Pre-nasal Vowel Raising in Tehrani Persian

by

Koorosh Ariyaee

A thesis submitted in conformity with the requirements
for the degree of Master of Arts

Department of Linguistics
University of Toronto

Pre-nasal Vowel Raising in Tehrani Persian

Koorosh Ariyaee

Master of Arts

Department of Linguistics
University of Toronto

2018

Abstract

This study is an attempt to provide an account of a phenomenon where pre-nasal /ɑ/ raises to [u] in Tehrani Persian. Importantly, /ɑ/ does not raise in all pre-nasal environments. Hence, this research tries to explain why pre-nasal raising does not always occur. In the literature, this phenomenon is investigated from various historical (Sadeghi 2001, Miller 2011), sociolinguistic (Modaressi 1978, Jahangiri 1980) and phonological (Kahn and Bernstein 1981, Kalbasi 2001, Bakhtiari 2008, Rees 2008 and Rohany 2012) viewpoints in order to identify the factors which influence raising. However, the common problem with their analyses is that a part of the data remains unanalyzed or their generalizations encounter counterexamples. In this study, I show that the reason for these problematic analyses is that pre-nasal raising is a multidimensional phenomenon; that is, there are phonological, morphological and lexical factors that block raising in different lexical items. As a result of this multidimensional analysis, a unified account is presented that can cover both the generalizations and the exceptions of the former studies.

Keywords: pre-nasal vowel raising, Tehrani Persian, formal Persian

Dedication

To Sahar

Acknowledgement

It did not seem to be such a difficult task to acknowledge people who have aided me to write the thesis. Now that I am writing these lines, I see how tough it is to express my deep gratitude. I thank my supervisor, Keren Rice! During the intense one-year MA program, I could not wait for my weekly meetings with Keren. The meetings nourished and developed my thesis in its best way, and at the same time filtered out stress, anxiety and disappointment. Her untiring and patient revisions, her meticulous corrections and simultaneously comprehensive comments made this draft possible. She was always one step ahead of me, enlightening my path and directing me on the right track. I could not ask for more!

I am also grateful to my second reader, Arsalan, whom I know before I joined U of T linguistics department. It is a great privilege for me to have meetings with him which have been highly constructive. He always viewed the data from an angle which was unique and creative. His insightful comments on my thesis have enhanced the quality of my work to a great extent.

I would like to extend my gratitude to the linguistics department of U of T. I have never felt alone at the department, as its members have built a friendly and attentive atmosphere. It is a great pleasure for me to be part of this fantastic department.

I also would like to thank my family members who did their best to keep my life as peaceful as possible so that I could concentrate on my studies. Even though I am thousands of miles away from them, their warm virtual presence never let me feel lonely or dejected.

I am also grateful to my consultants for their constant help in verifying the data. Without their help I could not develop the database.

Last but not least, my profound gratefulness goes to my soulmate, Sahar. Without her selfless devotion and passionate support, I was least likely to pass this year successfully. Our

linguistic and non-linguistic discussions are always great sources of inspiration and hope for me. She never gave up on me when I felt collapsed; that is how I could make it!

Table of Contents

CHAPTER ONE

INTRODUCTION

1.1. Introduction

Persian, a southwestern Iranian language (Windfuhr 2009:2), is spoken in Iran, Afghanistan and Tajikistan. According to Modaressi (1978:9), there are three major dialects, Tehran Persian (Iran), Kabul Persian (Afghanistan) and Tajik Persian (Tajikistan). In this study, the focus is on Tehran (Tehrani) Persian[1].

1.2. Formal Persian and Tehrani Persian

Following Modaressi (1978), Tehrani Persian is close to the language of literature, instruction and broadcasting. It is the most prestigious variety/accent spoken in Iran and has significant influence on other varieties/accents of Persian.

Tehrani Persian (hereafter TP) is used in daily spoken communications, in informal audio/video/text message exchanges and in media (except news). Despite its wide and broadening usage, Tehrani Persian is not used all the time. For instance, in administrative texts, literary and academic books and writings, news and, generally speaking, in official contexts "formal Persian" is used[2].

[1] Standard (spoken) Persian, informal Persian and vernacular (colloquial) Persian are other names used for this variety of Persian in the literature. I use the cover term Tehrani Persian.

[2] For discussion of whether contemporary Persian is diglossic, see Perry (2003) and Perry et al. (2012).

1.3. Goal of the study

There are lexical, phonological, morphological and syntactic differences between TP and formal Persian (FP hereafter)[3]. One of these differences is that in TP *a* raises to *u* when it appears pre-nasally, whereas this pre-nasal *a* > *u* alternation does not occur in formal Persian, as shown in (1):

(1) *pre-nasal [a] raising in TP*

FORMAL PERSIAN	TEHRANI PERSIAN	GLOSS
aram	arum	peaceful
xane	xune	home
penhan	penhun	hidden
badam	badum	almond

1.4. Research question

Critical for this research in TP is that *a* does not always raise to *u* when it appears before a nasal consonant, as shown in (2):

(2) *sample of words without pre-nasal raising*

TEHRANI PERSIAN				GLOSS
/ʃam/	→	[ʃam]	*[ʃum]	supper
/rajgan/	→	[rajgan]	*[rajgum]	free of charges
/damæn/	→	[damæn]	*[dumæn]	skirt
/zanu/	→	[zanu]	*[zunu]	knee
/emkan/	→	[emkan]	*[emkun]	possibility
/ævamel/	→	[ævamel]	*[ævumel]	factors
/tulani/	→	[tulani]	*[tuluni]	lengthy
/edame/	→	[edame]	*[edume]	continuation
/sazman/	→	[sazman]	*[sazmun]	system
/damdar/	→	[damdar]	*[dumdar]	stockbreeder
/mobleman/	→	[mobleman]	*[moblemun]	sofa
/livan/	→	[livan]	*[livun]	glass

[3] For discussion of the differences between formal and Tehrani Persian, see Hodge (1957), Kalbasi (2001) and Windfuhr (2009), among others.

Given these examples, the research question of this study is to identify what factors block pre-nasal *a* from raising.

1.5. Structure of the thesis

In order to identify the factors that block raising, I created a database of 1782 words that contain an *aN* (*N* = *n* and *m*) sequence. In chapter 3, I introduce the sources of my data and the general format of the database. In addition, I demonstrate how the filtering system works, or how one can spot the exceptions and interaction of different factors in the same words, among other functions of this database.

The data in this database reveal that there are different factors which block pre-nasal raising. The identified factors which prevent *a* > *u* alternation are categorized into three classes of phonological, morphological and lexical factors. Each of these factors are investigated in a separate chapter. In chapter 4, I discuss the phonological factors which block pre-nasal raising. I show that the factors which prevent raising in words with an *am* sequence differ from the blocking factor in words which contain an *an* sequence. In chapter 5, I address the morphological factors which influence pre-nasal raising; for instance, I address the blocking of pre-nasal raising in morphologically complex environments as well as the role of inflectional and derivational affixes in pre-nasal raising. In chapter 6, I focus on lexical items in which raising fails to apply even though the morphological and phonological conditions are met. Chapter 7 addresses the interaction of blocking factors which coincide in the same word. In addition, in this chapter I present a quantified description of the database and show what percentage of the database is occupied by which blocking factors. Finally, in chapter 8, as a conclusion, I present the outcome of this study.

In chapter 2, I review the related literature on pre-nasal raising in Tehrani Persian. In that chapter I highlight the significant outcomes of each work. I also challenge the analyses in the literature by raising questions that are addressed in the following chapters.

Given this multidimensional analysis, I claim that blockage of pre-nasal raising in TP is systematically under the influence of different factors which can intersect. The contribution of the analysis in this study to the literature is that it encompasses the former analyses in the literature (which sometimes contradict each other); moreover, it accounts for the data that are assumed to be exceptions in the literature; that is, the exceptions in the literature are now part of the generalizations proposed in this study.

1.6. Limitations

An important factor which interacts with pre-nasal raising is frequency. In order to measure the role of frequency, a corpus which is built based on spoken data from Tehrani Persian is required; however, currently there is not any such corpus of Tehrani Persian. Considering my limited time, it was not possible for me to build such a corpus. The available corpora are built based on formal Persian whose data are not helpful for this study since pre-nasal raising does not occur in formal Persian. That is why in this study the role of frequency is not investigated as a separate factor. The only possibility to examine the frequency of lexical items in the database was to trust the judgements of my consultants. For future studies, a corpus analysis of pre-nasal raising is in order.

CHAPTER TWO

LITERATURE REVIEW

2.1. Introduction

In this chapter, I review the analyses in the literature on pre-nasal raising in Tehrani Persian (TP) based on the chronology of the publication of the works. The majority of these works propose various analyses for the issue of pre-nasal raising. The major blocking factors they identified are phonological, lexical and morphological.

2.2. Hodge (1957)

Hodge (1957), in his article "Some Aspects of Persian Style", identifies several factors that highlight the difference between formal and informal Persian[4].

2.2.1. A phonological blocking factor

Hodge claims that one significant difference between formal and informal Persian is that in informal speech, *a* alternates with *u* when it appears next to /m/ and /n/, as in (3):

(3) *formal vs informal pronunciation of pre-nasal a in the same word*

Formal	**Informal**	
/nɑn/	[nun]	'bread'
/hæmɑm/	[hæmum]	'bath'

There are two problems with Hodge's analysis. First, he posits that *a* alternates with *u* when *a* is next to a nasal consonant. This claim implies that even if the nasal consonant appears before *a*, it alternates with *u*. However, *a* alternates with *u* if and only if it appears before a nasal consonant

[4] In this study formal and informal Persian are entitled formal and Tehrani Persian, respectively (see the introduction chapter).

(see the phonology chapter). In the following TP data the nasal consonant precedes *a*, but *a* does not raise to *u*:

(4) *list of words with Na sequence*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/ʃenɑ/	[ʃenɑ]	*[ʃenu]	swim
/tæmɑs/	[tæmɑs]	*[tæmus]	contact
/mɑhi/	[mɑhi]	*[muhi]	fish
/ɑʃnɑ/	[ɑʃnɑ]	*[ɑʃnu]	familiar
/kenɑr/	[kenɑr]	*[kenur]	next to
/ʃomɑl/	[ʃomɑl]	*[ʃomul]	north
/mɑst/	[mɑst]	*[must]	yogurt

Second, Hodge does not mention environments in which *a* does not alternate with *u*. Hence, the blocking factors that prevent alternation of *a* to *u* remain unexamined. For discussion concerning the blocking factors, see chapters four, five and six.

2.3. Modaressi (1978)

Modaressi (1978), in a sociolinguistic analysis of Modern Persian, examines some linguistic and sociolinguistic features of Tehrani Persian, showing that linguistic variation in TP is controlled by both linguistic and non-linguistic factors. He examines the effects of social parameters (social class, age, sex, style, and ethnicity) and linguistic constraints on the probability of application of pre-nasal vowel raising.

2.3.1. Modaressi's sample

Modaressi collected a sample of around 800 lexical items containing an *aN* sequence, mostly chosen from the Amid Dictionary. This sample of words with an *aN* sequence is divided into two parts: words in which pre-nasal raising frequently applies constitute 50.5% of the sample. The rest

of the sample (49.5%) contains forms which do not favor raising even though there are not any structural blocking factors.

Modaressi claims that the most influential factor in pre-nasal raising is the frequency of use of a word; thus, the higher the frequency, the higher the probability of pre-nasal raising. Modaressi clarifies that infrequent words are excluded from his list of examples.

2.3.2. Phonological analysis

Modaressi posits that if pre-nasal *a* in a simple word undergoes raising, it normally would raise when it is a part of a compound form, /xɑne/ 'house', for instance, favors pre-nasal vowel raising and surfaces as *xune*. There are compounds such as /ketɑb-xɑne/ 'library' in which pre-nasal *a* raises: [ketɑb-xune]. There are cases in which both components of a compound contain an *aN* sequence; then in both words pre-nasal *a* raises, as in /divɑne-xɑne/ which surfaces as [divune-xune] 'mental hospital'.

However, raising does not always occur; for instance, if an *aN* sequence precede *u*, raising is blocked: /zɑnu/ → *[zunu] 'knee'. This implies that the language does not create vowel harmony in a word. This observation is consistent with my findings discussed in the phonology chapter.

Modaressi posits that monosyllabic words like *vam* 'loan', *bang* 'cry' and *gam* 'step' disfavor raising (noting exceptions including *nan* 'bread' and *an* 'that').

However, there are highly frequent words in TP which are monosyllabic but undergo raising, as shown in (5):

(5) *list of monosyllabic words with pre-nasal raising*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/d͡ʒɑn/	[d͡ʒun]	soul
/dɑn/	[dun]	seed
/rɑn/	[run]	leg

For discussion of blocking factors in monosyllabic words, see the phonology chapter. For discussion of factors that block raising in morphologically derived words whose bases are monosyllabic words, see the morphology chapter.

Modaressi notes that there are words of pure Persian origin in which raising is blocked. He attributes this to the position of the *aN* sequence, suggesting that raising does not occur if the sequence is word-initial. Thus, raising in *amade* 'ready' and *amar* 'statistics' is blocked.

In chapter five, I claim that the position of the *aN* sequence does not prevent or trigger pre-nasal raising. Furthermore, I argue that raising in words like *amar* and *amade* is inhibited due to a syllable boundary constraint.

Modaressi suggests that there are forms containing two pre-nasal *a* vowels where both raise: *xaneman* → *xunemun* 'family home'. Nevertheless, he speculates that pre-nasal raising is not usually applicable to both vowels in the same lexical item. In such cases, pre-nasal raising occurs in only one of them which is, most frequently, the last one which is also the stressed one:

(6) *sample of words with two aN sequences in each word*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/samaN/	[samun]	order
/amijane/	[amijune]	common

I maintain that if raising in an *am* sequence does not occur, it is due to the syllable boundary blocking factor. For discussion on factors that block raising in an *am* sequence, see the phonology chapter. Also, in the phonology chapter, I argue that stress does not have any role in the blockage or triggering of raising.

2.3.3. Morphological analysis

Modaressi identifies morphological factors that influence raising. For instance, he notes that raising does not occur if *a* and the following nasal consonant are in two separate morphemes: /nɑ-mærd/ → *[numærd] ‘coward’.

Modaressi (1978:77) adds that if raising occurs in the infinitival form of a verb, the inflected forms also undergo raising. Similarly, “pre-nasal vowel raising rule also usually operates in the derivational forms of a morpheme which favor the rule itself”, giving these forms (7):

(7) *sample of derivations with pre-nasal raising whose non-derived forms favor raising*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/zæbɑn/	[zæbun]	tongue
/zæbɑn-i/ tongue-adj marker	[zæbuni]	verbal
/zæbɑn-e/ tongue-diminutive marker	[zæbune]	clapper’

However, there are lexical items whose non-derived forms undergo raising, but their derived forms disfavor raising, as shown in (8):

(8) *sample of derivations with blockage of pre-nasal raising whose non-derived forms favor raising*

NON-DERIVED FORM		DERIVED FORM		
/ɑrɑm/ →	[ɑrum] ‘peaceful’	/ɑrɑm-gɑh/ → peaceful-place	[ɑrɑmgɑh]	*[ɑrumgɑh] ‘cemetery’
/ɑn/ →	[un] ‘that	/ɑn-sɑn/→ end-similar	[ɑnsɑn]	*[unsɑn][5] ‘that way’

I claim that if pre-nasal *a* in a simple lexical item undergoes raising, it does not necessarily guarantee that the derived forms of that word undergo raising as well (for discussion, see the morphology chapter).

[5] *ansun* and *unsun* are not acceptable either.

Modaressi explains that affixes containing an *an* sequence undergo pre-nasal raising if the morphologically derived word is not a recent coinage, as shown in (9):

(9) *Modaressi's list of derivational affixes with an aN sequence*

9.1. *-an* attaches to nouns, adjectives and verb stems
   a. /nɑl-ɑn/→ [nɑlun] 'moaning'

9.2. *-ane* attaches to nouns and adjectives and changes them into adjectives and adverbs, respectively:
   a. /mærd-ɑne/ → [mærdune] 'manly'

9.3. *-estan* attaches to proper or common nouns to form a noun meaning 'place of…'
   a. /lor-estɑn/ → [lorestun] 'Lorestan'

9.4. *-man* is attached to roots (imperative) to form abstract nouns
   a. /ʃad-mɑn/ → [ʃɑdmun] 'happy'

9.5. Plural possessive pronoun suffixes *-eman* '1pl', *-etan* '2pl' and *-eʃan* '3pl'
   a. /ketɑb-eʃɑn/ → [ketɑbeʃun] 'their book'[6]

9.6. *-dan* meaning 'container, holder' is added to nouns
   a. gol-dɑn/ → [goldun] 'vase'

9.7. *-dan*[7] meaning 'one's specialty' can be added to nouns
   a. tɑrix-dɑn/ → [tɑrixdun] 'historian'

9.8. *-an* (transitivizer) is an infix that attaches to roots and functions in a group of verbs (mostly transitive) as a causative morpheme
   a. /res-ɑn/ → [resun] 'give a ride'

However, there are words which contain the transitivizer *-an*, shown in (9.8), but pre-nasal raising does not occur, as in (10):

(10) *sample of words with transitivizer -an without pre-nasal raising*

1. /kut͡ʃ-ɑn-dæn/ → [kut͡ʃɑndæn], *[kut͡ʃundæn] 'cause to migrate'
   migrate-CAUS-INF

[6] This example is not part of Modaressi's study.

