 't.s.'; दग्धे (dar'peś),
در پیش (*d-r-p-Y-š*) /dərpeʃ/ 'in front, forwards' ≤ P پیش (*d-r p-Y-š*)
/dərpiʃ/ 't.s.'; जा-बजा (*jā-bajā*), جا بجا (*j-A-b-j-A*) /dʒabədʒa/ 'here and
there; everywhere' ≤ P جا بجا (*j-A b-j-A*) /dʒa bædʒa/ 't.s.'; वापस
(*vāpas*), واپس (*V-A-p-s*) /vapəs/ 'back, behind' ≤ P واپس (*V-A-p-s*)
/vapəs/ 't.s.'; जेर (*zer*), زیر (*z-Y-r*) /zer/ 'under, below' ≤ P زیر (*z-Y-r*)
/zer/ ~ /zir/ 't.s.';

(iii) adverbs of time, e.g.: आइंदा (*āindā*), آینده (*Ā-Y^c-n-d-h*)
/a.inda/ 'in future' ≤ P آینده (*Ā-Y-n-d-h*) /ajændə/ 'coming; ensu-
ing, next'; बतारीख (*batārīkh*), تاریخ (*b-t-A-r-Y-x*) /bətarix/ 'on the
day of (in citing dates)' ≤ P به (*b-h*) /bə/ 'with' + P [≤ A] تاریخ
(*t-A-r-Y-x*) /tarix/ 'date'; फौरन (*fauran*), فوراً (*f-V-r-Ā*) /faurən/ 'at
once, immediately' ≤ P [≤ A] فوراً (*f-V-r-Ā*) /faurən/ 't.s.'; हमेशा
(*hameśā*), همیشه (*h-m-Y-š-h*) /həmeʃa/ 'always, constantly' ≤ P همیشه
(*h-m-Y-š-h*) /həmiʃə/ 't.s.'; पस (*pas*), پس (*p-s*) /pəs/ 'after, behind'
≤ P پس (*p-s*) /pəs/ 't.s.'; पेशतर (*peś'tar*), پیشتر (*p-Y-š-t-r*) /peʃtər/
'formerly, before' ≤ P پیشتر (*p-Y-š-t-r*) /peʃtər/ 'prior, more for-
ward'; देर (*der*), دیر (*d-Y-r*) /der/ 'delayed, late' ≤ P دیر (*d-Y-r*) /der/
~ /dir/ 't.s.';

(iv) adverbs of frequency, e.g.: अक्सर (*ak'sar*), اکثر (*A-k-š-r*)
/aksər/ 'usually, often' ≤ P [≤ A] اکثر (*A-k-š-r*) /aksər/ 't.s.'; अव्वलन
(*avvalan*), اولاً (*A-V-l-Ā*) /av:ələn/ 'first, first of all' ≤ P [≤ A] اولاً
(*A-V-l-Ā*) /æv:ælən/ 't.s.'; रोजाना (*rozānā*), روزانه (*r-V-z-A-n-h*)
/rozana/ 'daily' ≤ P روزانه (*r-V-z-A-n-h*) /rozana/ 't.s.'; साल-बसाल
(*sāl-basāl*), سال بسال (*s-A-l b-s-A-l*) /sāl bəsāl/ 'year by year, annu-
ally' ≤ P سال بسال (*s-A-l b-s-A-l*) /sāl bəsāl/ 't.s.'; मुदाम (*mudām*),
مدام (*m-d-A-m*) /mudam/ 'continuously; always' ≤ P [≤ A] مدام
(*m-d-A-m*) /mudam/ 'perpetual, continual, perennial'; मुकरर
(*mukarrar*), مکرر (*m-k-r-r*) /mukər:ər/ 'a second time, again' ≤ P
[≤ A] مکرر (*m-k-r-r*) /mukər:ər/ 't.s.';

(v) adverbs of certainty, e.g.: अलबत्ता (*al'battā*), البتّة (*A-l-b-tt-h*) /albət:a/ 'certainly, undoubtedly' ≤ P [≤ A] البتّة (*A-l-b-tt-h*) /albət:a/ 't.s.'; बेशक (*beśak*), بيشك (*b-Y-š-k*) /befək/ 'undoubtedly, certainly' ≤ P بی شک (*b-Y-š-k*) /bi fæk:/ 't.s.'; ग़ालिबन (*gālīban*), غالباً (*ġ-A-l-b-Ā*) /galībən/ 'most probably' ≤ P [≤ A] غالباً (*ġ-A-l-b-Ā*) /qalībən/ 't.s.'; शायद (*śāyad*), شاید (*š-A-Y-d*) /ʃajəd/ 'perhaps, probably' ≤ P شاید (*š-A-Y-d*) /ʃajəd/ 't.s.'; यकीनन (*yaqīnan*), يقيناً (*Y-q-Y-n-Ā*) /jəqinən/ 'certainly, assuredly' ≤ P [≤ A] يقيناً (*Y-q-Y-n-Ā*) /jəqinən/ 't.s.'; ज़रूर (*zarūr*), ضرور (*z-r-V-r*) /zərur/ 'certainly, of course' ≤ P [≤ A] ضرور (*z-r-V-r*) /zərur/ 'necessary, indispensable';

(vi) adverbs of degree, e.g.: विलकुल (*bil'kul*), بالكل (*b-A-l-k-l*) /bilkul/ 'entirely, completely; quite' ≤ P [≤ A] بالكل (*b-A-l-k-l*) /bilkul/ 't.s.'; करीबन (*karīban*), قريباً (*q-r-Y-b-Ā*) /qəribən/ 'approximately' ≤ P [≤ A] قريباً (*q-r-Y-b-Ā*) /qəribən/ 't.s.'; लबालब (*labālab*), لبالب (*l-b-A-l-b*) /ləbaləb/ '(full) to the brim' ≤ P لبالب (*l-b-A-l-b*) /ləbaləb/ 't.s.'; मुतलक (*mut'laq*), مطلق (*m-t-l-q*) /mutlək/ 'wholly, absolutely' ≤ P [≤ A] مطلق (*m-t-l-q*) /mutlək/ 'absolute, entire'; नज़दीक (*naz'dīk*), نزدیک (*n-z-d-Y-k*) /nəzdik/ 'near, approximately, almost' ≤ P نزدیک (*n-z-d-Y-k*) /nəzdik/ 't.s.'; तकरीबन (*taq'rīban*), تقريباً (*t-q-r-Y-b-Ā*) /təqribən/ 'approximately, about; nearly' ≤ P [≤ A] تقريباً (*t-q-r-Y-b-Ā*) /təqribən/ 't.s.'; ख़ुसूसन (*khusūsan*), خصوصاً (*x-š-V-š-Ā*) /xususən/ 'particularly, especially' ≤ P [≤ A] خصوصاً (*x-š-V-š-Ā*) /xususən/ 't.s.'; सिर्फ़ (*sirf*), صرف (*š-r-f*) /sirf/ 'only, merely' ≤ P [≤ A] صرف (*š-r-f*) /sirf/ 'mere, only, alone';

(vii) other adverbs, that cannot be assigned to any of the aforementioned categories, e.g.: हनोज़ (*hanoz*), هنوز (*h-n-V-z*) /hənoz/ 'still, yet' ≤ P [≤ A] هنوز (*h-n-V-z*) /hənoz/ ~ /hənuz/ 't.s.'; मसलन (*maslan*), مثلاً (*m-š-l-Ā*) /məsələn/ 'for example' ≤ P [≤ A] مثلاً (*m-š-l-Ā*) /məsələn/ 't.s.'; वग़ैरह (*vagairah*), وغيره (*V-ġ-Y-r-h*) /vəɣairəh/ 'and the rest, and so on' ≤ P [≤ A] وغيره (*V-ġ-Y-r-h*) /vəɣairəh/ 't.s.'.

4.2.5. Borrowed Persian and Arabic numerals and words indicating number or quantity in Hindustani are not very numerous and occur relatively rarely. They can be grouped as follows:

(i) the cardinal numbers, usually acting as parts of compounds (although they can also appear independently), e.g.: P یک /jæk/ 'one' in H [$\leq$ P] एकजा (*yak'jā*), یکجا (*Y-k-j-A*) /jækɖʒa/ 'in one place, together'; P دو /do/ ~ /du/ 'two' in H [$\leq$ P] दुसाला (*dusālā*), دو ساله (*d-V-s-A-l-h*) /dusala/ 'of two years, biennial'; P سه /sih/ 'three' in H [$\leq$ P] सितार (*sitār*), سه تار (*s-t-A-r*) /sitar/ 'three-stringed guitar'; P چهار /tʃæhar/ 'four' in H [$\leq$ P] चहारदीवारी (*cahār'dīvārī*), چهار دیواری (*č-h-A-r d-Y-V-A-r-Y*) /tʃəhardivari/ 'an enclosing wall'; P پنج /pæɖʒ/ 'five' in H [$\leq$ P] पंजाब (*pañjāb*), پنجاب (*p-n-j-A-b*) /pæɖʒab/ 'five rivers: the Punjab'; P شش /ʃæʃ/ 'six' in H [$\leq$ P] ششدار (*śas'dār*), ششدر (*š-š-d-r*) /ʃəʃdər/ 'six-doored'³⁹⁵, astonished'; P ده /dəh/ 'ten' in H [$\leq$ P] दहा (*dahā*), ده (*d-h-A*) /dəha/ 'the first ten days of the month of Muḥarram'; P هزار /hæzar/ 'thousand' $\geq$ H हजार (*hazār*), هزار (*h-z-A-r*) /hæzar/ 't.s.', which is the most assimilated Persian numeral in Hindustani;

(ii) the following borrowed ordinal numbers appear infrequently instead of the regular Hindustani ones: H अव्वल (*avval*), اول (*A-V-l*) /əv:vəl/ 'first' $\leq$ P [$\leq$ A] اول (*A-V-l*) /æv:vəl/ 't.s.'; H दोयम (*doyam*), دوم (*d-V-Y-m*) /dojəm/ ~ /dujəm/ 'second'³⁹⁶ $\leq$ P دویم (*d-V-Y-m*) /dujum/ 't.s.'; H सानी (*sānī*), ثانی (*š-A-n-Y*) /sani/ 'second' $\leq$ P [$\leq$ A] ثانی (*š-A-n-Y*) /sani/ 't.s.'; H सालिस (*sālis*), ثالث (*š-A-l-š*) /salis/ 'third' $\leq$ P [$\leq$ A] ثالث (*š-A-l-š*) /salis/ 't.s.';

(iii) the borrowed fractional numbers include: H निस्फ (*nisf*), نصف (*n-š-f*) /nisf/ 'a half' $\leq$ P [$\leq$ A] نصف (*n-š-f*) /nisf/ 't.s.'; H नीम (*nīm*), نیم (*n-Y-m*) /nim/ 'half' $\leq$ P نیم (*n-Y-m*) /nim/ 't.s.'; H दांग

³⁹⁵ Viz. 'six-door' or the moment in a board game where one's piece cannot move, at which point the a player becomes perplexed.

³⁹⁶ Note that the more common Persian form دوم (*d-V-m*) /dovum/ (modern P /dov:om/) practically does not occur in Hindustani (it is noted exclusively in Platts's dictionary).

(*dāmg*), دانگ (*d-A-n-g*) /*dang*/ 'a sixth part'³⁹⁷ ≤ P دانگ (*d-A-n-g*) /*danug*/ ~ /*dang*/ 'the fourth part of a drachm; a sixth of anything'; H سولس (*suls*), ثلث (*ṣ-l-ṣ*) /*suls*/ 'one third'³⁹⁸ ≤ P [≤ A] ثلث (*ṣ-l-ṣ*) /*suls*/ 't.s.';

(iv) the collective numbers are most frequently used in their nominalized forms and their secondary (metaphorical) meaning: H पंजा (*pañjā*), پنج (*p-n-j-h*) /*pəndʒa*/ 'an aggregate of five'³⁹⁹ ≤ P پنجه (*p-n-j-h*) /*pəndʒæ*/ 'the palm of the hand with the five fingers, or the sole of the foot with the five toes...'; H हफ्ता (*haftā*), هفته (*h-f-t-h*) /*hæfta*/ 'a week; hebdomad' ≤ P هفته (*h-f-t-h*) /*hæftæ*/ 't.s.'; H हजारहा (*hazār'hā*), هزارها (*h-z-A-r-h-A*) /*həzarha*/ 'thousands' ≤ P هزار ها (*h-z-A-r h-A*) /*həzarha*/ 't.s.'; H मुसद्दस (*musaddas*), سدس (*m-s-d-s*) /*musəd:əs*/ 'composed of six; hexagonal; hexameter' ≤ P [≤ A] مسدس (*m-s-d-s*) /*musæddæs*/ 't.s.';

(v) among the loanwords there is one distributive numeral: H निसफानिसफ (*nisfānisf*), نصفانصف (*n-ṣ-f-A-n-ṣ-f*) /*nisfanisf*/ 'half and half; by halves; in halves, (divided) in two' ≤ P نصفانصف (*n-ṣ-f-A-n-ṣ-f*) /*nisfanisf*/ 't.s.'. The categorically similar Persian word یکایک (*Y-k-A-Y-k*) /*jəkajək*/ 'one by one; suddenly; one opposed to another' has lost its distributive character in Hindustani and restricted its meaning to the adverbial 'suddenly, all of a sudden';

(vi) the following Arabic words (synonymous to genuinely Indic /*bar*/) are used for creating numeral adverbs ('once', 'twice', etc.): H दफ्ता (*dafā*), دفعة (*d-f-^c-h*) /*dəfa*/ 'time, occassion' ≤ P [≤ A] دفعة (*d-f-^c-ī*) /*dæf'æ(t)*/ 't.s.'; H मरतबा (*mar'tabā*), مرتبة (*m-r-t-b-h*) /*mərtəba*/ 'time, occassion' ≤ P [≤ A] مرتبة (*m-r-t-b-ī*)

³⁹⁷ Cf. other meanings: 'a minute weight; the equivalent of six *rattī* seeds; side, quarter; direction'.

³⁹⁸ STCD 221; Maddāh 707.

³⁹⁹ Cf. other meanings: 'a five (in cards, at dice); the hand with the fingers; paw, claw (of an animal); fig. 'clutch, grasp; power; the toe of a shoe; impression of open hand; rays'.

/mærtæbæ(t)/ 't.s.'; H नौबत (*naubat*), نوبت (*n-V-b-t*) /naubæt/ 'time, turn' ≤ P [≤ A] نوبة (*n-V-b-ī*) /naubæ(t)/ 't.s.';

(vii) the words indicating quantity (quantifiers), customarily categorized as adverbs or adjectives, include: H बस (*bas*), بس (*b-s*) /bæs/ 'enough; ample; very much; too much' ≤ P بس (*b-s*) /bæs/ 'a great number, many, more; very much...'; H चंद (*cand*), چند (*ĉ-n-d*) /ʃænd/ 'some, a few; several' ≤ P چند (*ĉ-n-d*) /ʃænd/ 't.s.'; H कसीर (*kasīr*), کثیر (*k-ṣ-Y-r*) /kəsir/ 'many, much; numerous' ≤ P [≤ A] کثیر (*k-ṣ-Y-r*) /kəsir/ 't.s.'; H ज्यादा (*zyādā*), زیادہ (*z-Y-A-d-h*) /zjada/ 'more; very many' ≤ P [≤ A] زیاده (*z-Y-A-d-ī*) /zjidadæ/ 'more; too much, excessively'; H बाज़ (*bāz*), بعض (*b-ĉ-z*) /baz/ 'some, certain, various' ≤ P [≤ A] بعض (*b-ĉ-z*) /baʔz/ 'more; too much, excessively'; H कम (*kam*), کم (*k-m*) /kəm/ 'little, few, small (of amount or quantity)' ≤ P کم (*k-m*) /kæm/ 't.s.'; H खुर्दा (*khurdā*), خرّ (*x-r-d-h*) /xurda/ ~ खुदरा (*khud'rā*) /xudra/ 'small (in size or quantity)' ≤ P خرده (*x-r-d-h*) /xurdæ/ 't.s.'.

4.2.6. Four Persian pronouns can be found among the analyzed loanwords: (i) the interrogative pronoun कि (*ki*), ک (k-h) /kɪ/ 'who?; which?; what?' ≤ P که (*k-h*) /kɪ/ 't.s.'; (ii) the reflexive pronoun खुद (*khud*), خود (*x-V-d*) /xud/ 'self (myself, himself, itself, &c.)' ≤ P خود (*x-V-d*) /xud/ 't.s.'; (iii) the indefinite pronoun फ़लā (*falām*), فلان (*f-l-A-n*) /fəlā/ ~ /fulā/ 'so and so; certain' ≤ P [≤ A] فلان (*f-l-A-n*) /fulan/ ~ /fælan/ 'he, that man, such a one, a certain'; (iv) the personal pronoun हू (*hū*), ه (*h-V*) /hu/ 'he; he is' ≤ P هو (*h-V*) /hu/ (≤ A هو (*h-V*) /huvæ/) 't.s.'.

4.2.7. There is a relatively numerous (approx. 25) group of borrowed Persian and Arabic conjunctions in Hindustani, comprising both coordinating and subordinating conjunctions. Coordinating conjunctions (coordinators) link two or more items (words, main clauses, or sentences) of equal syntactic importance. This group includes: (i) the connective conjunction

व (*va*), , (*V*) /və/ ~ /o/ 'and'⁴⁰⁰ ≤ P [≤ A] و (*V*) /væ/ ~ /o/ 't.s.'; (ii) the disjunctive conjunction⁴⁰¹ या (*yā*), یا (*y-A*) /ja/ 'or; either' ≤ P یا (*Y-A*) /ja/ 't.s.'; (iii) conjunctions used to express an alternative: वरना (*var'nā*), ورنه (*V-r-n-h*) /vərna/ 'otherwise, or else' ≤ P ورنه (*V-r-n-h*) /væɾ næ/ (≤ P و اگر نه (*VA-g-r n-h*) /væ ægær næ/) 't.s.'; खाह (*khvāh*), خواه (*k-V-A-h*) /x^(w)ah/ 'either, or; whether' ≤ P خواه (*k-V-A-h*) /x^(w)ah/ 't.s.'; (iv) adversative conjunctions: मगर (*ma-gar*), مگر (*m-g-r*) /mægər/ 'but, however' ≤ P مگر (*m-g-r*) /mægær/ 't.s.'; लेकिन (*lekin*), لیکن (*l-Y-k-n*) /lekɪn/ 'but, however' ≤ P [≤ A] لیکن (*l-Y-k-n*) /lekɪn/ 't.s.';⁴⁰² H बल्कि (*balki*), بلکه (*b-l-k-h*) /bəlki/ 'but, rather; moreover' ≤ P [≤ A] بلکه (*b-l-k-h*) /bəlki/⁴⁰³ 't.s.'; तहम् (*tāham*), تا هم (*t-A-h-m*) /tahəm/ 'yet, nevertheless' ≤ P تا هم (*t-A h-m*) /ta hæm/ 't.s.'; (v) conclusive (illative) conjunctions: चुनाँचे (*cunāmce*), چنانچه (*ĉ-n-A-n-ĉ-h*) /tʃunāŋʃe/ 'so, therefore' ≤ P چنانچه (*ĉ-n-A-n-ĉ-h*) /tʃunāŋʃi/ 't.s.'; लिहाजा (*lihāzā*), لہذا (*l-h-A-ȳ-A*) /lihaza/ 'therefore, hence' ≤ P لہذا (*l-h-ȳ-A*) /lihæza/ 't.s.'; यानी (*yānī*), یعنی (*Y-^c-n-Y*) /jani/ 'it meant: that is, i.e.' ≤ P یعنی (*Y-^c-n-Y*) /jæ'ni/ 't.s.'.

⁴⁰⁰ The conjunction /və/ may indicate hendiadys (= one through two) joining two independent words used rhetorically to replace the usual combination of independent word and its modifier. /və/ is then pronounced as an enclitic /o/ (often marked in the Nagari script), e.g.: दिलो दिमाग (*dilo di-māg*), دل و دماغ (*d-l V d-m-A-g*) /dilodimaɡ/ 'inner being, ambition' (lit. 'heart and mind'). Copulative constructions with -o- are discussed in 4.5.2.-3.

⁴⁰¹ The conjunction /ki/ is also sometimes used as disjunctive in the meaning 'or', *vide infra* 4.2.9 (iii).

⁴⁰² Note that /mægər/ 'but, except' and, in some cases, also /lekɪn/ 'but, however' are used as exceptive conjunctions. Cf. Platts 1909: 389.

⁴⁰³ In modern Persian the conjunction که (*k-h*) (also as a part of a compound) is realized phonetically as [kɛ]. The notation /ki/ given here represents its older pronunciation which correlates better with the traditional Hindustani utterance of this loanword (reflected in the Nagari notation). Note that in contemporary Hindustani/Urdu the pronunciation leans towards [kɛ].

4.2.8. Subordinating conjunctions (subordinators) join an independent clause with a dependent clause. They usually occur in correlative constructions, e.g.: /əgər... to.../ 'if... then...', /ja... ja.../ 'either... or...', although in colloquial speech correlatives are frequently deleted. In such constructions the subordinating conjunction marks the subordinate clause and the correlative marks the main clause. Conjunctions of this type may express: (i) possibility or supposition: अगर (*agar*), آگر (*A-g-r*) /əgər/ 'if; in case' ≤ P اگر (*A-g-r*) /ægær/ 't.s.'; (ii) concession: अगरचे (*agar'ce*), اگرچه (*A-g-r-č-h*) /əgərʃe/ 'although; even if' ≤ P اگرچه (*A-g-r-č-h*) /ægærʃi/ 't.s.'; हालाँकि (*hālāmki*), हालाँکه (*h-A-l-A-n-k-h*) /halākɪ/ 'even though' ≤ P حال آن که (*h-A-l Ā-n k-h*) /hal an kɪ/ 't.s.'; गो (*go*), گو (*g-V*) /go/ 'although; notwithstanding (that)' ≤ P گو (*g-V*) /go/ (present stem of گفتن (*g-f-t-n*) /guftæn/ 'speak');⁴⁰⁴ गोकि (*goki*), گوکي (*g-V k-h*) /go kɪ/ 'although'⁴⁰⁵ ≤ P گو که (*g-V k-h*) /go kɪ/ 't.s.'; हर-चंद (*har-cand*), هر چند (*h-r-č-n-d*) /hərʃænd/ 'although; however much; as often as' ≤ P هر چند (*h-r-č-n-d*) /hær ʃænd/ 't.s.'; (iii) resemblance: गोया (*goyā*), گویا (*g-V-Y-A*) /goja/ 'so to speak: as if, as though' ≤ P گویا (*g-V-Y-A*) /goja/ ~ /guja/ 'thus, in this manner, as you would say'; (iv) a cause or reason: चूँकि (*cūmki*), چونکه (*č-V-n-k-h*) /ʃūkɪ/ 'since, because'⁴⁰⁶ ≤ P چونکه (*č-V-n-k-h*) /ʃuŋkɪ/ 't.s.'; बशर्ते कि (*bašarte ki*), بشرطیکه (*b-š-r-ʔ-Y-k-h*)⁴⁰⁷ /bəʃarte kɪ/ 'if; in case, on condition that, provided that' ≤ P بشرطی که (*b-š-r-ʔ-Y-k-h*) /bəʃærti kɪ/ 't.s.'; (v) purpose or consequence: कि (*ki*), که (*k-h*) /kɪ/

⁴⁰⁴ In Persian this conjunction occurs only in the compound form /go kɪ/ – see the next example.

⁴⁰⁵ In Hindustani this form is archaic and used sparsely.

⁴⁰⁶ The conjunction /ʃūkɪ/ appears when the subordinate clause precedes the main clause and is paired with the correlative /ɪslɪje/, e.g. /ʃūkɪ vo bimar hē ɪslɪje vo nəhī aja/ 'Because he is ill, he did not come'.

⁴⁰⁷ Note the joint spelling of this and other compound conjunctions in Hindustani notation in Perso-Arabic script. In Persian all semantically distinct units are traditionally written separately.

‘in order to’ ≤ P که (*k-h*) /kɪ/ ‘t.s.’; ताकि (*tāki*), तकि (*t-A-k-h*) /takɪ/
 ‘so that, in order that’ ≤ P ता که (*t-A k-h*) /ta kɪ/ ‘t.s.’; काश (*kāś*), کاش
 (*k-A-š*) /kaʃ/ ‘would that’ ≤ P کاش که (*k-A-š*) /kaʃ/ ‘would to God!’ ×
 کاشکه (*k-A-š-k-h*) /kaʃkɪ/ ‘may it happen!’.

4.2.9. The complementizer /kɪ/ is used commonly to introduce a wide variety of dependent clauses, e.g.: (i) it can introduce the objects of verbs of speaking, knowing, thinking, wishing, etc., e.g. /mē dʒanta hū kɪ.../ ‘I know that...’; (ii) it occurs at the end of various phrases expressing possibility, necessity, etc., e.g. /mumkɪn he kɪ.../ ‘it is possible that...’; (iii) it can introduce both coordinate and subordinate clauses, replacing other conjunctions, like: /ja/ ‘or’, e.g.: /unse puʃhə kɪ vo aēge kɪ nəhī/ ‘Ask him if he is coming or not’ (Naim 1999 I: 164); /dʒəb/ ‘when’, e.g.: /mē dʒane hi vala thə kɪ vo mere pas aji/ ‘I was just about to go when she came to me’ (Shapiro 1989: 92); /kjōkɪ/ ‘because’, e.g.: /bol kɪ səʃ zinda he əb tək/ ‘Speak, because truth still lives’ (Schmidt 1999: 226).⁴⁰⁸

4.2.10. Persian and Arabic interjections in Hindustani are represented by approximately thirty words, divided into two groups. The first of them comprises ‘genuine’ interjections, or indeclinable particles expressive of emotion, e.g.: H आफ़रीन (*āfʹrīn*), آفرین (*Ā-f-r-Y-n*) /afrin/ ~ /afrī/ ‘bravo! well done!’ ≤ P أفیرین (*Ā-f-r-Y-n*) /afirin/ ~ /afrin/ ‘praise; bravo!’; H आमीन (*āmīn*), آمین (*Ā-m-n*) /amin/ ‘so be it, amen’ ≤ P [≤ A] آمین (*Ā-m-n*) /amin/ ‘t.s.’; H नौज़ (*nauz*), نعوذ (*n-^c-V-z*) /na.ʊz/ ‘God, preserve us!, God forbid!’ ≤ P [≤ A] نعوذ بالله (*n-^c-V-z b-A-l-l-h*) /naʊzʊbɪllah/ ‘we seek refuge in God’; H आबाश (*śābās*), شایش (*š-A-b-A-š*) /ʃabəʃ/ ‘well done! splendid!’ ≤ P شاباش (*š-A-b-A-š*) /ʃabəʃ/ ‘t.s.’; H वाह (*vāh*), وا (*V-A*) /va(h)/ ‘splendid!, wonderful!; goodness!; ah!, o!, alas!’ ≤ P واه (*V-A-h*) /vah/ ‘t.s.’; H या (*yā*), یا (*Y-A*) /ja/ ‘o!, oh!’ ≤ P

⁴⁰⁸ More examples presenting various syntactic functions of this conjunction see Schmidt 1999: 225-227; Naim 1999 I: 163-165.