[7] *-dan* is not an affix; rather, it is a present tense verb stem meaning 'know', as in *mi-dan-æm* 'I know' or *dan-estæn* 'to know'. Thus, a word like *tarixdan* is a compound.

2. /godɑz-ɑn-dæn/ → [godɑzɑndæn], *[godɑzundæn] ‘cause to melt’
melt-CAUS-INF

For further discussion on the relationship of affixation with pre-nasal raising, see chapter five.

2.3.4. Lexical blocking factors

Modaressi identifies words in which pre-nasal raising sounded odd to speakers[8]. Forms in (11) are of this kind.

(11) *Modaressi’s sample of words in which raising sounds odd*

TEHRANI PERSIAN				GLOSS
/ævɑm/	→	[ævɑm]	*[ævum]	masses
/bohtɑn/	→	[bohtɑn]	*[bohtun]	accusation
/ensɑn/	→	[ensɑn]	*[ensun]	human

In the lexical factors chapter, I examine the reason for the blockage of raising in such words.

In addition, Modaressi notes that forms of pure Persian origin, proper names, Arabic [broken] plural forms, and non-Arabic loanwords disfavor raising.

Modaressi shows that words that are used in formal contexts such as broadcasting, lectures, and literature also disfavor raising. Some of these words have informal equivalents that replace them in informal contexts. Words such as /bigɑne/ ‘stranger’, /bɑmdɑd/ ‘morning’, /hengɑm/ ‘time’, /dɑneʃ/ ‘science’ are replaced in informal speech by /qæribe/, /sobh/, /moqeʔ/ and /elm/, respectively, and former are not part of the TP lexicon.

Modaressi also identifies social factors of style, age, education and gender that are involved with pre-nasal raising. He finds that style has the most significant role in inhibition of raising; he argues that the more formal the style, the more blockage of raising. Also, the younger the speakers,

[8] Modaressi writes that this classification of his sample is based on the judgement of two native speakers of Tehrani Persian.

the more pre-nasal raising is observed. The more educated the Tehrani speakers, the lower the percentage of *u* realization. Finally, Modaressi observes that female speakers tend to raise pre-nasal *a* more than male speakers. I will not examine these factors, although further study is in order.

To conclude, Modaressi provides detailed information on phonological, morphological and lexical factors that promote and inhibit pre-nasal raising.

2.4. Jahangiri (1980)

Jahangiri (1980) studies pre-nasal raising and its co-variation with social and stylistic parameters. In addition, he notes that there are phonological, lexical and morphological factors that influence pre-nasal raising

2.4.1. Phonological factors

Jahangiri claims that the position of an *aN* sequence in a word does not influence pre-nasal raising; on the other hand, he attributes the observation that there is less raising in medial position to stress and syllabification. See chapter four on the influence of stress on raising. Regarding the syllable boundary, he does not clarify what he means by syllable boundary. Which syllable boundary is he talking about? The boundary between *a* and *n*, *a* and *m* or both? This is an important distinction. As I argue in chapter four, the *am* sequence is sensitive to the syllable boundary whereas the *an* sequence is not. The other observation that Jahangiri makes is that the *an* sequence favors raising more than the *am* sequence. He does not clarify why this happens. In chapter four, I propose that the *am* sequence encounters more phonological blocking factors than the *an* sequence, so *an* raises in more environments than *am*.

2.4.2. Lexical factors

Jahangiri finds that raising is blocked if the *aN* sequence is in a recent coinage, a learned item which is not so common in informal speech, titles, formal names and loanwords. Unlike Jahangiri, I postulate that Arabic loans pattern different from non-Arabic loans in terms of pre-nasal raising (see chapter six).

2.4.3. Morphological factors

Also, like Modaressi, Jahangiri finds that pre-nasal raising is morpheme-bounded; raising is blocked if *a* and *N* are not in the same morpheme (see chapter four).

2.4.4. Exceptions

Jahangiri notes that there are exceptions to pre-nasal raising that do not appear to have an explanation. He adds that these irregularities include words like *ælʔan* 'now', *qorʔan* 'Koran' and *dæbirestan* 'high school' in which raising is blocked even though there are not any factors that should block its occurrence (see chapter six).

2.4.5. Sociolinguistic factors

Jahangiri investigates the covariation of pre-nasal raising with the extralinguistic factors of age, gender, style and social class of his informants, finding that pre-nasal raising is sensitive to social class and style. The higher the informants' social class, the less raising occurs; also, the more formal the style, the less raising is observed. Unlike Modaressi (1978), Jahangiri posits that there is no noticeable gender difference; that is, both male and female Tehrani speakers raise *a* to the same degree, except for those male native speakers with primary school education who raise *a* more than female speakers with the same level of education. Finally, he notices is there is a high

range of irregularity in reading style. He suggests that it could be due to the writing system which clearly distinguishes between *a* and *u*.

2.5. Kahn and Bernstein (1981)

Kahn and Bernstein (1981), in an article on raising in colloquial Persian, mention the same blocking factors for raising that have been discussed already, including non-Arabic loans (which they entitle European loans), morpheme boundary and learned Arabic words. They further claim that the "preponderance of the lexical items and tokens of lexical items where *aN* goes to *uN* either have *aN* in a stressed syllable in final or penultimate position. It is rare to find a word, even a commonly used word, where /ɑn/ → [un] in the antepenult. For example, /hɑmele/ 'pregnant' never becomes *[humele], /dɑneʃkæde/ 'college, faculty' never becomes *[duneʃkæde], and /rɑnænde/ 'driver' does not become *[runænde]. The form /rɑnændegi/ 'driving' never becomes *[runændegi], indicating that the further *aN* is from the end of the word, the less likely it is to change to *uN*" (Kahn and Bernstein 1981:135).

The problem with this analysis is that Kahn and Bernstein, like Modaressi (1978) and Jahangiri (1980), do not distinguish between the factors that block raising in an *am* sequence from those in an *an* sequence, and hence propose an overgeneralized phonological analysis for both environments. As mentioned earlier, in this chapter, I argue that neither stress pattern nor position of *aN* sequence in a word influences raising; see chapter four.

Unlike Kahn and Bernstein, I propose a different blocking factor that prevents raising in *daneʃkæde* and *ranænde* in chapter five. This proposed blocking factor nullifies the role of stress as well.

Kahn and Bernstein classify common nouns in two groups based on their frequency of occurrence and style. One group tends to favor raising, as in /tæmɑm/ → [tæmum] 'end' and

/xijɑbɑn/ → [xijɑbun] ‘street’, among others. Yet, another group of words resists raising, as in (12):

(12) *Kahn and Bernstein’s list of common nouns with blockage of pre-nasal raising*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
a) /kɑ.mel/	[kɑ.mel]	*[ku.mel]	perfect
b) /hɑ.mele/	[hɑ.mele]	*[hu.mele]	pregnant
c) /dɑneʃd͡ʒu/	[dɑneʃd͡ʒu]	*[duneʃd͡ʒu]	university student
d) /emtehɑn/	[emtehɑn]	*[emtehun]	examination
e) /dæbestɑn/	[dæbestɑn]	*[dæbestun]	primary school
f) /ælʔɑn/	[ælʔɑn]	*[ælʔun]	now

In the end, they conclude that words in the class of common words which do not change are either Arabic borrowings (12.f) or are terms related to education (12.c) or both (12.d).

In chapter six, I argue that there are systematic patterns that block raising in the abovementioned examples. For instance, raising in (12.c) and (12.d) is not inhibited due to their being terms related to education. Rather, (12.c) disfavors raising as it contains a blocking affix (see chapter five). (12.d) does not raise since it is an Arabic triliteral infinitival form; see chapter six. Hence, based on the analyses that I suggest, there are different blocking factors that prevent raising in (12.c), (12.d) and (12.e), and thus I posit that it is sheer chance that they are related to education.

In the phonology chapter, I propose an analysis which explains why raising in (12.a) and (12.b) is blocked. This analysis shows that inhibition of raising in (12.a) and (12.b) is not exceptional; rather, there is a pattern which systematically blocks raising.

2.6. Sadeghi (2001)

Sadeghi (2001) writes a reply to Kahn and Bernstein (1981). His approach is historical. He claims that pre-nasal raising is an ancient alternation belonging to the pre-Islamic era. He posits that the

a>*u* alternation stopped being productive ca. 12th century AD. He details that if in contemporary TP pre-nasal *a* raises in a word, it is because raising in that word occurred before the 12th century, and thus the word has been used with a raised *a* up to contemporary times. Words which are pronounced without raising in contemporary TP entered the TP lexicon after the 12th century. Hence, only those loanwords that are borrowed into Persian before this time undergo raising; loanwords that are borrowed after the 12th century resist raising. He presumes that 12th century is the time when pre-nasal raising stopped being productive since Mongolian and Turkish loanwords that entered Persian lexicon after the Mongol invasion in 12th century resist raising. He adds that recently coined words and words that are reused from archaic and formal Persian sources disfavor raising as well.

Sadeghi's viewpoint is a diachronic approach whereas my approach in this study is synchronic. So, in order to be able to critique his work, a historical linguistic study is required. To do so, questions such as the following should be answered. First, we should answer whether pre-nasal raising remains active. Pisowicz (1985) claims that a tendency to raise *a* preceding *n* and *m* in Persian was observed since the beginning of the 15th century and became clearer in the 17th century. The raising at that time was to both *o* and *u* based on transcriptions of Persian by non-native speakers (cited in Rohany 2012:115). Even if Sadeghi's claim turns out to be true, then we should check if Arabic loans follow his generalization; since there might be Arabic loans that have entered Persian before 12th century, but raising does not occur in them; likewise, there could be words that have entered TP lexicon from formal Persian before 12th century yet raising is not observed. The other point is about the generalization on recent coinages. There are recently coined words like *saxteman* 'building', *atæʃ neʃan* 'fire fighter' and *atæʃ feʃan* 'volcano', which undergo raising. This shows that the role of frequency as an influential factor might have been neglected.

2.7. Rohany (2012)

Rohany approaches pre-nasal raising topic from a synchronic viewpoint.

Rohany posits that "some names of cities may show raising although the version without raising is more common; others do not show raising" (Rohany 2012:116), as in (13):

(13) *Rohany's sample of cities with pre-nasal raising*

1.	/tehrɑn/	→	[tehrun]	'capital of Iran, a city'
2.	/lɑhid͡ʒɑn/	→	[lɑhid͡ʒɑn], *[lɑhid͡ʒun] (cf. bɑdemd͡ʒɑn ~ bɑdemd͡ʒun 'eggplant')	'a city'

Rohany adds that pre-nasal *a* in the suffix *-estan*, a location suffix, shows raising in some common nouns and not it in others, as in (14):

(14) *Rohany's sample of words with pre-nasal raising in -estan suffix*

a. ɢæbr 'grave' + -estɑn → [ɢæbrestɑn] ~ [ɢæbrestun] 'graveyard'
b. kudæk 'child' + -estɑn → [kudækestɑn] ~ *[kudækestun] 'kindergarten'

She does not address the question of whether blockage of raising in the suffix *-estan* is systematic or random. I attempt to answer this question in the morphology chapter.

Rohany adds that when *-estan* is used in the name of countries, it does not show raising, as in (15).

(15) *Rohany's sample of words with blockage of raising in name of countries containing -estan suffix*

a. /tɑd͡ʒik + -estɑn/ → [tɑd͡ʒikestɑn] 'Tajikistan' *[tɑd͡ʒikestun]
b. /mɑdʒɑr + -estɑn/ → [mɑd͡ʒɑrestɑn] 'Hungary' *[mɑd͡ʒɑrestun]

In addition, Rohany notes that names of individuals (*sanaz* → **sunaz* 'a name for girls), loanwords (*departeman* → **departemun* 'department'), foreign proper names (*tam* → **tum* 'Tom') and names of countries (*tajvan* → **tajvun* 'Taiwan') do not undergo raising.

Rohany also writes that although pre-nasal raising is very common, it has some exceptions, and adds that some of these exceptions could be idiosyncratic since no particular reason or pattern is observed for them, as in (16):

(16) *Rohany's sample of words in which raising fails to occur for no reason*

a. /onvɑn/ → *[onvun] 'title'
b. /xɑme/ → *[xume] 'cream'

I also posit that pre-nasal raising has some exceptions; however, I do not consider (16.a) and (16.b) to be exceptions. In chapter six I show that a group of Arabic loans, including *onvan*, following a certain root-and-pattern of Arabic morphology, resist raising; moreover, blockage of raising in (16.b) is not patternless; raising is inhibited due to the syllable boundary blocking factor discussed in chapter four.

2.8. Miller (2011)

Miller (2011:1389) refers to Kahn and Bernstein (1981), who note that the lexical items most commonly undergoing the change to [un] are deictics and pronominal suffixes and that segments in such closed class forms resist changes undergone in other parts of the language. Based on this claim, he suggests to turn the problem around by proposing that perhaps the underlying form is /u/ which optionally changes to [ɑ] in higher registers, under the influence of orthography. In the phonology chapter, I show that the underlying form is /ɑ/ which surfaces either as [ɑ] or [u].

2.9. Rees (2008)

Rees (2008:218) posits that pre-nasal raising is under the influence of factors such as frequency and register effects. He specifically highlights the influence of frequency: "Iranian does mandate the change in all highly frequent tokens and disallows it in infrequent tokens" (Rees 2008:221).

I postulate that frequency and register are two influential factors; nevertheless, I disagree that all highly frequent words undergo raising since there are highly frequent lexical items in which pre-nasal *a* does not raise. In (17) some highly frequent words are listed in which raising is inhibited:

(17) *sample of frequent words in TP with blockage of raising*

	UNDERLYING FORM	SURFACE FORM		GLOSS
a)	/dɑmæn/	[dɑmæn]	*[dumæn]	skirt
b)	/xɑm/	[xɑm]	*[xum]	raw
c)	/estekɑn/	[estekɑn]	*[estekun]	teacup
d)	/bærɑ-m/ for-me	[bærɑm]	*[bærum]	for me

For the discussion on blockage of raising in (17.a) and (17.b), see chapter four; blocking factors of (17.c) and (17.d) are discussed in chapters six and five, respectively.

It is worth mentioning that Kalbasi (2001), Mahmoodi-Bakhtiari (2008), Windfuhr (2009:417) and Miller (2012; 2014), among others, briefly discuss pre-nasal raising, but since the main focus of their analyses is not pre-nasal raising, they do not discuss this *a~u* alternation in depth, and the factors they mention are mostly drawn from the literature surveyed in this chapter. Hence, what they discuss does not differ from the literature that I have reviewed.

2.10. Conclusion

This chapter reviews the analyses of pre-nasal raising in the literature. One goal of this chapter was to highlight the strong points of each work and at the same time critique them by raising questions that I address in the following chapters.

I will survey the phonological, morphological and lexical factors that inhibit raising. While sociolinguistic factors are clearly significant, as the literature shows, they are beyond the scope of this work.

CHAPTER THREE

DATABASE

3.1. Introduction: why a database?

As we have seen in chapter 2, while the raising of *a* to *u* in Tehrani Persian is widely discussed in the literature, there are numerous exceptions to it. While there have been many attempts to identify the factors that promote and inhibit raising, the literature is contradictory, and much of the work is based on small amounts of data. One goal of my research is to examine the various factors that have been identified as blocking raising using as large a database as possible. The goal of this chapter is to describe the database and its structure.

The chapter is organized as follows. I begin with a short discussion of the source of the data. I then describe the structure of the database, and how it is useful for filtering the data to examine different factors involved in promoting and inhibiting raising.

3.2. Source of data

The major source of data for this work is the Dehkhoda Online Dictionary (DOD) which is named after its founder, Ali Akbar Dehkhoda (1879-1955), an Iranian literary and political figure. This dictionary is the most comprehensive Persian dictionary to date. As of May 2018, the dictionary contains 343,466 entries. The pronunciation of each entry is indicated through diacritic features. If an entry is a loanword, the source language of that entry is mentioned beside the definition of that entry. If a loanword has a Persian equivalent, that equivalent is also indicated for that entry.

According to the Dehkhoda Lexicon Institute and the International Center for Persian Studies, the dictionary, besides vocabulary, contains proverbs, metaphors, literary figures of speech and compound structures in Persian. For each vocabulary item, there is a definition. In

addition, in order to exemplify how that word has been used in Persian literature, an excerpt of a Persian literary text (poetry or prose) which contains that word is included in that entry. There are also historical and geographical entries related to territory of Iran, Islam and some other non-Islamic nations. Various linguistic, literary, philosophical and religious topics are discussed. Scientific and technological terminologies are introduced. In other words, this dictionary has evolved into an encyclopedia (http://icps.ut.ac.ir/).

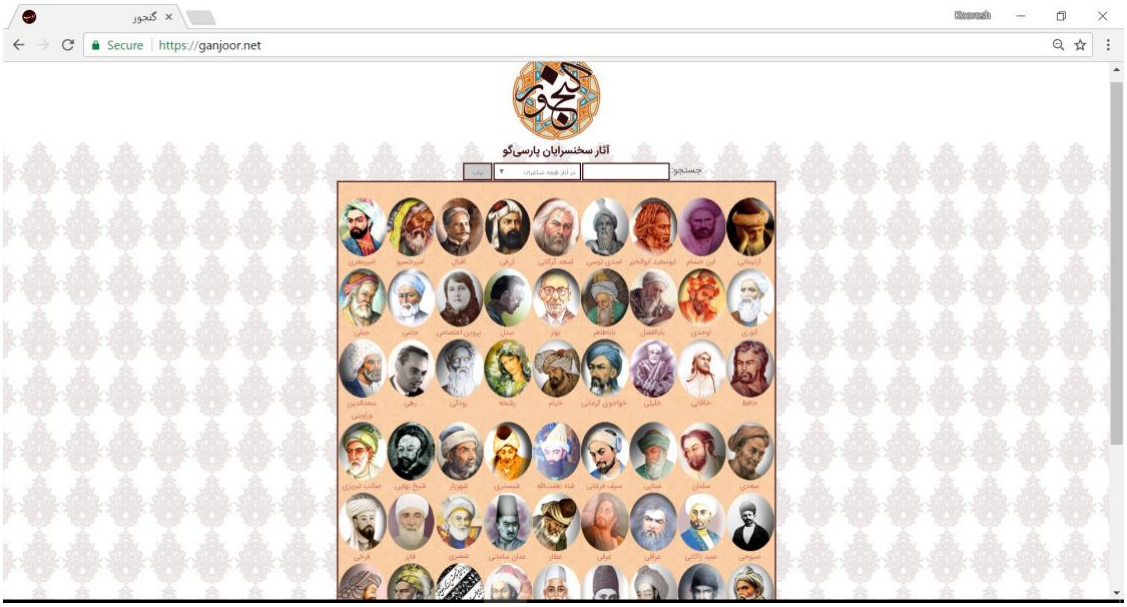
The complete electronic version of the DOD dictionary is accessible at http://www.jasjoo.com. This website is the source of my data. Below, in (18), is a screenshot of the first page of DOD on the website:

(The screenshot is on the next page)

(18) *first page of Dehkhoda Online Dictionary*

In addition to DOD, I have used the Moein Persian Dictionary, which is considered to be one of the most comprehensive Persian dictionaries, and the Ganjoor database. I have referred to the Moein Dictionary for further definition and exemplification of some words which are not fully explained in DOD. Moreover, I used this dictionary for finding the source language of some loanwords which are used in Persian whose source language is not mentioned in DOD. The Ganjoor database is a collection of literary works written by major and established Persian speaking literary figures. The database includes poems written in 9th century AD until contemporary time; however, only those contemporary poets and writers have been included in the database who are classicist. When I wanted to check whether a lexical item was recently coined, I searched this database (in Persian). Or when I wanted to find out since when a lexical item has been used in Persian literature, or whether a word has been used in a certain period of time, this database, besides DOD and Moein Dictionary, has been my source of investigation. This database has this option to search a word in the work of a single writer or in the whole database altogether. Below in (19) is a screenshot of the main page of Ganjoor website:

(19) *first page of the Ganjoor website*



3.3. The Database

The database, organized in Microsoft Excel version 2016, includes all items from DOD with an *aN* sequence. The database is organized into seven columns, with each column including important information to understand pre-nasal *a* raising and how it applies.

In the first column, entitled "word list", the data are organized alphabetically. I have been able to collect 1780 words from DOD which contain either an *an* or an *am* sequence. There are a number of words like *daman*[9] 'skirt' that contain both *am* and *an* sequences, and such words are listed once.

The second column, labelled "Raising", provides information if *a* in the *aN* sequence raises to *u* or not. If *a* raises to *u*, then the cell in the "Raising" column is filled with "Yes"; otherwise, it is filled with "No". Raising or non-raising in the words of the database is based on the judgement of four highly educated thirty-year-old Tehrani Persian speakers (one male and three females).

The third column, "blocking factors", is the most complex. I identify nine factors that have been argued to play a role in whether raising is allowed or not. These include the following lexical factors: is the word a proper name, is the word recently coined, is the word formal Persian, is the word a loanword. I determined whether a word is recently coined or not in the following way: those words which are coined recently as Persian equivalent of loanwords are considered to be in this category, like *bæsamæd* 'frequency' which has been coined to replace the French word *fréquence.* Also, those words which have been coined in contemporary Iran after technological, bureaucratic and educational changes are considered to be newly coined; for instance, *samane* 'system', *bajgani* 'archives', *pasox name* 'answer sheet', *zæban ʃenas* 'linguist', *bæxʃ name* 'circular', *omran* 'civil engineering field of study' are newly coined words. Also, those words that

[9] *daman* is an archaic Persian word. In contemporary Persian it is pronounced as *damæn.*

existed in old Persian literature and have been revived in contemporary Tehrani Persian with a slight change in the original meaning of the word are considered in this subcategory as well; for instance, words like *dæbestan,* 'primary school' and *dæbirestan* 'high school' are used in old Persian literature as places for children to go to school, but the former in contemporary Persian means school for primary level children and the latter means high school. The status as formal Persian was determined as follows: words that are not used in the daily speech of Tehrani Persian speakers (infrequent words in daily speech) are categorized as formal Persian words. The majority of these words have equivalents in TP; for instance, formal words like *rajgan* 'free of charges', *ærmæqan* 'gift', *dehqan* 'peasant' and *færjam* 'end' have these vernacular equivalents, as shown below in (20):

(20) *sample of formal Persian words with TP equivalents*

FORMAL PERSIAN	TEHRANI PERSIAN	GLOSS
rɑjgɑn	moft	free of charges
ærmæqɑn	kɑdo	gift
dehqɑn	keʃɑværz	peasant
færd͡ʒɑm	ɑxær	end

Finally, status as a loanword: according to Kemmer (2017) a foreign word is a lexical item that most speakers do not know, and if they hear it, they think it is from another language. On the other hand, a loanword (borrowing) is a foreign word that becomes conventionalized (the process through which a loanword adapts to the phonological inventory of the borrowing language). The focus of this study is on loanwords; for instance, *livan* 'glass', *anlajn* 'online' and *kamjon* 'truck' are loanwords that are borrowed from Russian, English and French, respectively. In Chapter six I will explain why I have not considered Arabic loanwords in this category.

In addition to lexical factors, there are phonological factors that block raising. Some of these are drawn from the literature. In other cases, inspection of a small database led me to identify factors that have not been discussed in the literature but that seemed to be important, and these were included as well, to some degree to allow their testing. The phonological factors are dissimilating [*a*], blocking affixes, monosyllabic *am*, and [*am*] not in the same syllable. See chapter four for discussion of these factors. There is also one morphological factor, "[*aN*] not in the same morpheme" which is discussed in chapter four as well.

In the row in which the information about each word is indicated, the title of that blocking factor is written in its appropriate cell; for instance, *a* in a word like *ba-nofuz* 'influential' does not change into *u*. I attribute this to the fact that the *an* sequence is not in the same morpheme; hence, in the row belonging to the lexical item *banofuz* under the "[α-N]" which means *a* and *N* are not in the same morpheme, "not in the same morpheme" is written. It is worth mentioning that in the database there are some words in which *a* does not change into *u* for several reasons, and all the reasons for disallowing pre-nasal raising are indicated in the database. For example, in a word like *ba-mæze* 'cute' raising is blocked. *a* and *m* are not in the same morpheme; in addition, they are not in the same syllable. As a result, there are two blocking factors at play. In Chapter seven, I address how the blocking factors interact with each other.

The "Part of speech" column follows the "Blocking factors" column. In this column, the part of speech of each lexical item is indicated. The reason for including this column is to see whether there is any relationship between the part of speech and raising. This topic is discussed in chapter five.

The next column gives the gloss of the lexical item in the first column. In the last column, the source of the data is indicated. Below, in (21), we see an unfiltered format of the database:

(21) *unfiltered format of the database*

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	word list	raising	[ɑN]	blocking factor												P of Sp	gloss	source
2				proper name	Arabic broken pl	Arabic infinit	MDW	[ɑ-N]	newly coined	dissimilating [ɑ]	blocking affix	standard Persia	loanword	monosyllabic [ɑm]	[ɑ.m]			
1246	natæmam	yes	am													adjective	endless	DOD
1247	natævan	no	an									standard Persian				adjective	unable	DOD
1248	natævani	no	an				MDW					standard Persian				noun	disability	DOD
1249	navdan	yes	an													noun	drain, waterpipe	DOD
1250	negæhban	yes	an													noun	guard	DOD
1251	negæhbani	yes	an													noun	to guard	DOD
1252	nej ænban	yes	an													noun	name of a musical intrument	DOD
1253	nejestan	no	an									standard Persian				noun	canebrake	DOD
1254	neʃan	yes	an													noun	sign	DOD
1255	neʃandæn	yes	an													verb	cause to sit	DOD
1256	neʃane	yes	an													noun	sign	DOD
1257	neʃani	yes	an													noun	address	DOD
1258	nezam	no	am						newly coined word			standard Persian				noun	government, regime	DOD
1259	nezami	no	am												[am] not	adjective	army man	DOD
1260	nezami	no	am	proper name											[am] not	noun	Nezami	DOD
1261	nijam	no	am									standard Persian				noun	scabbard	DOD
1262	niknam	no	am									standard Persian				adjective	honored, reputable	DOD
1263	niknami	no	am				MDW					standard Persian				noun	honor, reputation	DOD
1264	niruje entezami	no	am						newly coined word						[am] not	noun	disciplinary	DOD
1265	nisan	no	an									standard Persian				noun	name of a month	DOD
1266	njam	no	am									standard Persian				noun	scabbard	DOD
1267	noqsan	no	an			Arabic infinitive						standard Persian				noun	defect	DOD
1268	novambr	no	am	proper name												noun	November	DOD
1269	now d͡ʒame	no	am									standard Persian			[am] not	noun	one who has new clothes	DOD
1270	nowbæran	no	an	proper name												noun	name of a city	DOD
1271	nowbonjan	no	an	proper name												noun	name of a region in Fars pro	DOD
1272	nowdan	no	an	proper name												noun	name of a village	DOD
1273	nurani	no	an								blocking affix					adjective	luminous	DOD
1274	oman	no	an	proper name												noun	Oman	DOD
1275	omran	no	an			Arabic infinitive			newly coined word							noun	civil engineering, reconstruct	DOD
1276	onfævan	no	an									standard Persian				noun	bloom, prime	DOD
1277	onvan	no	an			Arabic infinitive										noun	title	DOD
1278	oqjanus	no	an							dissimilating [ɑ]						noun	ocean	DOD
1279	orjan	no	an			Arabic infinitive						standard Persian				adjective	naked	DOD
1280	osjan	no	an			Arabic infinitive						standard Persian				noun	riot	DOD
1281	osman	no	an	proper name												noun	male name	DOD
1282	osmani	no	an	proper name												noun	Ottomanian	DOD
1283	ostan	no	an						newly coined word							noun	province	DOD

Sheet1

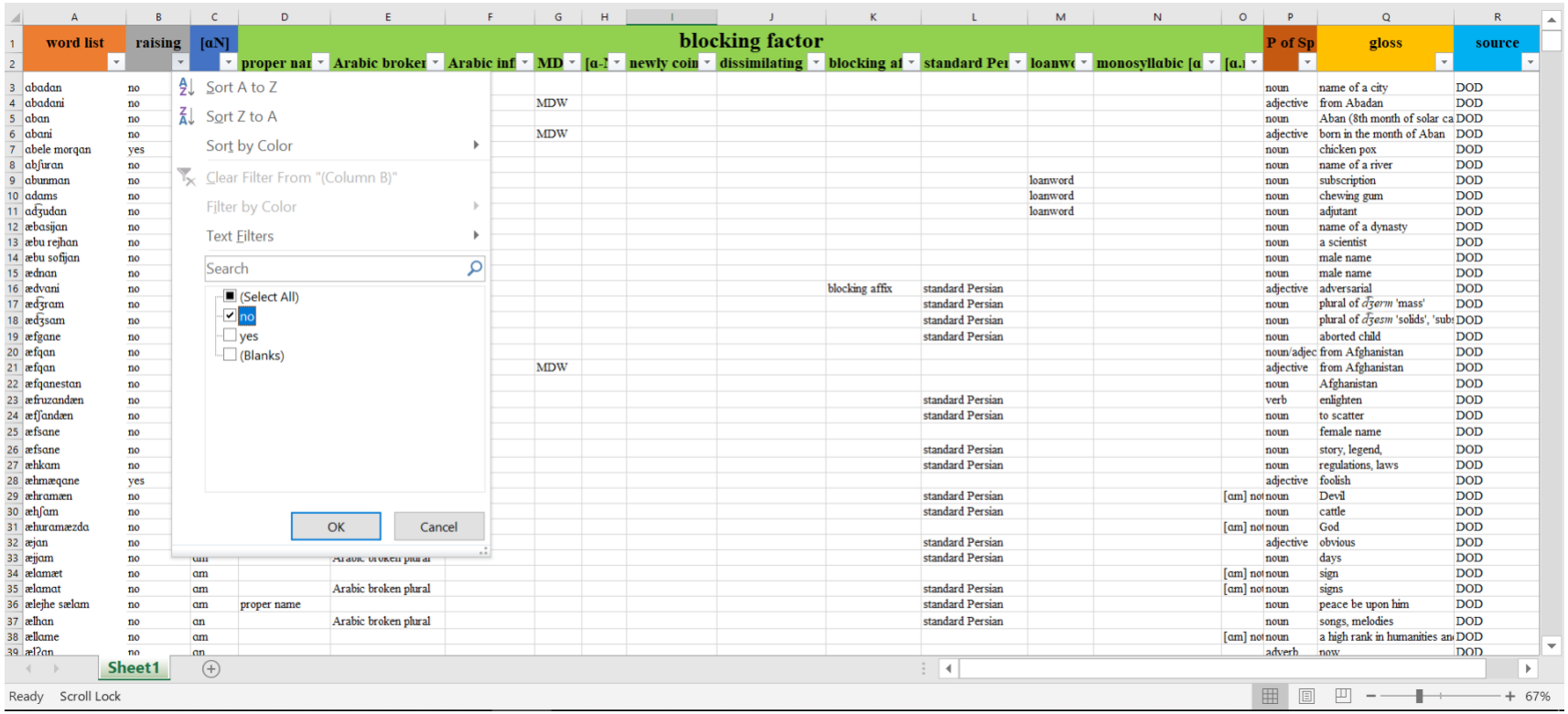
3.3.1. Filtering

One advantage to working with Microsoft Excel is that it allows for filtering of the data to look at forms that are relevant to specific questions. In order to work with the filter option, first we select the header of a column that we want to filter. Then under the "Home" section, in the "editing" part we click on "Sort & Filter". After that, from the options that show up, we click on the "Filter" option. Automatically a drop-down arrow for the column that we wanted to filter will appear. By clicking on the arrow, the menu of filters will show up. Then we can choose from among the options in the menu by clicking on them.

As an example, in the "Raising" column we can filter this column by choosing the "No" option in the menu. Selecting "No" means that the data will be filtered so that only lexical items will be shown in which raising is blocked, as displayed in (22).