یا (Y-A) /ja/ 't.s.'; H जिंदाबाद (*zindābād*), زنده باد (z-n-d-A-b-A-d) /zindabad/ 'may (he) live: long live...!' ≤ P زنداباد (z-n-d-A-b-A-d) /zindabad/ 't.s.'. The second (more numerous) group consists of substantives (nouns and adjectives) which are used in the function of interjections, e.g.: H अल्लाह-ताला (*allāh-tālā*), تالای الله (t-^c-A-l-ā A-l-l-h) /əllah.tala/ ~ /tala.əllah/ 'God the All-high' ≤ P [≤ A] تالای الله (t-^c-A-l-ā A-l-l-h) /ta'la.əllah/ 't.s.';⁴⁰⁹ H अलविदा (*al'vidā*), الوداع (A-l-V-d-A-^c) /əlvɪdɑ/ 'good-bye, adieu'⁴¹⁰ ≤ P [≤ A] الوداع (A-l-V-d-A-^c) /əlvædɑ/ 't.s.'; H अफ़सोस (*aʃsos*), افسوس (A-f-s-V-s) /əʃsos/ 'alas!, ah!, dear!; dear dear!'⁴¹¹ ≤ P افسوس (A-f-s-V-s) /əʃsos/ 'what a pity!'; H आदाब (*ādāb*), آداب (Ā-d-A-b) /adab/ 'good morning (etc.), good bye'⁴¹² ≤ P [≤ A] آداب (Ā-d-A-b) /adab/ 'civilities, good manners...'; H बंदगी (*bandagī*), بندگی (b-n-d-g-Y) /bəndəgi/ 'greetings!'⁴¹³ ≤ P بندگان (b-n-d-g-Y) /bəndægi/ 'servitude; compliment'; H बावेली (*bāvelī*), وایلاه (V-A V-Y-l-A-h) /baveli/ ~ /vavela/ 'alas!, ah!'⁴¹⁴ ≤ P [≤ A] وایلاه (V-A V-Y-l-A-h) /va vailah/ 't.s.'; H हैफ़ (*haif*), حیف (ḥ-Y-f) /haif/ 'ah!, alas!; what a pity!'⁴¹⁵ ≤ P [≤ A] حیف (ḥ-Y-f) /haif/ 't.s.'; H ख़ैर (*kḥair*), خیر (x-Y-r)

⁴⁰⁹ Other interjections employing the name of God include वल्लाह (*val-lāh*), والله (V-A-l-l-h) /vallah/ ~ /və.əllah/ 'by God!'; अल्लाह-अल्लाह (*allāh-allāh*), الله الله (A-l-l-h A-l-l-h) /əllah.əllah/ 'good heavens!' (of surprise, approval, &c.), अल्लाहो-अक़बार (*allāho-ak'bar*), الله اکبر (A-l-l-h-A-k-b-r) /əllaho.əkbar/ 'God is great!' (an exclamation of surprise, wonder, resignation, &c.).

⁴¹⁰ Also m. 'farewell'.

⁴¹¹ Also m. 'sorrow, regret'.

⁴¹² Also m. pl. 'greetings; compliments, salutation, rules or etiquette, good manners, mores'.

⁴¹³ Also f. 'devotion, worship; salutation'.

⁴¹⁴ Also m. 'turmoil; creating turmoil; weeping and wailing', although it seems that in this case the genuine exclamation was subjected to substantivization and not *vice versa*.

⁴¹⁵ Also m. 'sorrow, regret; shame'.

/xair/ 'indeed! really!'⁴¹⁶ ≤ P [≤ A] خیر (x-Y-r) /xair/ 't.s.'; H तोबा (tobā), توبہ (t-V-b-h) /toba/ 'good heavens!, heaven forbid!'⁴¹⁷ ≤ P [≤ A] توبہ (t-V-b-ī) /taubæ/ ~ /tobæ/ 'repenting'; H दुरुस्त (durust), درست (d-r-s-t) /durust/ 'right! exactly!'⁴¹⁸ ≤ P [≤ A] درست (d-r-s-t) /durust/ 'entire, complete...'; H मुबारक (mubārak), مبارک (m-b-A-r-k) /mubaræk/ 'blessings!, congratulations!'⁴¹⁹ ≤ P [≤ A] مبارک (m-b-A-r-k) /mubaræk/ 'blessed; happy'.

4.2.11. In the analyzed material one can find less than ten genuine Persian and Arabic prepositions which in some contexts retain their prepositional character and precede the noun they govern.⁴²⁰ These are: H अज़ (az), از (A-z) /æz/ 'from, of' ≤ P از (A-z) /æz/ 't.s.'; H ब (ba), به (b-h) /bə/⁴²¹ 'with, by, in; upon' ≤ P به (b-h) /be/ 't.s.'; H बा (bā), با (b-A) /ba/ 'with, by; possessing' ≤ P बा (b-A) /ba/ 't.s.'; H बे (be), بے (b-e) /be/ 'without'⁴²² ≤ P بی (b-Y) /bi/ 't.s.'; H पर (bar), پر (b-r) /bær/ 'on, upon, up, above' ≤ P بر (b-r) /bær/ 't.s.'; H दर (dar), در (d-r) /dær/ 'in, on, at, by, concerning' ≤ P در (d-r) /dær/ 't.s.'; H फ़ी (fī), فی (f-Y) /fi/ 'in, into; per; as for, as of' ≤ P [≤ A] فی (f-Y) /fi/ 't.s.'; H मय (may), مع (m-^c) /maj/ 'together

⁴¹⁶ Also f. 'good, goodness, virtue, health, happiness, safety, welfare'; adv. 'all right, very well'; adj. 'good, best'; pref. 'well'.

⁴¹⁷ Also f. 'abjuring sin; repentance, penitence, adjuration, renunciation, giving up; conversion, reiteration of faith'.

⁴¹⁸ Also adj. 'right; correct; precise, exact; on good condition or order; sound (of health); fit, proper'.

⁴¹⁹ Also f. 'congratulation, felicitation'; adj. 'lucky, fortunate, auspicious, blessed'.

⁴²⁰ Prepositions are not very common in Hindustani and they are used predominantly in Persian phrases or Persianized style of speech.

⁴²¹ Although this preposition can stand alone, more often than not it joins the first letter of the following word, e.g.: बदिल (badil), بدل (b-d-l) /bədil/ 'heartily, cordially'.

⁴²² This preposition in Hindustani is classified more often as a negative prefix, vide *infra* 4.2.14.(i).

(with), in addition' $\leq P [\leq A]$ مع (m^{-c}) /ma/ ~ /maʔ/ 't.s.'; H ता ($t\bar{a}$), त (t-A) /ta/ 'until, as far as, as long as' $\leq P$ ता (t-A) /ta/ 't.s.'.

4.2.12. Most Persian and Arabic words which, due to their grammaticalization, can serve as prepositions in Persian (these are adverbs and nouns chiefly), take the function of postpositions in Hindustani. These are always complex postpositions and the borrowed element follows the grammatical postposition /ke/ (and /ki/ in some cases). Contrary to their Persian counterparts, postpositions occur after a noun they govern. Here are a few examples of such formations: H (के) अंदर (*andar*), (के) अन्दर (*A-n-d-r*) /ændər/ 'within, inside' $\leq P$ اندر (*A-n-d-r*) /ændər/ 't.s.'; H (के) बाद (*bād*), (के) बाद (*b-A-d*) /bad/ 'post, after' $\leq P$ بعد ($b^{-c}-d$) /bad/ 't.s.'; (की or के) वजाय (*bajāy*), (की or के) بجای (*b-j-A-Y^o-e*) /bədʒaj/ ~ /bədʒa.e/ 'instead' $\leq P$ بجای (*b-j-A-Y*) /bədʒaj/ 't.s.'; H (के) दरमियान (*darmiyān*), (के) درمیان (*d-r-m-Y-A-n*) /dærmijā/ 'in the middle of, between, among' $\leq P$ درمیان (*d-r-m-Y-A-n*) /dærmijan/ 't.s.'; H (के) सिवा (*sivā*), (के) سوا (*s-V-A*) /sivā/ 'except (for), apart (from)' $\leq P$ [$\leq A$] سوا (*s-V-A-Y*) /sivaj/ 't.s.'; H (की) तरफ़ (*taraf*), (की) طرف (*t-r-f*) /tərəf/ 'in the direction (of), towards, to' $\leq P$ [$\leq A$] طرف (*t-r-f*) /tərəf/ 't.s.'; H (के) ज़रिया (*zariyā*), (के) زیره (*z-r-Y^c-h*) /zarija/ 'by means (of), through the agency or intervention (of); in or by virtue'⁴²³ $\leq P$ [$\leq A$] ذریعة (*z-r-Y^c-ī*) /zariʔat/ 't.s.'.

4.2.13. Numerous Persian and Arabic verbal stems and infinitives have been transferred to Hindustani in the form of nominal units – verbal nouns and verbal adjectives (*vide supra* 4.2.1. (iii-iv)). They function as common nominals and many of them easily form conjunct verbs, adjoining the most popular Hindustani verbs.⁴²⁴ However, there is also a small group of sim-

⁴²³ A noun signifying 'means; agency, cause; source; connection, influence (with one)' in the function of a postposition.

⁴²⁴ This is not such a rare device and various languages use it to adapt alien verbs to their morphosyntactic systems, cf.: "(...) it has been noted that verbs are crosslinguistically less easily borrowed than nouns and if the

ple Persian and Arabic verbs borrowed by Hindustani. They take the Hindustani infinitive suffix *-na*⁴²⁵ (which joins either the past stem or the present stem of a Persian verb) and follow Hindustani verbal paradigms.⁴²⁶ Cf. e.g.: H *خاریدنا* (*khariḏ'nā*), *خارید* (*x-r-Y-d-n-A*) /*xəridnə*/ 'to buy' ≤ P *خرید* (*x-r-Y-d*) /*xərid*/, past stem of P *خریدن* (*x-r-Y-d-n*) /*xəridæn*/ 'to buy'; H *گزرنا* (*guzar'nā*), *گزر* (*g-z-r-n-A*) /*guzəmə*/ 'to pass, &c.' ≤ P *گذر* (*g-z-r*) /*guzær*/, present stem of P *گذشتن* (*g-z-š-t-n*) /*guzæštæn*/ 'to pass, to pass by'; H *آزماینا* (*āz'mānā*), *آزمای* (*Ā-z-m-A-n-A*) /*azmə*/ 'to test &c.' ≤ P *آزمای* (*Ā-z-m-A*) /*azmə*/, present stem of P *آزمودن* (*Ā-z-m-V-d-n*) /*azmudæn*/ 'to try, to prove'; H *فرماینا* (*far'mānā*), *فرمای* (*f-r-m-A-n-A*) /*fərmə*/ 'to order' ≤ P *فرمای* (*f-r-m-A*) /*fərmə*/, present stem of P *فرمودن* (*f-r-m-V-d-n*) /*fərmudæn*/ 'to order'.

4.2.14. The nominal system of Persian is quite rich in derivational affixes (Perry 2007: 1007ff.). Therefore, it should not be surprising that a large amount of Perso-Arabic formative units have been assimilated and have supplied the stock of genuine Hindustani word-formatting elements. Generally, Perso-Arabic affixes which have been adopted to and are productive⁴²⁷ in Hindustani can be categorized in three groups:

need for borrowing does arise, many languages instead borrow a nominal form of the verb and employ a native 'all-purpose' verb such as *do*, *make* as a means of turning that form into the equivalent of a verb" (Hock 1991: 386). Persian commonly employed the same pattern to accommodate a large amount of Arabic deverbal forms.

⁴²⁵ The suffix *-na* forms regular verbs also from borrowed nominal parts of speech: nouns, adjectives and participles, as well as from adverbs, cf. Kuczkiewicz-Fraś 2003b: 82.

⁴²⁶ Actually, such verbs operate in the same manner as regular and genuine Hindustani verbs – they can be both transitive and intransitive, they can also create causative forms, according to the Hindustani verb system (ibid.).

⁴²⁷ The indicator of productivity of a formant borrowed from another language is whether it does or does not create hybrid words, i.e. new lexical

(i) prefixes: H बे- (*be-*), -॒ (b-*e-*) /be/ 'in-, un-, im-, dis-, &c.' ≤ P -بی (*b-Y*) /bi/ ~ /be/ 'without; when prefixed to nouns it is equivalent to *in, un, im, ir, less, &c.*'; H ना- (*lā-*), -॒ (l-*A-*) /la/ 'not, there is not; without' ≤ P [≤ A] -॒ (l-*A*) /la/ 't.s.'; H ना- (*nā-*), -॒ (n-*A-*) /na/ 'not; un-, in-, dis-, &c.; without' ≤ P -॒ (n-*A*) /na/ '(before nouns) a negative or privative part., equivalent to *in, un, &c.*';

(ii) suffixes: forming diminutives: H -चा (*-cā*), -॒ (-*č-h*) /tʃa/ ≤ P -چہ (-*č-h*) /tʃæ/ (modern P /tʃɛ/; forming agent nouns (names of professions): H -ची (*-cī*), -॒ (-*č-Y*) /tʃi/ ≤ P [≤ T] -چی (-*č-Y*) /tʃi/; forming names of receptacles: H -दान (*-dān*) ~ -दानी (*-dānī*), -॒ (-*d-A-n*) ~ -دانی (-*d-A-n-Y*) /dan/ ~ /dani/ ≤ P -دان (-*d-A-n*) /dan/; forming nouns of agency: H -गार (*-gar*) ~ -गार (-*gār*), -॒ (-*g-r*) ~ -॒ (-*g-A-r*) /gər/ ~ /gar/ ≤ P -گر (-*g-r*) ~ -گار (-*g-A-r*) /gær/ ~ /gar/; forming nouns of occupation (= 'master' or 'keeper'): H -बन (*-ban*) ~ -वन (-*van*), -॒ (-*b-n*) ~ -॒ (-*V-n*) /bən/ ~ /vən/⁴²⁸ ≤ P -بن (-*b-n*) /bæn/; forming adjectives and adverbs (joining substantives): H -वार (-*vār*), -॒ (-*V-A-r*) /var/ ≤ P -وار (-*V-A-r*) /var/; forming abstract nouns (= '-ity', '-ism'): H -इयत (-*iyat*), -॒ (-*Y-t*) /iɐt/ ≤ P [≤ A] -یت (-*Y-t*) /iɐt/⁴²⁹

items with bases belonging to the system of the borrowing language. For details see Kuczkiewicz-Fraś 2003b.

⁴²⁸ The difference is purely allophonic, connected with the tendency towards neutralization of /v/ and /b/ observed in Hindustani.

⁴²⁹ Enumerated here are only those Perso-Arabic suffixes which are irrefutably productive in Hindustani. Additionally, there are also suffixes jointly inherited from the Indo-Iranian stage by both Hindustani and Persian, e.g.: the very productive abstract suffix -ई (-*ī*), -॒ (-*Y*) /i/, deriving abstract nouns ('-ness') from nouns and adjectives or the similarly common suffix -कार (-*kār*), -॒ (-*k-A-r*) /kar/ used in both languages to make agent nouns. For obvious reasons none of them can be treated as a borrowed item. Not taken into consideration here are also those Perso-Arabic suffixes which occur in Hindustani exclusively as part of borrowed Perso-Arabic lexical units, thus their productivity in Hindustani is not confirmed, e.g.: the

(iii) 'half-suffixes'⁴³⁰ – this is the most numerous group of borrowed Perso-Arabic formatives, comprising words which have their own meaning but very rarely appearing as independent lexical units. They can be classified either as second members of compounds or as regular suffixes, modifying the meaning of the word they join. In the group of nineteen Perso-Arabic half-suffixes which are productive in Hindustani (see Kuczkiewicz-Fraś 2003b: 92-102) all but two⁴³¹ are deverbal forms of various Persian verbs – most often their present stems – which, in Persian, are used extensively in composition and acting as determining components. Two half-suffixes that are most commonly used in Hindustani are: -دار (-*dār*), دار (-*d-A-r*) /dar/ forming nominals with the meaning 'having, holding; possessor' ≤ P دار (*d-A-r*) /dar/, the present stem of P داشتن (*d-A-š-t-n*) /daštæn/ 'to have, to hold, to possess', and -باز (-*bāz*), باز (-*b-A-z*) /baz/ 'playing; player' ≤ P باز (*b-A-z*) /baz/ present stem of P باغتن (*b-A-ġ-t-n*) /baytæn/ 'to play &c.'. ⁴³²

popular suffix -آباد (-*ābād*), آباد (-*A-b-A-d*) /abad/, forming names of places, or the suffix -مند (-*mand*), مند (-*m-n-d*) /mænd/ used to create adjectives having the sense 'endowed with..., possessed of...'.
⁴³⁰ This term is an English equivalent of the apellation 'polusuffixsȳ', used to describe this category of formatives by Russian linguists (e.g. Barannikov & Barannikov 1956; Beskrovnȳi 1960a, 1960b; Zograf 1960, 1961).

⁴³¹ These two are: -خانه (-*khānā*), خانه (-*x-A-n-h*) /xana/ 'house' ≤ P خانه (-*x-A-n-h*) /xanæ/ 't.s.', forming mostly masculine names of places and -شاه (-*šāh*), شاه (-*š-A-h*) /jah/ 'king' ≤ P شاه (-*š-A-h*) /jah/ 't.s.', creating masculine neologisms, with some negative shade of meaning.

⁴³² As in the case of suffixes, there are numerous Perso-Arabic half-suffixes which occur in Hindustani as part of Perso-Arabic compound loanwords, e.g.: -گاه (-*gāh*), گاه (-*g-A-h*) /gah/, in Persian forming nouns of location or -بخش (-*bakhs*), بخش (-*b-x-š*) /bæxʃ/, creating adjectives having the meaning 'giving'. Since they have not become productive elements of the

4.3. Adaptation of Borrowings to the Hindustani Morphological System

4.3.1. As far as gender, number and case are concerned, Perso-Arabic loanwords have been almost wholly assimilated in the basic structure of the Hindustani language. There are two numbers, singular and plural. The rules for plural formation in Hindustani generally apply to all the Perso-Arabic loanwords. Unlike Persian, there is a two way gender system in Hindustani and all inanimate objects are classified either as masculine or feminine. Similarly, Persian does not have case inflections. Perso-Arabic loanwords are fully assimilated into Hindustani, usually taking all the case inflections and categories.

4.3.2. In Hindustani, masculine or feminine gender is mandatorily assigned to all borrowed nouns. It is a prefatory but at the same time absolutely crucial stage of adaptation of a loanword since all its further assimilation to the system (e.g. inflection for number or the declension pattern) and its functioning in the morphosyntactic tissue of the borrowing language (i.e. conditioning the form of the parts of speech affected by gender agreement – nouns, verbs, postpositions and adjectival modifiers) depend greatly on its prior gender assignment. In Hindustani, the grammatical categories of number and case are strictly connected with gender, which affects count distinction and declension of nominals as well as conjugation of verbs.

4.3.3. Perso-Arabic loanwords are assigned a gender mostly as per their word structure, i.e. on the basis of their word-final vowel or consonant, or on the basis of their semantics, or both. In Hindustani, lexical categories that carry the value of gender (or change their form according to the gender of the noun they refer

to) can be marked or unmarked and this quality also applies to the borrowed words. Thus, Perso-Arabic loan nouns which end with /a/⁴³³ are marked for the masculine gender and those which end with /i/, noted by the suffix -ई (-ī), -ی (-Y)⁴³⁴ are marked for the feminine gender.⁴³⁵ They can be further classified as either direct (nominative) or oblique (non-nominative, usually signalled by a postposition) and take appropriate endings which replace the marking suffixes /a/ and /i/ according to the following paradigm:⁴³⁶

⁴³³ This group comprises loan nouns which originally, in Persian or Arabic, ended with (i) *alif* (ا -); (ii) the 'silent' *h* (ه -); (iii) *ain* (ع -), e.g.: (i) H सौदा (*saudā*), سودا (*s-V-d-A*) /sauda/, m. 'trade; goods' ≤ P سودا (*s-V-d-A*) /sauda/ 't.s.'; (ii) H अफसाना (*af'sānā*), افسانه (*A-f-s-A-n-h*) /əfsana/, m. 'chemistry; alchemy' ≤ P افسانه (*A-f-s-A-n-A*) /æfsanæ/ 't.s.'; (iii) H नफ़ा (*naḥfā*), نفع (*n-f-'*) /nəfa/, m. 'profit' ≤ P [≤ A] نفع (*n-f-'*) /næfæ/ 't.s.'.

⁴³⁴ The ending /ija/, noted -इया (-iyā), -ی (-Y-A), which in Hindustani is traditionally counted as another determinant of the category of marked feminine nouns (but rare and occurring exclusively in genuine Hindustani words) is intentionally omitted here in order to avoid confusion, since all but one Perso-Arabic nouns ending with /ija/ are masculine. Cf. e.g.: H दरिया (*dariyā*), دریا (*d-r-Y-A*) /dərja/, m. 'a river; the sea' ≤ P دریا (*d-r-Y-A*) /dærja/ 't.s.'; H कीमیا (*kīmiyā*), کیمیا (*k-Y-m-Y-A*) /kimija/, m. 'chemistry; alchemy' ≤ P [≤ A] کیمیا (*k-Y-m-Y-A*) /kimija/ 't.s.'; and the only exception H दुनिया (*duniyā*), دنیا (*d-n-Y-A*) /donija/, f. 'world, people' ≤ P [≤ A] دنیا (*d-n-Y-A*) /donja/ 't.s.'.

⁴³⁵ For example: H मस्ज़ी (*mar'zī*), مرضی (*m-r-ṣ-ī*) /mərzi/, f. 'pleasure' ≤ P مرضی (*m-r-ṣ-ī*) /mərzi/ 't.s.'. The count of this group is rather low but significantly enhanced by the very productive suffix /i/, forming feminine abstract nouns, e.g.: H अहमकी (*ah'makī*), احمی (*A-h-m-q-Y*) /əhmaki/, f. 'stupidity' ≤ P احمی (*A-h-m-q-Y*) /æhmaki/ 't.s.' ≤ P [≤ A] احمق (*A-h-m-q*) /æhmak/ 'foolish; a fool'.

⁴³⁶ A certain number of borrowed masculine nouns ending with /a/ remain unchanged in the singular oblique case, e.g.: H खुदा (*khudā*), خدا (*x-d-A*) /xuda/, m. 'god' ≤ P خدا (*x-d-A*) /xuda/ 't.s.'; H रहनुमा (*rah'numā*), رہنما (*r-h-n-m-A*) /rəhnuma/, m. 'a guide' ≤ P رہنما (*r-h-n-m-A*) /ræhnæma/ 't.s.'.

		direct	oblique
masculine	singular	/a/: -आ (-ā), - (-A), - (-h)	/e/: -ए (-e), - (-e)
	plural	/e/: -ए (-e), - (-e)	/ō/: -ओं (-om), - (-om)
feminine	singular	/i/: -ई (-ī), - (-Y)	/i/: -ई (-ī), - (-Y)
	plural	/ijā/: -इयों (-iyām), - (-Y-A-n)	/ijō/: -इयों (-iyom), - (-Y-o-n)

Table 8. Scheme of inflexion of marked Perso-Arabic loanwords

Consequently, masculine forms of inflecting adjectives end with /a/ in the singular direct and /e/ in the plural direct, singular oblique and plural oblique cases. Feminine inflecting adjectives always end with /i/, regardless of case or number. Borrowed adjectives ending with consonants or all vowels except /a/ and /i/ and a group of several dozen invariable adjectives ending with /a/ (like H दाना (*dānā*), دان (d-A-n-A) /dana/ 'wise' ≤ P دانا (d-A-n-A) /dana/ 't.s.', मना (*manā*), من (m-n-) /māna/ 'forbidden, prohibited' ≤ P منع (m-n-) /māna/ 'forbidding' or H کھافا (*kḥafā*), خف (x-f-A) /xāfa/ 'angry' ≤ P خفا (x-f-A) /xāfa/ 't.s.'), do not change their form independently of the case or number of the noun they are qualifying.

4.3.4. Nouns which do not end with described suffixes are unmarked. Among the Perso-Arabic loanwords this category is significantly larger in number than the one comprising marked nouns. Most of the unmarked loan nouns end with a consonant but this group also includes feminine suffixless loanwords ending with /a/, masculine suffixless loanwords ending with /i/ (both types discussed in 4.3.5.) and several nouns (feminine and masculine) ending with /u/ (discussed in 4.3.6). They are inflected by attaching adequate endings according to the following paradigm:

		Direct	Oblique
masculine	Singular	-C or /i/: -ی (-ī), -ی (-Y) or /u/: -و (-ū), -و (-V)	Ø
	plural	Ø	/ð/: -ائ (-om), -ائ (-on)
feminine	singular	-C or /a/: -آ (-ā), -آ (-A) or /u/: -و (-ū), -و (-V)	/ē/: -ئ (-em), -ئ (-Y-n)
	plural	Ø	/ð/: -ائ (-om), -ائ (-on)

Table 9. Scheme of inflexion of unmarked Perso-Arabic loanwords

4.3.5. A number of Perso-Arabic loanwords which end with /a/, are assigned the feminine gender in Hindustani. Analogously, a number of nouns which originally end with /i/ in Persian and Arabic, are assigned the masculine gender in Hindustani. This is not at all a rare phenomenon and, at the same time, somewhat confusing. It often results in the erratic classification of these words in terms of gender chosen by various speakers and, consequently, entails contrasting differences in syntax. In the analyzed material there are more than sixty feminine nouns ending with /a/. They are both of Persian and Arabic origin and in their source languages they end with: (i) *alif* (ا -); (ii) the 'silent' *h* (ه -);⁴³⁷ (iii) *alif* + *hamzah* (ء -). Cf. the following examples: (i) H जा (*jā*), ج (j-A) /dʒa/, f. 'place' ≤ P جا (j-A) /dʒa/ 't.s.'; H हवा (*havā*), ه (h-V-A) /həva/, f. 'air; wind' ≤ P [≤ A] هوا (h-V-A) /hæva/ 't.s.'; (ii) H बेवा (*bevā*), ب (b-Y-V-h) /beva/, f. 'a widow' ≤ P بیوه (b-Y-V-h) /bivæ/ 't.s.'; H आशिका (*āśikā*), ع (ʿ-A-š-q-h) /aʃiqa/, f. 'sweetheart,

⁴³⁷ Which is a remnant of *tāʾ marbūʿah* – the Arabic feminine ending, *vide supra* 2.1.5.(ii).

misstress' ≤ P [≤ A] عاشقة (^c-A-š-q-ī) /ʔaʃiqæ/ 't.s.'; (iii) H दवा (*davā*), دواء (*d-V-A*) /dava/, f. 'medicine' ≤ P [≤ A] دواء (*d-V-A*-^c) /dæva/ 't.s.'; H वफ़ा (*vaḥfā*), وفاء (*V-f-A*) /vəfa/, f. 'faithfulness (to a commitment)' ≤ P [≤ A] وفاء (*V-f-A*-^c) /væfa/ 't.s.'. Within the analyzed material, the group of Perso-Arabic masculine nouns ending with /i/ comprises about 80 lexemes, e.g.: H आदमी (*ād'mī*), آدمی (*Ā-d-m-Y*) /admi/, m. 'a man' ≤ P آدمی (*Ā-d-m-Y*) /adæmi/ 't.s.'; H दर्ज़ी (*darzī*), درزی (*d-r-z-Y*) /dərzi/, m. 'a tailor' ≤ P درزی (*d-r-z-Y*) /dærzi/ 't.s.'; H नबी (*nabī*), نبی (*n-b-Y*) /nəbi/, m. 'a prophet' ≤ P [≤ A] نبی (*n-b-Y*) /næbi/ 't.s.'.⁴³⁸

4.3.6. Among the unmarked loan nouns in Hindustani there is also a group of nearly forty lexemes, both feminine and masculine, ending with /u/. All except six⁴³⁹ are genuinely Persian and many of them are very common words, e.g.: H अब्रू (*ab'rū*), ابرو (*A-b-r-V*) /əbru/, f. 'the eyebrow' ≤ P ابرو (*A-b-r-V*) /æbru/ 't.s.'; H आलू (*ālū*), آلو (*A-l-V*) /alu/, m. 'a plum' ≤ P الو (*A-l-V*) /ælu/ 't.s.'; H बू (*bū*), بو (*b-V*) /bu/, f. 'smell, scent' ≤ P بو (*b-V*) /bu(j)/ 't.s.'; H जादू (*jādū*), جادو (*j-A-d-V*) /dʒadu/, m. 'magic, a spell' ≤ P جادو (*j-A-d-V*) /dʒadu/ 't.s.'. A few masculine compound nouns formed by adding /go/ 'speaking' (active stem of the Persian verb /guftæn/ 'to speak') to a noun (e.g.: H किस्सागो (*kissāgo*), قصه گو (*q-ṣ-h*

⁴³⁸ It includes also agent nouns formed by adding the suffix -ची (-cī), چّی (-č-Y) /tʃi/, e.g.: H एलची (*el'cī*), ایلچی (*A-l-č-Y*) /elʃi/, m. 'envoy' ≤ P [≤ T] الجی (*A-l-č-Y*) /ilʃi/ 't.s.', as well as those that emerged in the process of nominalization of adjectives created by the productive adjectival suffix -ی (-ī), ی (-Y) /i/, e.g.: H कैदी (*kaidī*), قیدی (*q-Y-d-Y*) /qaidi/, m. 'a prisoner, a captive' ≤ P قیدی (*q-Y-d-Y*) /qaidi/ 'imprisoned; a prisoner' ≤ P [≤ A] قید (*q-Y-d*) 'confinement, imprisonment, &c.'.