(22) *how to filter the database to show the words in which raising is blocked*

(the screenshot is not the next page)

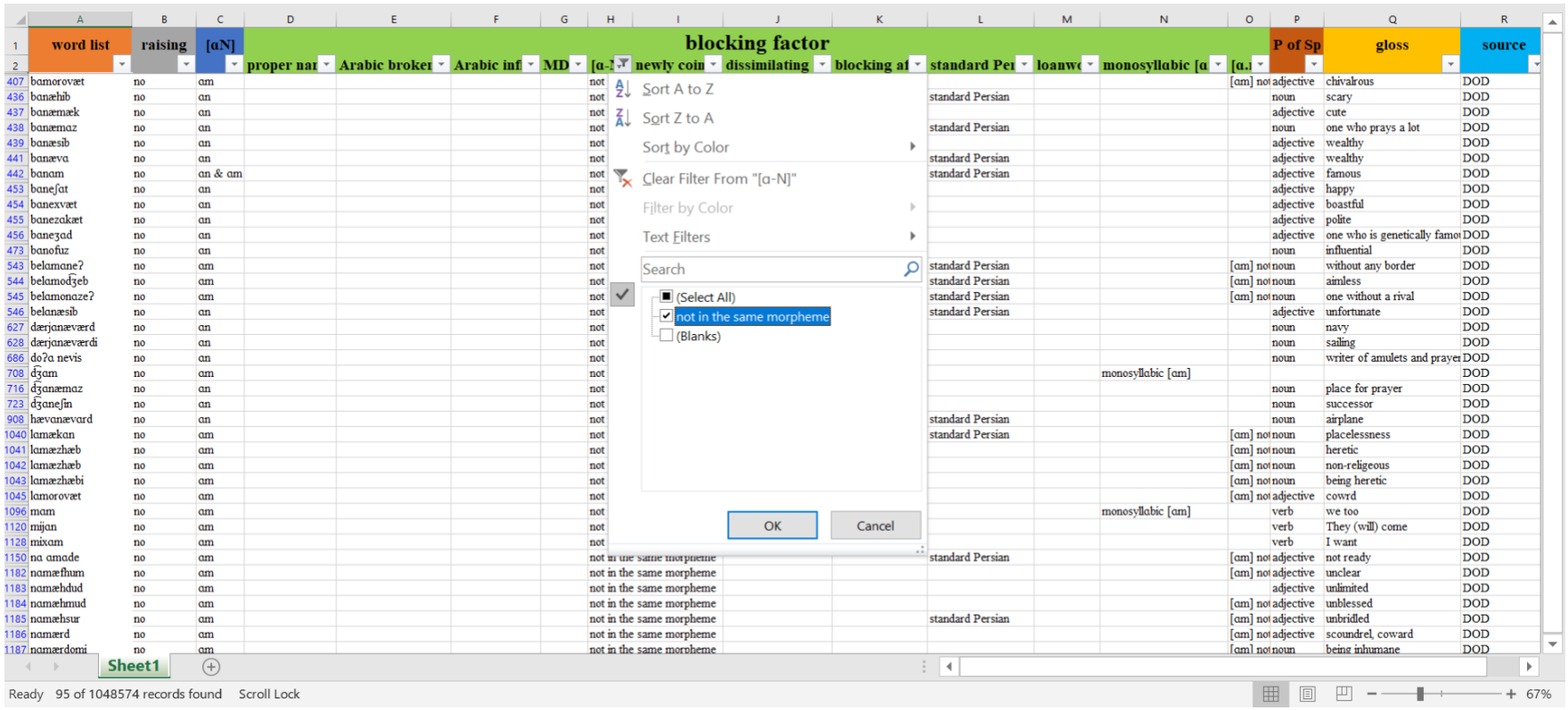


	word list	raising	[aN]	proper nar	Arabic broker	Arabic inf	MD	[a-	newly coin	dissimilating	blocking af	standard Per	loanwo	monosyllabic [a	[a.	P of Sp	gloss	source
3	abadan	no														noun	name of a city	DOD
4	abadani	no					MDW									adjective	from Abadan	DOD
5	aban	no														noun	Aban (8th month of solar ca	DOD
6	abani	no					MDW									adjective	born in the month of Aban	DOD
7	abele morqan	yes														noun	chicken pox	DOD
8	abʃuran	no														noun	name of a river	DOD
9	abunman	no											loanword			noun	subscription	DOD
10	adams	no											loanword			noun	chewing gum	DOD
11	adʒudan	no											loanword			noun	adjutant	DOD
12	æbasijan	no														noun	name of a dynasty	DOD
13	æbu rejhan	no														noun	a scientist	DOD
14	æbu sofjan	no														noun	male name	DOD
15	ædnan	no														noun	male name	DOD
16	ædvani	no									blocking affix	standard Persian				adjective	adversarial	DOD
17	ædʒram	no										standard Persian				noun	plural of *dʒerm* 'mass'	DOD
18	ædʒsam	no										standard Persian				noun	plural of *dʒesm* 'solids', 'subs	DOD
19	æfgane	no										standard Persian				noun	aborted child	DOD
20	æfqan	no														noun/adjec	from Afghanistan	DOD
21	æfqan	no					MDW									adjective	from Afghanistan	DOD
22	æfqanestan	no														noun	Afghanistan	DOD
23	æfruzandæn	no										standard Persian				verb	enlighten	DOD
24	æʃʃandæn	no										standard Persian				noun	to scatter	DOD
25	æfsane	no														noun	female name	DOD
26	æfsane	no										standard Persian				noun	story, legend,	DOD
27	æhkam	no										standard Persian				noun	regulations, laws	DOD
28	æhmæqane	yes														adjective	foolish	DOD
29	æhramæn	no										standard Persian			[am] no	noun	Devil	DOD
30	æhʃam	no										standard Persian				noun	cattle	DOD
31	æhuramæzda	no													[am] no	noun	God	DOD
32	æjan	no										standard Persian				adjective	obvious	DOD
33	æjjam	no	am		Arabic broken plural							standard Persian				noun	days	DOD
34	ælamæt	no	am												[am] no	noun	sign	DOD
35	ælamat	no	am		Arabic broken plural							standard Persian			[am] no	noun	signs	DOD
36	ælejhe sælam	no	am	proper name								standard Persian				noun	peace be upon him	DOD
37	ælhan	no	an		Arabic broken plural							standard Persian				noun	songs, melodies	DOD
38	ælname	no	am												[am] no	noun	a high rank in humanities an	DOD
39	æl?an	no	an													adverb	now	DOD

Then, we can narrow the data more by adding another filter to the data that have already been filtered; for instance, we can click on the top-down arrow which belongs to the "not in the same morpheme". As shown in (23), we only click on the box which belongs to "not in the same morpheme" and take out the check mark which belongs to the "Blanks". In this way we have narrowed down the data to those words that block the raising and *aN* is not in the same morpheme.

(23) *how to filter the database to words with a and N in separate morphemes*

(the screenshot is on the next page)


word list
raising
[ɑN]
blocking factor
proper nar
Arabic broker
Arabic inf
MD
[ɑ-N]
newly coin
dissimilating
blocking af
standard Per
loanwo
monosyllabic [ɑ
[ɑ..]
P of Sp
gloss
source
Sort A to Z
Sort Z to A
Sort by Color
Clear Filter From "[ɑ-N]"
Filter by Color
Text Filters
Search
(Select All)
not in the same morpheme
(Blanks)
OK
Cancel
Sheet1
Ready
95 of 1048574 records found
Scroll Lock
67%

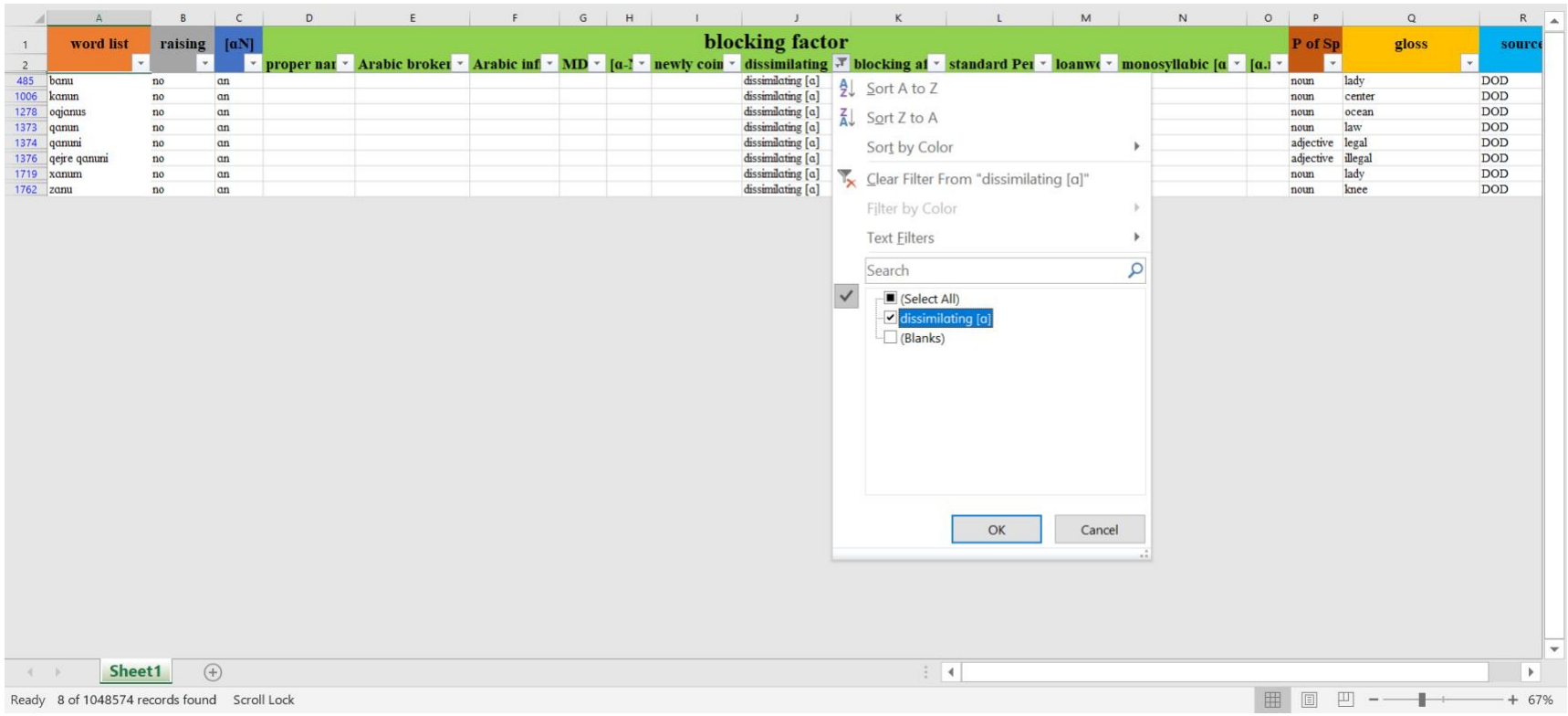
This method of filtering does not keep those lexical items in the list that resist raising only due to *aN* being in two separate morphemes; rather, all those words in which raising is blocked for this reason and other possible factors remain in the list of filtered data. That is why a word like *bameknæt* 'wealthy' in which raising is blocked for three different factors (*aN* not in the same morpheme, *am* not in the same syllable, being a formal word) is in the list of filtered data.

In order to take the file back to its original form without any filters, it is enough to click on the "Filter" option in the "Editing" section of "Home".

One of the benefits of using this method of filtering is that we can spot exceptional cases in which *a* shouldn't raise to *u* due to one of the blocking factors but nevertheless we see raising in that word. In order to find these exceptions, first, in the raising column we click on the top-down arrow and choose "yes" so that all words that allow the raising would be selected. Then, we pick any of the blocking factors by clicking on the top-down arrow which is beside the header of that blocker. After clicking on the arrow, if in the menu we encounter only two options of "Select All" and "Blanks", it signifies that that blocking factor does not have any exceptions; for instance, below, in (24), we see that in all those words that "dissimilating [*a*]" as the blocking factor has blocked the *a>u* raising, there are no exceptions.

(24) *how to filter the database to TP words with [anu] sequence*

(the screenshot is on the following page)


blocking factor
word list
raising
[ɑN]
proper nar
Arabic broker
Arabic inf
MD
newly coin
dissimilating
blocking af
standard Per
loanwo
monosyllabic [ɑ
P of Sp
gloss
source
485 banu no an dissimilating [ɑ] noun lady DOD
1006 kanun no an dissimilating [ɑ] noun center DOD
1278 oqjanus no an dissimilating [ɑ] noun ocean DOD
1373 qanun no an dissimilating [ɑ] noun law DOD
1374 qanuni no an dissimilating [ɑ] adjective legal DOD
1376 qejre qanuni no an dissimilating [ɑ] adjective illegal DOD
1719 xanum no an dissimilating [ɑ] noun lady DOD
1762 zanu no an dissimilating [ɑ] noun knee DOD
Sort A to Z
Sort Z to A
Sort by Color
Clear Filter From "dissimilating [ɑ]"
Filter by Color
Text Filters
Search
(Select All)
dissimilating [ɑ]
(Blanks)
OK
Cancel
Sheet1
Ready 8 of 1048574 records found Scroll Lock
67%

On the other hand, despite the generalization that pre-nasal raising is blocked in proper names, in (25) it is evident that there are proper names that allow raising and it is because in the top-down arrow next to the header of proper names column besides “Select All” and “Blanks”, we have the third option of “proper names” which means that there are some proper names in the already filtered data that allow raising. In this way, the exceptional cases of proper names that allow the *a>u* raising are easily spotted. The reason that these exceptional cases exist is discussed in Chapter six.

(25) *how to find the exceptions of a blocking factor in the database*

(the screenshot is on the following page)

	word list	raising	[αN]	blocking factor												P of Sp	gloss	source
				proper na	Arabic broke	Arabic inf	MD	[α-	newly coin	dissimilating	blocking af	standard Per	loanwo	monosyllabic [α	[α.			
785	esfæhan	yes	an	proper name												noun	name of a city in Iran	DOD
911	hæmedan	yes	an	proper name												noun	Hamedan (name of a city)	DOD
956	iran	yes	an	proper name												noun	Iran	DOD
1015	kaʃan	yes	an	proper name												noun	name of a city in Iran	DOD
1023	kerman	yes	an	proper name												noun	name of a city in Iran	DOD
1025	kermanʃah	yes	an	proper name												noun	name of a city in Iran	DOD
1034	kordestan	yes	an	proper name												noun	name of a province in Iran	DOD
1384	qolam	yes	am	proper name												noun	male name	DOD
1395	ræmezan	yes	an	proper name												noun	male name	DOD
1491	ʃæʔban	yes	an	proper name												noun	male name, name of a month	DOD
1544	ʃemiran	yes	an	proper name												noun	name of region in Tehran	DOD
1625	tehran	yes	an	proper name												noun	capital city of Iran	DOD
1732	xorasan	yes	an	proper name												noun	name of a province in Iran	DOD

The other point regarding the filtering option is that when a column has been filtered, a filter symbol appears beside the header of that column. In order to change the filtering type of a header or remove the filter from a column it is enough to click on the filter symbol, and then the options appropriate to the filtering of that certain column appear.

3.4. Summary

In this chapter, I have provided a brief overview of the structure of the database, and of how filtering works in order to allow us to see which of the blocking factors are exceptionless and which allow for exceptions. In the chapters that follow, I look at each of the blocking factors in more depth.

CHAPTER FOUR

PHONOLOGICAL FACTORS

4.1. Introduction

This chapter is an analysis of the pre-nasal raising in TP from a phonological viewpoint. The research question of this study is a phonological one; that is, *a* raises to *u* when *a* is immediately followed by a nasal consonant. As we have seen in chapter 2, while the process is phonological, it does not apply whenever its structural description is met. I examine why the *a~u* alternation is a raising phenomenon and not a lowering one; I then show that pre-nasal raising occurs if and only if both *a* and the nasal consonant are in the same morpheme. Following that, I present my analysis regarding the phonological factors that influence pre-nasal raising, focussing on the different patterning of *am* and *an* sequences. I begin with the blocking factors that inhibit raising of *a* followed by *m.* Then, I discuss the blocking factor of *a* preceding *n*.

Before, the discussion of the phonological blocking factors, it is worth mentioning that this chapter addresses and discusses those words in which pre-nasal raising is inhibited merely by a phonological blocker. There are words in which raising fails to occur due to morphological or lexical factors. For the discussion of non-phonological blocking factors, see chapters five and six.

4.2. Why pre-nasal raising (*a*→ *u*) and not lowering (*u* → **a*)

In the literature, it is assumed that the *a~u* alternation is a raising phenomenon (*a* → *u*) and not a lowering one (*u* → *a*). Nevertheless, Miller (2011) refers to Kahn and Bernstein (1981) who argue that the lexical items that mostly undergo pre-nasal raising are deictics and pronominal suffixes and that segments in such closed class forms resist changes undergone in other parts of the

language. Based on this generalization, Miller (2011:1389) concludes that "perhaps the underlying forms contain *u*, and these are optionally changed to *a* in higher registers, influenced by the presence of alef ا in the orthography." Miller does not elaborate on this claim.

I follow the majority of the literature, postulating that the underlying form is *a* and not *u* since there are actively used lexical items in formal and Tehrani Persian which contain *uN* sequence, but the pre-nasal *u* in these words never lowers to *a*. The following table in (26) contains a few examples that support my supposition. I am not aware of other blocking factors that might play a role in the patterning of such words.

(26) *sample of words with [uN] sequence in the underlying form*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/mærhum/	[mærhum]	*[mærhɑm]	departed
/niʃgun/	[niʃgun]	*[niʃgɑn]	pinch
/sotun/	[sotun]	*[sotɑn]	pillar
/xun/	[xun]	*[xɑn]	blood
/mæskuni/	[mæskuni]	*[mæskɑni]	residential
/birun/	[birun]	*[birɑn]	outside
/bedune/	[bedune]	*[bedɑne]	without
/bum/	[bum]	*[bɑm]	canvas
/mum/	[mum]	*[mɑm]	wax
/mæhkum/	[mæhkum]	*[mæhkɑm]	convict
/ɑbrumænd/	[ɑbrumænd]	*[ɑbrɑmænd]	honorable
/pune/	[pune]	*[pɑne]	spearmint
/mæʔlum/	[mæʔlum]	*[mæʔlɑm]	obvious

These examples show that, unlike words with an *an* sequence, in words with an *un* sequence there is not the option for the pre-nasal *u* to lower to *a* even in highly formal contexts.

4.3. Pre-nasal raising and morpheme boundary

Before I begin the discussion on *a* raising in *aN* sequence, it is necessary to mention that, pre-nasal *a* raising is sensitive to the presence of a morpheme boundary. Following Modaressi (1978:83),

Jahangiri (1980:81), Kahn and Bernstein (1981:136-7) and Miller (2011:1387), I postulate that pre-nasal vowel *a* raises to *u* only when the vowel and its following nasal consonant are both in the same morpheme; otherwise, raising is blocked. In (27), raising fails to occur as *a* and *N* are not in the same morpheme:

(27) *sample of words with [a] and [N] in separate morphemes*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/bæra-m/ for-me	[bæram]	*[bærum]	for me
/dust-a-m/ friend-PL-1SG.POSS	[dustam]	*[dustum]	my friends
/ba-mæze/ with-taste	[bamæze]	*[bu-mæze]	tasty
/na-mærd/ NEG-man	[namærd]	*[nu-mærd]	treacherous
/mi-xa-n/ PROG-want-they	[mixan]	*[mixun]	they want.