⁴³⁹ Four of them are genuinely Arabic nouns originally ending with ^cain (ع-), like H शुरू (*śurū*), شروع (*š-r-V*-^c) /ʃuru/, m. 'beginning' ≤ P [≤ A] شروع (*š-r-V*-^c) /ʃuru/ 't.s.'; one (/ʃaqu/ 'knife') comes from Turkic and one (/əɾəstu/ 'Aristotle') from Greek (through Arabic).

g-V) /kis:ago/, m. 'a story-teller'), in Hindustani would inflect by case and number like unmarked masculine nouns ending with /u/.

4.3.7. As stated before (4.3.3.), meaning is another factor responsible for assigning gender to loanwords (irrespective of their gender in the source language – if there is such a category in the source language at all). Therefore, in the absence of other decisive determinants, the analogy to synonyms already existing in the borrowing language may be paramount. Thus P [$\leq$ A] /kitab/ 'a book' (masculine in Arabic) takes the feminine gender in Hindustani, similarly to H [$\leq$ S] /pustak/ 't.s.', and P /jadi/ 'marriage', even though recorded in dictionaries as feminine, is commonly assigned the masculine gender in colloquial use⁴⁴⁰ – analogously to its genuine Indic counterpart H [$\leq$ S] /vivah/ 't.s.'.

4.3.8. Grammatical gender may or may not correspond to natural gender.⁴⁴¹ For example, the Persian noun /mærd/ 'a man' becomes masculine in Hindustani, thus remaining in obvious grammatical agreement with its natural gender. However, its second derivative P /mærdangi/ meaning 'manhood, manliness' (= the trait of being manly) is assigned the feminine gender in Hindustani, similarly to other derivatives formed by the suffix /-gi/. This is not the only case when the morphological structure of a loanword determines which gender it will be assigned to. There are a few general and usually applicable rules: (i) Arabic loanwords ending with the suffix -ات (-at), -ة (-t) /æt/ and -يـات (-iyat), -ية (-Y-t) /ijæt/ are always feminine in Hindustani, e.g.: H اکساریات (*ak'sariyat*), اکثریت (*A-k-ṣ-r-Y-t*) /æksarijat/, f. 'greater part, majority' $\leq$ P [$\leq$ A] اکثريت (*A-k-ṣ-r-Y-t*) /æksarij:at/ 't.s.'; (ii) Per-

⁴⁴⁰ Bhatia attributes this phenomenon to a "weakening of grammatical gender in the preference for the natural gender" (Bhatia 1992: 174).

⁴⁴¹ Although, in the case of nouns denoting animate beings, it usually does (cf. Weinreich 1953: 45). See also interesting works by Kira Hall, studying the relationship between gender in linguistics and gender in society in the South Asian context (e.g. Hall 2002).

sian names of places formed by adding the suffix -गाह (-gāh), ग-ह (-g-A-h) /gah/ are always feminine, e.g.: H कबरगाह (*kab'r'gāh*), قبرگاه (*q-b-r-g-A-h*) /qəb(ə)rgah/, f. 'burial-ground' ≤ P قبرگاه (*q-b-r-g-A-h*) /qəbrgah/ 't.s.'; (iii) Persian names of places formed by adding the suffix -(i)ستان (-i)stān), -ستان (-s-t-A-n) /(i)stan/ are always masculine, e.g.: H रेगिस्तान (*registān*), ريگستان (*r-g-s-t-A-n*) /registan/, m. 'desert' ≤ P ريگستان (*r-g-s-t-A-n*) /rægistan/ 't.s.'; (iv) Persian abstract nouns formed by adding the suffix -گی (-ī), -ی (-Y) /i/ or -گی (-gī), گ-ی (-g-Y) /gi/ are always feminine, e.g.: H मरदानगी (*mar'dān'gī*), مردانگی (*m-r-d-A-n-g-Y*) /mærdangi/, f. 'manliness' ≤ P مردانگی (*m-r-d-A-n-g-Y*) /mærdangi/ 't.s.'.

4.3.9. It cannot be denied that gender is the grammatical category which gives users of Hindustani (including native-speakers) a lot of trouble. Even educated individuals, for whom language is a matter of creative work, admit to use of dictionaries in this regard. But since instances of assigning grammatical gender (especially in regard to inanimate nonus⁴⁴²) vary significantly within the Hindustani linguistic continuum, depending on geographical region, language register or social background of the speaker (cf. Bhatia 1992; Hall 2002), one has to acknowledge that the development of normative principles or established rules that would regulate this question is hardly possible.⁴⁴³

⁴⁴² This remark by no means relates exclusively to the loanwords. Problems with erratic grammatical gender pertain to all categories of Hindustani nouns.

⁴⁴³ As evidenced by dictionaries, e.g.: H [≤ P ≤ A] /baqi/ 'remining; reminder, rest' is categorized either as adj. & f. (McGr. 719) or adj. & m. (STCD 95), H [≤ P] /pjaz/ 'onion' as f. but also m. (McGr. 651) and H [≤ P] /xudkəʃt/ 'land cultivated by its owner' as f. & m. (McGr. 239), adj. & m. (Platts 495) or adj. & f. (STCD 293) – to enumerate only a few samples.

4.4. Elements of Arabic and Persian Morphology in Hindustani

4.4.1. In Persian, where gender is not marked, sex is distinguished either (i) lexically or (ii) by the use of an appropriate qualifier. Even though Perso-Arabic loanwords are generally adapted to the Hindustani gender system (as described in the previous chapter), both the strategies are occasionally also found in Hindustani:

(i) examples of lexical distinction between the sexes: H **ज़न** (*zan*), زن (*z-n*) /zən/ 'a woman' ≤ P زن (*z-n*) /zæn/ 't.s.' vs. H **مرد** (*mard*), مرد (*m-r-d*) /mærd/ 'a man' ≤ P مرد (*m-r-d*) /mærd/ 't.s.'; H **والیدا** (*vālidā*), والدہ (*V-A-l-d-h*) /valida/ 'mother' ≤ P [≤ A] والدہ (*V-A-l-d-h*) /validæ/⁴⁴⁴ 't.s.' vs. H **والید** (*vālid*), والد (*V-A-l-d*) /valid/ 'father' ≤ P [≤ A] والد (*V-A-l-d*) /valid/ 't.s.'; H **پری** (*parī*), پری (*p-r-Y*) /pəri/ 'a fairy' ≤ P پری (*p-r-Y*) /pæri/ 't.s.' vs. H **دِی** (*dev*), دی (*d-Y-V*) /de.o/ 'a demon' ≤ P دیو (*d-Y-V*) /dev/ ~ /div/ 't.s.'; H **خاتون** (*khātūn*), خاتون (*x-A-t-V-n*) /xatun/ 'lady' ≤ P [≤ T] خاتون (*x-A-t-V-n*) /xatun/ 't.s.' vs. H **خواجه** (*khvājā*), خواجا (*x-V-A-j-A*) /xʷadʒa/ 'man of distinction' ≤ P خواجه (*x-V-A-j-h*) /xadʒæ/ 't.s.'.

(ii) as the qualifiers used to differentiate between the sexes in Persian, the words **نر** (*n-r*) /nær/ 'male' and **ماده** (*m-A-d-h*) /madæ/ 'female' are employed in relation to animals, e.g.: H **نار-شیر** (*nar-šēr*), نر شیر (*n-r š-Y-r*) /nær fer/ '(he) lion' ≤ P شیر نر (*š-Y-r n-r*) /ʃiře nær/ 't.s.' vs. H **مادہ-شیر** (*mādāšēr*), مادہ شیر (*m-A-d-h š-Y-r*) /mada fer/ 'lioness' ≤ P شیر مادہ (*š-Y-r m-A-d-h*) /ʃiře madæ/ 't.s.'. The words **مرد** (*m-r-d*) /mærd/ 'man' and **زن** (*z-n*) /zan/ 'woman' are used for humans, e.g.: H **مرد-گدا** (*mard-gadā*), مرد گدا (*m-r-d g-d-A*) /mærd gæda/ 'beggar-man' ≤ P مرد گدا (*m-r-d g-d-A*) /mærdæ gæda/

⁴⁴⁴ As a matter of fact, this is the lexicalized Arabic feminine form, characterized by the addition of the suffix *-h* (imperceptible in Class. P and pronounced as /ɛ/ in modern P) to the masculine form.

't.s.' vs. H ज़न-गदा (*zan-gadā*), زن گدا (*z-n g-d-A*) /zən gəda/ 'beggar-woman' ≤ P زن گدا (*z-n g-d-A*) /zæne gæda/ 't.s.'.⁴⁴⁵

4.4.2. Perso-Arabic loanwords occasionally form the vocative by adding the Persian vocative termination ۱- /a/ to the nominative (۱- /ja/, with a euphonic /j/, when the nominative ends with /a/), e.g.: H जाना (*jānā*), جانا (*j-A-n-A*) /džana/ 'o soul!' ≤ P جان (*j-A-n*) /džan/ 'soul &c.'; H दोस्ता (*dostā*), دوستا (*d-V-s-t-A*) /dosta/ 'o friend!' ≤ P دوست (*d-V-s-t*) /dost/ 'friend'; H खुदाया (*khudāyā*), خدايا (*x-d-A-Y-A*) /xudaja/ 'o God!' ≤ P خدا (*x-d-A*) /xuda/ 'God'. Sometimes the vocative is also formed by prefixing a Persian interjection آ (Ā-Y) /aj/ or ۱ (Y-A) /ja/ 'oh!', e.g.: آئی (Ā-y b-č-h) /aj bæf:a/ 'o son! o child!'.

4.4.3. There is no separate morphological system for the comparison of adjectives in Hindustani, but Perso-Arabic adjectives can take Persian, synthetic forms,⁴⁴⁶ made by means of the suffixes ۱- (-t-r) /tær/ for comparative and ۱۱- (-t-r-Y-n) /tærin/ for superlative, added to the positive, e.g.: H बेहतर (*beh'tar*), بهتر (*b-h-t-r*) /behtær/ 'better' and बेहतरीन (*beh'tarīn*), بهترین (*b-h-t-r-Y-n*) /behtærin/ 'the best' ≤ P به (*b-h*) /bi:h/ 'good'; H कमतर (*kam'tar*), کمتر (*k-m-t-r*) /kəmtær/ 'smaller' and کمترین (*kam'tarīn*), کمترین (*k-m-t-r-Y-n*) /kəmtærin/ 'the smallest' ≤ P کم (*k-m*) /kæm/ 'small'. Perso-Arabic adverbs can follow the same paradigm of comparison, e.g.: H पेशतर (*peš'tar*), پیشتر (*p-Y-š-t-r*) /peʃtær/ 'earlier' and पेशतरीन (*peš'tarīn*), پیشترین (*p-Y-š-t-r-Y-n*) /peʃtærin/ 'foremost' ≤ P پیش (*p-Y-š*) /piʃ/ 'before'.

⁴⁴⁵ Note that in Hindustani the *izāfat* is not used. In Persian compounds of this type the word order is reversed, e.g.: P /gæda-mærd/ 'beggar-man', P /gæda-zæn/ 'beggar-woman'.

⁴⁴⁶ Similarly, loanwords from Sanskrit (*tatsama*) have morphologically derived comparative and superlative forms, taking Sanskrit suffixes -तः (-tar) /tær/ for comparative and -तम (-tam) /tæm/ for superlative forms. However, these forms occur mostly in educated speech and texts written in a high style (Kachru 2006: 65).

4.4.4. Generally, plurals of Perso-Arabic loanwords in Hindustani are formed according to Hindustani grammatical rules. But in some cases (chiefly in Hindustani/Urdu and in colloquial Hindustani) borrowed words can form the plural in the same way as in Persian (i) or in Arabic (ii):

(i) Persian plural termination ان - (-*A-n*) /an/ (with the variants -یان (-*Y-A-n*) /jan/, -گان (-*g-A-n*) /gan/ and -وان (-*V-A-n*) /van/ when a noun ends with a vowel or the ‘silent’ *h*) is used with nouns denoting humans, higher animals and some other entities by convention (Perry 2007: 980),⁴⁴⁷ e.g.: H मज़दूरान (*maz’dūrān*), مزدوران (*m-z-d-V-r-A-n*) /məzduran/ ~ /məzdurā/ ‘workers’ ≤ P مزدور (*m-z-d-V-r*) /muzdur/ ‘worker’; H درختان (*darakhtān*), درختان (*d-r-x-t-A-n*) /dərəxtan/ ~ /dərəxtā/ ‘trees’ ≤ P درخت (*d-r-x-t*) /dɪræxt/ ‘tree’; H दानायान (*dānāyān*), دانايان (*d-A-n-A-Y-A-n*) /danajan/ ~ /danajā/ ‘sages’ ≤ P دانا (*d-A-n-A*) /dana/ ‘sage’; H बंदगान (*bandagān*), بندگان (*b-n-d-g-A-n*) /bəndəgan/ ~ /bəndəgā/ ‘servants’ ≤ P بنده (*b-n-d-h*) /bəndæ/ ‘servant’; H वानुवान (*banuvān*), بانوان (*b-A-n-V-A-n*) /banuvān/ ~ /banuvā/ ‘ladies’ ≤ P بانو (*b-A-n-V*) /banu/ ‘lady; Mrs.’. The Persian termination ها - (-*h-A*) /ha/, traditionally forming plurals of nouns denoting things inanimate and abstracts, occurs less frequently in Hindustani,⁴⁴⁸ e.g.: H सालहा (*sāl’hā*), سال ها (*s-A-l h-A*) ~ سالها (*s-A-l-h-A*)⁴⁴⁹ /salha/

⁴⁴⁷ This method of forming plurals was very common in Hindustani until the beginning of 20th c. Beg points out that in its nasalized realization (/ā/), it was the most important plural termination in old Dakkani in both direct and oblique cases (Beg 1988: 127).

⁴⁴⁸ In contrast to Persian, where this termination is universally applicable – especially in modern everyday Persian there is a tendency to add /ha/ also to the nouns denoting people (cf. Perry 2007: 980; Mace 2003: 40-41).

⁴⁴⁹ The termination /ha/ can be written jointly with a noun or can be written detached; however, when a noun ends with ‘silent’ *h* it is always written separately.

'years' ≤ P سال (*s-A-l*) /sal/ 'a year'; H खानाहा (*khānāhā*), خانه (*x-A-n-h h-A*) /xanaha/ 'houses' ≤ P خانہ (*x-A-n-h*) /xanæ/ 'a house'.

(ii) several Arabic loanwords can form their plurals according to the rules of Arabic grammar – either by adding terminations to the singular forms (so-called sound plural) or by altering the pattern of consonants and vowels inside the singular form (so-called broken plural). Sound plurals are formed by adding the termination ين (-*Y-n*) /in/⁴⁵⁰ or ات (-*A-t*) /at/⁴⁵¹ e.g.: H नाज़िरिन (*nāzirīn*), ناظرین (*n-A-ẓ-r-Y-n*) /nazirin/ 'spectators' ≤ A ناظر (*n-A-ẓ-r*) /nazir/ 'a seer, observer'; H हर्फ़ात (*harfāt*), حرفات (*h-r-f-A-t*) /hərfat/ 'letters' ≤ A حرف (*h-r-f*) /hærf/ 'a letter of the alphabet'.⁴⁵² Broken plurals can be created by the lengthening of the noun vowel -/ā/- in nouns ending with -/āt/, e.g.: H हालात (*hālāt*), حالات (*h-A-l-A-t*) /halat/ 'circumstances; facts' ≤ A حالت (*h-A-l-t*) /halæt/ 'particular state, condition'; H आदात (*ādāt*), عادات (*ʿ-A-d-A-t*) /ada/ 'customs' ≤ A عادة (*ʿ-A-d-ī*) /ʾadæt/ 'custom'. It can also be formed irregularly, e.g.: H अख़बार (*akh'bār*), اخبار (*A-x-b-A-r*) /əxbar/ 'news; a newspaper'⁴⁵³ ≤ P [≤ A] خبر (*x-b-r*) /xæbær/ 'news, information'; H औकात (*aukat*), اوقات (*A-V-q-A-t*) /auqat/ 'times, &c.' ≤ P [≤ A] وقت (*V-q-t*) /væqt/ 'time'; H औलिया (*auliyā*), اوليا (*A-V-l-Y-A*) /aulija/

⁴⁵⁰ Actually, in classical Arabic, the regular plural masculine has two cases and is formed by adding the termination -ون (-*V-n*) /un/ to the singular for the nominative and -ين (-*Y-n*) /in/ for the remaining cases. In Persian, as well as in Hindustani, only the latter is used for creating plural.

⁴⁵¹ Which in classical Arabic is the termination forming the regular (sound) plural of feminine nouns.

⁴⁵² A few words of Persian origin can also form the plural through the addition of the termination /at/, e.g.: H देहात (*dehāt*), دیہات (*d-Y-h-A-t*) /dehat/ 'villages' ≤ Class. P دیہ (*d-Y-h*) /dih/ (modern P /deh/) 'a village'; H कागज़ात (*kāgazāt*), کاغذات (*k-A-ġ-ẓ-A-t*) /kacəzat/ 'papers, documents' ≤ P کاغذ (*k-A-ġ-ẓ*) /kayæz/ 'paper'.

⁴⁵³ Note that in Hindustani /əxbar/ is basically used in the singular and means 'a newspaper'. It is hardly ever used to denote the plural meaning of 'news'.

‘saints’ ≤ P [≤ A] ولی (*V-l-Y*) /væli(j)/ ‘a saint, &c.’. For some particular terms there exists the dual, formed by adding the Arabic termination ين (-*Y-n*) /ain/,⁴⁵⁴ but it is of very rare appearance, e.g.: H तरफ़ैन (*tarafain*), طرفین (*t-r-f-Y-n*) /təɾəfain/ ‘both sides (in the lawsuit)’ ≤ P [≤ A] طرف (*t-r-f*) /təɾæf/ ‘side, direction’; H इमामैन (*imāmain*), امامین (*A-m-A-m-Y-n*) /imamain/ ‘both the imams (i.e. Hasan and Hussein)’ ≤ P [≤ A] امام (*A-m-A-m*) /imam/ ‘an imam, a leader (in religious matters)’; H वालिदैन् (*vālidain*), والدین (*V-A-l-d-Y-n*) /valɪdain/ ‘parents’ ≤ P [≤ A] والد (*V-A-l-d*) /valɪd/ ‘father’.⁴⁵⁵

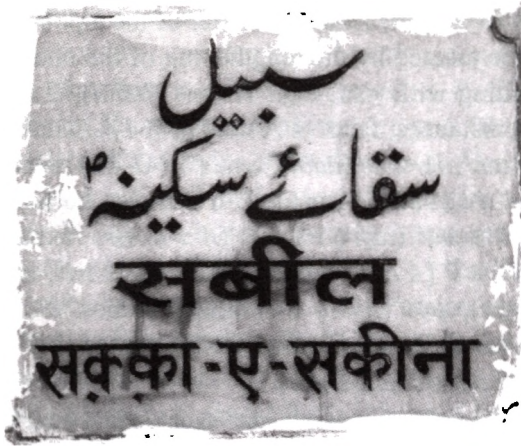


Fig. 12. A marble tablet with a digraphic legend indicating the reservoir of drinking water. The information reads in Nagari: *sabīl saḡkā-e-sabīnā*, and in Perso-Arabic script: *s-b-Y-l s-q-A-Y-e s-k-Y-n-h*. The *izāfat* construction is clearly marked in both the scripts. Šāh Najaf Imāmbāra, Lucknow. (Photo by the author, 2011)

⁴⁵⁴ Which in classical Arabic is used to form not nominative duals.

⁴⁵⁵ One must realize that these kinds of devices used to form the plural are, in practice, employed only by some educated speakers familiar with the inflections borrowed from Arabic and Persian. Generally, such plural forms appeared in Hindustani as ‘ready made’ lexical borrowings and “they have no ‘organic connection’ with the language” (Beg 1988: 120).

4.5. Persian and Arabic Morphosyntactic Constructions in Hindustani

4.5.1. The simplicity of Persian morphology and syntax as well as the high congruence of Persian and Hindustani morphological and syntactic systems reinforced the adaptation of Persian elements in Hindustani also at this level. As a result, several characteristic Persian phrasal constructions can be found in Hindustani. They generally apply to Persian and Arabic loanwords but their appearance and function in Hindustani influenced the development of some features that incorporate also indigenous components. Here the formation of Hindustani complex postpositions modelled on Persian compound prepositions⁴⁵⁶ or the creation of Hindustani phrase-verbs copying the Persian pattern⁴⁵⁷ can be mentioned. A great deal of these new items are direct loan-translations from Persian, often further adapted by the substitution of native lexis (including Sanskrit *tatsama*-s). Hindustani sentence structure has also been influenced by the Persian pattern of forming complex sentences, especially through the use of borrowed Persian conjunctions (cf. 4.2.8.-4.2.9.). They al-

⁴⁵⁶ E.g.: P بطور (*b-t-V-r*) /bætaure/ 'by way (of), like' → (loanblend) H के तौर पर (*ke taur par*), كے طور پر (*k-e t-V-r p-r*) /ke taur par/ → (further adaptation in Hindustani/Hindi – Perso-Arabic loanword is substituted by a Sanskrit *tatsama*) के प्रकार (*ke prakār*) /ke prakar/ 't.s.' (example qtd. after Shackle & Snell 1990: 72).

⁴⁵⁷ E.g.: P شروع کردن (*š-r-V-^c k-r-d-n*) /ʃuru[?] kærdæn/ 'to begin' → (loanblend) H शुरू करना (*šurū kar'nā*), شروع کرنا (*š-r-V-^c k-r-n-A*) /ʃuru kærna/ → (further adaptation in Hindustani/Hindi – the Perso-Arabic loanword is substituted by a Sanskrit *tatsama*) आरंभ करना (*ārambh kar'nā*) /arāmb^h kærna/ 't.s.' (note that /arāmb^h/ is itself a deverbal noun ← S √*ārabh* 'to begin').

lowed expressing clause linkages more sophisticated than clauses in parallel parataxis, traditionally preferred by Hindustani.⁴⁵⁸

4.5.2. Persian copulative phrases formed by the copula *-o/-* are very common in Hindustani. They can consist either of two nominals of the same (or very close) meaning, e.g.: बंदोबस्त (*bandobast*), بندوبست (*b-n-d V b-s-t*) /bəndobəst/ 'organisation, arrangement' (both the words meaning 'tie; tied &c.');

दिलो-जान (*dilo-jān*), دل و جان (*d-l V j-A-n*) /dɪlodʒan/ '(with ppn. /se/) willingly; heartily' (both the words meaning 'heart, soul') or of two elements of opposite signification, e.g.: कमो-बेश (*kamo-beś*), کم و بیش (*k-m V b-Y-ś*) /kəməbeʃ/ 'more or less' (the first word means 'less', the second means 'more');

खरीदो-फ़रोख्त (*kharīdo-farokht*), خرید و فروخت (*x-r-Y-d V f-r-V-x-t*) /xəridofəroxʈ/ 'trade, commerce' (the first word means 'buying', the second means 'selling').

Some compounds of this type have fresh or idiomatic meaning, e.g.: पसोपेश (*pasopeś*), پس و پیش (*p-s-V-p-Y-ś*) /pəsopeʃ/ 'hesitation' (lit. 'behind and before');

आबोहवा (*ābohavā*), آب و هوا (*Ā-b V h-V-A*) /abohəvə/ 'climate; (fig.) atmosphere' (lit. 'water and air').

4.5.3. In the case of compounds borrowed as a whole entity from Persian, the copula *-o/-* may vanish in Hindustani, e.g.: আব-দানা (*āb-dānā*), آب دانہ (*Ā-b d-A-n-h*) /abdana/ 'livelihood; lot, destiny' ≤ P آب و دانه (*Ā-b V d-A-n-h*) /abodanæ/ 'means of subsistence'. The two members of a compound can also be connected by the letter *-a/-*, e.g.: सरापा (*sarāpā*), سراپا (*s-r-A-p-A*) /sərapa/ 'from head to foot: throughout, totally' (cf. P سراپا (*s-r-A-p-A*) /sərapa/ 'totally, from head to foot, cap-à-pié' and P سر و پا (*s-r V p-A*) /səropa/ 'all, the whole') or by the preposition *ba* (*ba*), به (*b-h*) /bə/, e.g.: हरफ़-बहरफ़ (*har'f-bahar'f*), حرف به حرف (*h-r-f b-h h-r-f*) /hərf

⁴⁵⁸ Syntactic functions performed by these conjunctions are being replicated in Hindustani/Hindi by such *tatsama*-s as: यदि (*yadi*) /jədi/ 'if' or यद्यपि (*yadyapi*) /jədjəpi/ 'although', which Shackle & Snell describe as "deliberately introduced loan-conversions from the Persian conjunctions" (ibid.).

bəhərʃ/ 'word for word, literally' ≤ P حرف به حرف (*h-r-f b-h h-r-f*) /hərʃ bəhərʃ/ 't.s.'.