If in a compound word the final segment of the first lexical item ends in *a* and the second word begins with a nasal consonant, then in that compound raising is blocked due to the same factor, as shown in (28):

(28) *sample of compound words with [a] and [N] in separate morphemes*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/d͡ʒa neʃin/ place sit	[d͡ʒa neʃin]	*[d͡ʒu neʃin]	successor
/pa mal/ foot rub	[pa mal]	*[pu mal]	stampeded
/doʔa nevis/ spell write	[doʔa nevis]	*[doʔu nevis]	witch-doctor
/dærja næværd/ sea surf	[dærja næværd]	*[dærju næværd]	sailor

According to the database, without any exception, raising is blocked in words where *a* and the nasal consonant are not within the same morpheme. In other words, pre-nasal raising is sensitive to the presence of a morpheme boundary.

This phenomenon is contrary to what Kiparsky (1973, 1993), McCarthy (2003a) and Burzio (2011) define as derived environment effect (DEE). McCarthy (2003a:21) states that a derived environment effect is "a process that takes place only when its conditions are crucially met by virtue of material from two different morphemes." Likewise, Burzio (2011:2089) states that "in all cases, "environment" refers to some phonological context. Such environment or context can be "derived" in a phonological sense, by virtue of some phonological process having applied to obtain it, or in a morphological sense, by virtue of it being the result of the combination of morphemes or other morphological operation." Regarding morphologically derived environments, Burzio states that the context application of a phonological process spans across morphemes, as in *criti[s]-ism*, where the velar of *criti[k]* softens before the *i*, across a morpheme boundary. Finnish assibilation displays a similar pattern where assibilation turns *t* to *s* before *i*, but only when the latter belongs to a different morpheme, as in (29.a) and (29.b). However, assibilation is not found in (29.c) as *t* and *i* are within the same morpheme:

(29) *Finnish t > s alternation in morpheme boundary*

a. halut-a	'want-INF'	halus-i 'want-PST'	
b. tilat-a	'order-INF'	tilas-i 'order-PST'	
c. äiti	'mother'	*äisi	(Kiparsky 1973a, 1993)

In addition, Inkelas (2014:244) posits that "the parade example of a morphologically derived environment is that in which the trigger and target belong to different morphemes, and in which the relevant alternation neutralizes a phonemic contrast." TP stands out as unusual with respect to DEE, as raising in TP is found only in non-derived environments.

4.4. *a* raising before *n* within a morpheme

As discussed earlier, when *a* appears before a nasal consonant, it raises to *u*. In this section I address the environments in which *a* raises to *u* when it appears before *n*. After that, I show in what environment raising fails to occur in an *an* sequence.

As shown in (30), pre-nasal *a* in an *an* sequence undergoes raising in monosyllabic and multisyllabic lexical items. In multisyllabic words, regardless of the position of the *an* sequence, stress pattern and the number of syllables, *a* raises to *u*. "In Persian for nouns and adjectives, the main stress goes on the final syllable of the word" (Kahnemuyipour 2003:334).

(30) *sample of monosyllabic and multisyllabic words with pre-nasal raising*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/ɑn/	[un]	that
/rɑn/	[run]	leg
/ɑsɑn/	[ɑsun]	easy
/færɑvɑn/	[færɑvun]	a lot
/xɑnevɑde-gi/ family-adj marker	[xunevɑdegi]	familial

4.4.1. Phonological blocking of *an* raising within a morpheme

Following Modaressi (1978:83), I claim that if an *an* sequence is immediately followed by *u*, the pre-nasal vowel does not raise. In Modaressi's (ibid) terms "a postnasal *u* prevents the raising of the pre-nasal *a*. The language seems not to allow too much vowel harmony in one word". That is why in words like *zanu* 'knee' and *qanun* 'law' *a* does not raise to *u.*"

It is worth noting that if *a* raising in this environment is blocked it is not because the sequence of *uCu* sequence is prohibited. There are words with a *uCu* sequence like *kutule* 'dwarf' and *buxur* 'fumigation' that are actively used in TP. Moreover, in TP when there is a string *oCu* in a word, *o* can harmonize with the following /u/ and raise to *u*; so words like *d͡ʒonun* 'madness', *doruq* 'lie',

boluz ‘blouse’ and *qorub* ‘dusk’ can be pronounced as *d͡ʒunun*, *duruq*, *buluz* and *qurub* in TP. These examples, especially *d͡ʒunun*, validate that *unu* sequence is not a phonologically invalid sequence.

According to the database, there are only three lexical items *xanum*[10], *zanu* and *qanun* in TP in which raising is blocked merely due to the dissimilating *a* factor. Lexical items in which raising is inhibited due to the dissimilating *a* factor cover 0.31% of the whole data. Morphologically derived forms like *qanun-i* ‘legal’ or *xanum-ane* ‘womanly’ are excluded (see chapter five). For discussion on the blockage of raising in words like *banu* ‘lady’, *kanun* ‘center’ and *oqjanus* ‘ocean’, which are not categorized in this list, see chapter six.

4.5. *a* raising before *m* within a morpheme

As shown in the Introduction chapter, when *a* precedes *m* in TP, it raises to *u*, provided that they are in the same morpheme, as in (31):

(31) *sample of words with pre-nasal raising whose [a] and [m] are in the same morpheme*

a. /bɑdɑm/	→	[bɑdum]	‘almond’
b. /hæmɑm/	→	[hæmum]	‘bath’

4.5.1. Phonological blocking of *am* within a morpheme

Pre-nasal vowel raising does not occur in all lexical items with an *am* sequence. As mentioned before, in the literature the factors that influence raising of an *am* sequence are not distinguished from the blocking factors that inhibit raising in an *an* sequence. In this section I show that there are two blocking factors that inhibit raising in an *am* sequence that do not block raising in an *an* sequence.

[10] x*anum* is the TP pronunciation of ‘lady’; its formal Persian pronunciation is *xanom.*

First, pre-nasal raising in monosyllabic words with an *am* sequence is prohibited. Thus, raising in words in (32) fails to occur:

(32) *blockage of raising in monosyllabic words with [am] sequence*[11]

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/ʃɑm/	[ʃɑm]	*[ʃum]	supper
/dɑm/	[dɑm]	*[dum]	cattle
/vɑm/	[vɑm]	*[vum]	loan
/rɑm/	[rɑm]]	*[rum]	tame
/xɑm/	[xɑm]	*[xum]	raw

From among the whole data in the database, there are 17 monosyllabic lexical items with an *am* sequence and the pre-nasal *a* does not raise in these words. These 17 words cover 0.90% of the whole items in the database. Within this 0.90%, there are some words in which raising is blocked by more than one factor. For the discussion on the interface of this phonological blocker with other blocking factors, see chapter seven.

Modaressi (1978:108) posits that pre-nasal *a* in monosyllabic words disfavors raising. On the other hand, Jahangiri (1980:82) categorizes lexical items like *xam* 'raw', *kam* 'palate, desire' and *dam* 'trap' as cases of irregularity in which raising is blocked for no reason. I assume that *a* > *u* raising fails to occur in monosyllabic words in which *a* precedes *m* and not *n* since pre-nasal raising in monosyllabic words with an *an* sequence is not blocked, as shown in (33):

(33) *pre-nasal raising in monosyllabic words with [an] sequence*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/ɑn/	[un]	that
/dɑn/	[dun]	seed
/rɑn/	[run]	leg
/d͡ʒɑn/	[d͡ʒun]	soul
/nɑn/	[nun]	bread

[11] Except for *bam* 'roof'. For a list of exceptions see chapter seven.

The second factor that blocks raising in an *am* sequence is that in multisyllabic lexical items if *a* and *m* are not in the same syllable, raising fails to occur, as shown in (34):

(34) *blockage of raising in multisyllabic words whose [a] and [m] are in separate syllables*[12]

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/æ.lɑ.mæt/	[æ.lɑ.mæt]	*[æ.lu.mæt]	sign
/æm.mɑ.me/	[æm.mɑ.me]	*[æm.mu.me]	turban
/ɑ.mɑr/	[ɑ.mɑr]	*[u.mɑr]	statistics
/ɑ.mɑs/	[ɑ.mɑs]	*[u.mɑs]	swell
/ɑ.mel/	[ɑ.mel]	*[u.mel]	factor
/ɑ.min/	[ɑ.min]	*[u.min]	amen
/ɑ.mɑ.de/	[ɑ.mɑ.de]	*[u.mɑ.de]	ready
/bær.nɑ.me/	[bær.nɑ.me]	*[bær.nu.me]	plan
/nɑ.me/	[nɑ.me]	*[nu.me]	letter
/bɑ.mi.je/	[bɑ.mi.je]	*[bu.mi.je]	okra
/fæ.rɑ.muʃ/	[fæ.rɑ.muʃ]	*[fæ.ru.muʃ]	forget
/hɑ.me.le/	[hɑ.me.le]	*[hu.me.le]	pregnant
/he.d͡ʒɑ.mæt/	[he.d͡ʒɑ.mæt]	*[he.d͡ʒu.mæt]	bloodletting
/d͡ʒɑ.med/	[d͡ʒɑ.med]	*[d͡ʒu.med]	solid
/d͡ʒɑ.meʔ/	[d͡ʒɑ.meʔ]	*[d͡ʒu.meʔ]	comprehensive
/kɑ.mel/	[kɑ.mel]	*[ku.mel]	complete
/mɑ.mɑ/	[mɑ.mɑ]	*[mu.mɑ]	midwife
/nɑ.mus/	[nɑ.mus]	*[nu.mus]	reputation
/ʃæ.hɑ.mæt/	[ʃæ.hɑ.mæt]	*[ʃæ.hu.mæt]	audacity
/ʃɑ.mel/	[ʃɑ.mel]	*[ʃu.mel]	inclusion
/xɑ.muʃ/	[xɑ.muʃ]	*[xu.muʃ]	off
/zɑ.men/	[zɑ.men]	*[zu.men]	guarantor
/dɑ.mæn/	[dɑ.mæn]	*[du.mæn]	skirt

Hence, unlike Rohany's (2012:116) claim, blockage of raising in a word like *xa.me* 'cream' is not exceptional; an *am* sequence raises to *um* in words that are multisyllabic whose *a* and *m* are in the same syllable, as shown in (35):

[12] Except for *da.mad* 'groom' and *a.mædæn* 'to come' in which raising is not blocked. For a list of exceptions see chapter seven.

(35) *pre-nasal raising in multisyllabic words with [am] sequence*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/ɑrɑm/	[ɑrum]	peaceful
/nɑ-ɑrɑm/ not-peaceful	[nɑʔɑrum]	hysterical
/dævɑm/	[dævum]	durability
/bɑ-dævɑm/ with-durability	[bɑdævum]	durable
/bɑdɑm/	[bɑdum]	almond
/hæmɑm/	[hæmum]	bath
/kodɑm/	[kodum]	which
/tæmɑm/	[tæmum]	end
/nɑ-tæmɑm/ not-end	[nɑtæmum]	incomplete
/hæɾɑm/	[hærum]	haram

Pre-nasal *a* in multisyllabic words with an *an* sequence is not sensitive to the presence of a syllable boundary. So, if *a* and *n* are not in the same syllable, raising is not inhibited, as in (36):

(36) *pre-nasal raising in multisyllabic words with [a] and [n] in separate syllables*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/æs.rɑ.ne/	[æs.ru.ne]	afternoon tea
/xɑ.ne/	[xu.ne]	home
/bæ.hɑ.ne/	[bæ.hu.ne]	excuse
/bæ.tɑ.ne/	[bæ.tu.ne]	filler, putty
/t͡ʃɑ.ne/	[t͡ʃu.ne]	chin
/di.vɑ.ne/	[di.vu.ne]	crazy
/d͡ʒɑ.ne.vær/	[d͡ʒu.ne.vær]	animal
/d͡ʒu.ʃɑn.de/	[d͡ʒu.ʃun.de]	decoction
/lɑ.ne/	[lu.ne]	nest
/mær.dɑ.ne/	[mær.du.ne]	manly
/mɑh.jɑ.ne/	[mɑh.ju.ne]	monthly
/ne.ʃɑ.ni/	[ne.ʃu.ni]	address
/pi.ʃɑ.ni/	[pi.ʃu.ni]	forehead
/raz.jɑ.ne/	[rɑz.ju.ne]	fennel
/sæl.mɑ.ni/	[sæl.mu.ni]	barber shop

/ʃɑ.ne/	[ʃu.ne]	shoulder
/ʃir.vɑ.ni/	[ʃir.vu.ni]	tin roof
/xæ.zɑ.ne/	[xæ.zu.ne]	treasury
/xɑ.ne.vɑ.de/	[xu.ne.vɑ.de]	family
/sob.hɑ.ne/	[sob.hu.ne]	breakfast
/t͡ʃe.rɑ.qɑ.ni/	[t͡ʃe.rɑ.qu.ni]	decoration with lamps
/zæ.nɑ.ne/	[zæ.nu.ne]	womanly

On the other hand, it is not clear whether the postnasal *u* which blocks raising in words with an *an* sequence when it immediately appears after *n* is a blocking factor for an *am* sequence or not. The reason is that in all lexical items in which there is the *amu* sequence, *a* and *m* are in separate syllables. Words like *xa.mush* 'off' or *ka.ra.muz* 'trainee', among others, are examples that verify this observation. The reason for this type of syllabification is that in Persian every syllable must have an onset; thus, it is not possible to syllabify *xamush* as *xam.ush* and *karamuz* as *kar.am.uz*. That is why in words with an *amu* sequence *a* is always the final segment of one syllable and *m* is always the onset of the following syllable. Hence, there is always the syllable boundary as a blocking factor that prevents raising in multisyllabic words with an *amu* sequence.

Monomorphemic native Persian words with *am.Cu* sequence like *gam.bu* 'fat' and *bam.bul* 'deceit' whose *am* sequence is in the same syllable and is followed by *u* resist raising. This implies that vowel harmony does not occur in such words. In the database there are not any monomorphemic words with *anCu* sequence. So, it is not possible to ascribe this generalization to words with an *an* sequence.

4.6. The phonological analysis in this study vs the analyses in the literature

As discussed in the literature review chapter, there are various analyses that attempt to identify the phonological factors that block pre-nasal raising. That is, for one phonological phenomenon there are different and even sometimes contradictory analyses which face counterexamples. To the

extent that this study identifies the blocking factors correctly, it illuminates that the data in earlier analyses pattern systematically, and the proposed analysis accounts for the phonological patterning.

Given this, contrary to Modaressi's (1978) claim that due the word-initial position of the *aN* sequence raising is blocked in (37), I claim that in these words raising fails to occur since *a* and *m* are not in the same syllable.

(37) *sample of blockage of raising in am-initial words*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/ɑ.mɑr/	[ɑ.mɑr]	*[u.mɑr]	statistics
/ɑ.mix.tæn/	[ɑ.mix.tæn]	*[u.mix.tæn]	to mix
/ɑ.mɑ.de/	[ɑ.mɑ.de]	*[u.mɑ.de]	ready

The same analysis identifies blockage of raising in words like *hamele* 'pregnant'. Kahn and Bernstein (1980) claim ***humele* is unacceptable as *a* and *m* are not in a stressed syllable (recall that stress is generally final in TP). They posit that pre-nasal *a* raising occurs in stressed syllables in final or penultimate position. However, I claim that the fact that *a* and *m* are not in the same syllable is the source of blockage of raising in this lexical item. In addition, in (38) there are words whose pre-nasal *a* raises to *u* even though the *aN* sequence is not in a stressed syllable.

(38) *words with pre-nasal raising whose aN sequence is in an unstressed syllable*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/xɑ.ne.vɑ.dé/	[xu.ne.vɑ.dé]	family
/d͡ʒɑ.ne.vǽr/	[d͡ʒu.ne.vǽr]	animal

Jahangiri (1980:79) observes that there is less raising when *a* appears before *m* compared to when *a* appears before *n*. He assumes the reason for this difference could be the stress pattern. But the above analysis shows that stress pattern does not influence pre-nasal raising.

Miller (2011:1387) posits that an inhibiting factor for raising of *a* may be dissimilation. For example, when two raising environments occur in a word, usually only the final one, which is stressed, undergoes raising:

(39) *Miller's sample of words with two aN sequence in each word*

a. /sɑmɑn/ → [sɑmun] 'order',
b. /ɑmijɑne/ → [ɑmijune] 'colloquial'

Again, I posit that stress is not influential in pre-nasal raising. In (39.a) raising in *an* sequence is observed because there are not any blocking factors which prevent raising; however, in the *am* sequence *a* does not raise since *a* and *m* are not in the same syllable. Hence, the same analysis is applicable here.

In (39.b), the factor which prohibits raising in the *am* sequence is a morphological one. For discussion on blockage of raising in (39.b), see the morphological factors chapter.

4.7. Summary

In this chapter, I have attempted to provide an explanation for pre-nasal raising phenomenon and identify the phonological blocking factors; in addition, this analysis provides an account for the data that are irregularities or counterexamples to the generalizations proposed in the literature.

In the following chapters I discuss the non-phonological factors that influence pre-nasal raising in TP. These factors are classified into two main categories, morphological and lexical factors. In fact, the lexical items in which raising is blocked due to non-phonological factors are exceptions to the phonological generalizations, with raising blocked in unexpected environments.

CHAPTER FIVE

MORPHOLOGICAL FACTORS

5.1. Introduction

This chapter addresses the of role of morphology in *a* > *u* raising in Tehrani Persian. The main goal is to examine what morphological factors exist that influence raising; in other words, it is an attempt to show that pre-nasal raising is not solely a phonological process; rather, there are environments in which raising is influenced by morphology. One of those was discussed in chapter four – it only applies within a morpheme. Additional morphological factors are addressed in this chapter.

The structure of this chapter is as follows. First, I discuss inflectional affixes and pre-nasal raising. Following that, I address derivational affixes and exceptional suffixes. After that, I consider words where pre-nasal raising is sensitive to a syllable boundary and investigate whether re-syllabification after affixation influences raising in these words. The final section is dedicated to the topic of homophonous affixes and their role in pre-nasal raising.

5.2. Inflectional affixes and pre-nasal raising

In this section I discuss the patterning of pre-nasal raising in suffixes with an *aN* sequence.

5.2.1. Pre-nasal raising in inflectional affixes with an *aN* sequence

In Persian there are three categories of inflectional affixes that contain an *aN* sequence: the plural personal pronouns, the plural marker and the causativizer.

5.2.1.1. Plural personal pronouns

Pre-nasal *a* in the plural personal pronouns *-man* (1PL), *-tan* (2PL), *-ʃan* (3PL), undergoes raising, as shown in (40):

(40) *pre-nasal raising in plural personal pronouns*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/kelɑs-emɑn[13]/ → classroom-our	[kelɑs-emun]	our classroom
/bɑq-etɑn/ → garden-your.PL	[bɑq-etun]	your (pl) garden
/pedær-eʃɑn/ → father-their	[pedær-eʃun]	their father

5.2.1.2. Plural maker

The other inflectional affix which contains an *aN* sequence is the plural marker. In formal Persian *-an* and *-ha* are plural markers[14] (41.a and 41.b), but in TP *-a* is used as the plural marker (41.c) unless the pluralized word is vowel-final. In that circumstance, *-ha* is used as the plural maker, as shown in (41.d):

(41) *plural markers in FP and TP*

a. kif-hɑ	b. zæn-ɑn	c. mɑʃin-ɑ	d. bæste-hɑ
bag-PL	woman-PL	car-PL	'package-PL'
'bags'	'women'	'cars	'packages'

The pre-nasal *a* in *-an* plural marker affix does not raise because this suffix is used only in formal Persian; in other words, this plural marker does not exist in TP.

[13] When the plural personal pronouns attach to a vowel-final word, the *e* vowel in the suffix is deleted: *baba-tun* and not **baba-etun* 'your (pl) father' because hiatus is not tolerated in Persian. Hence, the underlying form of these personal pronouns could be *-eman*, *-etan*, and *-eʃan*. Note that the underlying forms are determined by the forms in FP.

[14] Arabic plural markers like *-at* as in *sæbzi-d͡ʒ-at* 'vegetables', *-in* as in *moʔælem-in* 'teachers' and *-un* as in *enqelaby-un* 'revolutionaries' are usually used in formal Persian, and not in TP.

5.2.1.3. Causativizer

The final inflectional affix which contains an *aN* sequence is the causative marker *-an*. In this suffix pre-nasal *a* undergoes raising as long as the causativized form of the verb is used in TP (42); otherwise, raising is inhibited in that affix[15] since the inflected form is formal Persian, as in (43):

(42) *pre-nasal raising in causative marker -an*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/ræqs + -ɑn/ →	[ræqsun]	cause to dance
/puʃ + -ɑn/ →	[puʃun]	cause to wear
/t͡ʃærx + -ɑn/ →	[t͡ʃærxun]	cause to turn
/t͡ʃek + -ɑn →	[t͡ʃekun]	cause to drip
/dæv + -ɑn/ →	[dævun]	cause to run
/d͡ʒonb + -ɑn/ →	[d͡ʒonbun]	cause to move
/fæhm + -ɑn/ →	[fæhmun]	cause to understand
/d͡ʒuʃ + -ɑn/ →	[d͡ʒuʃun]	cause to boil
/neʃɑn + -ɑn/ →	[neʃun]	cause to sit
/pær + -ɑn/ →	[pærun]	cause to jump/fly
/pɑʃ + -ɑn/ →	[pɑʃun]	cause to splash
/qælt + -ɑn/ →	[qæltun]	cause to roll
/rænd͡ʒ + -ɑn/ →	[rænd͡ʒun]	cause to bother
/ʃurɑn + -ɑn/ →	[ʃurun]	cause to rebel
/suz + -ɑn/ →	[suzun]	cause to burn
/xɑb + -ɑn/ →	[xɑbun]	cause to sleep
/xor + -ɑn/ →	[xorun]	cause to eat

(43) *blockage of pre-nasal raising in causative marker -an*

UNDERLYING FORM	SURFACE FORM		GLOSS
/kut͡ʃ + -ɑn/ →	[kut͡ʃɑn]	* [kut͡ʃun]	cause to migrate
/ræhɑ + -ɑn/ →	[ræhɑn]	* [ræhɑʔun]	to free somebody
/ʃekɑf + -ɑn/ →	[ʃekɑfɑn]	* [ʃekɑfun]	cause to split
/bærq + -ɑn/ →	[bærqɑn]	*[bærqun]	cause to glitter
/bɑvær + -ɑn/ →	[bɑværɑn]	*[bɑværun]	cause to believe
/ ru-j[16] + -ɑn/ →	[rujɑn]	*[rujun]	cause to grow

[15] Unfortunately, I have not been able to find a verb stem with an *aN* sequence which can take causativizer *-an* so that I could observe how that verb behaves in the environment where there are two *aN* sequences in one verb.
[16] The glide *j* is epenthesized to avoid hiatus.

5.2.2. Influence of inflectional affixes on raising of the base

This section addresses the influence of inflectional affixes on pre-nasal raising in words with an *aN* sequence. First, the role of plural personal pronouns is discussed; then the plural markers and finally the influence of the causativizer is addressed.

The main claim in this section is that these affixes do not have a role in raising. If pre-nasal *a* in the base undergoes raising, the inflected form raises as well; similarly, when an uninflected lexical item resists raising, pre-nasal *a* in the inflected form does not raise either.

5.2.2.1. Role of personal pronouns

If a word hosts one of the personal pronouns and that word contains an *aN* sequence, the pre-nasal *a* in the stem raises to *u* provided that the *aN* sequence in that word meets the phonological, morphological and lexical requirements for raising. For instance, in (44.a), pre-nasal *a* in the host and in the affix raise, since there are not any morphophonological or lexical restrictions for raising. In (44.b), the pre-nasal *a* in personal pronoun undergoes raising, but raising in the stem is blocked as *a* and *m* are not in the same syllable. Likewise, in (44.c), the personal pronoun's pre-nasal *a* raises, but the pre-nasal *a* in the base resists raising since the lexical item is a non-Arabic loanword (for discussion on loanwords, see chapter six):

(44) *bases with an sequence hosting plural personal pronouns*

UNDERLYING FORM	SURFACE FORM		GLOSS
/t͡ʃu.pɑ.n-e.tɑn/→ shephard-your.PL	[t͡ʃu.pu.ne.tun]	----------	your (pl) shepherd
/dɑ.mæ.n-ɑ.-ʃɑn/→ skirt-PL-their	[dɑ.mæ.nɑ.ʃɑn]	*[du.mæ.nɑ.ʃun]	their skirts
/li.vɑ.n-e.mɑn/→ glass-our	[li.vɑ.ne.mɑn]	*[li.vu.ne.mun]	our glass

5.2.2.2. Role of plural markers

Like personal pronouns, the TP plural marker *-a* does not have any influence on blockage of raising in its host. If the base is a TP lexical item that meets the morphophonological requirements of pre-nasal raising, then with or without the plural marker, raising occurs. In (45.a), since the host of the plural marker is a monosyllabic word with an *am* sequence, raising is blocked. And in (45.b) pre-nasal *a* does not raise as the *an* sequence is preceded by *u.* In (45.c) there are not any blocking factors for raising and hence *a* raises to *u*:

(45) *sample of words with aN sequence hosting TP plural marker -a*

UNDERLYING FORM	SURFACE FORM		GLOSS
a) /vɑm-ɑ/→ loan-PL	[vɑmɑ]	*[vumɑ]	loans
b) /qɑnun-ɑ/→ law- PL	[qɑnuna]	*[qununa]	laws
c) /mehmɑn-ɑ/→ guest-PL	[mehmun-ɑ]		guests

The database does not have any causativized verb which contains an *an* sequence, except for the causative marker; hence, the influence of the causativizer on the patterning of the pre-nasal *a* in the base is not observable.

5.3. Interim conclusion

As shown above, inflectional affixes do not have any influence on blocking pre-nasal raising; in fact, if a lexical item is used in TP, then its inflected forms undergo raising. Other inflectional affixes like tense markers, the imperative marker and the negative do not play a role in the occurrence or prohibition of this phenomenon. The examples in (46) verify this claim:

(46) *sample of words with aN sequence hosting different inflectional affixes*

	TP LEXICAL ITEMS	FORMAL PERSIAN LEXICAL ITEMS	
PAST TENSE MARKER *-d*	/xɑn-**d**/ → [xund] read-PST 'read'	/setɑn-**d**/ → [setɑnd] get-PST 'ignited (tr)'	*[setund]
IMPERATIVE MARKER *be-*	/**be**-xɑn/ → [bexun] IMP-read 'You (sg) read!'	/**be**-setɑn/ → [besetɑn] IMP-get 'You (sg) get!'	*[bestun[17]]
NEGATIVE MARKER *næ-*	/**næ**-xɑn/ → [**næ**xun] NEG-read 'You (sg) do not read!'	/**næ**-setɑn/ → [næsetɑn] NEG-get 'You (sg) do not get!'	*[næsetun]

5.4. Derivational affixes

In this section I discuss the role of derivational affixes in allowing or prohibiting pre-nasal raising. My main claim is if a morphologically derived word is a lexical item that is not used in formal Persian and is used by TP speakers, then pre-nasal *a* is allowed to raise to *u.* But if the result of affixation is a word which does not exist in TP lexicon (i.e. it is a formal word), raising is blocked There are non-formal morphologically derived words which are used by TP speakers in which raising is blocked. These lexical items are either recent coinages or non-Arabic loanwords hosting Persian affixes. These cases will be investigated later in this chapter.

This section is divided into three subsections. The first subsection addresses the role of derivational affixes that do not contain an *aN* sequence on pre-nasal raising in hosts that contain an *an* sequence. The second subsection investigates the patterning of pre-nasal *a* in derivational affixes when they attach to hosts that do not contain an *aN* sequence. The last subsection is about derived forms whose base and the affix both contain an *aN* sequence. As mentioned above, regardless of the position of *aN* sequence, what determines occurrence or prohibition of pre-nasal

[17] This verb is acceptable in Isfahani accent of Persian but is not used in TP.

raising is that if the morphologically derived is a part of TP lexicon, raising in that lexical item is not inhibited.

The derivational affixes that are discussed in this chapter are listed in (47):

(47) *list of derivational affixes investigated in this study*

AFFIX	FUNCTION	PREFERRED BASE	DATA	GLOSS
-ɑn	adj/adv marker	verb	/d͡ʒuʃ-ɑn/ boil-ing	boiling
-ɑne	adj/adv marker	noun/adj	/mɑh-ɑne/ month-ly	monthly
-ɑni	adj/noun marker	noun	/nur-ɑni/ light-adj marker	luminous
-ænde	noun marker	verb	/xɑn-ænde/ sing-agentive	singer
bɑ-	adj marker	noun	/bɑ-mohæbæt/ with-kindness	kind
-bɑn	noun marker	noun	/bɑq-bɑn/ garden-agentive	gardener
bi-	negative marker	noun/adj	/bi-nɑm/ without-name	anonymous
-dɑn	noun marker	noun	/gol-dɑn/ flower-place	vase
-e	noun marker	noun	/dɑmæn-e/ skirt-noun maker	domain
-eʃ	noun marker	verb	/dɑn-eʃ/ know-noun marker	science
-estɑn	noun marker	noun	/ʃæhr-estɑn/ city-place	town
-gɑh	temporal/location adverb	noun	/ʃɑm-gɑh/ supper-temp. adv	nighttime
-gɑne	adj/adv marker	noun/adj	/d͡ʒodɑ-gɑne/ separate-ly	separately
-gær	adj/noun marker	noun	/toqjɑn-gær/ rebellion-maker	rebellious
-gi	noun marker	adj	/divɑne-gi/ crazy-noun marker	craziness
-i	adj marker	noun	/hejvɑn-i/ animal-like	animal-like
-i	noun marker	adj	/d͡ʒævɑn-i/ young-noun marker	youth
-ijæt	noun marker	noun	/tæmɑm-ijæt/ end-noun marker	entirety
-mænd	possession marker	noun	/dɑneʃ-mænd/ science-possessor	scientist
-mɑn	noun marker	verb	/sɑz-mɑn/	building

			build-noun marker	
nɑ-	negative marker	adj/verb	/nɑ-dɑn/ not-know	idiot
-vær	adj marker	noun	/nɑm-vær/ name-noun marker	famous
-vɑr	adv marker	adj/noun	/divɑne-vɑr/ crazy-ly	crazily

The adjective/noun marker *-an* is homophonous with the causative marker *-an.* For discussion of this suffix, see section 5.7 in this chapter.

5.4.1. Bases with an *aN* sequence hosting affixes without an *aN* sequence

5.4.1.1. Bases that undergo raising before and after affixation

As mentioned above, if a word with an *aN* sequence which is used in TP hosts a derivational affix, and as a result of that affixation the derived lexical item is a word which is used in TP, then pre-nasal *a* raises to *u* (48).

(48) *list of words with pre-nasal raising before and after hosting a derivational affix*

NON-DERIVED FORM IN TP		DERIVED FORM IN TP	
/d͡ʒævɑn/ →	[d͡ʒævun] ‘young’	/d͡ʒævɑn-i/ → young-noun marker	[d͡ʒævun i] ‘youth’
/bɑdɑm/ →	[bɑdum] ‘almond’	/bɑdɑm-i/ → almond-adj marker	[bɑdumi] ‘almond-like’
/gerɑn/ →	[gerun] expensive’	/gerɑn + -i/ → expensive-noun marker	[geruni] ‘expensiveness’
/hærɑm/	[hærum] ‘haram’	/hærɑm-i/ haram-adj marker	[hærumi] ‘misbegotten’
/dærmɑn/ →	[dærmun] ‘cure’	/bi-dærmɑn/ → without-cure	[bidærmun] ‘incurable’
/dævɑm/ →	[dævum] ‘duration’	/bi-dævɑm/ → without-duration	[bidævum] ‘unsubstantial’
/dɑne/ →	[dune] ‘seed’	/bi-dɑne/ → without-seed	[bidune] ‘seedless’

/dændɑn/ →	[dændun] ‘tooth’	/bi-dændɑn/ → without-tooth	[bidændun] ‘toothless’
/d͡ʒɑn/ →	[d͡ʒun] ‘soul’	/bi-d͡ʒɑn/ → without-soul	[bid͡ʒun] ‘dead’
/zæbɑn/ →	[zæbun] ‘tongue’	/zæbɑn-e/ → tongue-noun marker	[zæbune] ‘flame’
/dændɑn/ →	[dændun] ‘tooth’	/dændɑn-e/ → tooth-noun marker	[dændune] ‘cog’
/neʃɑn/ →	[neʃun] ‘sign’	/neʃɑn-e/ → sign-noun marker	[neʃune] ‘indication’
/pejmɑn/ →	[pejmun] ‘oath’	/pejmɑn-e/ → oath-noun marker	[pejmune] ‘chalice’
/rævɑn/→	[rævun] ‘flowing’	/rævɑn-e/ → flowing-noun marker	[rævune] ‘delegated’
/virɑn/ →	[virun] ‘ruin’	/virɑn+-e/ → ruined-noun marker	[virune] ‘ruined place’

5.4.1.2 Bases that undergo raising before affixation and resist raising after affixation

There are words in which pre-nasal *a* undergoes raising, but if an affix attaches to them, raising is blocked. In (49) there are words with *aN* sequence in which *a* is allowed to raise, but after affixation, pre-nasal raising is blocked because the derived lexical item is a word that is used in formal Persian and not in TP:

(49) *sample of words with pre-nasal raising before affixation but with blockage of raising after hosting an affix*

NON-DERIVED FORM IN TP		DERIVED FORM IN FP			
/hejvɑn/ →	[hejvun] ‘animal’	/hejvɑn-i/ → animal-adj marker	[hejvɑni]	*[hejvuni][18]	animal-like
/ɑn/ →	[un] ‘that’	/ɑn-gɑh/ → that-temporal adv	[ɑn-gɑh]	*[ungɑh]	then

[18] Note that raising in this word occurs when it is a term of endearment, as in *axej hejvuni!*

/tæmɑm/[19] →	[tæmum] ‘end’	/tæmɑm-ijæt/ → end-noun marker	[tæmɑmijæt]	*[tæmumijæt]	entirety
/ɑrɑm/ →	[ɑrum] ‘calm’	/ɑrɑm-gɑh/ → calm-place	[ɑrɑm-gɑh]	*[ɑrum-gɑh]	cemetery
/sɑmɑn/ →	[sɑmun] ‘order’	/sɑmɑn-e/ → system-noun marker	[sɑmɑne]	*[sɑmun-e]	system

Samane does not undergo raising as it is a newly coined word and also not used in TP (the loanword *sistem* ‘system’ is used in TP). In addition, there are rare cases in which a base’s pre-nasal *a* resists raising before affixation, but after affixation the *a* undergoes raising, as shown in (50):

(50)

a. /zæmɑn/ → [zæmɑn] ‘time’ *[zæmun], but
b. /zæmɑn + -e/ → [zæmun-e] ‘era’

Note that *zæmane* is more colloquial than *zæman*.