4.5.4. In some phrases borrowed as a whole from Arabic, the Arabic definite article ال (*A-l*) may occur. Its aim is to connect the determined and the determining words. The determining word is placed after the determined one and the joining copula *-al/-* changes in pronunciation according to the neighbouring phones.⁴⁵⁹ This group of borrowings generally comprises two kinds of phrases: Arabic possessive phrases, e.g.: ईदुज्जोहा (*īduzzohā*), عيد الاضحي (^c-*Y-d A-l-A-z-h-ā*) /iduzzoha/⁴⁶⁰ 'Eid-ul-Azha, sacrificial festival'; دارलहुकूमत (*dāruḥukūmat*), دارالحکومت (*d-A-r-A-l-h-k-V-m-t*) /darulhukumat/ '«abode of government»: capital city' and Arabic prepositional phrases, e.g.: बिल्कुल (*bil'kul*), بالکل (*b-A-l-k-l*) /bilkul/ 'entirely, completely; quite'; फ़िलहाल (*fil'hāl*), فی الحال (*f-Y A-l-h-A-l*) /filhal/ 'at the present time; for the time being'.

4.5.5. The *izāfat* ('annexation, suppletion') construction, is a characteristic structure of the Persian noun phrase which, with varying frequency, is also used in Hindustani (i.e. the more Persianized register of the language, the more commonly *izāfat* phrases are employed).⁴⁶¹ It comprises a head noun linked to a

⁴⁵⁹ The vowel /a/ of the article is assimilated to the final vowel of the preceding noun (i.e. the nominative suffix *-u/*, the genitive suffix *-i/* or the accusative suffix *-a/*). The consonant /l/ of the article is assimilated in pronunciation to the following letter if it is any of the so-called 'sun' letters (i.e. dental or alveolar consonants: ت (*t*), ث (*ṭ*), د (*d*), ذ (*ḏ*), ر (*r*), ز (*z*), س (*s*), ش (*ṣ*), ص (*ṣ*), ض (*ḏ*), ط (*ṭ*), ظ (*ḏ*), ل (*l*), ن (*n*)). In Perso-Arabic script, regardless of assimilation in utterance, it is always written *-l/-*, while the Nagari notation closely follows the pronunciation.

⁴⁶⁰ This pronunciation is regarded as colloquial (cf. STCD 441), in contrast to the more Arabicized form /idalad'ha/.

⁴⁶¹ Generally *izāfat* constructions engage only Persian and Arabic words, with one exception: in the place of the governed word there may stand a non Perso-Arabic proper noun. Therefore, one can say e.g.: /fahre

following modifier by the unstressed enclitic *-e/* (*-/je/* after a vowel) and has two principal functions: it indicates a possessive relationship between two nouns or the qualification of a noun by an adjective. The determined word (governing the construction) is placed first – in contrast to all similar formations in Hindustani. *Izāfat* phrases joining two nouns correspond to genuinely Indic possessive phrases formed by the posposition */ka/*, e.g.: हुकूमते पाकिस्तान (*hukūmate pākistān*), حکومت پاکستان (*h-k-V-m-t p-A-k-s-t-A-n*) */hukumate pakistan/* ‘the Government of Pakistan’ vs. पाकिस्तान की हुकूमत (*pākistān kī hukūmat*), پاکستان کی حکومت (*p-A-k-s-t-A-n k-Y h-k-V-m-t*) */pakistan ki hukumat/* ‘Pakistan’s government’;⁴⁶² दर्दे दिल (*darde dil*), درد دل (*d-r-d d-l*) */dārde dīl/* ‘heartache: grief, anguish’ vs. दिल का दर्द (*dil kā dard*), دل کا درد (*d-l k-A d-r-d*) */dīl ka dārd/* ‘t.s.’. *Izāfat* phrases joining a noun and an adjective follow the reverse order of words from native Hindustani phrases of this type where an adjective invariably precedes a noun, e.g.: इश्क़े हकीकी (*iškē hakīkī*), عشق حقیقی (*‘-š-q h-q-Y-q-Y*) */iʃqe hākiki/* ‘divine love, spiritual love’ vs. हकीकी इश्क़ (*hakīkī išk*), حقیقی عشق (*h-q-Y-q-Y ‘-š-q*) */hākiki iʃq/* ‘t.s.’.

4.5.6. Some formations with *izāfat* borrowed from Persian are fixed proper names or have idiomatic meaning, e.g.: वज़ीरे-आज़म (*vazīre-āzam*), وزیر اعظم (*V-z-Y-r A-‘-z-m*) */vəzire.azəm/* ‘Prime Minister’; दीवाने-खास (*dīvāne-khās*), دیوان خاص (*d-Y-V-A-n x-A-š*) */divanexas/* ‘Cabinet, Privy Council; Privy Council Chamber’; तपेदीक (*tapedīk*), تپ دی (*t-p d-q*) */tapedīq/* ‘tuberculosis’; लख्ते-दिल (*lakhte-dil*), لخت دل (*l-x-t d-l*) */læxtedīl/* ‘«a piece of (one’s) liver or heart»; a darling child’.

dīli/ ‘the city of Delhi’. The governing noun, however, must always be Persian or Arabic.

⁴⁶² Note the difference in connotation – the phrase with *izāfat* is a proper name of the office in its official capacity, while the other one is a common noun.

4.5.7. In contemporary spoken Hindustani there is an increasing tendency to omit the *izāfat*, especially when the governing word ends with a single consonant following a short vowel (cf. Schmidt 1999: 247). This tendency is observed not only in pronunciation but also in both scripts – the enclitic *izāfat* vowel after a consonant is hardly ever marked with *kasrah* in Perso-Arabic script and it is also gradually disappearing from Nagari notation, following the oral practice, e.g.: H अहलकार (*ah'l'kār*), اهل کار (*A-h-l-k-A-r*) /əhlkar/ ‘clerk; an agent, officer’ ≤ P اهل کار (*A-h-l-k-A-r*) /əhlkar/ ‘t.s.’; H मीर-मजलिस (*mīr maj'lis*), میر مجلس (*m-Y-r m-j-l-s*) /mir mæɖʒlis/ ‘president, chairman’ ≤ P میر مجلس (*m-Y-r m-j-l-s*) /mir mæɖʒlis/ ‘t.s.’.

4.5.8. It may happen that the order of words is reversed and the governed noun is placed first. In juxtapositions of this type the *izāfat* is not used, e.g.: शेरे बबर (*šere bab'r*), شیر ببر (*š-Y-r b-b-r*) /ʃere bəbr/ ‘lion’ → बबर शेर (*bab'r šer*), ببر شیر (*b-b-r š-Y-r*) /bəbr ʃer/ ‘t.s.’. Very often the result of such a reversal takes the form of a compound – either possessive/attributive (corresponding to Sanskrit *bahuvrihi*), e.g.: कदे आदम (*kade ādam*), قد آدم (*q-d Ā-d-m*) /qəde adəm/ ‘of a man’s height’ → आदमकद (*ādam'kad*), آدم قد (*Ā-d-m-q-d*) /adəmqəd/ ‘t.s.’, or determinative (corresponding to Sanskrit *tatpuruṣa*), e.g.: हालते मौजूदा (*hālate maujūdā*), حالت موجوده (*ḥ-A-l-t m-V-j-V-d-h*) /haləte maudʒuda/ ‘present circumstances’ → मौजूदा-हालत (*maujūdā-hālat*), حالت موجوده (*m-V-j-V-d-h ḥ-A-l-t*) /maudʒuda halət/ ‘t.s.’.⁴⁶³

4.5.9. A specific reminder of the *izāfat* in Hindustani are syntagmatic structures created by reversing the order of a noun and its qualifier in a possessive phrase with the genitive postposition /ka/. The order of elements in such a syntagm imitates the *izāfat* pattern: the determined noun governing the whole phrase is

⁴⁶³ See also very interesting observations regarding the possible influence of the *izāfat* construction on forming compounds used as poetic devices in Ritter 2004: 432-433.

placed before the determining element, which may be a noun with posposition (= genitive) or a possessive pronoun. Structures of this kind, not uncommon in Hindustani, may have a highlighting or emphasising purpose,⁴⁶⁴ e.g.: एक दाना अनाज का (*ek dānā anāj kā*), اک دانا اناج کا (*A-k d-A-n-A A-n-A-j k-A*) /*ek dana ənadʒ ka*/ ‘necessary minimum; lit. grain for bread’, instead of /*ənadʒ ka ek dana*/; बात एक रात की (*bāt ek rāt kī*), بات اک رات کی (*b-A-t A-k r-A-t k-Y*) /*bat ek rat ki*/ ‘A tale of one night’ (a title of a movie) instead of /*ek rat ki bat*/; बाप तेरा (*bāp terā*), باپ تیرا (*b-A-p t-Y-r-A*) /*bap tera*/ ‘your father’ instead of /*tera bap*/; इज़ारबंद की ढीली (*izār'band kī dhīlī*), ازاربند کی دھیلی (*A-z-r-b-n-d k-Y dh-Y-l-Y*) /*izarbənd ki dhīli*/, f. ‘a woman of easy virtue, an unchaste woman; lit. (having) a loose trouser string’.⁴⁶⁵

4.5.10. An analogous phenomenon of inversion may occur in Hindustani complex postpositions which originate from Persian denominative prepositions, which, in Persian, are connected with the following determined noun by the *izāfat*.⁴⁶⁶ Such in-

⁴⁶⁴ Cf. Zograf 1960: 236. Such ‘improper genitive connection’ (German *Uneigentliche Genitivverbindung*) occurs in other languages too. In Arabic it is known as a ‘false *izāfat*’ or ‘«unreal annexation»’ (‘*idāfa ġayr ḥaqīqiyya*’ (إضافة غير حقيقية) because adjectives do not properly occur as the first element of an annexation construction, hence this is not a true possessive relationship’) (Badawi 2004: 111).

⁴⁶⁵ According to Platts (45), this metaphorical syntagm functions as a feminine noun (or, more strictly, a nominalized adjective); cf. McGr. इज़ारबंद का ढीला (*izār'band kā dhīlā*), adj. ‘lecherous’. This example faithfully copies the other *izāfat* pattern, placing an adjectival qualifier after the determined noun.

⁴⁶⁶ Some of these postpositions still have a trace of the *izāfat* in their rhyme, e.g.: H के (की) बजाय (*ke (kī) bajāy*), بجاى (کے) (*k-e (k-Y) b-j-A-Y^c-e*) /*ke (ki) bædʒaj/ ~ /ke (ki) bæja.e/* ‘instead, in place of’ ≤ P بجاى (*b-j-A-Y*) /*bædʒajə/*, prep. ‘in lieu; lit. in place (of)’; H के सिवाय (*ke sivāy*) ~ के सिवा (*ke sivā*), سوا کے (*k-e s-V-A-Y^c-e*) ~ سوا (*k-e s-V-A*) /*ke sivaj/ ~ /ke siva.e/ ~ /ke*

verted postpositional phrases imitate the *izāfat* pattern by placing one part of a complex postposition (a semantic modifier) before the modified noun, e.g.: *बगैर खौफ के* (*bagair khauf ke*), *بغیر خوف کے* (*b-ġ-Y-r x-V-f k-e*) /bəɣair xauf ke/ 'without fear', instead of /xauf ke bəɣair/. Inversion of this type, affecting mainly complex postpositions formed on the basis of borrowed words, may sometimes have influence on indigenous Hindustani phrases,⁴⁶⁷ e.g.: *मार बीमारी के* (*māre bīmārī ke*), *مارے بیماری کے* (*m-A-r-e b-Y-m-A-r-r-Y k-e*) /mare bimari ke/ 'because of the illness', instead of /bimari ke mare/.

4.5.11. As can be seen from the considerations presented above, it is not only vocabulary or phonology that are affected by the extensive language contact, but also the – much less vulnerable to foreign influences – spheres of morphology and syntax.⁴⁶⁸ Therefore, the still generally prevailing opinion, formulated by Grierson nearly a century ago:

Only in the Urdū of Musalmāns do we find the Persian order of words in a sentence, and there has been no other introduction of Persian construction, except that which is actually borrowed as a construction with Persian words. Nor are arabic [*sic!*] words inflected except by purists, but they have to conform to the grammatical system of their host. (Grierson 1920: 74-75)

sivā/ 'except (for), apart (from)' ≤ *سواى* (*s-V-A-Y*) /sivajɛ/, prep. 'besides, except'.

⁴⁶⁷ Zograf enumerates three: /ke mare/ 'because of, on account of', /ke samne/ 'before, in front of' and /ke bina/ 'without' (Zograf 1960: 237-238).

⁴⁶⁸ As Lehmann put it: "Extensive bilingualism may bring about the importation of syntactic patterns more widely than in a few phrases. But when foreign patterns or sounds occur in borrowed words, they are generally modified in keeping with the patterns of the borrowing language (...). An extended period of influence may, however, lead to different results" (Lehmann 1992: 17).

seems to be not quite convincing and the alleged “strong native intuition against the use of foreign constructions” (ibid.) does not find confirmation in the linguistic practice. Trying to support his opinion, Grierson informs that “Hindū writers class a dialect as Urdū, not on the basis of its vocabulary, but on that of the order of words employed” and quotes as an example “a well-known Hindī work, (...) called *Kahānī Ṭhēṭh Hindī-mē*, or «Tales in Pure Hindi»”⁴⁶⁹ which was supposed not to contain a single Persian word, and yet “Hindū writers class it as Urdū on account of the order of the words” (ibid.). Instead, the quoted observation confirms the phenomenon Grierson does not want to admit: Persian (and Arabic) constructions have left a lasting mark on the Hindustani system. This mark is hard to get rid of, even if one tries to purposely ‘purify’ his way of expression by excluding obvious foreign words and constructions. Some phrases which, on the surface, are not recognized as ‘alien’ may still prove to be modelled on rejected not-indigenous patterns. The situation described by Grierson clearly confirms the undeniable vitality and importance of Perso-Arabic components – including those which are ostensibly less susceptible to adaptation in an alien linguistic system.

⁴⁶⁹ Here Grierson mentions the tale known better as *Rānī Ket'kī kī kahānī* by an Urdu poet Inṣā Allāh Xān (d. 1818). McGregor specifies more indications visibly assigning this literary experiment rather to the Persian literary and linguistic tradition: “the use of rhyming elements, strings of verbs, and inversions, and colloquialisms (...)” (McGregor 1984: 64).

5. Semantics and Lexicon

لسان عربی و زبان فارسی که قریب الفهم و کثیر الاستعمال باشد
و روزمره دهلی که میرزایان هند و فصیحان رند در محاوره دارند منظور⁴⁷⁰
(Hātim 1756: 40)

5.1. Observations on the Morphosemantic Taxonomy of Perso-Arabic Borrowings in Hindustani

5.1.1. Haugen's classification (Haugen 1950), based on a comparison of the original pattern with its imitation in the recipient language, remains one of the best-known and most commonly followed taxonomies of loans.⁴⁷¹ It distinguishes between loanwords, where both form and meaning are copied, or assimilated; loanblends, where the meaning is borrowed, but only part of the form;⁴⁷² loanshifts, where the meaning is borrowed, but the form is native;⁴⁷³ and loan translations, or calques, where the

⁴⁷⁰ "[Words from] the Arabic and Persian languages which are intelligible and commonly used [should be used]. And the daily usage of the gentlemen of India and the correct users of language and their idiom is acceptable". Transl. and qtd. after Rahman 2010: 91.

⁴⁷¹ Cf. e.g. Hock 1991; Crystal 2008; Haspelmath 2008; Haspelmath & Tadmor 2009; Matras 2009; Durkin 2009.

⁴⁷² Loanblends are sometimes called hybrids (e.g. Myers-Scotton 2006: 219), which is generally a broader term, referred to in etymology and historical linguistics (Kuczkiewicz-Fraś 2003: 17ff.).

⁴⁷³ Which Hock (1991: 398) labels as "the polar opposite of lexical adoption [when] a foreign concept is borrowed without its corresponding linguistic form and without the introduction of a new word into the borrowing language". A typical loan shift is the Hindustani word /māgni/ 'en-

morphemes in the borrowed word are translated item by item. Since it is the first of these categories, viz. loanwords, that remain the basic object of inquiry of this study, and the last one, viz. calques, which are a specific form of adaptation of the former group – they will both be given main attention in the course of the following considerations.

5.1.2. Loanwords, i.e. lexical items which show the borrowing of a word form together with its associated meaning (or a component of its meaning), constitute the basic stock of linguistic material analyzed in the former chapters. Along with the already described changes and adaptations in the realms of orthography, phonology and morphosyntax, loanwords can be (and, actually, they very often are) subject to semantic modification. In most cases this modification boils down to the reduction or narrowing of a loanword's meaning, since, as a rule, polysemantic words are adopted only in one or two of their denotations. It happens quite often that a lexical item is borrowed with one of its less significant and/or less frequent meanings and thus the meaning of the borrowed word in the target language becomes distant or different from the basic meaning of the same word in the source language. Speakers of Persian would probably be greatly astonished to hear that in Hindustani /*rasta*/ means 'a road, a street', since they never use this word in such a sense (although it is given in Persian dictionaries as one of the further meanings, cf.: "راسته *rāsta*, Dexterous, right-handed; upright, just, true; a market-place; **a level road**" (Stein. 563); "(1) Noun 1. A row; a series of shops; a section or part of the market-place. Ex. راسته کفشدوز ها *the shoemakers' row*. 2. *Zool. & Bot.* A group. 3. *R[arely]. A (level) road*" (Hayyim 1911). However, in Hindustani, /*rasta*/ meaning 'a road' is a commonly used noun.

gagement, matrimonial suit', coined to name a custom not known in India before the arrival of Muslims and to replace the Persian term /*xastgari*/.

5.1.3. Change of the basic meaning in a loanword in relation to the meaning of the source word occurs quite commonly in the analyzed material, here are a few more examples: P [≥ A] /mausim/ 'time, season' ≥ H /maʊsəm/ 'weather'; P [≥ A] /hæva/ 'air, weather' ≥ H /həvə/ 'wind'; P /xərab/ 'ruined, destroyed' ≥ H /xərab/ 'bad', P /sazɪʃ/ 'accord' ≥ H /sazɪʃ/ 'conspiracy'. Apart from restriction in the number of denotations, the meaning of a loanword can also shift from more general to a more particular one. This is the case of many Arabic loanwords which are associated with Islam by speakers of Hindustani. Therefore, their meaning is restricted only to the narrow religious dimension⁴⁷⁴ The opposite semantic change, viz. extension or widening of meaning, can also be observed, e.g.: in Hindustani P /rah/ 'way, road' gets the additional, metaphorical meaning of 'journey, progress' (although, at the same time, it loses several secondary meanings compared to its Persian source word); P /ʃufæ/ 'an ingot of gold or silver' gets generalized in Hindustani to H /ʃofa/ 'mixing, dilluting', etc.

5.1.4. A separate group of words which can be distinguished in the analyzed material are doublets, or pairs of (different) words which have a common origin and display similarities of form and meaning. In the case of Perso-Arabic loanwords in Hindustani, this category can be further differentiated into two subgroups: (i) doublets consisting of a borrowed word and a native word, and (ii) doublets existing within the boundaries of the borrowed lexicon.

⁴⁷⁴ As Qutbuddin observes, "Many Arabic words which have more general meanings in Middle Eastern Arabic take on a religious connotation in India. *Ṣaḥīfa*, which can mean several things in Arabic, including a leaf in a notebook, a page, a newspaper, or a prayer book, signifies here the last sense only. *Ziyārat*, which means visit, connotes here a visit to the shrine of a saint. *Majlis*, which means sitting or assembly, indicates here a religious assembly. (...) In many cases, Arabic words are automatically considered sacred by virtue of their being Arabic" (Qutbuddin 2007: 328-329).

Given the close genetic relationship between Persian and Hindustani, the occurrence of doublets of the former type, originating in the common Indo-Iranian source, is easy to predict and illustrate, e.g.: H [≥ P] /əndər/ 'within, inside' vs. H [≥ S] /əntər/ 'interior, midst'; H [≥ P] /ənguʃt/ 'the finger' vs. H [≥ S] /ənguʃtʰ/ 'the thumb'; H [≥ P] /ʃəhar/ 'four' vs. H /ʃar/ 'four', etc.⁴⁷⁵ Doublets of the second type comprise, first of all, originally Arabic etyma with the feminine ending (in Arabic written as *tāʾ marbūʿah*) which, in Persian, can be lexicalized more or less definitively as either *-æt/* or *-æ/* (cf. Perry 1995). Hindustani usually adopted only one of the two forms but there are also cases where both the variants are in use, e.g.: H [≥ A] /əkidət/ ~ /əkida/ f. 'faith, belief'; H [≥ A] /alət/ ~ /ala/ 'implement'; H [≥ A] /ɪfarət/ ~ /ɪfara/ m. 'sign'; H [≥ A] /ɪzafət/ f. 'increment' ~ /ɪzafa/, m. 'increase'.⁴⁷⁶ Other cases of doublets include pairs of words that can differ in: (i) the length of the vowel, e.g.: H [≥ P] /rəh/ ~ /rah/ 'a road, way'; H [≥ P] /əʃar/ ~ /aʃar/ 'pickle'; H [≥ P] /surax/ ~ /surax/ 'hole, opening'; (ii) the quality of the vowel, e.g.: H [≥ P] /rozina/ ~ /rozana/ 'daily pay or allowance'; (iii) phonological adaptation, e.g.: H [≥ A] /nemət/ ~ /nijamət/ 'affluence, wealth'; (iv) morphological adaptation (adding of an inflexional termination), e.g.: H [≥ A] /qərz/ ~ /qərza/ 'a debt, a loan'; (v) grammatical or lexical category in the source language, e.g.: H [≥ A] /hal/ ~ /halət/ 'state, condition' (masculine vs. feminine); H [≥ A] /halət/ ~ /halat/ 'situation' (singular vs. plural); H [≥ A] /nuqs/ ~ /nuqsan/ 'loss' (two differently created gerunds); (vi) general form of the words, e.g.: H [≥ A] /izad/ ~ /zjada/ 'increase, addition'; H [≥ P] /jəxni/ ~ /əxni/ 'a meat stew'.

⁴⁷⁵ Such doublets occur among basic words, *vide infra* 5.2.9.(v) and also include other morphological units, for example derivational affixes, like P *-kar/* ~ *-gar/* and H *-kar/*, creating nouns of agency, or P *-stan/* and H *-stʰan/*, productive suffixes forming names of places.

⁴⁷⁶ As can be seen, in Hindustani they can be assigned the non-feminine gender too. On their graphical notation cf. *supra* 2.1.5. and 2.2.17.

The interesting phenomenon of Perso-Arabic doublets in Hindustani certainly deserves further scrutiny and separate comprehensive study, but that would exceed the intended scope of this book.

5.1.5. A specific method of dealing with borrowed lexical items is calquing, a process which – contrary to the simple transfer of alien items to the linguistic tissue of the recipient language – presupposes a certain familiarity with the donor language and its grammatical structure.⁴⁷⁷ Many examples of calquing, or producing loan translations can be found among conjunct verbs, which are prevalent in Hindustani. They reproduce the metaphorical meaning of Persian compound verbs by means of novel combinations of native elements which correspond to the meaning and structure of the component parts forming the foreign, morphologically complex expression. Verbal calques often evolve from verbal loanblends which combine a (foreign in this case) nominal part and a native operational verb, according to the following four-step pattern: P خواب دیدن (*x-V-A-b d-Y-d-n*) /xab didæn/ ‘to dream’ → H (a loanblend where P verb is replaced by matching H verb of the same meaning) خواب देखना (*khvāb dekʰnā*), خواب دکھا (*x-V-A-b d-kʰ-n-A*) /xʷab dekʰna/ ‘t.s.’ → (a loan translation where a *tadbʰava* word is put in place of P nominal part) सपना देखना (*sap'nā dekʰnā*), सपना دکھا (*s-p-n-A d-kʰ-n-A*) /səpna dekʰna/ ‘t.s.’ → (further nativization of a calque by substituting the native word with a Sanskrit *tatsama*) स्वप्न देखना (*svapn dekʰnā*) /svəpn dekʰna/ ‘t.s.’ (the latter form is used only in the highbrow register of Hindustani/Hindi). However, not every case of translating a foreign complex verb results in the creation of a full calque – sometimes the process does not go beyond the stage of forming a loanblend, e.g.: P شرط بستن (*ṣ-r-! b-s-t-n*) /ʃært bæstæn/ ‘to stipulate, to bet’ → H शर्त बाँधना (*śart bāndʰnā*),

⁴⁷⁷ Actually, for calquing it is a crucial condition, since this process affects only morphologically complex units. Minimum knowledge of the donor language and its structure is necessary to recognize the morphological complexity of a given foreign item (cf. Hock 1991: 400).

شرط باندھنا (*ṣ-r-ṭ b-A-n-dʰ-n-A*) /ʃərt bādʰna/ 't.s.'; P قسم خوردن (*q-s-m x-V-r-d-n*) /qəsəm xurdæn/ 'to swear' → H कसम खाना (*kasam kʰānā*), قسم کھانا (*q-s-m kʰ-A-n-A*) /qəsəm kʰana/ 't.s.' → सौगंद खाना (*saugand kʰānā*), سونگند کھانا (*x-V-A-b d-d-n*) /saugand kʰana/⁴⁷⁸ 't.s.'.

5.1.6. Loan translations are not necessarily limited to structures in which the component parts of a morphologically complex unit are independent lexical items. Calques can also involve compounds and derivatives, e.g.: P بازوبند (*b-A-z-V-b-n-d*) /bazubænd/ 'armlet' → H भुजबंद (*bʰuj'band*), بھجند (*bʰ-j-b-n-d*) /bʰujbænd/ 't.s.' → H भुजबंध (*bʰuj'bandʰ*), بھجندہ (*bʰ-j-b-n-dʰ*) /bʰujbændʰ/ 't.s.'; P بکتر پوش (*b-k-t-r-p-V-ṣ*) /bæktærpoʃ/ 'wearing an armour' → H कवचधारी (*kavac'dʰārī*), کوچهاری (*k-V-č-dʰ-A-r-Y*) /kəvəʃdʰari/ 't.s.'. In recent times calquing has often been employed by proponents of linguistic purism as a device for replacing foreign Perso-Arabic, and therefore unwelcome, expressions with "politically correct" loan translations usually based on loanwords from Sanskrit, e.g.: P بی دم (*b-Y d-m*) /bidəm/ 'breathless, lifeless' → H बेदम (*bedam*), بی دم (*b-e-d-m*) /bedəm/ 't.s.' → H/Hindi प्राणहीन (*prāṇ'hīn*) 'breathless'; P کاریگر (*k-A-r-Y-g-r*) /karigær/ 'artisan' → H कारीगर (*kārīgar*), کارگر (*k-A-r-Y-g-r*) /karigær/ 't.s.' → H/Hindi शिल्पकार (*śilp'kār*) 'craftsman'. The results of such manipulation are difficult to assess without researching to what extent, if at all, these artificially created linguistic novelties are being accepted and used among the speakers. Although, over time, as Myers-Scotton put it, "ideas as to what is «politically correct» may change the popularity of a loan translation" (Myers-Scotton 2006: 218).

⁴⁷⁸ Note that /saugand/ 'an oath' is also a loanword from Persian. However, puritanically oriented speakers of highbrow Hindustani/Hindi commonly nativize this noun (both in pronunciation and in writing!) by aspirating the final consonant and, as a result, it becomes phonetically close to the Sanskrit *tatsama* /saugandʰ/ 'fragrance; perfume'. The verb thus created: /saugandʰ kʰana/, which is used in the sense 'to swear', would translate literary (and nonsensically) to 'to eat perfume' instead of 'to eat an oath'.