5.4.1.3. Bases that resist raising before and after affixation

There are formal Persian words with an *aN* sequence that host derivational affixes, and it is unusual for pre-nasal *a* in these words to undergo raising after affixation. The reason is that usually derivations from a formal Persian word belong to FP lexicon, and hence these derived forms, like their bases, are not used in TP. For instance, examples in (51) are non-derived forms in which raising is blocked before they host an affix. In (52) the same words of (51) are listed with an affix attached to them, but pre-nasal *a* in the base still resists raising:

(51) *sample of non-derived forms of words with blockage of pre-nasal raising*

UNDERLYING FORM	SURFACE FORM		GLOSS
/færd͡ʒɑm/ →	[færd͡ʒɑm]	*[færd͡ʒum]	end
/legɑm/ →	[legɑm]	*[legum]	harness

[19] For discussion of the influence of affixation on re-syllabification of words with an *am* sequence, see section 5.6. of this chapter.

/tɑzjɑne/ →	[tɑzjɑne]	*[tɑzjune]	whip
/toqjɑn/ →	[toqjɑn]	*[toqjun]	riot
/dæhɑn/ →	[dæhɑn]	*[dæhun]	mouth
/kæræn/ →	[kæræn]	*[kærun]	limit
/ɑstɑn/ →	[ɑstɑn]	*[ɑstun]	threshold
/hengɑm/ →	[hengɑm]	*[hengum]	moment
/mɑnænd/ →	[mɑnænd]	*[munænd]	resembling
/xæræm/ →	[xæræm]	*[xærum]	‘elegant walk’

(52) *sample of derived forms of words in (50) with blockage of pre-nasal raising*

UNDERLYING FORM	SURFACE FORM		GLOSS
/bi-færd͡ʒɑm/ → without-end	[bifærd͡ʒɑm]	*[bifærd͡ʒum]	fruitless
/bi-legɑm/ → without-harness	[bilegɑm]	*[bilegum]	harnessless
/tɑzjɑne-vɑr/ → whip-like	[tɑzjɑnevɑr]	*[tɑzjunevɑr]	whip-like
/toqjɑn-gær/ → rebel-adj marker	[toqjɑngær]	*[toqjungær]	rebellious
/dæhɑn-e/ → mouth-noun marker	[dæhɑne]	*[dæhune]	mouth
/kæræn-e/ threshold-noun marker	[kærane]	*[kærune]	threshold
/ɑstɑn-e/ → threshold-noun marker	[ɑstɑne]	*[ɑstune]	threshold
/bi-hengɑm/ → without-moment	[bihengɑm]	*[bihengum]	untimely
/bi-mɑnænd/ → without-similarity	[bimɑnænd]	*[bimunænd]	unique
/xæræm-ænde/→ elegant walk-agentive	[xæræmænde]	*[xærumænde]	one who walks elegantly

5.4.2. Bases without an *aN* sequence hosting affixes containing an *aN* sequence

There are words which do not contain an *aN* sequence, but they host derivational affixes that have pre-nasal *a*. In these cases, the aforementioned generalization is still valid; if the result of affixation is a word which is used by TP speakers, then raising is allowed; otherwise, the pre-nasal *a* does

not raise. The examples below from (53) to (59) illustrate this. The derived words that are in the first column are used in TP and hence pre-nasal *a* in the affixes raises. In the second column, those examples are listed in which raising is blocked as these are formal words that are not used in TP:

(53) *TP and FP words hosting -estan suffix*

TP WORDS: ɑ → u		FORMAL PERSIAN WORDS: ɑ → *u			
/ʃæhr-estɑn/→ city-place	[ʃæhrestun] 'town'	/ræz-estɑn/→ vine-place	[ræzestɑn]	*[ræzestun]	vineyard
/qæbr-estɑn/→ grave-place	[qæbrestun] 'cemetery'	/bot-estɑn/→ temple-place	[botestɑn]	*[botestun]	idols temple
/gol-estɑn/→ flower-place	[golestun] 'flower garden'	/næhɑl-estɑn/→ treelet-place	[næhɑlestɑn]	*[næhɑlestun]	treelet garden
/tɑk-estɑn/→ vine-place	[tɑkestun] 'vineyard'				

(54) *TP and FP words hosting -an 'adjective adverb marker' suffix*

TP WORDS: ɑ → u		FORMAL PERSIAN WORDS: ɑ → *u			
/xænd-ɑn/ → laugh-adj/adv m[20]	[xændun] 'laughing'	/d͡ʒæh-ɑn/ → hop-adj/adv m	[d͡ʒæhɑn]	*[d͡ʒæhun]	hopping
/keʃ-ɑn/[21] → drag-adj/adv m	[keʃun] 'dragging'	/bæhɑr-ɑn/ → spring-adj/adv m	[bæhɑrɑn]	*[bæhɑrun]	spring time
/dæv-ɑn/[22] → run-adj/adv m	[dævun] 'running'	/ʃɑm-gɑh-ɑn/ → supper-time-adj/adv m	[ʃɑmgɑhɑn]	*[ʃɑmgɑhun][23]	nighttime

(55) *TP and FP words hosting -ane 'adjective/adverb marker' suffix*

TP WORDS: ɑ → u		FORMAL PERSIAN WORDS: ɑ → *u			
/zæn-ɑne/ → woman-ly	[zænune] 'womanly'	/bɑno-v[24]-ɑne/ → woman-ly	[bɑnovɑne]	*[bɑnovune]	womanly
/mærd-ɑne/ → man-ly	[mærdune] 'manly'	/gostɑx-ɑne/ → impudent-ly	[gostɑxɑne]	*[gostɑxune]	impudently

[20] *adj/adv m* is the abbreviated form of *adjective/adverb marker*.
[21] As in *kesh-an kesh-an* 'dragging'
[22] As in *dæv-an dæv-an* 'while running'
[23] **ʃumgahan* and **ʃumgahun* are unacceptable.
[24] Here *v* seems to be an epenthetic consonant to avoid hiatus.

/æsr-ɑne/ → evening-ad/adv m	[æsrune] ‘afternoon tea’	/ɑgɑh-ɑne/ → know-adj/adv m	[ɑgɑhɑne]	*[ɑgɑhune]	knowingly
/ɑʃeq-ɑne / → lover-ly	[ɑʃequne] ‘amorously’	/ɑref-ɑne/ → mystic-ly	[ɑrefɑne]	*[ɑrefune]	gnostic
/div-ɑne/ → beast-like	[divune] ‘crazy’	/kur kur-ɑne/ → blind blind-ly	[kur kurɑne]	*[kur kurɑne]	blindly
/mɑh-i-ɑne/ → month-INDF-ly	[mɑhjune] ‘monthly’	/mobtædi-ɑne/ → amateur-ly	[mobtædjɑne]	*[mobtædjune]	amateurly
/sɑl-i-ɑne/ → year-INDF-ly	[sɑljune] ‘annual’	/moluk-ɑne/ → king-ly	[molukɑne]	*[molukune]	kingly
/ʃæb-ɑne/ → night-ly	[ʃæbune] ‘nightly’	/sær-ɑne/ → head-adj/adv m	[særɑne]	*[særune]	per capita
/sobh-ɑne/ → morning-adj/adv m	[sobhune] ‘breakfast’	/sufi-ɑne/ → Sufi-like	[sufjɑne]	*[sufjune]	Sufi-like
		/tefl-ɑne/ → child-ish	[teflɑne]	*[teflune]	childish
		/d͡ʒæsur-ɑne/ → bold-ly	[d͡ʒæsurɑne]	*[d͡ʒæsurune]	boldly
		/dælir-ɑne/ → bold-ly	[dælirɑne]	*[dælirune]	boldly
		/motehæver-ɑne/ → bold-ly	[motehæverɑne]	*[motehæverune]	boldly

(56) *TP and FP words hosting -ban ‘agentive’ suffix*

TP WORDS: ɑ → u		FORMAL PERSIAN WORDS: ɑ → *u			
/ɑsjɑb-bɑn/ → mill-agentive	[ɑsjɑbun] ‘miller’	/bɑd͡ʒ-bɑn/ → tax-agentive	[bɑd͡ʒbɑn]	*[bɑd͡ʒbun]	tax collector
/dærvɑze-bɑn/ → goal-agentive	??[dærvɑzebun] ‘goalkeeper’	/deʒ-bɑn / → fortress-agentive	[deʒbɑn]	*[deʒbun]	fortress guard
/bɑq-bɑn/ → garden-agentive	[bɑqbun] ‘farmer’	/mærz-bɑn/ → border-agentive	[mærzbɑn]	*[mærzbun]	border guard
/bɑd-bɑn/ → wind-agentive	[bɑdbun] ‘sail’	/miz-bɑn/ → table-agentive	[mizbɑn]	*[mizbun]	host
/dær-bɑn/ → door-agentive	[dærbun] ‘doorman’	/poʃt-i-bɑn/ → back-agentive	[poʃtibɑn]	*[poʃtibun]	school consultant
/nærde-bɑn/ → bar-agentive	[nærdebun] ‘ladder’	/sɑr-bɑn/ → cattle-agentive	[sɑrbɑn]	*[sɑrbun]	caravan guide
/mehr(æ)-bɑn/ →	[mehr(æ)bun] ‘kind’	/ʃotor-bɑn →	[ʃotorbɑn]	*[ʃotorbun]	camel rider

kindness-agentive		camel-agentive			
/negæh-bɑn/ → guard-agentive	??[negæhbun] ‘guard’	/pɑrk-bɑn/ → park-agentive	[pɑrkbɑn]	*[pɑrkbun]	parking lot guard
/pɑs(e)-bɑn/ → watch-agentive	[pɑs(e)bun] ‘guard’	/sængær-bɑn/ → trench-agentive	[sængær-bɑn]	*[sængærbun]	trench guard
/sɑje-bɑn/ → shade-agentive	[sɑjebun] ‘shade’				

(57) *TP and FP words hosting -dan ‘place’ suffix*

TP WORDS: ɑ → u		FORMAL PERSIAN WORDS: ɑ → *u			
/ɑʃqɑl-dɑn/ → garbage-place	[ɑʃqɑldun] ‘wastebasket’	/ɑtæʃ-dɑn/ → fire-place	[ɑtæʃdɑn]	*[ɑtæʃdun]	fire-box
/bæt͡ʃe-dɑn/ → child-place	?[bæt͡ʃedun] ‘womb’	/tir-dɑn/ → arrow-place	[tirdɑn]	*[tirdun]	arrow case
/t͡ʃerɑq-dɑn/ → lamp-place	[t͡ʃerɑqdun] ‘lamp holder’	/toxm-dɑn/ → egg-place	[toxmdɑn]	*[toxmdun]	ovary
/qælæm-dɑn/ → pen-place	[qælæmdun] ‘pencil holder’	/zeh-dɑn/ → stomach-place	[zehdɑn]	*[zehdun]	womb
/kɑh-dɑn/ → straw-place	[kɑhdun] ‘straw-rick’				
/næmæk-dɑn/ → salt-place	[næmækdun] ‘salt shaker’				
/ʃekær-dɑn/ → sugar-place	[ʃekærdun] ‘sugar shaker’				
/gol-dɑn/ → flower-place	[goldun] ‘vase’				

(58) *TP and FP words hosting -gane ‘adjective/adverb marker’ suffix*

TP WORDS: ɑ → u		FORMAL PERSIAN WORDS: ɑ → *u			
/bæt͡ʃe-gɑne/ → child-adj marker	[bæt͡ʃegune] ‘childish’	/jek-gɑne/ > /jegɑne/ → one-adj marker	[jegɑne]	*[jegune]	unique
/d͡ʒodɑ-gɑne/ → separate-ly	[d͡ʒodɑgune] ‘separately’	/t͡ʃænd-gɑne/ → multi-adj marker	[t͡ʃændgɑne]	*[t͡ʃændgune]	manifold
		/cardinal number-gɑne/ → cardinal number-adj m	[cardinal number-gɑne]	*[cardinal number-gune]	----

(59) *TP and FP words hosting -man 'noun marker' suffix*

TP WORDS: ɑ → u		FORMAL PERSIAN WORDS: ɑ → *u			
/sɑxt(e)-mɑn/ → built-noun marker	[sɑxt(e)mun] 'building'	/goft(e)-mɑn/ → said-noun marker	[goft(e)mɑn]	*[goft(e)mun]	dialogue
/t͡ʃid(e)-mɑn/ → decorated-noun marker	[t͡ʃid(e)mun] 'design'	/sɑz-mɑn/ → built-noun marker	[sɑzmɑn]	*[sɑzmun]	organization
/zɑj-mɑn/ → birth give-noun marker	[zɑjmun] 'childbirth'	/jɑd-mɑn/ → remembrance-noun marker	[jɑdmɑn]	*[jɑdmun]	memorial
		/pors(e)-mɑn/ → question-noun marker	[pors(e)mɑn]	*[pors(e)mun]	questioning
		/duxt(e)-mɑn/ → sewed-noun marker	[duxt(e)mɑn]	*[duxt(e)mun]	sewing

From the examples of the above table it can be deduced that no affix blocks raising, it is only a matter of formality; if the derived word is a formal Persian lexical item, then raising is blocked; otherwise, pre-nasal *a* is allowed to raise.

Before moving to the next subsection, regarding the above data I should clarify that proper names that contain affixes with an *aN* sequence like *æfqan-estan* 'Afghanistan', *mæd͡ʒar-estan* 'Hungary', *bælut͡ʃ-estan* 'Baluchistan', *engel-estan* 'England', etc. are not listed above. The reason is that these proper names are not formal words; they are lexical items that are used in TP, but they resist raising because they are proper names (see chapter six). Moreover, some of them are recently coined proper names (like *mæd͡ʒar-estan* 'Hungary', *pak-estan* 'Pakistan' and *ærmæn-estan* 'Armenia') which means there are two blocking factors that disallow raising in a single lexical item.

Similarly, there are newly coined words that contain one of the above affixes like *-estan* as in *færhæng-estan* 'academy' and *-ane* as in *raj-ane* 'computer' which are recent coinages and are not listed above. Pre-nasal *a* in these words does not raise as they are newly coined lexical items. Hence, if raising in them is blocked, it does not necessarily mean they are formal words; rather, they are used in TP, but with pre-nasal *a* which does not raise because of not being recently coined.

For further discussion regarding pre-nasal raising in relation to recent coinages and proper names see chapter six.

5.4.3. Bases with an *aN* sequence hosting affixes with an *aN* sequence

We have seen cases of words with an *aN* sequence in the stem and cases of words with an *aN* in affix. Another possibility is that both the stem and the affix contain an *aN* sequence. In this regard, again, no matter which component part of the derived word contains *aN* sequence, raising depends on the formality of the morphologically derived lexical item. Examples in (60) show that it does not matter if an *aN* sequence in the base before affixation undergoes raising or not. If a derived form is part of the formal Persian lexicon and not used or infrequent in TP, raising is blocked:

(60) *derived forms with two aN sequences; one in the base and another in the affix*

UNDERLYING FORM	UNGRAMMATICAL FORM	GRAMMATICAL FORM	GLOSS
/d͡ʒɑn/	----------	[d͡ʒun]	soul
/d͡ʒɑn-ɑne/ soul-adj marker	*[d͡ʒun-ɑne] *[d͡ʒɑn-une] *[d͡ʒun-une]	[d͡ʒɑnɑne]	extreme
/qæhremɑn/	*[qæhremun]	[qæhremɑn]	hero
/qæhremɑn-ɑne/ hero-adj marker	*[qæhremunɑne] *[qæhremɑnune] *[qæhremunune]	[qæhremɑnɑne]	heroic
/kɑm/	*[kum]	[kɑm]	wish
/nɑ-kɑm/ NEG-wish	*[nɑkum]	[nɑkɑm]	unfruitful
/nɑ-kɑm-ɑne/ NEG-wish-adv marker	*[nɑkumɑne] *[nɑkɑmune] *[nɑkumune]	[nɑkɑmɑne]	unfruitfully
/ɑm/	*[um]	[ɑm]	common, typical
/ɑm-i/ commoner-adj marker	*[umi]	[ɑmi]	commoner
/ɑm-i-ɑne/ commoner-adj marker-adj marker	*[umjɑne] *[umjune]	[ɑmjune]	vulgar

amjune is a word which is used in TP; thus, pre-nasal *a* in the affix *-ane* undergoes raising. The reason that pre-nasal *a* in the base does not raise is that it is a monosyllabic word whose pre-nasal *a* precedes *m*. For further discussion regarding monosyllabic words with an *am* sequence, see the discussion on affixation in monosyllabic words with *am* sequence in section 5.6.1. of this chapter. Also, for more information on double *aN* sequence in a lexical item, see the Phonology chapter.