5.2. Concluding Remarks: Typological Classifications of the Perso-Arabic Lexicon in Hindustani

5.2.1. Lexical borrowing is always the most apparent occurrence of contact between languages. As a result of the long-lasting liaison between Persian and Hindustani, that took place at various levels of life and culture and was strongly marked by such favourable factors as frequent bilingualism of the speakers, the prestigious position for centuries maintained by Persian in South Asia and the natural permissiveness for incorporating foreign linguistic elements evinced by Hindustani – the lexicon of the latter contains not hundreds but thousands of Persian, Arabic and, to a lesser extent, also Turkic lexemes.⁴⁷⁹ This large and varied stock may be classified according to various criteria: the criterion of origin, the criterion of pathway (i.e. how a loanword gets into the language), the functional criterion (based on the examining goals and motivation for borrowing) and, finally, the degree of assimilation criterion. In the following paragraphs, Perso-Arabic loanwords in Hindustani will be analyzed respectively in relation to these typological benchmarks.

5.2.2. From the standpoint of etymology, the analyzed lexical material is not homogenous, consisting overwhelmingly of lexemes originating from two main sources of borrowing, namely Persian and Arabic, supplemented by a small number of genuinely Turkic⁴⁸⁰ and a few originally Greek words. The group

⁴⁷⁹ The dictionary of Perso-Arabic loanwords (Kuczkiewicz-Fraś 2008), on which this study is based, contains approximately 5500 lemmata.

⁴⁸⁰ The number of Turkic loanwords is somewhat problematic. Otto Spies listed 135 words in his article (Spies 1955), labelling them as Turkic loanwords in Hindustani ("Türkisches Sprachgut im Hindūstānī"). However, even though his collection undeniably includes words of Turkic origin, one can have serious doubts as to their actual presence in Hindustani. Spies based his choice mostly on Platts's dictionary, known for including a

of Arabic loanwords comprises slightly more than one-third of the analyzed vocabulary (ca. 1800 lexemes), not counting hybrid derivatives and compounds (i.e. created by means of Persian devices), which are regarded as Persian. This data corresponds quite well with the estimated percentage of Arabic vocabulary in Persian.⁴⁸¹ Arabic origin of loanwords is also relatively easy to establish (orthography being helpful in many a case). However, determining the strict etymology of loanwords which are traditionally labelled as “Persian” may sometimes be quite problematic, especially when dealing with the oldest part of the lexicon which already in the early New Persian stage (i.e. around 10th-11th c.) developed and spread geographically, mixing Iranian and non-Iranian elements.⁴⁸² From the 11th c. onwards the proportion of Arabic words in Persian increased (Jazayery 1970). Additionally, the close relations between Turkic languages and Persian,⁴⁸³ which probably also started at that time, and, most of all, the interpenetration of their vocabulary, are responsible for the sometimes dubious and erratic etymological qualification of many a word, including some loanwords in Hindustani. Spies

lot of lexemes used exclusively in the highly Persianized/Arabicized register of Hindustani. Many Turkic words given on his list are not only absent from dictionaries of modern Hindi and Urdu, but at the same time, would remain incomprehensible for the average speaker of Hindustani.

⁴⁸¹ A dictionary-based analysis shows that there are approximately 8000 Arabic loanwords in current use in Persian, which is about forty per cent of an everyday literary vocabulary (consisting of ca. 20 000 words, not counting derivatives and compounds) (Perry 2002: 229).

⁴⁸² Bo Utas enumerates the participants of this mixed linguistic milieu: “Soghdian and varieties of Persian, Parthian and other north-west dialects, Khwarezmian, Bactrian and the Saka languages, as well as non-Iranian languages, either Indo-European, like Tokharian, or non-Indo-European, especially Turkic” (Utas 2006: 246).

⁴⁸³ Utas calls them “symbiotic” (ibid.), similarly Golden (2002: 18): “...symbiosis of Turk and Iranian reaches deep into the past”.

draws attention to this problem, observing that “a number of words which are labelled in dictionaries as Turkic, are actually not Turkic, but Persian and Arabic; (...) there are, however, a number of Turkic words not recognized as Turkic; they have not been incorporated directly into Urdu, but through Persian.”⁴⁸⁴ Examples of the latter can be found on his list, e.g.: H ≤ T /ata/ ‘father’ recognized in Platts as ≤ P آتا (*Ā-t-A*) /ata/ ‘father’; H ≤ T /apa/ ‘elder sister’ explained in Platts as ≤ H आपा (*āpā*) ‘t.s.’ [≤ S आत्मा (*ātmā*) ‘soul; self; the mind’] (probably after Shak. 16) or H ≤ T /sauyat/ ‘a rich present’, also recognized as ≤ P. A flagship word, commonly used in Hindustani (and in many other languages, including English, to which it was transferred) and equally commonly but incorrectly regarded to be of Turkic⁴⁸⁵ origin, is H कुली (*kulī*), قلی (*q-l-Y*) /quli/ ‘a porter; a labourer; (arch.) a slave’, unanimously classified as ≤ T by most Hindustani lexicographic works,⁴⁸⁶ e.g.: McGr. 209, STCD 468, Platts 794 or MHK I 561.

5.2.3. It may not be easy to properly identify the source language of loanwords, but it is much more difficult to definitively indicate the way through which they got transmitted into Hindustani. Generally, two patterns can be distinguished: direct loan, when the elements of the donor language are borrowed directly

⁴⁸⁴ It should be remembered that the very term “Turkic” (similarly to “Persian”) is ambiguous in this context and would require further clarification as to indicate which language(s) or dialect(s) of the Turkic family was actually the source of the loans (cf. Kuczkiewicz-Fraś 2001).

⁴⁸⁵ “Dabei stellt sich heraus, daß eine Reihe als türkisch in den Wörterbüchern bezeichneter Wörter nicht türkisch, sondern persisch und arabisch sind; (...) Andererseits sind aber eine Anzahl türkischer Wörter nicht als türkisch erkannt; sie sind nicht unmittelbar ins Urdu, sondern über das Persische aufgenommen worden” (Spies 1955: 323).

⁴⁸⁶ The etymology of this word is uncertain but it is most probably genuinely Indian, cf. Yule & Burnell 1996: 249-251.

from the donor language and indirect loan, when the elements of the donor language are carried to the borrowing language via another language, in which they are often subject to some adaptation before – in the already changed form – they find their way to the target language. In the case of Perso-Arabic loanwords the situation is more convoluted because there are three donor languages which, as a result of mutual relations and dependencies, started to share a good part of the vocabulary derived from all three of them (Arabic being the least influenced by either Persian or Turkic). Additionally, each of the three could theoretically function as the direct donor, since throughout the centuries, they were all represented and used in South Asia (this was described in detail in Chapter 1.2.). Moreover, Persian was the intermediary for Arabic and Turkic loanwords that had been previously adapted, at various levels, to its system.⁴⁸⁷ Its enormous prestige in India as well as the genetic relatedness to Hindustani and, consequently, the similarity of their linguistic structures, made Persian the main source of borrowings – both indigenous and those taken over beforehand as a result of experiencing the massive influx from Arabic. Therefore, Arabic components are believed to have reached India not directly, but through the medium of Persian and common words of Arabic origin are customarily treated as belonging to the shared Persian component of Hindustani. This pattern is, in many cases, well documented by linguistic evidence which indicates changes that occurred in the intermediate (Persian) stage, as also shown in previous chapters of this study. However, it should not be accepted uncritically and without exception because it may be very difficult to prove the actual mediation in borrowing if there has been no change in

⁴⁸⁷ Turkic, which was an important medium of communication among some of the newcomers, especially in the times of the Delhi Sultanate (13th-15th c.) could, to some extent, play a similar role for vocabulary derived from Arabic and Persian.

word form or meaning in the intermediate language.⁴⁸⁸ In this situation we cannot exclude the chance of borrowing partly via an intermediary and partly directly from the source language (Durkin 2009: 140), and the historical facts may speak in favour of such a possibility. Since the appearance of Islam in South Asia, Arabic has been used by its followers as the language of liturgy in its various aspects⁴⁸⁹, taught in *maktab*-s and *madrasa*-s, employed for nomenclature, inscriptions⁴⁹⁰ and religious writings (Qutbuddin 2007; Hamidullah 1954), even though that use never reached a large scale.⁴⁹¹ But for the time being it has heavily influenced the Hindustani/Urdu script (e.g. by retaining, not so rarely, elements which in Persian disappeared long ago, like

⁴⁸⁸ E.g.: A نور (*n-V-r*) /nur/ 'light' [≥ P نور (*n-V-r*) /nur/ 't.s.'] ≥ H नूर (*nūr*), نور (*n-V-r*) /nur/ 't.s.']; A يتيم (*Y-t-Y-m*) /jætīm/ 'orphan' [≥ P يتيم (*Y-t-Y-m*) /jætīm/ 't.s.'] ≥ H यतीम (*yaṭīm*), يتيم (*Y-t-Y-m*) /jætīm/ 'orphan'; A تريخ (*t-r-Y-x*) /tærix/ 'date' [≥ P تريخ (*t-r-Y-x*) /tærix/ 't.s.'] ≥ H तरिख (*tarīkh*), تريخ (*t-r-Y-x*) /tærix/ 't.s.'.

⁴⁸⁹ Like recitation of the *Qurʾān*, mandatory ritual prayer (*ṣalāh*), litanies (*tasbiḥ*) and prose prayers (*duʿāʾ*) as well as chanting Arabic religious poetry (for particulars see Qutbuddin 2007: 318-320).

⁴⁹⁰ Including the digraphical convention of distinguishing original Arabic phrases from Hindustani texts written in *nastaʿlīq* by retaining their original form in *nasx*.

⁴⁹¹ One should bear in mind that during the Muslim reign in India, Persian gained the status of a sacred language for Muslims. "The prayers were said in Arabic but books on religion were mostly composed in Persian. Moreover, both Sufis and Mullas delivered their lessons and sermons in Persian with the result that many Persian terms, such as *Khudā* (God), *Paighambar* (Prophet), *namāz* (prayer), *ruza* (fasting), *bihisht* (paradise), and *duzakh* (hell), replaced even the religious terminology of Arabic for the main pillars of Islam and are still in use in South Asia, Central Asia and Turkey" (Syed 2012: 291). Cf. similar fact reported by Jazayery concerning the beginnings of Islam in Persia: "Even religious rituals, including the daily prayers, were at first conducted in Persian, at least in some regions" (Jazayery's 1970: 115).

tāʾ marbūṭah or *ʾalif maqṣūrah*) Moreover, several features of Arabic grammar (*vide supra* 4.4.-4.5.) still appear in the formal style of Hindustani, which, in addition, draws much of its learned and abstract vocabulary from Arabic (Shackle & Snell 1990: 46). All these points prove the inaccuracy of statement, prevalent since Grierson's times, that Arabic elements in Hindustani resulted exclusively (or with a few insignificant exceptions) from indirect loan.⁴⁹²

5.2.4. A small part of the analyzed Perso-Arabic vocabulary can be classified as internationalisms or international words, commonly defined as expressions "attested in a number of unrelated languages or language families, sharing a similar orthographic or phonetic shape and a partial or identical semantic field" (Wexler 2009: 77). Regardless of the theoretical divagations concerning the phenomenon as such,⁴⁹³ among the collected loanwords one can indicate without hesitation many words, mainly Arabic in origin,⁴⁹⁴ which found their way to a multitude of languages across the world – and among them also to Hindustani. Many such internationalisms are words culturally and semantically

⁴⁹² Cf.: "The influence of the religion of Islām has opened another door for the entry of Arabic, and a few words have been imported on the west coast from Arab traders; but in the main the Arabic element in all the Indo-Aryan Vernaculars came in with Persian, and as a part of that language." (Grierson 1920: 74)

⁴⁹³ The number and kinds of required "unrelated languages" still remain the subject of scholarly debate, moreover some internationalisms are more 'global' than others, which may be restricted to certain geographic regions.

⁴⁹⁴ Not without reason, the Arabic language is mentioned among the most influential languages in the history of cultural (and linguistic) interference, cf. Sapir's opinion: "There are just five languages that have had an over-whelming significance as carriers of culture. They are classical Chinese, Sanskrit, Arabic, Greek and Latin" (Sapir 1921: 207). On Arabic as an international language see Cannon 1994.

connected with Islam and traditions of the Arab people, like: /muslim/ 'a Muslim', /məsdʒid/ 'a mosque', /ərəb/ 'Arab', /ʃaɪx/ 'sheikh' or /gəzəl/ 'a ghazal', etc. But there are also words denoting common objects and phenomena, i.a.: /əksir/ ~ /ɪksir/ 'elixir', /jasmin/ 'jasmin', /kimija/ 'chemistry', /ladʒvərd/ 'lazarite', /nəzir/ 'nadir', /səndəl/ 'sandal', /tarif/ 'tariff', /xaki/ 'khaki' or /zafran/ 'saffron'. Interestingly, some Arabic internationalisms, like 'alcohol' or 'algebra' have been re-introduced to Hindustani only lately, through the medium of English, which additionally confirms their international character.⁴⁹⁵

5.2.5. Another possible classification connected with the manner of linguistic penetration of Perso-Arabic loanwords is the division into oral vs. literary borrowings. The former were caused by the necessity of quotidian communication and originated in close mutual relations between local people and the newcomers who brought in a different culture and a different language. Oral loanwords include, in all probability, names of items previously unknown in India (for example names of tailored clothes, e.g.: /qəmis/ 'shirt, tunic', /padʒama/ or /ʃalvar/ 'trousers'; food, e.g.: /qorma/ 'a kind of stew'; /birjani/ 'name of dish made of meat and rice', /səbz/ 'vegetables', /əndʒir/ 'fig'; arms, e.g.: /top/ 'a gun, cannon', /bənduq/ 'rifle', /bəktər/ 'armour'), names of professions and occupations (often also new, e.g.: /dərzi/ 'tailor', /qəsa.i/ ~ /qəs.ab/ 'butcher', /dʊkandar/ 'shopkeeper', /ahəŋgər/ 'blacksmith'), terms connected with trade (e.g.: /xərid/ 'buying', /fərox/ 'selling'; /bazar/ 'a market-place', /saʊdagər/ 'a merchant', /tərazu/

⁴⁹⁵ Nagari script depicts Anglicization of these words, cf. ऐल्कोहल (*ailkohāl*) /ailkohal/ 'alcohol' (Awasthi & Awasthi 1981: 26; Sharma 2006: 19) vs. ائكل (*A-l-k-h-l*) /əlkohul/ 't.s.' (Haqee 1995: 34); ऐल्जेब्रा (*ailjebrā*) /aildʒebra/ 'algebra' (Awasthi & Awasthi 1981: 27) vs. الجبر (*A-l-j-b-r-A*) /əldʒibra/ 't.s.' (Haqee 1995: 34). They are not quoted by former dictionaries, only Platts records جبر و مقابله (*j-b-r V m-q-A-b-l-h*) /dʒəbromuqabələ/ 'algebra' (Platts 375).

'pair of scales'), numerous crafts (/nəq:ʌʃi/ 'painting', /zərdozi/ 'gold embroidery'; /qəla.igəri/ 'a business of tinning'), and with the army (e.g.: /faʊdʒ/ 'army', /dʒəŋg/ 'fight, battle', /tirəndəz/ 'an archer'). Borrowings of this kind were the first effect of the inter-linguistic and intercultural clash of the Muslim and Indian communities manifesting itself in a vernacular – they must have begun to appear in Hindustani probably as early as the 11th-12th c. Loan-words of the second kind appeared somewhat later and were caused by the activity of the growing number of bilingual intellectuals in the premodern and early modern period.⁴⁹⁶ Of course, the category 'intellectuals' comprises, apart from the literary intelligentsia, also scholars, clergy or government officials who operated at the courts of Muslim rulers (i.e. 14th-18th c.). Thus, among borrowings of literary origin we find a large number of terms (often originating in Arabic) referring to different fields of science (e.g.: /ɪlm/ 'science', /hɪkmət/ 'wisdom, knowledge', /tɪb/ 'medicine'⁴⁹⁷, /hɪsəb/ 'arithmetic', /məqsum/ *arith.* 'divisor', /hasil/ *math.* 'total'), literature (e.g.: /ədəb/ 'literature, belles-lettres', /mərsija/ 'elegy', /məsnəvi/ 'a narrative poem', /dastan/ 'a tale, story'), philosophy and religion (e.g.: /din/ 'religion', /axləq/ 'morality, ethics', /iman/ 'belief (in God)', /kafir/ 'infidel', /qəzija/ 'syllogism'), terms concerning administration and law (e.g.: /ədalət/ 'court of law', /a.in/ 'regulation, statut', /ɪnsaf/ 'justice', /xɪrəɖʒ/ ~ /xərəɖʒ/ 'tax, tribute', /məhsul/ 'revenue', /təhsil/ 'district'), words referring to the ruler and his court (e.g.: /badʃah/ 'monarch', /divan/ 'a royal court', /məh(ə)l/ 'palace', /vəzir/ 'vizier, minister', /xan/ 'lord, prince', /əʃraf/ 'aristocracy'). Their penetration of Hindustani was particularly widespread during the full bloom of the Delhi and Lakhnau schools of poetry in the 18th and 19th c. Many loanwords belonging to this group are now used mainly in the formalized

⁴⁹⁶ It must be remembered that the first India-born poets who composed their verse in Persian appeared as early as the 11th c. (*vide supra* 1.3.4.).

⁴⁹⁷ On Perso-Arabic medical terminology see Kuczkiewicz-Fraś 2003a.

highbrow register of Hindustani/Urdu, forming the greater part of its lexical basis, but dozens can also be found in literary and vernacular Hindustani.

5.2.6. The most universal and most common distinction of motivations for lexical borrowing is traditionally based on the binary division between need and prestige⁴⁹⁸ or, in other words, necessary and unnecessary (luxury) borrowings.⁴⁹⁹ The first category is generally accepted to be the basic cause of lexical innovation. Among Perso-Arabic loanwords in Hindustani there are numerous cases of simple lexical importations of names for new objects and terms which followed culturally motivated borrowings. The arrival of Muslims in India caused cultural, economic and intellectual revolution pertaining to all walks of life:

With the Muslim rule in India came a new religion, a new polity, a new government, a new social structure, a new set of etiquettes, a new way of life, a new dress system, new food habits, new furniture, a new philosophy, and a new literature. (Abidi & Gargesh 2008: 115)

Indians needed names for all these novelties and since using ready-made designations is easier and more economical than describing things afresh (Weinreich 1953: 57) – they adopted alien things and ideas along with their alien designations. Information provided by the words and forms obtained in this way supple-

⁴⁹⁸ The historical and cultural aspects of both these reasons in the Indian context were already discussed in the introductory Chapter, in paragraphs 1.2.5. (need) and 1.2.9. (prestige).

⁴⁹⁹ A functionally similar distinction has been proposed by Carol Myers-Scotton (2006) who contrasts “cultural borrowings”, which fill the gaps in the borrowing language and encompass words for new objects or concepts, with “core borrowings”, which duplicate existing elements in the borrowing language’s inventory. The latter may occur either in the speech of bilinguals who regularly use both the languages, or due to the “sheer magnetism of the dominant culture of the donor language” (ibid., 215).

ments other sources on early modern history of South Asia and often provides valuable data regarding the then Indo-Muslim society and culture.⁵⁰⁰

5.2.7. Need as the motivation for borrowing is a category that is both objective and easy to comprehend category. However, other reasons why words introduced as alternative lexicalizations for already existing concepts win the competition remain more volatile and difficult to define.⁵⁰¹ The choice of a speaker or writer who opts for particular loanword is the decisive factor in the process of unnecessary borrowing. This choice is fuelled by special motivation, which may concern various dimensions of reality and may be connected with prestige and social dominance

⁵⁰⁰ One of many typical spheres which, in the Middle Ages, underwent radical transformation under the prepotent cultural dominance of the newcomers was the art of writing. The Perso-Arabic /*kitab*/ became the symbolic embodiment of the new era. It denoted a book brought from Arabia and Persia in its modern paper shape, and soon completely replaced local Indic terms used for indigenous manuscripts (/pothi/, /granthə/), which usually had the form of palm leaves or pieces of birch bark tied in the middle with a string. An entire group of loanwords connected with books and writing made their permanent home in Hindustani, e.g.: /*kacəz*/ 'paper', /*qələm*/ 'pen', /*qələmdan*/ 'pen-holder', /*qələmkar*/ 'writer, painter', /*soxta*/ 'blotting paper', /*s(i)jahi*/ 'ink', /*dəvət*/ 'inkpot', /*təxti*/ 'writing-board', /*məsauda*/ 'draft, manuscript', /*dəftər*/ 'archive', /*kitab-xana*/ 'library', /*katib*/, /*muhər.ir*/, /*munfi*/ all meaning 'writer, clerk', /*dəbir*/ 'writer, secretary', /*musan:if*/ 'writer, author', /*hərf*/ 'letter, word', /*xət*/ 'a letter, handwriting', /*məktub*/ 'letter, epistle', /*nama*/ 'letter, document', /*pərjə*/ 'piece of paper, note', /*ruq:a*/ 'a note', /*imla*/ 'orthography', /*xətkəfi*/ 'calligraphy', /*xuʃnəvis*/ 'calligrapher', /*təhrir*/ 'writing', and many others. They functioned as new names for recently adopted objects and conceptions or supplanted old genuine terms which gradually fell into disuse along with their designates.

⁵⁰¹ There are many common loanwords of everyday use which have, in the course of time, become ubiquitous and have totally supplanted existing native Indic terms, e.g.: /*nəmək*/ 'salt', /*həzar*/ 'thousand', /*əngur*/ 'a grape', /*tir*/ 'an arrow', /*kəman*/ 'bow', /*dəftər*/ 'office', and many others.

assigned to the donor language.⁵⁰² The desire to be associated with speakers of a dominant language and gain socially from its prestige is an important mechanism in contact-induced language change (Thomason & Kaufman 1988: 44ff). This was also true for the contact between Persian and Hindustani: the mechanism proved to be effective in many spheres, and above all in literature, where its influence was, and still is, particularly appreciable. Since the times of Kabīr (*fl.* 15th c.) and Tulsīdās (*fl.* 16th c.) Indian authors freely mixed foreign (i.e. Persian and Arabic) lexemes into their writings, initially probably because these were just commonly used words.⁵⁰³ But in the course of time Persian aesthetic standards gained such a great reputation that it came to be recognized as the only effective language to express cultural accomplishments. Deficiency in elegant self-expression in Persian meant cultural failure (Alam 1998: 330). Therefore, even those writers (and poets in particular) who did not compose in Persian, tended to follow the desired pattern in terms of stylistics and rhetorical devices, and encrusted their compositions with Persian or Arabic synonyms of indigenous words and phrases. Thus, they would rather prefer ≤ P /ʃəʃm/, /dida/, /nəzər/ or

⁵⁰² According to Tariq Rahman, in mediaeval Indo-Muslim society the linguistic reality was categorized in a hierarchical manner: “the language of cities is superior to that of rural areas; written language is superior to spoken; the language of certain élites is superior to that of ordinary people, etc.” (Rahman 2010: 92). But these were actually the Persian speaking newcomers who soon began to constitute urban-oriented elites and took control over many appreciated spheres of life, including the art of writing and producing books (cf. fn. 500 on page 242). Bearing this in mind, the set of preferences enumerated in the above quotation may be brought down to one general conclusion: Persian was commonly regarded as superior to any other language or dialect spoken in India.

⁵⁰³ ≤ P /nærgis/ ‘narcissus’, fig. ‘the eye of the mistress’.

/nərgis/⁵⁰⁴ to the indigenous /ākʰ/ 'the eye', ≤ A /səɖʒəndʒəl/, /a.ina/ or ≤ P /ʃiʃa/ to H /dərpən/ 'mirror', or one of many elaborate epithets meaning 'beautiful',⁵⁰⁵ to the genuine Hindustani adjective /sundər/. Employing foreign words of greater prestige was the easiest method to imitate ideal literary models and, for this reason, it was commonly used even by mediocre and amateur writers. Given the great esteem enjoyed by the poetic art in India, where every second person considers himself a connoisseur and every third – a poet, it is not surprising that these kind of luxury borrowings dispersed widely in the society and percolated even to the lowest, uneducated strata of society.⁵⁰⁶

5.2.8. In terms of the degree of assimilation, the situation of Perso-Arabic loanwords in Hindustani is highly complex. The naturally multidimensional character of the process of assimilation is, in this case, further complicated by heterogenous sources and origins of borrowings (Persian, Arabic, Turkic) on the one hand and the inherent diverseness of Hindustani on the other. The process of assimilation takes place at all levels of the language system, including orthography, phonetic form, morphological struc-

⁵⁰⁴ On the use of Perso-Arabic vocabulary by Indian writers see Bahri 1960: 70ff.

⁵⁰⁵ E.g.: ≤ P /gulbədən/ lit. 'having a body like the rose', /dɪlbər/ lit. 'captivating the heart', /dɪlkəʃ/ lit. 'heart-attracting', /xubsurət/ lit. 'of beautiful face or form', /maʃuqana/ 'lovely, like a beloved', /numa.ʃi/ lit. 'for show or display', /pərizad/ 'lit. 'fairy-born', /nəmkin/ lit. 'salty', fig. 'charming, attractive, beautiful'; ≤ A /həsin/ 'beautiful' was also one of the favourites.

⁵⁰⁶ And from the 20th c. onwards another transmission channel has turned out to be extremely propitious for broadcasting Perso-Arabic vocabulary among South Asians and sustaining its vitality, namely, Indian popular cinema, which crosses all the social, cultural and political boundaries. Film songs which continue the vivid tradition of Hindustani *ghazal* in its musical and poetic dimensions are nowadays the most effective disseminators and preservers of Persian and Arabic borrowings.

ture, lexical category, and the meaning and usage of a borrowed word. Traditionally, according to the degree of assimilation, loanwords are divided into subcategories of fully assimilated, partly assimilated and non-assimilated (e.g. Holden 1976). It can be stated with a high degree of certainty that borrowings of the last type (non-assimilated) practically do not occur in the analyzed material, for the reason that all the discussed loanwords have undergone graphemic adaptation – they all have a standardized form of notation in Nagari along with their notation in Perso-Arabic script which often remains unchanged in comparison to its form in the source language.⁵⁰⁷ From this point of view, all Perso-Arabic loanwords should be treated by rule as units assimilated to the Hindustani linguistic system. Moreover, most loanwords show at least a minimal degree of acceptance at one or another level and this entails the immediate change of one or more qualities of a word. Regardless of how minute the change, in every case it is a sign of assimilation. The multilevel processes of assimilation to which Perso-Arabic loanwords have been subject in Hindustani are discussed respectively in the previous chapters of this study. The analyses demonstrated that apart from orthography, the fullest integrity can be observed in the sphere of morphosyntax: although it may happen that a loanword follows

⁵⁰⁷ Bearing this in mind, only those Arabic words and ready-made phrases which function as quotations (casuals, cf. p. 51) both in speech and in writing and whose written form remains graphically unchanged (*vide supra* fn. 490 on p. 237) may be considered truly non-assimilated loanwords form. The most common of them is the fossilized graphemic representation of the word Allāh (الله) which even in the Unicode Standard system has got its unique codepoint, automatically changing the sequence of letters (*A-l-l-h*) into the encoded presentation of this word.

the rules of its source language morphology, it is only one of the possibilities and never the only applicable rule.⁵⁰⁸

5.2.9. Another issue connected with the subject of the assimilation of loanwords is borrowability of basic or so-called “core vocabulary” items which are generally regarded as particularly resistant to change (replacement or loss), and, therefore, very rarely (or never) borrowed.⁵⁰⁹ A cursory comparison of the analyzed linguistic material with the Swadesh list of one hundred basic words,⁵¹⁰ conducted primarily out of curiosity, without former planning, and, therefore obviously lacking the necessary methodical scrutiny, revealed quite surprising results: (i) lexical equivalents originating in Persian (mostly) or Arabic exist in Hindustani for 51 out of the hundred words included on the list; (ii) in twelve cases there is more than one Perso-Arabic equivalent current in Hindustani; (iii) in most cases, Perso-Arabic loanwords in Hindustani function interchangeably with indigenous Indic words (*tadbhava* or *tatsama*); their use, or lack of it, in the speech act probably depends entirely on the speaker’s choice, although it can be assumed that some Perso-Arabic loanwords are used more fre-

⁵⁰⁸ For example, in Hindustani the genuinely Arabic noun /ħarf/ f. ‘a letter’ may form the Arabic sound plural /ħarfāt/ ‘letters’, but more regularly it will take Hindustani plural terminations for both its direct (/ħarfē/) and oblique (/ħarfō/) form. Moreover, the plural can be multiplied by a speaker unaware of the fact that /ħarfāt/ already is a plural. Thus the forms /ħarfātē/ in direct and /ħarfātō/ in oblique are not completely unthinkable possibilities.

⁵⁰⁹ Cf.: “(...) basic vocabulary tends to resist change most successfully, even in the face of heavy prestige” (Hock & Joseph 1996: 385), a similar opinion can be found in Thomason 2001: 71-72 and Field 2002: 6. The core vocabulary is most commonly defined as high frequency vocabulary, relatively small in number, constituting the vast majority of what people say in normal communication and containing, i.e., terms for kinship relations, fundamental activities, basic body parts, colours, substances, etc.

⁵¹⁰ Original final Swadesh list published in 1971 (see References).

quently than indigenous words;⁵¹¹ (iv) however, it seems that in six cases Perso-Arabic loanwords completely (or almost completely) replaced native lexemes; these are: H $\leq$ P /ʃərbi/ 'grease', H $\leq$ P /gərdən/ 'neck', H $\leq$ P /kələdʒa/ 'liver', H $\leq$ P /lal/ 'red', H $\leq$ P /səfed/ 'white', H $\leq$ P /gər(ə)m/ 'hot'; (v) there are three pairs of doublets derived from the same, Indo-Iranian source: H $\leq$ P /jək/ vs. H $\leq$ S /ek/ 'one', H $\leq$ P /sər/ vs. H $\leq$ S /sır/ 'head', H $\leq$ P /barıf/ vs. H $\leq$ S /bərsat/ 'rain'; (vi) out of 18 words denoting names of state or action, enumerated on the list, two can be expressed in Hindustani by conjunct or compound verbs; these two verbs have popular synonyms formed by means of borrowed lexemes: 'die' = H /dəm nıkəlna/ $\leq$ P /dəm/ 'breath, life', H /guzər jana/ $\leq$ P /guzər/ 'passage; life'; H /xətm hona/ $\leq$ P [$\leq$ A] /xətm/ 'finished' and 'kill' = H /qətıl kərna/ $\leq$ P [$\leq$ A] qətıl/ 'killing'.

5.2.10. There is no doubt that both the question of the degree of loanwords' integrity with the borrowing language system and this language's tendency towards replacing items belonging to the core vocabulary ought to be analyzed taking into account a number of circumstances which influence both the processes. Such factors as genetic relatedness of the languages in contact, oral way of borrowing, high frequency and historical persistence of a loanword, can smooth and accelerate the process of assimilation of both cultural and core linguistic items. In the case of Hindustani, another crucial agent comes into play: a multitude of conversional styles, dependent on speakers' geographical, social, cultural, religious etc. environment. This diversity of styles and registers results in the lack of one satisfactory answer to the question concerning the degree of assimilation of loanwords. This matter would require further scrutiny based on extensive

⁵¹¹ E.g.: H $\leq$ P /xun/ 'blood' is probably more frequently used than H $\leq$ S /rəkt/; H $\leq$ P [$\leq$ A] /admi/ 'man' is probably more common than H $\leq$ S /poruʃ/; H $\leq$ P /barıf/ 'rain' is probably more common than H /bərsat/.

and detailed research conducted among individual groups of speakers. The author remains with the hope that the tentative conclusions presented here will encourage further meticulous and systematic studies on Perso-Arabic borrowings in Hindustani.

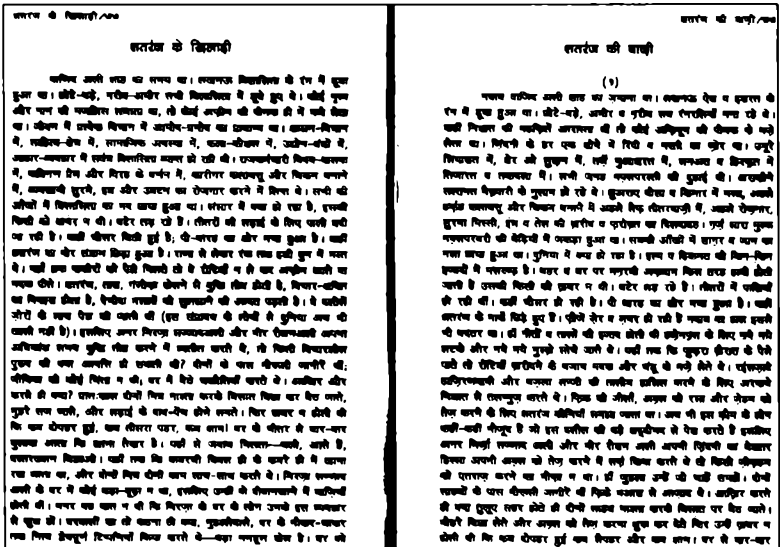


Fig. 13. "Bilingual" Hindi (left) and Urdu (right) edition of Premchand's short story *The Chess Players* in Nagari script. The differences between both texts, mainly lexical and syntactic, start from the title: *Sātrañj ke khilārī* (Hindi) vs. *Sātrañj kī bāzī* (Urdu). (Source: *Samakṣ: Muṁśī Premchand kī bis hindī-urdu kahāniyom kā samāntar pāṭh*. 2002. Ed. Ālok Rāy, Muṣṭak Alī. Ilāhābād: Haṁs Prakāśan)

References

- Abbi, Anvita, R. S. Gupta, and Ayesha Kidwai, eds. 2001. *Linguistic Structure and Language Dynamics in South Asia: Papers from the Proceedings of SALA XVIII Roundtable*. MLBD Series in Linguistics 15. Delhi: Motilal Banarsidass.
- Abbi, Anvita, Imtiaz Hasnain, and Ayesha Kidwai. 2004. *Whose Language is Urdu?* Heidelberg Papers in South Asian and Comparative Politics (HPSACP) 24. South Asia Institute, Department of Political Science, University of Heidelberg. [Available online: <http://hpsacp.uni-hd.de/>]
- Abidi, Sayyid Amir Hasan, and Ravinder Gargesh. 2008. Persian in South Asia. In: Kachru et al. 2008. 103-120.
- Abū'l-Faẓl, 'Allāmī. *The Ā'in-i Akbarī*. 1873. Vol. I, transl. by H. Blochmann, ed. D. C. Phillott. Vol. II & III, transl. by H. S. Jarret. Calcutta: Asiatic Society. [Repr. 1993]
- Adle, Chahryar. 1982. *Art et société dans le monde iranien*. Bibliothèque iranienne 26. Paris: Editions Recherche sur les civilisations.
- Ahmad, Aziz. 2002. *Studies in Islamic Culture in the Indian Environment*. New Delhi: Oxford University Press. [1st ed. 1964]
- Ahmad, Imtiaz. 1967. The Ashraf and Ajlaf Categories in Indo-Muslim Society. *Economic and Political Weekly*. Vol. 2, No. 19. 887+889-891.
- Ahmad, Rizwan. 2008. Scripting a New Identity. The Battle for Devanagari in Nineteenth Century India. *Journal of Pragmatics*. Vol. 40. 1163-1183.
- Ahmad, Rizwan. 2011. Urdu in Devanagari. Shifting Orthographic Practices and Muslim Identity in Delhi. *Language in Society*. Vol. 40. 259-284.
- Ahmad, Rizwan. 2012. Hindi is Perfect, Urdu is Messy. The Discourse of Delegitimation of Urdu in India. In: Jaffe, Alexandra, Jannis Androutsopoulos, Mark Sebba, and Sally Johnson, eds. *Orthography as Social Action: Scripts, Spelling, Identity and Power*. Language and Social Processes [LSP] 3. Boston/Berlin: De Gruyter. 103-134.
- Ahmad, Zafar. 2000. *Islam and Muslims in South Asia*. Delhi: Authorspress.
- Alam, Muzaffar. 1998. The Pursuit of Persian: Language in Mughal Politics. *Modern Asian Studies*. Vol. 32, No. 2. 317-349.
- Alam, Muzaffar. 2003. The Culture and Politics of Persian in Precolonial Hindustan. In: Pollock 2003. 131-198.
- Ali, A. Yusuf. 1917. The Importance of Hindustani. *Bulletin of the School of Oriental Studies*. Vol. 1, No. 1. 109-111.
- Ali, A. Yusuf. 1920. Note on Urdu Orthography. *Bulletin of the School of Oriental Studies*, Vol. 1, No. 3. 29-34.

- Amritavalli, R., and K. A. Jayaseelan. 2007. India. In: Simpson, Andrew, ed. *Language and National Identity in Asia*. Oxford: Oxford University Press. 55-83.
- Aneesh, A. 2010. Bloody Language: Clashes and Constructions of Linguistic Nationalism in India. *Sociological Forum*. Vol. 25, No. 1. 86-109.
- Apte, Mahadev L. 1974. Pidginization of a Lingua Franca: A Linguistic Analysis of Hindi-Urdu Spoken in Bombay. In: Southworth, Franklin C., and Mahadev L. Apte, eds. *Contact and Convergence in South Asian Languages. International Journal of Dravidian Linguistics*. Vol. 3, No. 1-2. Trivandrum: Dept. of Linguistics, University of Kerala. 21-41.
- Austin, Granville. 1966. *The Indian Constitution: Cornerstone of a Nation*. Oxford: Clarendon P.
- Awasthi, Suresh, and Induja Awasthi, eds. 1981 [2001]. *Chambers English-Hindi Dictionary*. New Delhi [etc.]: Allied Publications Limited.
- Ayres, Alyssa. 2009. *Speaking Like a State: Language and Nationalism in Pakistan*. Cambridge [etc.]: Cambridge University Press.
- Āzād, Muḥammad Ḥusain. 2001. *Āb-e ḥayāt. Shaping the Canon of Urdu Poetry*. Translated and edited by Frances Pritchett, in association with Shamsur Rahman Faruqi. New Delhi: Oxford University Press.
- Bābur. 1922. *The Bābur-nāma in English (Memoirs of Bābur)*. Translated from the original Turki Text of Zahiru'd-dīn Muḥammad Bābur Pādshāh Ghāzī by Annette Susannah Beveridge. Vol. II. London: Luzac & Co.
- Badawi, Elsaid, Mike G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A Comprehensive Grammar*. London-New York: Routledge.
- Bahri, Hardev. 1960. *Persian Influence on Hindi*. Allahabad: Bharati Press Publication.
- Bahri, Hardev. 1989. *Learners' Hindi-English dictionary – Siksarṭhi Hindi-Angrejhi sabdakosa*. Delhi: Rajapala. [Available online: <http://dsal.uchicago.edu/dictionaries/bahri/>]
- Baig, Bushra, M. Kumar, and Sujoy Das. 2012. Rule Based Hindi to Urdu Transliteration System. *Journal of Emerging Trends in Computing and Information Sciences*. Vol. 3, No. 8. 1200-1204.
- Bailey, Thomas Grahame. 1922. The Pronunciation of Urdu and Hindi. *Bulletin of the School of Oriental Studies*. Vol. 2, No. 3. 539-547.
- Bailey, Thomas Grahame. 1926. The Use and Meaning of the Term Khaṭī Bolī. *Journal of the Royal Asiatic Society of Great Britain and Ireland*. No. 4. 717-723.
- Bailey, Thomas Grahame. 1938. *Studies in North Indian Languages*. Farnham: Lund Humphries.
- Balić, Smail. 1978. *Die Kultur der Bosniaken, Supplement I: Inventar des bosnischen literarischen Erbes in orientalischen Sprachen*. Vienna: Adolf Holzhausens.

- Barannikov, A. 1936. Modern Literary Hindī. *Bulletin of the School of Oriental Studies*. Vol. 8, No. 2/3. Indian and Iranian Studies: Presented to George Abraham Grierson on His Eighty-Fifth Birthday, 7th January, 1936. 373-390.
- Barannikov, Alekseĭ Petrovich, and Petr Alekseevich, Barannikov. 1956. *Khindustani (khindi i urdu)*, Moskva: Izdatel'stvo Literatury na Inostrannykh Jazykakh.
- Barz, Richard, and Yogendra Yadav. 1993. *An Introduction to Hindi and Urdu*. New Delhi: Munshiram Manoharlal Publishers.
- Bayevsky, Solomon. 1999. Farhang-e Jahāngīrī. *Encyclopaedia Iranica*. Vol. IX. [Available online: <http://www.iranicaonline.org/articles/farhang-e-jahangiri>]
- Beames, John. 1872. *A Comparative Grammar of the Modern Aryan Languages of India*. Vol. 1-3. London: Trübner & Co.
- Beg, Mirza Khalil Ahmad. 1988. *Urdu Grammar: History and Structure*. Series in Indian Languages & Linguistics (SILL) 18. New Delhi: Bahri Publications.
- Beg, Mirza Khalil Ahmad. 1996. *Sociolinguistic Perspective of Hindi and Urdu in India*. Series in Indian Languages & Linguistics (SILL) 22. New Delhi: Bahri Publications.
- Bentinck, William. 1835. Governor General William Bentinck's Resolution for Provision for English Education. In: Chaudhary 2009. 505-506.
- Beskrovnĭi, Vasilĭi Matveevich. 1960a. Morfologicheskoe gibrinoe slovo-obrazovanie v khindi. In: Beskrovnĭi 1960b. 118-151.
- Beskrovnĭi, Vasilĭi Matveevich, ed. 1960b. *Khindi i urdu. Voprosy leksikologii i slovoobrazovaniya*. Moskva: Izdatel'stvo vostochnoi literatury.
- [Besk.] = Beskrovnĭi, Vasilĭi Matveevich, ed. 1972. *Khindi-russkii slovar'*. Vol. I-II. Moskva: Izdatel'stvo "Sovetskaya Ėnciklopediya".
- Bhatia, Tej K. 1987. *A History of the Hindi Grammatical Tradition. Hindi-Hindustani Grammar, Grammarians, History and Problems*. Leiden: Brill.
- Bhatia, Tej K. 1992. Variation in Hindi. Parameters for a Pan-dialectal Grammar. In: Dimock et al. 1992. 163-197.
- Bhatia, Tej K., and William C. Ritchie. 2004. Bilingualism in South Asia. In: Bhatia, Tej K., and William C. Ritchie, eds. *The Handbook of Bilingualism*. Malden, MA [etc.]: Blackwell Publishing. 780-807.
- Bhatia, Tej K., and Kazuhiko Machida. 2008. *The Oldest Grammar of Hindustānī. Contact, Communication and Colonial Legacy*. Vol. 1-3. Tokyo: Research Institute for Languages and Cultures of Asia and Africa. Tokyo University of Foreign Studies.
- Bhatty, Zarina. 1996. Social Stratification Among Muslims in India. In: Srinivas, M.N., ed. *Caste: Its Twentieth Century Avatar*. New Delhi: Viking. 244-262.
- Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt and Company. [Repr. New Delhi: Motilal Banarsidass, 2005]

- Bodewitz, Hendrik Wilhelm. 1994-95. Ketelaar and Millius and Their Grammar of Hindustani. *Bulletin of the Deccan College Post-graduate & Research Institute*. Vol. 54-55. 123-131
- Bodrogligeti, András. 1971. *The Persian Vocabulary of the Codex Cumanicus*. Budapest: Akadémiai Kiadó.
- Brass, Paul R. 1994. *The Politics of India since Independence*. Cambridge: Cambridge University Press. [2nd ed.]
- Brass, Paul R. 2005. *Language, Religion and Politics in North India*. Lincoln, NE: iUniverse. [1st ed. 1974]
- Bruijn, Johannes Thomas Pieter de. 2012. Sabk-i Hindī. *Encyclopaedia of Islam*. Second Edition. Brill Online. [Available online: http://referenceworks.brillonline.com/entries/encyclopaedia-of-islam-2/sabk-i-hindi-SIM_6377]
- Bubenik, Vit. 1996. *The Structure and Development of Middle Indo-Aryan Dialects*. Delhi: Motilal Banarsidass.
- Burton-Page, John. 1957. Compound and Conjunct Verbs in Hindi. *Bulletin of the School of Oriental and African Studies*. Vol. 19. 469-478.
- Busch, Allison. 2010. Hidden in Plain View: Brajbhasha Poets at the Mughal Court. *Modern Asian Studies*. Vol. 44, Nr 2. 267-309.
- Byrne, Darragh, and Max Kölbel, eds. 2009. *Arguing About Language*. London: Routledge.
- Canfield, Robert L. 1991. Introduction: the Turko-Persian Tradition. In: Canfield, Robert L., ed. *Turco-Persia in Historical Perspective*. Cambridge [etc.]: Cambridge University Press. 1-34.
- Cannon, Garland H. 1994. *The Arabic Contributions to the English Language: An Historical Dictionary*. Collab. Alan S. Kaye. Wiesbaden: Harrassowitz.
- Cannon, Garland H. 2000. Problems in Studying Loans. *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society*. Vol. 25, No. 1. 326-336.
- Cannon, Garland H., and Alan S. Kaye. 2001. *The Persian Contributions to the English Language: An Historical Dictionary*. Wiesbaden: Otto Harrassowitz.
- Cardona, George, and Dhanesh Jain, eds. 2003. *The Indo-Aryan Languages*. London and New York: Routledge.
- Caturvedi, Mahendra. 1970. *A Practical Hindi-English Dictionary*. Delhi: National Publishing House. [Available online: <http://dsal.uchicago.edu/dictionaries/caturvedi/>]
- Census of India*. Available online: <http://www.censusindia.gov.in>.
- Chand, Tara. 1944. *The Problem of Hindustani*. Allahabad: Indian Periodicals Ltd. [Available online: <http://www.columbia.edu/itc/mealc/pritchett/00urduhindilinks/tarachand/tarachand.html>]
- Chang, Charles B. 2012. Phonetics vs. Phonology in Loanword Adaptation: Revisiting the Role of the Bilingual. In: *Proceedings of the 34th Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on*

- Information Structure*. Berson, Sarah, et al., ed. Berkeley, CA: Berkeley Linguistics Society. 61-72.
- Chatterji, Suniti Kumar. 1922. Phonetic Transcriptions from Indian Languages. *Bulletin of the School of Oriental Studies*. Vol. 2, No. 2. 177-195.
- Chatterji, Suniti Kumar. 1931. Calcutta Hindustani: A Study of a Jargon Dialect. In: Chatterji 1972. 204-256.
- Chatterji, Suniti Kumar. 1949. India: Language Problems. In: Chatterji 1972. 108-132.
- Chatterji, Suniti Kumar. 1960. *Indo-Aryan & Hindi: Eight Lectures Originally Delivered in 1940 before the Gujarat Vernacular Society, Ahmedabad*. Calcutta: Firma K. L. Mukhopadhyay. [Repr. 1969]
- Chatterji, Suniti Kumar. 1972. *Select Papers. Āṅgla-Nibandha-Chayana*. Vol. 1. New Delhi [etc.]: People's Publishing House.
- Chaudhary, Shreesh. 2009. *Foreigners and Foreign Languages in India: A Sociolinguistic History*. Delhi [etc.]: Cambridge University Press India under the imprint of Foundation Books.
- Chernyššev, Vladimir Aleksandrovich. 1971. Nekotorŷe chertŷ bombeiskogo govora khindustani (na materiale sovremennoĭ prozy khindi). In: Dvoryankov, Nikolai Aleksandrovich, ed., *Indiiskaya i iranskaya filologiya, voprosŷ dialektologii*. Moskva: Izdatel'stvo "Nauka". 121-141.
- Choudhury, Monojit, Anupam Basu and Sudeshna Sarkar. 2004. A Diachronic Approach for Schwa Deletion in Indo Aryan Languages. *Proceedings of the Workshop of the ACL Special Interest Group on Computational Phonology (SIGPHON)*. Barcelona: Association for Computations Linguistics.
- Ciancaglini, Claudia A. 2008. *Iranian Loanwords in Syriac*. Beiträge zur Iranistik 28. Wiesbaden: Dr. Ludwig Reichart Verlag.
- Cohn, Bernard S. 1985. The Command of Language and the Language of Command. In: Guha, Ranjit, ed. *Subaltern Studies IV: Writings on South Asian History and Society*. Delhi: Oxford University Press. 276-329.
- Corriente, Federico. 2008. *Dictionary of Arabic and Allied Loanwords: Spanish, Portuguese, Catalan, Galician and Kindred Dialects*. Leiden: Brill.
- Crystal, David. 2008. *A Dictionary of Linguistics and Phonetics*. 6th edition. Malden, MA [etc.]: Blackwell.
- Dalmia, Vasudha. 2003. The Locations of Hindi. *Economic and Political Weekly*. Vol. 38, No. 14. 1377-1384.
- Danecki, Janusz. 1971. Fonologiczna adaptacja słownictwa arabskiego w języku urdu. *Acta Philologica*. Vol. 3. 89-95.
- Das, Rahul Peter. 2005. Kaste. In: Pietruschka, Ute. 2005. *Gemeinsame kulturelle Codes in koexistierenden Religionsgemeinschaften. Leucorea-Kolloquium 2003*. Hallesche Beiträge zur Orientwissenschaft 38 (2004). Halle: Universität Halle. 75-116.
- Datt, Ruddar 1968. Hindi as Link Language: A Comment. *Economic and Political Weekly*. Vol. 3, No. 38. 1445.

- Deshpande, Anirudh. 2000a. Hindustani in India. *Economic and Political Weekly*. Vol. 35, No. 15. 1240-1242.
- Deshpande, Anirudh. 2000b. Hindi's Decimation. *Economic and Political Weekly*. Vol. 35, No. 34. 3067-3068.
- Dimock, Edward C., Braj B. Kachru, and Bhadriraju Krishnamurti. 1992. *Dimensions of Sociolinguistics in South Asia. Papers in Memory of Gerald Kelly*. New Delhi: Oxford & IBH.
- Diringer, David. 1948. *The Alphabet. A Key to the History of Mankind*, New York: Philosophical Library.
- Doerfer, Gerhard. 1963-75. *Türkische und mongolische Elemente im Neupersischen*. Vol. I-IV. Wiesbaden, Franz Steiner Verlag.
- Du Jarric, Pierre. 1611. *L'histoire des choses plus memorables advenue tant és Indes Orientales qu'autres pays de la decouverte des Portugais*. (...) Valanciennes: Chez Jean Veruliet. [available online: http://daten.digital-sammlungen.de/bsb00076626/image_7]
- Du Jarric, Pierre. 1926. *Akbar and the Jesuits: an Account of the Jesuit Missions to the Court of Akbar*. Translated with introduction and notes by Charles Herbert Payne. London: George Routledge & Sons. [Repr. London: Routledge, 2004]
- Durkin, Philip. 2009. *The Oxford Guide to Etymology*. Oxford-New York: Oxford University Press.
- Eaton, Richard M., ed. 2003. *India's Islamic Traditions, 711-1750*. Delhi: Oxford University Press. [1st ed. 1982]
- Ejaz, Madiha. 2004. Phonological Behavior of Multiple Aspirated Consonants in Urdu. *Axbar-e Urdu*. Vol. 20, No. 12. 140-149.
- Elizarenkova, Tat'jana Jakovlevna. 1988. Morphophonology in Hindi. In: Andronov, Mikhail Sergeevich, and Bhaktiprasad Mallik, eds. *Linguistics: a Soviet approach*. Indian Journal of Linguistics. 367-386.
- Elwell-Sutton, Laurence Paul. 1976. *The Persian Metres*. Cambridge-New York: Cambridge University Press.
- Emeneau, Murray Barnson. 1956. India as a Linguistic Area. *Language*. Vol. 32, No. 1. 3-16.
- Ernst, Carl W. 1992. *Eternal Garden. Mysticism, History, and Politics at a South Asian Sufi Center*. New York: State University of New York Press.
- Encyclopædia Britannica*. Available online: <http://www.britannica.com>.
- Ethnologue: Languages of the World*. See: Lewis, M. Paul, ed. 2009.
- Everaert, Christine. 2010. *Tracing the Boundaries between Hindi and Urdu: Lost and Added in Translation between 20th Century Short Stories*. Brill's Indological Library 32. Leiden-Boston: Brill.
- Fallon, S. W. 1879. *A New Hindustani-English Dictionary, with Illustrations from Hindustani Literature and Folk-lore*. Banaras, London: Trubner and Co.
- Faruqi, Shamsur Rahman. 2001. *Early Urdu Literary Culture and History*. Delhi: Oxford University Press.

- Faruqi, Shamsur Rahman. 2003. A Long History of Urdu Literary Culture, Part I: Naming and Placing a Literary Culture. In: Pollock 2003. 805-863.
- Faruqi, Shamsur Rahman. 2004. A Stranger in the City: The Poetics of *Sabk-e Hindi*. *Annual of Urdu Studies*. Vol. 19. 1-93.
- Ferguson, Charles A. 1992. South Asia as a Sociolinguistic Area. In: Dimock et al. 1992. 25-36.
- [Ferg.] = Fergusson, John. 1773. *A Dictionary of the Hindostan Language, in Two Parts: I, English and Hindostan, II, Hindostan and English.... to which is Prefixed a Grammar of the Hindostan Language*. London: T. Cadell.
- Field, Fredric W. 2002. *Linguistic Borrowing in Bilingual Context*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Fishman, Joshua A. 1973. *Language and Nationalism. Two Integrative Essays*. Rowley, Mass.: Newbury House Publishers.
- Forbes, Duncan. 1846. *A Grammar of the Hindūstānī Language in the Oriental and Roman Character (...)*. London: Wm. H. Allen & Co.
- [Forbes] = Forbes, Duncan. 1858. *A Dictionary, Hindūstānī and English: Accompanied by a Reversed Dictionary, English and Hindūstānī*. London. [Repr. Lahore: Sang-e-meel, 2002]
- Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*. No. 100. 25-50.
- Friedlander, Peter G. 2006. Hindustani Textbooks from the Raj. *Electronic Journal of Foreign Language Teaching*. Vol. 3, Suppl. 1. 39-56. [Available online: <http://e-flt.nus.edu.sg/v3sp2006/friedlander.pdf>]
- Gandhi, Mohandas, Karamchand. 1965. *Our Language Problem*. Edited and published by Anand T. Hingorani. Bombay: Bharatiya Vidya Bhavan.
- Gellner, Ernest. 1983. *Nations and Nationalism*. Ithaca, NY: Cornell University Press.
- Gilchrist, John Bortwick. 1787-90. *A Dictionary, English and Hindoostanee, in which the Words are Marked with their Distinguishing Initials, as Hinduwee, Arabic, and Persian*. 2 vols. Calcutta: Stuart & Cooper (vol. 1); Cooper & Upjohn (vol. 2).
- Gilchrist, John Bortwick. 1796. *A Grammar of the Hindoostanee Language; or, Part Third of Volume First, of a System of Hindoostanee Philology*. Calcutta: Chronicle Press.
- Gilchrist, John Bortwick. 1798. *The Oriental Linguist: an Easy and Familiar Introduction to the Popular Language of Hindoostan (Vulgarly but Improperly Called the Moors), Comprising the Rudiments of That Tongue, with an Extensive Vocabulary, English and Hindoostanee, and Hindoostanee and English (...)*. Calcutta: Ferris and Greenway.
- Gilchrist, John Bortwick. 1826. *Dialogues, English and Hindoostanee; for Illustrating the Grammatical Principles of the Strangers' East Indian Guide, and to Promote the Colloquial Intercourse of Europeans on the Most Indis-*

- pensable and Familiar Subjects with the Natives of India (...)*. London: Kingsbury, Parbury, and Allen.
- Golden, Peter B. 2006. Turks and Iranians: An Historical Sketch. In: Johanson & Bulut 2006. 17-38.
- Green, Nile. 2010. The Uses of Books in a Late Mughal Takiyya: Persianate Knowledge between Person and Paper. *Modern Asian Studies*. Vol. 44, Nr. 2. 241-265.
- Grierson, George Abraham, ed. 1916. *Linguistic Survey of India*. Vol. IX. *Indo-Aryan Family, Central Group*. Part I: *Specimens of Western Hindī and Pañjābī*. Calcutta: Superintendent Government Printing, India.
- Grierson, George Abraham. 1918. Indo-Aryan Vernaculars. *Bulletin of the School of Oriental Studies*. Vol. 1, No. 2. 47-81.
- Grierson, George Abraham. 1920. Indo-Aryan Vernaculars (Continued). *Bulletin of the School of Oriental Studies*. Vol. 1, No. 3. 51-85.
- Gumperz, John. 1958. Dialect Differences and Social Stratification in a North Indian Village. *American Anthropologist*. New Series. Vol. 60, No. 4. 668-682.
- Gumperz, John J., and Choudhri Mohammed Naim. 1960. Formal and Informal Standards in Hindi Regional Language Area. In: Ferguson, Charles A., and John J. Gumperz, eds. *Linguistic Diversity in South Asia. International Journal of American Linguistics*. Vol. 26, No. 3 (3). 92-118.
- Gumperz, John J. 1961. The Language Problem in South Asian Studies. *Asian Survey*. Vol. 1, No. 2. 28-31.
- Hadley, George. 1772. *Grammatical Remarks on the Practical and Vulgar Dialect of the Indostan Language, Commonly Called the Moors. With a Vocabulary, English and Moors. The Whole Calculated for the Common Practice in Bengal*. London: T. Cadell.
- Hakala, Walter Niels. 2010. *Diction and Dictionaries: Language, Literature, and Learning in Persianate South Asia*. Philadelphia: University of Pennsylvania.
- Hall, Kira. 2002. 'Unnatural' Gender in Hindi. In: Hellinger, Marlis, and Hadumod Bussman, eds. *Gender Across Languages. The Linguistic Representation of Women and Men*. Amsterdam: John Benjamins. 133-162.
- Hamidullah, Muhammad. 1954. Cultural and Intellectual History of Indian Islam. *Die Welt des Islams*. New Series. Vol. 3, Issue 3/4. 127-144.
- Haqee, Shanul Haq. 2011. *Oxford English-Urdu Dictionary*. Karachi: Oxford University Press. [8th ed.]
- Hashabeiky, Forogh. 2005. *Persian Orthography: Modification Or Changeover? (1850-2000)*. Acta Universitatis Upsaliensis. Studia Iranica Upsaliensia 7. Uppsala: Uppsala Universitet.
- Hasnain, S. Imtiaz, and K. S. Rajyashree. 2004. Hindustani as an Anxiety between Hindi-Urdu Commitment. *The Yearbook of South Asian Languages and Linguistics*. Vol. 7. 247-266.

- Haspelmath, Martin. 2008. Loanword Typology. Steps Toward a Systematic Cross-linguistic Study of Lexical Borrowability. In: Stolz, Thomas, Dik Bakker and Rosa Salas Palomo, eds. *Aspects of Language Contact. New Theoretical, Methodological and Empirical Findings with Special Focus on Romancisation Processes*. Empirical Approaches to Language Typology 35. Berlin-New York: Mouton de Gruyter. 43-62.
- Haspelmath, Martin. 2009. Lexical Borrowing: Concepts and Issues. In: Haspelmath & Tadmor 2009. 35-54.
- Haspelmath, Martin, and Uri Tadmor, eds. 2009. *Loanwords in the World's Languages. A Comparative Handbook*. Berlin: De Gruyter Mouton.
- Haugen, Einar. 1950. The Analysis of Linguistic Borrowing. *Language*. Vol. 26, No. 2. 210-231.
- Haugen, Einar. 1966. Dialect, Language, Nation. *American Anthropologist. New Series*. Vol. 68, No. 4. 922-935.
- Hātim, Šaīx Ẓuhūru 'd-Dīn. 1756. *Dīvān Zāda*. Compiled and edited by Ġulām Ḥusain Ẓulfiqār. Lahaur: Maktaba Xiyābān-e Adab. [1975]
- [Hayyim] = Hayyim, Sulayman 1934-1936. *New Persian-English Dictionary, Complete and Modern...* Vol. I-II. Teheran. [Available online: <http://dsal.uchicago.edu/dictionaries/hayyim/>]
- Heffernan, Kevin. 2005. Phonetic Similarity and Phonemic Contrast in Loanword Adaptation. *Toronto Working Papers in Linguistics*. Vol. 24. 117-123.
- Hemmerdinger, Bertrand. 1969. 158 noms communs grecs d'origine iranienne, d'Eschyle au grec moderne. *Byzantinoslavica* 30. 18-41.
- Henderson, Michael M. T. 1975. Diglossia in Kabul Persian Phonology. *Journal of the American Oriental Society*. Vol. 95, No. 4. 651-654.
- Henderson, Michael M. T. 1983. Four Varieties of Pashto. *Journal of the American Oriental Society*. Vol. 103, No. 3. 595-597.
- Hock, Hans Heinrich. 1991. *Principles of Historical Linguistics*. Berlin-New York: Mouton de Gruyter. [second revised and updated edition]
- Hock, Hans Heinrich, and Brian D. Joseph. 1996. *Language History, Language Change and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Trends in Linguistics. Studies and Monographs 93. Berlin-New York: Mouton de Gruyter.
- Hodgson, Marshall G. S. 1974. *The Venture of Islam*. Volume 1: *The Classical Age of Islam*. Chicago and London: The University of Chicago Press.
- Hoenigswald, Henry M. 1948. Declension and Nasalization in Hindustani. *Journal of the American Oriental Society*. Vol. 68, No. 3. 139-144.
- Hoenigswald, Henry M. 1960. *Language Change and Linguistic Reconstruction*. Chicago: University of Chicago Press.
- Hoffmann, Charlotte. 1991. *An Introduction to Bilingualism*. London and New York: Longmann.
- Holden, Kyril. 1976. Assimilation Rates of Borrowings and Phonological Productivity. *Language*. Vol. 52, No. 1. 131-147.

-
- Hughes, Geoffrey. 2000. *A History of English Words*. Oxford: Blackwell Publishers.
- International Phonetic Association (IPA). 1999. *Handbook of the International Phonetic Association: A Guide to the Use of the international Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Jain, Dhanesh. 2003. Sociolinguistics of the Indo-Aryan Languages. In: Cardona & Jain 2003. 46-65.
- Jazayery, Mohammad Ali. 1970. The Arabic Element in Persian Grammar: A Preliminary Report. *Iran*. Vol. 8. 115-124.
- Johanson, Lars, and Christiane Bulut, eds. 2006. *Turkic-Iranian Contact Areas. Historical and Linguistic Aspects*. Turcologica 62, Wiesbaden: Harrassowitz Verlag.
- Kachru, Braj B., Yamuna Kachru, and S. N. Sridhar, eds. 2008. *Language in South Asia*. Cambridge: Cambridge University Press.
- Kachru, Yamuna. 2006. *Hindi*. London Oriental and African Language Library 12. Amsterdam: John Benjamins Publishing Co.
- Kachru, Yamuna. 2008. Hindi-Urdu-Hindustani. In: Kachru et al. 2008. 81-102.
- Kahane, Henry. 1986. A Typology of the Prestige Language. *Language*. Vol. 62, No. 3 (September). 495-508.
- Karttunen, Klaus. 1995. The Name of India. In: *Cracow Indological Studies*. Vol. 1. *International Conference on Sanskrit and Related Studies. September 23-26, 1993 (Proceedings)*. Cracow: The Enigma Press. 151-163.
- Karve, D. D. 1968. Hindi as Link Language. *Economic and Political Weekly*. Vol. 3, No. 11. 447.
- Katyu, Markandey. 2012. Required, Two Tongues. *The Hindu*. September 20, 2012. [Online edition: <http://www.thehindu.com/opinion/op-ed/required-two-tongues/article3915475.ece>]
- Kelkar, Ashok Ramchandra. 1968. *Studies in Hindi-Urdu. I. Introduction and Word Phonology*. Poona: Deccan College.
- Kent, Roland G. 1950. *Old Persian. Grammar. Texts. Lexicon*. New Haven, Connecticut: American Oriental Society.
- Ketelaar 1698. See: Bhatia & Machida. 2008.
- Khan, Abdul Jamil. 2006. *Urdu/Hindi: An Artificial Divide: African Heritage, Mesopotamian Roots, Indian Culture & British Colonialism*. New York: Algora Publishing.
- Khan, Iqtidar H. [ca. 1999]. *A Contrastive and Comparative Study of Standard Urdu and Standard Hindi*. Aligarh: Aligarh Muslim University.
- Kidwai, Sadiq-ur-Rahman. 1972. *Gilchrist and the "Language of Hindoostan"*. New Delhi: Rachna Prakashan.
- King, Christopher R. 1974. *The Nagari Pracharini Sabha (Society for the Promotion of the Nagari Script and Language) of Benares 1893-1914. A Study in the Social and Political History of the Hindi Language*. A thesis submitted in

- partial fulfillment of the requirements for the degree of Doctor of Philosophy (History). University of Wisconsin.
- King, Christopher R. 1989. Forging a New Linguistic Identity: The Hindi Movement in Banaras, 1868-1914. In: Freitag, Sandria B., ed. *Culture and Power in Banaras: Community, Performance, and Environment, 1800-1980*. Berkeley [etc.]: University of California Press. 179-202.
- King, Christopher R. 1994. *One Language, Two Scripts: The Hindi Movement in Nineteenth Century North India*. New Delhi: Oxford University Press.
- Kłagisz, Mateusz. 2008. Perskie (dari) zapożyczenia oraz kalki językowe w języku paszto. *Studia Linguistica Universitatis Iagellonicae Cracoviensis*. Vol. 125. 79-102.
- Kłagisz, Mateusz. 2010. *Pismo i ortografia języka perskiego*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Koul, Ashok K. 2008. *Lexical Borrowings in Kashmiri*. Delhi: Indian Institute of Language Studies.
- Kuczkiewicz-Fraś, Agnieszka. 1997. Persian Elements in Contemporary Hindi and Their Genesis. *Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace Językoznawcze*. Vol. 118 (MCCX). Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego. 47-67.
- Kuczkiewicz-Fraś, Agnieszka. 2001. Turkic in India and Its Elements in Hindi. *Studia Turcologica Cracoviensia*. Vol. 8. 43-48.
- Kuczkiewicz-Fraś, Agnieszka. 2003a. Medical Terminology of Perso-Arabic Origin in Hindi. *Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace Językoznawcze*. Vol. 121 (MCCLXI). 153-172.
- Kuczkiewicz-Fraś, Agnieszka. 2003b. *Perso-Arabic Hybrids in Hindi: The Socio-linguistic and Structural Analysis*. New Delhi: Manohar.
- Kuczkiewicz-Fraś, Agnieszka. 2008. *Perso-Arabic Loanwords in Hindustani. Part I. Dictionary*. Kraków: Księgarnia Akademicka.
- Kuczkiewicz-Fraś, Agnieszka. 2013. History Preserved in Names. Delhi Urban Toponyms of Perso-Arabic Origin. In: Kuczkiewicz-Fraś, Agnieszka, ed. *Islamicate Traditions in South Asia. Themes from Culture & History*. New Delhi: Manohar. 57-80.
- Kymlicka, Will. 1995. *Multicultural Citizenship. A Liberal Theory of Minority Rights*. Oxford: Clarendon Press.
- Labov, William. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Language in Society 20. Oxford-Cambridge, Mass.: Blackwell.
- Labov, William. 2001. *Principles of Linguistic Change. Volume 2: Social Factors*. Language in Society 29. Oxford-Cambridge, Mass.: Blackwell.
- Labov, William. 2010. *Principles of Linguistic Change. Volume 3: Cognitive and Cultural Factors*. Language in Society 39. Oxford-Cambridge, Mass.: Blackwell.

- LaCharité, Darlene, and Carole Paradis. 2005. Category Preservation and Proximity versus Phonetic Approximation in Loanword Adaptation. *Linguistic Inquiry*. Vol. 36, No. 2. 223-258.
- Lambton, Ann K. S. 2000. *Persian Grammar: Including Key*. Cambridge: Cambridge University Press. [1st ed. 1953]
- Lazard, Gilbert. 1992. *A Grammar of Contemporary Persian*. Transl. from French into English by Shirley A. Lyon. Costa Mesa, Calif.: Mazda Publishers.
- Lehmann, Winfred P. 1992. *Historical Linguistics: An Introduction*. London-New York: Routledge.
- Lelyveld, David. 1993. Colonial Knowledge and the Fate of Hindustani. *Comparative Studies in Society and History*. Vol. 35, No. 4. 665-682.
- Lelyveld, David. 2001. Words as Deeds: Gandhi and Language. *Annual of Urdu Studies*. Vol. 16. 64-75.
- Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the World*. Sixteenth edition. Dallas, Tex.: SIL International. [Online edition: <http://www.ethnologue.com/>]
- Macaulay, Thomas Babington. 1835. Minute on Indian Education for India. In: Chaudhary 2009. 492-505.
- Macdonell, Arthur Anthony. 1900. *A History of Sanskrit Literature*. New York: D. Appleton and Company.
- Mace, John. 2003. *Persian Grammar: For Reference and Revision*. London-New York: RoutledgeCurzon.
- Maciuszak, Kinga. 2013. Persian Lexicography in South Asia. In: Kuczkiewicz-Fraś 2013. 339-368.
- MacKenzie, David Neil. 1971. *A Concise Pahlavi Dictionary*. London [etc.]: Oxford University Press. [Repr. (with corrections) 1986]
- [Maddāh] = Kham 'Maddāh', M. M. 1959. *Urdū-hindī śabd'koś* [Urdū-Hindī Dictionary]. Lak^hnau: Uttar Pradesh Hindī Samst^han. [8th ed. 1997]
- Maillo Salgado, Felipe. 1991/1998. *Los arabismos del castellano en la Baja Edad Media: consideraciones históricas y filológicas*. Salamanca: Universidad de Salamanca. [1st ed. 1983]
- Maillo Salgado, Felipe. 1996. *Vocabulario de historia árabe e islámica*. Madrid: Akal.
- Martynyuk, Stanislav. 2003. Statistical Approach to the Debate on Urdu and Hindi. *Annual of Urdu Studies*. Vol. 18. 1-13.
- Masica, Colin P. 1991. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- Masica, Colin P. 2005. *Defining A Linguistic Area. South Asia*. New Delhi: DC Publishers.
- Matin, Abdul, P. K. Mathur, and Imtiaz Hasnain. 2001. Hindi-Urdu Construct: Analyses of Antagonism. In: Abbi et al. 2001. 197-206.

- Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.
- Mathews, David. 2002. Urdu in India. *Annual of Urdu Studies*. Vol. 17. 155-167.
- Mathews, Peter H. 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford-New York: Oxford University Press.
- McGregor, Ronald Stuart. 1984. *Hindi Literature of the Nineteenth and Early Twentieth Centuries*. A History of Indian Literature. Vol. 8, fasc. 2. Wiesbaden: Otto Harrassowitz.
- [McGr.] = McGregor, Ronald Stuart. 1993. *The Oxford Hindi-English Dictionary*. New Delhi: Oxford University Press. [11th ed. 2003]
- McGregor, Ronald Stuart. 1999. *Outline of Hindi Grammar with Exercises*. Delhi: Oxford University Press. [3rd ed. revised and enlarged]
- McMahon, April M. S. 1994. *Understanding Language Change*. Cambridge: Cambridge University Press.
- [MHK] = Varmmā, Rām'candr. 1962. *Mānak hindī koś* [Standard Hindī Dictionary]. Vol. I-V. Prayāg: Hindī Sāhitya Sammelan. [2nd ed. 1991]
- Mohan, Peggy. 2000. Hindustani, Hindi and English in India. *Economic and Political Weekly*. Vol. 35, No. 19. 1672-1673.
- Monier Williams, Monier, ed. 1859. *Original Papers Illustrating the History of the Application of the Roman Alphabet to the Languages of India*. London: Longman, Brown, Green, Longmans, and Roberts. [Repr. New Delhi: Asian Educational Services, 2003]
- Montaut, Annie. 2004. *A Grammar of Hindi*. München: LINCOM EUROPA.
- Morton, A. H. 2012. The transmission of Persian Texts Compared to the Case of Classical Latin. In: Spooner & Hanaway 2012. 143-158.
- Myers-Scotton, Carol. 2006. *Multiple Voices. An Introduction to Bilingualism*. Malden, MA [etc.]: Blackwell.
- Naim, Choudhri Mohammed. 1999. *Introductory Urdu*. Vols. I-II. South Asia Language and Area Center, University of Chicago.
- Nair, Rami. 2001. Acoustic Correlates of Lexical Stress in Hindi. In: Abbi et al. 2001. 123-143.
- Narang, Gopi Chand, and Donald A. Becker. 1971. Aspiration and Nasalization in the Generative Phonology of Hindi-Urdu. *Language*. Vol. 47, No. 3. 646-667.
- Nayar, Baldev Raj. 1968. Hindi as Link Language. *Economic and Political Weekly*. Vol. 3, No. 6. 297-305.
- Nespital, Helmut. 1994. The Development of Literary Urdu in Delhi in the 17th and 18th Centuries with Regard to Changes of its Language Structure. In: Chitre, Dilip, Günther-Dietz Sontheimer, Heidrun Brückner, Anne Feldhaus, and Rainer Kimmig, eds. *Tender Ironies: A Tribute to Lothar Lutze*. New Delhi: Manohar Publishers. 308-325.
- Ohala, Manjari. 1983. *Aspects of Hindi Phonology*. Delhi: Motilal Banarsidass.

-
- Ohala, Manjari. 1999. Hindi. In: International Phonetic Association (IPA). 1999. 100-103.
- Oksaar, Els. 1996. The History of Contact Linguistics as a Discipline. In: Goebel, Hans et al., eds. *Kontaktlinguistik/contact linguistics/linguistique de contact: ein internationales Handbuch zeitgenössischer Forschung/An International Handbook of Contemporary Research/manuel international des recherches contemporaines*. Berlin/New York: Walter de Gruyter. 1-12.
- [OL] See: Gilchrist 1798.
- Orsini, Francesca. 2002. *The Hindi Public Sphere 1920-1940. Language and Literature in the Age of Nationalism*. New Delhi: Oxford University Press.
- Orsini, Francesca, ed. 2010. *Before the Divide: Hindi and Urdu Literary Culture*. New Delhi: Orient BlackSwan.
- Pandey, Pramod. 1990. Hindi schwa Deletion. *Lingua*. Vol. 82. 277-311.
- Pandey, Pramod. 2007. Phonology-Orthography Interface in Devanāgarī for Hindi. *Written Language & Literacy*. Vol. 10, No. 2. 139-156.
- Paul, Hermann. 1886. *Prinzipien der Sprachgeschichte*. Zweite Auflage. Halle: Max Niemeyer.
- Pattanayak, D. P. 2001. Multilingualism. In: Abbi et al. 2001. 62-69.
- Pelikan, Heike. 1987. Die "Grammatica Hindostanica" des Benjamin Schultze. *Quelle zur Erforschung des Dakhini des 18. Jahrhunderts*. Wissenschaftliche Beiträge 1987/35 (I 41). Halle (Saale): Martin-Luther-Universität Halle-Wittenberg.
- Perry, John R. 1984. Differential Assimilation of some Arabic Loanwords in Tajik and Uzbek. *Folia Slavica*. Vol. 7, No. 1-2. 268-82.
- Perry, John R. [1992]. "Persian-Arabic Bilingualism in the Evolution of New Persian". [Unpublished text to appear in: *Proceedings of the International Symposium on Bilingualism in the Iranian World* (Bamberg, July 1992); available online: http://nelc.uchicago.edu/sites/nelc.uchicago.edu/files/Perry_1992paper%20Pers-Arab%20biling.pdf].
- Perry, John R. 1995. Lexical Doublets as a Derivational Device in Persian: The Arabic Feminine Ending. *Acta Orientalia Academiae Scientiarum Hungarica*. Vol. XLVIII. 127-53.
- Perry, John R. 2001a. Introduction. *Asian Folklore Studies*. Vol. 60, No. 2. 191-202.
- Perry, John R. 2001b. The Historical Role of Turkish in Relation to Persian of Iran. Asatryan, Garnik, ed. *Iran & the Caucasus*. Research Papers from the Caucasian Centre for Iranian Studies, Yerevan. Vol. 5. Tehran: International Publications of Iranian Studies. 193-200.
- Perry, John R. 2002. Arabic Elements in Persian. *Encyclopaedia Iranica*. Vol. II. 229-236. [Last updated 2011]
- Perry, John R. 2005. Lexical Areas and Semantic Fields of Arabic Loanwords in Persian and Beyond. In: Csató, Éva Ágnes, Bo Isaksson, and Carina Jahani,

-
- eds. *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*. London and New York: RoutledgeCurzon. 97-109.
- Perry, John R. 2007. Persian Morphology. In: Kaye, Alan S., ed. *Morphologies of Asia and Africa*. Vol. I. Winona Lake, Ind.: Eisenbrauns.
- Perry, John R. 2012. New Persian: Expansion, Standardization, and Inclusivity. In: Spooner & Hanaway 2012. 70-94.
- Phillot, Douglas C. 1919. *Higher Persian Grammar for the Use of the Calcutta University, Showing Differences between Afghan and Modern Persian, with Notes on Rhetoric*. Calcutta: Calcutta University.
- Pisowicz, Andrzej. 1985. *Origins of the New and Middle Persian Phonological Systems*. Rozprawy Habilitacyjne nr 101. Kraków: Uniwersytet Jagielloński.
- [Platts] = Platts, John. 1884. *A Dictionary of Urdū, Classical Hindī and English*. London: Wm. H. Allen, & Co. [Repr. New Delhi: Manohar, 2000]
- Platts, John. 1909. *A Grammar of the Hindūstānī or Urdū Language*. London: Crosby Lockwood and Son. [Repr. of the fifth ed., Lahore: Sang-e-Meel Publications, 2002; 1st ed. 1870]
- Pokorny, Julius. 1959. *Indogermanisches Etymologisches Wörterbuch*. Vol. 1-2. Bern & München: Francke Verlag.
- Pollock, Sheldon, ed. 2003. *Literary Cultures in History. Reconstructions from South Asia*. Berkeley [etc.]: University of California Press.
- Poplack, Shana, and David Sankoff. 1988. Code-switching. In: Ammon, Ulrich, Norbert Dittmar and Klaus J. Mattheier, (eds.). *Sociolinguistics. An International Handbook of the Science of Language and Society*. Vol. 2. Berlin: Walter de Gruyter. 1174-1180.
- Pridmore, Fred. 1975. *The Coins of the British Commonwealth of Nations. Part 4 India. Volume 1: East India Company Presidency Series c. 1642-1835*. London: Spink & Son.
- Pritchett, Frances W., and Khaliq Ahmad Khaliq. 1987. *Urdu Meter. A Practical Handbook*. Madison: South Asian Studies Publication Series, University of Wisconsin-Madison. [Online edition revised 2003-2012: <http://www.columbia.edu/itc/mealac/pritchett/00ghalib/meterbk/>]
- Qureshī, Waheed. 1965. The Indian Persian. In: Dil, Anwar S. *Studies in Pakistani Linguistics*. Lahore: Linguistic Research Group of Pakistan. 39-52.
- Qutbuddin, Tahera. 2007. Arabic in India: A Survey and Classification of Its Uses, Compared with Persian. *Journal of the American Oriental Society*. Vol. 127, No. 3. 315-338.
- Rahman, Tariq. 2010. Urdu and the Muslim Identity: Standardization of Urdu in the Eighteenth and Early Nineteenth Centuries. *Annual of Urdu Studies*. Vol. 25. 83-107.
- Rahman, Tariq. 2011. *From Hindi to Urdu: A Social and Political History*. Karachi: Oxford University Press.

-
- Rahman, Tariq. 2013. The Learning of Persian in South Asia: the Curricula and Educational Institutions for Teaching Persian in South Asia. In: Kuczkiewicz-Fraś. 2013. 369-390.
- Rai, Alok. 2005. The Persistence of Hindustani. *Annual of Urdu Studies*. Vol. 20. 135-144.
- Rai, Alok. 2007. *Hindi Nationalism*. New Delhi: Orient Longman. [1st ed. 2001]
- Rai, Amrit. 1984. *A House Divided. The Origin and Development of Hindi/Hindavi*. Delhi [etc.]: Oxford University Press.
- Rao, P. Kodanda. 1968. Hindi as Link Language: A Comment. *Economic and Political Weekly*. Vol. 3, No. 13. 531-532.
- Rao, Vijayendra K. R. V. 1978. Many Languages, One Nation: Quest for an All-India Language. *Economic and Political Weekly*. Vol. 13, No. 25. 1025-1030.
- Ritter, Robert M., ed. 2003. *The Oxford Style Manual*. Oxford: Oxford University Press.
- Ritter, Valerie. 2004. The Language of Hariaudh's "Priyapravās": Notes toward an Archaeology of Modern Standard Hindi. *Journal of the American Oriental Society*, Vol. 124, No. 3 (July-September). 417-438.
- Rizvi, Saiyid Athar Abbas. 2000. *The Wonder that Was India. Volume II: A Survey of the History and Culture of the Indian Sub-continent from the Coming of the Muslims to the British Conquest 1200-1700*. New Delhi: Rupa & Co. [1st ed. 1987]
- Robinson, Francis. 1991. Perso-Islamic Culture in India from Seventeenth to the Early Twentieth Century. In: Canfield, Robert L., ed. 1991. *Turco-Persia in Historical Perspective*. Cambridge [etc.]: Cambridge University Press. 104-131.
- Rohany Rahbar, Elham. 2012. *Aspects of Persian Phonology and Morphophonology*. A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy. Toronto: Department of Linguistics, University of Toronto.
- Romaine, Suzanne. 1995. *Bilingualism*. Oxford: Blackwell Publishers. [1st ed. 1989]
- Rosen, Fritz [Friedrich]. 1898. *Modern Persian Colloquial Grammar, Containing a Short Grammar, Dialogues and Extracts from Nasir-Eddin Shahs Diaries, Tales, etc., and a Vocabulary*. London: Luzac & Co.
- [Rub.] = Rubinchik, Yuriĭ A., ed. 1983. *Persidsko-russkii slovar'*. Vol. I-II. Moskva: Russkii Yazyk. [2nd ed.]
- Russell, Ralph. 1980. *A New Course in Hindustani for Learners in Britain*. London: SOAS. [2nd ed. titled *A New course in Urdu and Spoken Hindi: for Learners in Britain*, 1986].
- Russell, Ralph. 1996. Some Notes on Hindi and Urdu. *Annual of Urdu Studies*. Vol. 11. 203-208.

- Russell, Ralph. 1999. Urdu in India since Independence. *Economic and Political Weekly*. Vol. 34, No. 1/2. 44-48.
- Rypka, Jan. 1968. History of Persian Literature up to the Beginning of the 20th Century. In: Jahn, Karl, ed. *History of Iranian Literature*. Dordrecht: D. Reidel Publishing Company.
- Šadrī Afšār, Ġulāmhūsayn, Nasrīn Ḥakamī, and Nastaran Ḥakamī. 1993. *Farhang-e mu‘āšir-e fārsī*. Tehrān: Kalamah Publishers [1414].
- Sakel, Janette. 2007. Types of Loan: Matter and Pattern. In: Matras & Sakel 2007. 16-29.
- Salomon, Richard. 2003. Writing Systems of the Indo-Aryan Languages. In: Cardona & Jain 2003. 67-103.
- Samiuddin, Abida, ed. 2007. *Encyclopaedic Dictionary of Urdu Literature*. Vol. 1. New Delhi: Global Vision Publishing House.
- [SHŚ] = *Samkṣipt hindī śabd’sāgar*. 1929. Vāraṇasī: Nāgarīpracārīṇī sabḥā. [11th ed., Vāraṇasī-Naī Dillī, 1998]
- Sankoff, Gillian. 1980. *The Social Life of Language*. Philadelphia: University of Pennsylvania Press.
- Sankoff, Gillian. 2002. Linguistic Outcomes of Language Contact. In: Chambers John K., Peter Trudgill, and Natalie Schilling-Estes, eds., *The Handbook of Language Variation and Change*. Oxford: Blackwell. 638-668.
- Sapir, Edward. 1921. *Language. An Introduction to the Study of Speech*. New York: Harcourt, Brace.
- Saran, Saraswati. 1969. *India's National Writing*. Delhi: Munshiram Manoharlal.
- Schimmel, Annemarie. 1973. *Islamic Literatures of India*. A History of Indian Literature. Vol. 8, fasc. 1. Wiesbaden: Otto Harrassowitz.
- Schimmel, Annemarie. 1975. *Classical Urdu Literature from the Beginning to Iqbāl*. A History of Indian Literature. Vol. 8, fasc. 3. Wiesbaden: Otto Harrassowitz.
- Schimmel, Annemarie. 2007. *Mystical Dimensions of Islam*. New Delhi: Yoda Press. [1st ed. 1975]
- Schmidt, Ruth Laila. 1999. *Urdu. An Essential Grammar*. New York: Routledge.
- Schmidt, Ruth Laila. 2003. Urdu. In: Cardona & Jain 2003. 286-350.
- Schultze, Benjamin. 1745. *Grammatica Hindostanica collectis in diuturna Inter Hindostanos commoratione in justum ordinem redactis Ac larga exemplorum luce parfusus regulis constans et missionariorum usui consecrata*. Halae Saxonum: In Typographia Instituti Judaici.
- Sebba, Mark, Shahrzad Mahootian, and Carla Johnsson, eds. 2012. *Language Mixing and Code-Switching in Writing. Approaches to Mixed-Language Written Discourse*. New York-Abington: Routledge.

- Shackle, Christopher. 1978. Approaches to the Persian Loans in the "Ādi Granth". *Bulletin of the School of Oriental and African Studies*. Vol. 41, No. 1. 73-96.
- Shackle, Christopher, and Rupert Snell. 1990. *Hindi and Urdu Since 1800: A Common Reader*. London: School of Oriental and African Studies, University of London.
- Shahabuddin, Syed. 2001. Urdu in India. *Economic and Political Weekly*. Vol. 36, No. 24. 2098.
- Shahabuddin, Syed. 2006. A Trinity without a Church: Urdu Language, Urdu Education in India, and Muslim Indians. In: Farouqui, Ather, ed. 2006. *Redefining Urdu Politics in India*. New Delhi: Oxford University Press. 165-192.
- [Shak.] = Shakespear, John. 1834. *A Dictionary, Hindūstānī and English: with a Copious Index, Fitting the Work to Serve, also, as a Dictionary of English and Hindūstānī*, 3rd ed. much enlarged. London: J. L. Cox and Son. [Repr. Lahore: Sang-e-meel, 2000]
- [Shak.A] = Shakespear, John. 1834. Additional of Words, Phrases and Meanings to *A Dictionary, Hindūstānī and English...* London: J. L. Cox and Son. [Repr. Lahore: Sang-e-meel, 2000]
- Shapiro, Michael C. 1989. *A Primer of Modern Standard Hindi*. Delhi: Motilal Banarsidass.
- Shapiro, Michael C. 2003. Hindi. In: Cardona & Jain 2003. 250-285.
- Shapiro, Michael C., and Harold F. Schiffman. 1981. *Language and Society in South Asia*. MLBD Series in Linguistics 1. Delhi: Motilal Banarsidass.
- Sharar, Abdul Halim. 1975. *Lucknow: The Last Phase of Oriental Culture*. Transl. and ed. by E. S. Harcourt and Fakhir Hussain. London: Paul Elek.
- Sharma, V. P. 2006. *The Rupa's English-Hindi Dictionary: Aṅgrezī-Hindī Śabd'koś*. New Delhi: Rupa & Co.
- Sherwani, Haroon Khan. 1985. *The Bahmanis of the Deccan*. New Delhi: Munshiram Manoharlal Publishers.
- Shukla, Shaligram. 2000. *Hindi Phonology*. LINCOM Studies in Indo-European Linguistics, 12. München: LINCOM EUROPA.
- Siddiqi, M. Atique. 1963. *Origins of Modern Hindustani Literature. Source Material: Gilchrist Letters*. Aligarh: Naya Kitab Ghar.
- Singh, Rajendra. 1995. *Linguistic Theory, Language Contact, and Modern Hindustani: The Three Sides of Linguistic Story*. American University Studies. Series XIII: Linguistics. Vol. 31. New York [etc.]: Peter Lang.
- Singh, Rajendra, and Rama Kant Agnihotri. 1997. *Hindi Morphology. A Word-Based Description*. Delhi: Motilal Banarsidass.
- Skaljic, Abdulah. 1966. *Turcizmi u Srpsko-Hrvatskom Jeziku*. Sarajevo: „Svjetlost” Izdavačko Preduzeće.
- Sommerstein, Alan H. 1977. *Modern Phonology*. London: Edward Arnold.
- Spear, Percival. 1978. *A History of India. Volume Two: From the Sixteenth Century to the Twentieth Century*. New Delhi: Penguin Books. [1st ed. 1965]

- Spies, Otto von. 1955. Türkisches Sprachgut im Hindüstānī. In: Spies, Otto, ed. *Studia Indologica. Festschrift für Willibald Kirfel zur Vollendung seines 70. Lebensjahres*. Bonner Orientalistische Studien Neue Serie 3. 321-343.
- Spies, Otto von, and Ernst Bannerth. 1945. *Lehrbuch der Hindüstānī-Sprache*. Leipzig [etc.]: Harrassowitz.
- Spooner, Brian, and William L. Hanaway, eds. 2012. *Literacy in the Persianate World: Writing and the Social Order*. Philadelphia: University of Pennsylvania Museum of Archaeology and Anthropology.
- Srivastava, Ravindra Nath, rev. 1969. Studies in Hindi and Urdu, I: Introduction and Word Phonology by Ashok R. Kelkar. *Language*. Vol. 45, No. 4. 913-927.
- Srivastava, Sushil. 2000. The Farce That is Hindi. *Economic and Political Weekly*. Vol. 35, No. 43/44. 3898-3899.
- Staal, Johan Frederik. 1967. *Word Order in Sanskrit and Universal Grammar*. Foundations of Language Supplementary Series 5. D. Reidel Publishing Company/Dordrecht-Holland.
- Staal, Johan Frederik. 1995. The Sanskrit of Science. *Journal of Indian Philosophy*. Vol. 23, No 1. 73-127.
- [STCD] = *Standard Twentieth Century Dictionary Urdū to English*. 1980. Compiled by B. A. Qureshi, revised and enlarged by A. Haq. Delhi: Educational Publishing House. [New Edition, Delhi, 1999]
- Steadman-Jones, Richard. 2007. *Colonialism and Grammatical Representation: John Gilchrist and the Analysis of the 'Hindustani' Language in the Late Eighteen and Early Nineteenth Centuries*. Publications of the Philological Society, 41. Oxford: Blackwell Publishing.
- [Stein.] = Steingass, Francis J. 1892. *A Comprehensive Persian-English Dictionary*. London: W. H. Allen. [Repr. Delhi: Manohar, 2006]
- Subbārāo, Kārumūri V. 2012. *South Asian Languages. A Syntactic Typology*. Cambridge: Cambridge University Press.
- Swadesh, Morris. 2006. *The Origin and Diversification of Language*. Ed. by Joel F. Sherzer. With a foreword and appendix by Dell Hymes. New Brunswick, NJ: Transaction Publishers. [1st ed. 1971]
- Syed, Muhammad Aslam. 2012. How Could Urdu Be the Envy of Persian (*rashk-i Fārsi*)! The Role of Persian in South Asian Culture and Literature. In: Spooner & Hanaway 2012. 279-310.
- Tadmor, Uri. 2009. Loanwords in the World's Languages: Findings and Results. In: Haspelmath & Tadmor 2009. 55-75.
- Tassy, Garçin de. 1875. *La langue et la littérature hindoustaniens en 1874: Revue annuelle*. Paris: Librairie orientale de Maisonneuve.
- Tassy, Garçin de. 2011. Hindustani Language and Literature. Tr. By S. Kamal Abdali. *Annual of Urdu Studies*. Vol. 26. 135-166.
- Teufik, Muftić. 2000. On Arabic Loanwords in the Serbo-Croat Language. *Prilozi za orijentalnu filologiju*. Vol. 50. 21-46.

- Thomason, Sarah Grey. 2001. *Language Contact. An Introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah Grey. 2006. Arabic in Contact with Other Languages. In: Versteegh, Kees et al., eds. *Encyclopedia of Arabic Language and Linguistics*. Vol. 2. Leiden: Brill.
- Thomason, Sarah Grey, and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley-Los Angeles-Oxford: University of California Press.
- Tietze, Andreas, and Gilbert Lazard. 1967. Persian Loanwords in Anatolian Turkish. *Oriens*. Vol. 20. 125-168.
- Tivārī, Mohan'lāl. 1969 [Sam. 2026]. *Hindī b'hāṣā par fār'sī aur aṅgrezī kā prab'āv*. Vārāṇasī: Nagarīpracārīṇī Sab'hā.
- Tristram, Hildegard. 2007. Why Don't the English Speak Welsh? In: Higham, Nicolas J., ed. *The Britons in Anglo-Saxon England*. Woodbridge: The Boydell Press. 192-214.
- Trivedi, Harish. 2004. The Anxiety of Hindustani. *The Yearbook of South Asian Languages and Linguistics*. Vol. 7. 267-276.
- Trivedi, Harish. 2013. Muslims and Hindus: Urdu and Hindi. In: Kuczkiewicz-Fraś 2013. 113-246.
- Trubetzkoy, Nikolai S. 1928. Proposition 16. *Actes de 1^{er} Congrès international de linguistes*. Vol. 17-18. Leiden: A. W. Sijthaff's Uitgevermaatschappij.
- Truschke, Audrey Angeline. 2012. *Cosmopolitan Encounters. Sanskrit and Persian at the Mughal Court*. Dissertation submitted in partial fulfillment of the Requirements for the degree of Doctor of Philosophy in the Graduate School of Arts and Sciences, Columbia University. [Available online: <http://hdl.handle.net/10022/AC:P:12951>]
- T'sou, Benjamin K. 1975. On the Linguistic Covariants of Cultural Assimilation. *Anthropological Linguistics*. Vol. 17, No. 9. 445-465.
- Turek, Wacław P. 2001. *Słownik zapożyczeń z języka arabskiego w polszczyźnie*. Kraków: Universitas.
- Turner, Ralph Lilley. 1962-66. *A Comparative Dictionary of Indo-Aryan Languages*. London: Oxford University Press.
- Utas, Bo. 2006. A Multiethnic Origin of New Persian? In: Johanson & Bulut 2006. 241-251.
- Van der Veer, Peter. 1994. *Religious Nationalism: Hindus and Muslims in India*. Berkeley: University of California Press.
- Van Olphen, Herman H. 1988. Religion and Language Varieties: the Case of Hindi-Urdu. In: Jazayery, Mohammad Ali, and Werner Winter, eds. *Languages and Cultures: Studies in Honor of Edgar C. Polomé*. Trends in Linguistics, Studies & Monographs 36. New York: Mouton de Gruyter. 741-747.
- Van Olphen, Herman H. 1994. Lexical Convergence in Urdu and Hindi. In: Marashi, Mehdi, ed. *Persian Studies in North America. Studies in Honor of Mohammad Ali Jazayery*. Bethesda, Maryland: Iranbooks. 141-149.

- Varmā, Dh̥tendr. 1933 [1995]. *Hindī b'hāṣā kā itiḥās*. Ilāhābād: Hindustānī Ekeḍemī.
- Vogel, Jean Philippe. 1936. Joan Josua Ketelaar of Elbing, Author of the First Hindūstānī Grammar. *Bulletin of the School of Oriental Studies, University of London*, Vol. 8, No. 2/3: Indian and Iranian Studies: Presented to George Abraham Grierson on His Eighty-Fifth Birthday, 7th January, 1936. 817-822.
- Watson, Janet C. E. 2007. *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press.
- Weinreich, Uriel. 1953. *Languages in Contact: Findings and Problems*. The Hague [etc.]: Mouton Publishers. [Repr. 1979]
- Weinreich, Uriel. 1958. On the Compatibility of Genetic Relationship and Convergent Development. *Word*. Vol. 14. 374-379.
- Wexler, Paul. 2009. Towards a Structural Definition of 'Internationalisms'. *Linguistics*. Vol. 7, No. 48. 77-92.
- Windfuhr, Gernot L. 1979. *Persian Grammar. History and State of its Study*. Trends in Linguistics. State-of-the-Art Report 12. The Hague: Mouton.
- Windfuhr, Gernot L. 1997. Persian Phonology. In: Kaye, Alan. S., ed. *Phonologies of Asia and Africa (Including the Caucasus)*. Vol. II. Winona Lake, Indiana: Eisenbrauns. 675-689.
- Windmiller, Marshall. 1954. Linguistic Regionalism in India. *Pacific Affairs*. Vol. 27, No. 4. 291-318.
- Wink, André. 1999. *Al-Hind. The Making of the Indo-Islamic World*. Vol. I: *Early Medieval India and the Expansion of Islam 7th-11th Centuries*. Vol. II: *The Slave Kings and the Islamic Conquest 11th-13th Centuries*. New Delhi: Oxford University Press.
- Yule, Henry, and Arthur C. Burnell. 1996. *Hobson-Jobson. The Anglo-Indian Dictionary*. Ware, Hertfordshire: Wordsworth Editions Ltd. [1st ed. 1886]
- Zaidi, Ali Jawad. 1993. *A History of Urdu Literature*. New Delhi: Sahitya Akademi.
- Zilli, Ishtiyāq Ahmad. 2000. Development of *Inshā* Literature till the End of Akbar's Reign. In: Alam, Muzaffar, Francoise Nalini Delvoye, Marc Gaborieau, eds. *The Making of Indo-Persian Culture: Indian and French Studies*. New Delhi: Manohar Publishers. 309-350.
- Zavos, John. 2000. *The Emergence of Hindu Nationalism in India*. New Delhi: Oxford University Press.
- Zograf, Georgii Aleksandrovich. 1960. Iranskie i arabskie élementy v urdu. In: Beskrovnýi 1960b. 152-244.
- Zograf, Georgii Aleksandrovich. 1961. *Khindustani na rubezhe XVIII-XIX vv.* Moskva: Izdatel'stvo vostochnoi literatury.

Index

- adaptation VIII, 1, 4, 60, 82, 140, 228
- addition 59, 64, 65, 82-83, 166, 169, 183
- adjectives 153, 187, 190, 191-192, 197, 201, 209, 211; comparison 215
- adjectivization 191
- adoption VIII, 4, 82, 227
- adverbs 187, 192, 194, 197, 203; comparison 215; numeral a. 196; of certainty 194; of degree 194; of frequency 193; of manner 192; of place 193; of time 193
- adverbialization 114
- alif-e kūāh* 62, 105
- 'alif maqṣūrah* XIII, 62, 79, 103, 105, 116, 162, 238
- allographs, allography 76, 77, 138, 165
- Amīr Xusrau Dihlavī 41
- anaptyxis 166, 167, 169
- anusvāra* 66, 72, 92, 93
- apocope 169
- apheresis 192
- approximation 176; phonetic 139; phonological 139
- Arabic script: adaptation in Persian 60-63, 123; used by other languages 59, 63
- assimilation 36, 111, 141-143, 147, 161, 207, 233, 244-247
- Avadhī 32, 52
- bahuvrihi* 223
- Balochi 63
- barī he (hā'ī hoṭṭī)* 91
- barī ye (yā-e ma'kūsī)* 76-77
- Bhojpurī 32, 52
- bilingualism 27, 30, 34, 36, 40-41, 45, 49, 51, 55, 60, 89, 140, 225, 233, 240, 241
- Brahui 63
- Braj b^hāṣā 10, 32, 52, 89, 179
- calque (loan translation) 219, 227-228, 231-232
- Cand Bar'dāī 41, 89
- candrabindu* 66, 93
- case 192, 207-209, 212, 217
- casuals (Gastworts) 51, 245
- č^hoṭī he (hā'ī havvaz)* 68, 72, 91, 103-104, 107, 116
- č^hoṭī ye* 76, 77
- code-switching 30, 36, 53
- College of Fort William 16, 19, 26
- complementizer 200
- compounds 19, 88, 137, 195, 198, 206, 215, 220, 232, 234; copulative 189; determinative 189; possessive/attribution 189, 223
- conjunctions 187, 200, 219-220; coordinating 197-198; subordinating 197, 199-200
- consonants: aspirated 68-71, 81, 136; clusters 63, 66, 69-71, 75, 92-94, 97-101, 144, 150, 166; conjunct s. ligatures; nasal 66, 72-73, 124; retroflex 65, 68-69, 134, 144; sibilants 134, 138, 179
- convergence area 5
- copula 188, 220-221
- core borrowings 241
- core vocabulary 246-247
- correlative clause 199
- cultural borrowings 42, 241
- Dakkani (Dakk^hinī, Dakhni) 11, 14, 17, 46, 216
- Dari 57, 89, 136, 150, 153
- debuccalization 179, 181
- defricativization 146, 176-177
- depalatalization 177, 179
- devoicing 178
- diphthongization 156, 158

- doublets 61, 229-231, 247
do-čašmī he (he-ye do čašm) 68
 dual 138, 218
 elision 169, 181-183
 epenthesis 83, 166-169
 euphonic vowel 166-169, 215
 Farsi 53, 57, 125, 131
fathah 58, 76, 102
 fluctuation (of vowels) 141
 fortition 146, 177
 gender 23, 138, 188, 192, 207-214, 230
 glide 79-80, 85, 104, 111, 129, 134, 160, 168, 182-183
 grammaticalization 203
 Gujarati 11, 51, 59
 Gujri 11
 Guru Nānak 52
hā-ī havvaz s. čhoī he
hā-ī hoīī s. barī he
hamzah XIII, 57, 60-61, 79-80, 85, 111, 134, 141, 160, 210
 haplology 136
harakāt 58
 hendiadys 198
 heterographs 87, 91, 96, 113
 heterophony, heterophones 76-78, 108, 110
he-ye do čašm s. do-čašmī he
hiatus 80, 84, 85, 104, 183
 Hindavi 41, 43, 48
 Hindi VII, 4, 6-11, 14-30, 32-33, 41, 47, 64-65, 114, 142, 152, 226
 Hindūsthānī 16
 Hirdu VII, 19, 24, 30
 homography, homographs 77-78, 80, 87, 113
 homophony, homophones 86-87, 95, 113
 interference 3, 34, 36, 40, 48, 55, 238
 interjections 187, 200-202, 215
izāfar 104, 138, 162, 172, 188, 215, 218, 221-225
jazm 58
 juxtaposition 103-107, 109-110, 115, 223
 Kabīr 52, 243
 Kashmiri 51, 59, 63
kasrah 58, 76, 104, 107, 172, 223
Kauṣītaki Brāhmaṇa 11
 K^haṛī bolī 9, 10, 12, 20, 28
kursī 60, 79-80
 labialization 111, 149
 language: contact 34-37, 40-41, 43, 48, 49-50, 140, 187, 225; interference s. interference
 lenition 147, 178, 181
 ligatures (conjunct cons.) 66, 69-71, 88, 95, 97-101, 106, 113-114, 167, 172
lingua franca VIII, 7, 12, 16-18, 22-23, 29, 41, 180
 linguistic area 5, 27
 linguistic borrowing 3-4, 34-36, 40, 42, 52, 228, 242
 loanblend 219, 231
 loan shift 227
 loan translation s. calque
madd 76, 103, 116
Mahāb^hārat 10
Mahāb^hāṣya 12
 Marathi 23, 41, 51, 64
 Mas'ūd Sa'd Salmān 41
 metanalysis 75
 metaplasm 82
 monophthongization 124, 130-131, 133, 136, 155
 multilingualism s. bilingualism
 Nagari VIII, XI, XIII-XIV, 4, 16, 18, 21, 26-27, 30, 33, 56, 64-75, 82, 84, 86, 89-120, 125, 138, 142, 147-148, 154, 166-167, 169, 172, 174-176, 179, 186, 198, 218, 221, 223, 239, 245, 248
 Nāmdev 52
 nasalization 68, 72-73, 84, 92-93, 150-151, 163, 183-184
nasta'liq 63, 68, 72, 88, 152, 237
nasx 63, 237
 nativization 3, 81-82, 140-141, 185, 231
 Nepali 64

- nouns 23, 153, 166, 187, 190-192, 201, 203, 205-213, 216-217, 222, 230
- number 23, 138, 188-189, 192, 195-196, 207, 209, 212
- numerals 87, 187, 195; cardinal 195; collective 196; distributive 196; fractional 195-196; ordinal 195
- nunation s. *tanvīn*
- nūn-e gunnah* 72-73, 152
- nuqṭa* 94
- omission 62-63, 82-83
- palatalization 151, 179
- Pañcaviṃśa Brāhmaṇa* 11
- Pāṇini 12
- Panjabi 32, 43-44, 51, 59, 63
- paragoge 169
- Pashto 12, 43-44, 51, 59, 63, 144
- past stem (of Persian verb) 204
- Patañjali 12
- pīṣ* s. *ṣammah*
- plural 188, 207, 209-210, 216, 218, 246; broken 217; sound 217, 246
- postpositions 24, 74, 137, 187-188, 203, 207-208; complex p. 203, 219, 224-225
- Prācya 11
- prepositions 137, 187-189, 192, 202-203; compound p. 219; denominative p. 224
- present stem (of Persian verb) 204
- pronouns 187, 197, 224
- Ṗṛt'vīrāj-rasau* 41, 89
- sabk-i hindī* 52-53
- Sanskrit 10, 12, 18-19, 23-24, 26, 52, 64-66, 92, 104-105, 114, 122, 133, 146, 148, 150, 187-188, 215, 219, 231-232, 238
- Saraiki 59
- schwa (neutral vowel) 65, 169; deletion (schwa syncope, schwa-dropping) 66, 83, 163, 166, 169-170, 173
- segmentalization 166
- Sindhi 43-44, 51, 59, 63
- spacing 74, 87-88
- Sprachbund 5
- substantivization 191, 201
- substitution 36, 82, 84, 140, 219
- Sufi 44-45, 47, 52, 237
- sukūn* 58, 113, 166
- 'sun' letters 221
- syncope 169, 171-172
- Śāraṣenī Apabhraṃś 12
- tā' marbūṭah* 61, 79, 84, 107, 114, 182, 210, 230, 238
- tadb'ava* 231, 246
- Tajiki 57, 121, 136
- tanvīn* (nunation) 58, 114
- tašdīd* 58, 83, 85, 102, 108, 110, 114, 174
- tatpuruṣa* 223
- tatsama* 66, 146, 150, 215, 219-220, 231-232, 246
- transposition 82, 84
- Tulsīdās 243
- Ṭūsī, Asadī 59
- Udīcya 11
- Urdu VII, 4-8, 10, 15-33, 44, 50, 59, 63-64, 123, 140-142, 144, 152, 180, 225-226, 234-235, 248
- verbal adjectives 190-191, 203
- verbal nouns 190, 203
- verbs 23, 106, 187, 189, 203, 204, 207; compound v. 189, 231, 247; conjunct v. 189, 203, 231, 247
- virāma* 66, 167
- visarga* 66
- vocative 188, 215
- voicing 144, 178
- vowel clusters 79-80, 102-107, 109-110, 156
- Xāliq bārī* 41
- yā-e ma'kūsī* s. *barī ye*
- yā-ma'ruṣ* 81
- yā-majhūl* 76, 81, 125
- zabar* s. *fatḥah*
- ṣammah* 58, 76, 106
- zer* s. *kasrah*

The detailed linguistic analysis is embedded in a broader socio-cultural and historical context of the Indian subcontinent (...) The book thus fills an important, yet until now widely neglected, gap in the history of Indo-Persian contacts and, without a doubt, it will be a precious source of information and reference both for the specialists in the field of Indo-Aryan and Iranian linguistics and for all those people who want to understand how different traditions, so fully represented in language, intertwine with one another.

Kinga Maciuszak, Jagiellonian University in Krakow

* * *

One and yet not one, two and yet not two — the complex relationship between Hindi and Urdu, as well as with that problematic entity labelled Hindustani, has often been commented upon, but never systematically studied on the basis of a single characteristic exhaustively followed through in all its ramifications. This is a lacuna which this study adroitly fills to a marked extent.

Rahul Peter Das, Martin Luther University of Halle-Wittenberg

* * *

Agnieszka Kuczkiewicz-Fraś, Ph.D. (2001) in Sociolinguistics, lectures at the Jagiellonian University in Krakow, Poland. Her main field of research covers Hindi-Urdu Linguistics as well as the Islamicate traditions of South Asia and Intercultural Communication. She also translates from Hindi and Urdu into Polish. Her main publications include: *Perso-Arabic Hybrids in Hindi. The Sociolinguistic and Structural Analysis* (New Delhi: Manohar, 2003); *Perso-Arabic Loanwords in Hindustani. Part I: Dictionary* (Krakow: Księgarnia Akademicka, 2008); *Umrão Džan Ada. Pamiętnik kurtyzany* (translation, edition & introduction, Krakow: Księgarnia Akademicka, 2011); *Islamicate Traditions in South Asia: Themes from Culture and History* (ed., New Delhi: Manohar, 2013). At present she is preparing a comprehensive study on Indo-Muslim courtesans (*tawa'if*) and their socio-cultural role in the Islamicate society of North India.



www.akademicka.pl