As a conclusion to this discussion of the interaction of derivational affixes with pre-nasal raising, I highlight what was presented at the beginning of this section concerning the position of the *aN* sequence and its relation with pre-nasal raising: contrary to Miller (2011) and Kahn and Bernstein (1981), I showed that the position of the *aN* sequence in a (derived) word does not influence pre-nasal raising. What determines raising in the word is whether that lexical item is a part of the TP lexicon or not (as well as the factors to be addressed in chapter six). If the result of affixation is a formal word which is only used in formal Persian, it means that that certain word is part of formal Persian lexicon and does not belong to TP.

5.5. Exceptional affixes

In this section I discuss the role of two derivational affixes of TP, *-ef* and *-ani*, which pattern differently from other derivational affixes concerning pre-nasal raising. Sadeghi (2001) argues that in *xan-ænde* 'singer' and *ran-ænde* 'driver' raising is blocked because these words are recent coinages. Similarly, Modaressi (1978:84), referring to words like *dan-ef-d͡ʒu* 'university student', posits that raising in newly coined words is inhibited. Following Modaressi and Sadeghi, I assume that recent coinages and their morphologically derived forms resist raising. Furthermore, as mentioned earlier, morphologically derived words that are used in formal Persian and not in TP resist raising (for further discussion about the interaction of pre-nasal raising with recent coinages,

non-Arabic loans and formal Persian words see chapter seven). The following morphologically derived data support this claim that FP words and recent coinages do not undergo raising.

(61) *sample of derived forms that are FP and/or recently coined words as a result of hosting adj. marker -ani or noun marker -eʃ*

UNDERLYING FORM	GRAMMATICAL FORM	UNGRAMMATICAL FORM	GLOSS	REASON FOR NOT RAISING
/æql-ɑni/	[æqlɑni]	*[æqluni]	logical	formal Persian
/væhd-ɑni/	[væhdɑni]	*[væhdɑni]	monotheistic	formal Persian
/ræb-ɑni/	[ræbɑni]	*[ræbuni]	celestial	formal Persian
/tæht-ɑni/	[tæhtɑni]	*[tæhtuni]	lower	formal Persian
/foq-ɑni/	[foqɑni]	*[foquni]	upper	formal Persian
/zolm-ɑni/	[zolmɑni]	*[zolmuni]	pitch-black	formal Persian
/næfs-ɑni/	[næfsɑni]	*[næfsuni]	sensual	formal Persian
/dɑn-eʃ/	[dɑneʃ]	*[duneʃ]	knowledge	formal Persian
/xɑn-eʃ/	[xɑneʃ]	*[xuneʃ]	understanding	recent coinage
/rɑn-eʃ/	[rɑneʃ]	*[runeʃ]	propulsion	recent coinage

However, the data in (62) and (63) cast doubt on Modaressi's and Sadeghi's claim since these are lexical items that are neither recent coinages nor formal Persian words; also, they are not non-Arabic loanwords. That is, for these morphologically derived words that are frequently used by TP speakers there are not any lexical or morphophonological blocking factors that prohibit pre-nasal *a* to raise; nevertheless, raising is inhibited. Thus, it can be concluded that *-ani* in (62) and *-eʃ* in (63) are exceptional affixes. When *-eʃ* attaches to a base with an *aN* sequence, raising in the derived form is blocked. Also, when the noun/adjective marker *-ani* attaches to a base, pre-nasal *a* in the affix does not raise even if the derived form is frequently used in TP. Hence, *-eʃ* is a blocking affix and *-ani* is a resistant suffix.

(62) *derived forms hosting -ani that are used in TP*

UNDERLYING FORM	GRAMMATICAL FORM	UNGRAMMATICAL FORM	GLOSS
/tul-ɑni/ length-adj marker	[tulɑni]	*[tuluni]	lengthy

/æzole-ɑni[25]/ muscle-adj marker	[æzolɑni]	*[æzoluni]	muscular
/nur-ɑni[26]/ light-adj marker	[nurɑni]	*[nuruni]	luminous
/æsæb-ɑni/ nerve-adj marker	[æsæbɑni]	*[æsæbuni]	angry

There are two points about *-ani* that I should highlight. First, among the words that host this *-ani* suffix, the only exception to pre-nasal raising is *t͡ʃeraqani* 'decoration with lamp' in which pre-nasal *a* is allowed to raise. Second, if pre-nasal *a* in *pæriʃani* 'anxiety' undergoes raising, it is because the stem *pæriʃ* hosts two affixes: adjective marker suffix *-an* and noun marker suffix *-i*, and not one *-ani* suffix. My evidence for this claim is that *-ani* attaches to nouns (like the nouns in (62)) whereas *pæriʃ* is an adjective.

Note that pre-nasal *a* in /ɑrɑm/ 'peaceful' raises and thus the surface form is [ɑrum]. *dan* 'know' is a present tense verb stem and is not used in its bare form, so [dun] does not exist in the TP lexicon.

(63) *derived forms hosting noun marker -eʃ that are used in TP*

UNDERLYING FORM	GRAMMATICAL FORM	UNGRAMMATICAL FORM	GLOSS
/ɑrɑm-eʃ/ peaceful-noun marker	[ɑrɑmeʃ]	*[ɑrumeʃ]	peace
/dɑn-eʃ-mænd/ know-noun marker-possessor	[dɑneʃmænd]	*[duneʃmænd]	scientist
/dɑn-eʃ-gɑh/ know-noun marker-place	[dɑneʃgɑh]	*[duneʃgɑh]	university
/dɑn-eʃ d͡ʒu/ know-noun marker seek	[dɑneʃ d͡ʒu]	*[duneʃ d͡ʒu]	university student
/dɑn-eʃ ɑmuz/ know-noun marker learn	[dɑneʃ ɑmuz]	*[duneʃ ɑmuz[27]]	student

[25] As in *ampule æzolani* 'ampule injected in muscles'
[26] As in *t͡ʃehreje nurani* 'luminous face'
[27] If *a* in present tense verb stem *a.muz* 'learn' does not raise, it is because *a* and *m* are not in the same syllable. (see the phonology chapter).

/dɑn-eʃ-kæde/ know-noun marker-place	[dɑneʃkæde]	*[duneʃkæde]	college, faculty

Regarding the data in (63), the last three examples can be categorized as recent coinages, but the point is that when a word with an *aN* sequence enters the TP lexicon and is used frequently, it is quite probable that the pre-nasal *a* in these lexical items would undergo raising. Given this, even though words like *daneʃd͡ʒu* can be regarded as frequently used words, the surface form in TP is not **duneʃd͡ʒu*; and the reason for this is that the derived word hosts the blocking affix *-eʃ*. In other words, pre-nasal raising in the above stems (*aram* and *dan*) occurs when they host other affixes, but not when they host *-eʃ*. The data in (64) show that these two stems host different derivational affixes and pre-nasal raising is not prohibited:

(64) *pre-nasal raising in "aram" and "dan" after hosting different affixes, except -eʃ*

UNDERLYING FORM IN TP	SURFACE FORM IN TP	GLOSS
/be ɑrɑm-i/ → with peaceful-noun marker	[be ɑrumi]	slowly
/nɑ-ɑrɑm/ → NEG-peaceful	[nɑʔɑrum]	restless
/nɑ-ɑrɑm-i/ → NEG-peaceful-noun marker	[nɑʔɑrumi]	turmoil
/nɑ-dɑn/ → NEG-know	[nɑdun]	ignorant person
/nɑ-dɑn +-i/ → NEG-know-noun marker	[nɑduni]	ignorance
/dɑn-est-æn/ → know-PST-INF	[dunestæn]	to know

Based on this generalization that frequency is able to trigger pre-nasal *a* raising in frequently used recently coined words, there is the possibility that *-ænde* could be categorized as a blocking affix as well. *ran-ænde* 'driver' and *xan-ænde* 'singer' are two highly frequent TP words in which pre-nasal raising is expected, but it does not occur. As a result, it can be assumed that the only

possible blocking factor could be the *-ænde* suffix, especially since pre-nasal *a* in these two stems undergoes raising when they host other affixes, as shown in (65):

(65) *pre-nasal raising in "xan" and "ran" after hosting different affixes, except -ænde*

UNDERLYING FORM IN TP	SURFACE FORM IN TP	GLOSS
/rɑn-d-e/ → drive-PST-adj marker	[runde]	driven, outcast
/mi-rɑn-æm/ → PROG-driv.PRS-1SG.SBJ	[mirunæm]	I am driving.
/rɑn-d-æn/ → drive-PST-INF	[rundæn]	to drive
/be-xɑn/ → IMP-read.PRS	[bexun]	You (sg) read!
/xɑn-d-æn/ → read-PST-INF	[xundæn]	to read

However, due to lack of sufficient number of frequently used words in TP with *aN* sequence that host *-ænde* suffix, this postulation cannot be confirmed.

The last point about blocking affixes is that the affix like *-gah* cannot be categorized as a blocking affix. Whenever a word with an *aN* sequence hosts *-gah*, the pre-nasal *a* in the base is prohibited from raising. The reason for this is that that derived form is a formal Persian word which is not used in TP, as shown in (66). On the contrary, *-eʃ* is categorized as blocking affixes since it blocks raising in words that are frequently used lexical items in TP.

(66) *sample of words hosting -gah*

Before hosting *-gah*			**After hosting *-gah***	
/ɑn/	→ [un]	'that'	/ɑn-gɑh/	→ [ɑngɑh], *[ungɑh] 'then'
/hæm-ɑn/ also-that	→ [hæmun]	'the same'	/hæm-ɑn-gɑh/ also-that-time	→[hæmɑngɑh], *[hæmungɑh] 'that time'
/ɑrɑm/	→ [ɑrum]	'peaceful'	/ɑrɑm-gɑh/ peaceful-place	→[ɑrɑmgɑh], *[ɑrumgɑh] 'cemetery'

It is worth noting that *-gah* in the first two examples is a temporal adverb, but in the third example it is a location adverb.

5.6. Re-syllabification after affixation

Words with an *am* sequence are particularly worthy of attention because of the extra phonological conditions on them, and these are the topic of this section. As explained in chapter four, if an *am* sequence occurs in a monosyllabic word, raising is blocked: *xam* → **xum* 'raw'. In addition, in multisyllabic words that contain an *am* sequence, *a* raises only if both *a* and *m* are in the same syllable: *da.mæn* → **du.mæn* 'skirt'. The point is that when a word with a final *am* sequence hosts a vowel-initial suffix, resyllabification occurs. Hence, the question is whether re-syllabification influences pre-nasal raising since *m* is now syllable-initial. In order to answer this question, first, I will discuss monosyllabic words that contain an *am* sequence as hosts of vowel-initial suffixes, and then I will discuss multisyllabic words.

5.6.1. Monosyllabic words with *am* sequence and affixation

As discussed above, monosyllabic words containing an *am* sequence disallow raising. In this part I will show that if these words host affixes, pre-nasal raising in them is still prohibited even though the derived and/or inflected word is now a multisyllabic lexical item. In these cases, blocking can be attributed to the same syllable constraint:

(67) *blockage of raising in morphologically derived monosyllabic words with [am] sequence*

Monosyllabic word before affixation			**Monosyllabic word after affixation**		
/ɑm/	→[ɑm] *[um]	'common'	/ɑm-i/ commoner-adj marker	→[ɑ.mi] *[u.mi]	'commoner'[28]
			/ɑm-e/ commoner-adj marker	→[ɑ.me] *[u.me]	'majority'

[28] For */am-j-ane/* → *[amjune]* 'vulgar', see the former section, the discussion on derivational affixes and pre-nasal raising.

/nɑm/	→[ɑm] *[num]	‘name’	/nɑm-vær/ →[nɑm.vær] *[num.vær] ‘famous’ name-adj marker
			/nɑm-i/ →[nɑ.mi] *[nu.mi] ‘famous’ name-adj marker
			/bi-nɑm/ →[bi.nɑm] *[bi.num] ‘nameless’ without-name
			/hæm-nɑm/ →[hæm.nɑm] *[hæm.num] ‘homonym’ same-name
/dɑm/	→[dɑm] *[dum]	‘cattle’	/dɑm-i/ →[dɑ.mi] *[du.mi] ‘of cattle’ cattle-adj marker
/ʃɑm/	→[ʃɑm] *[ʃum]	‘supper’	/ʃɑm-gɑh/ →[ʃɑm.gɑh] *[ʃum.gɑh] ‘nighttime’ supper-time

The monosyllabic words in (68) are formal words which are not used in TP, so after affixation, they still resist raising.

(68) *FP monosyllabic words and their morphologically derived forms that resist raising*

monosyllabic words before affixation	**Monosyllabic words after affixation**
/kɑm/ → [kɑm] *[kum] ‘prosperity’	/nɑ-kɑm/ → [nɑkɑm] *[nɑkum] ‘fruitless’ NEG-wish
	/bɑ-kɑm/ → [bɑkɑm] *[bɑkum] ‘fortunate’ with-wish
/gɑm/ → [gɑm] *[gum] ‘step’	/piʃ-gɑm/ → [piʃgɑm] *[piʃgum] ‘pioneer’ prior-step
	/hæm-gɑm/→[hæmgɑm] *[hæmgum] ‘companion’ same-step
/d͡ʒɑm/ → [d͡ʒɑm] *[d͡ʒum] ‘chalice’	/fær-d͡ʒɑm/ → [færd͡ʒɑm] *[færd͡ʒum] ‘end’

/rɑm/ → [rɑm] *[rum] ‘tame’ /rɑm-eʃ/ → [rɑm-eʃ] *[rum-eʃ] ‘music, peace’
tame-noun marker

As a result, it can be concluded that the number of syllables of words with an *am* sequence concerning pre-nasal raising is influential before affixation; i.e. the bare form (un-inflected and/or un-derived form) of the word is a criterion for this generalization.

5.6.2. Multisyllabic words with an *am* sequence and affixation

As discussed in the Phonology chapter, pre-nasal *a* which precedes *m* in multisyllabic words is allowed to undergo raising only if the *am* sequence is in the same syllable. In this section I investigate lexical items whose *a* and *m* segments are in the same syllable, but when a vowel-initial suffix is added, the *m* is in an onset. Despite this, raising occurs in the derived/inflected form. Examples in (69) illustrate:

(69) *pre-nasal raising in re-syllabified multisyllabic TP words with [am] sequence*

Before affixation		**After affixation**	
/ɑ.rɑm/	→ [ɑ.rum] ‘calm	/ɑ.rɑm + -i/ → [ɑ.ru.mi]	‘You (sg) are calm.’
		calm-COP.PRS.2SG	
/bɑ.dɑm/	→ [bɑ.dum] ‘almond’	/bɑ.dɑm + -i/ → [bɑ.du.mi]	‘almond-like’
		almond-like	
/ko.dɑm/	→ [kodum] ‘which’	/ko.dɑm + -ɑ/ → [ko.du.mɑ]	‘which ones?’
		which-PL	
/hæ.rɑm/	→ [hæ.rum] ‘haram’	/hæ.rɑm + -e/ → [hæ.ru.me]	‘It is haram.’
		haram COP.PRS.3SG	

Re-syllabification after affixation does not have any influence on words in which *a* and *m* are not in the same syllable before they host an affix since even after re-syllabification due to affixation, *a* and *m* are still in two separate syllables:

(70) *blockage of pre-nasal raising in re-syllabified multisyllabic TP words with [am] sequence*

Before affixation	**After affixation**
/dɑ.mæn/→ [dɑ.mæn], *[du.mæn] 'skirt'	/dɑ.mæn-e/→[dɑ.mæ.ne], *[du.mæ.ne] 'domain' skirt-noun marker
/ɑ.mɑ.de/→ [ɑ.mɑ.de], *[u.mɑ.de] 'ready'	/ɑ.mɑ.de-gi/→[ɑ.mɑ.de.gi], *[u.mɑ.de.gi] 'preparation' ready-noun marker
/ɑ.mɑr/→ [ɑ.mɑr], *[u.mɑr] 'stats'	/ɑ.mɑr-gær/→[ɑ.mɑr.gær], *[u.mɑr.gær] 'statistician' stats-noun marker
/xɑ.muʃ/→ [xɑ.muʃ], *[xu.muʃ] 'off'	/xɑ.muʃ-i/→[xɑ.mu.ʃi], *[xu.mu.ʃi] 'blackout' off-noun marker
/hɑ.me.le/→[hɑ.me.le], *[hu.me.le] 'pregnant'	/hɑ.me.le-gi/→[hɑ.me.le.gi], *[hu.me.le.gi] 'pregnancy' pregnant-noun marker
/sæ.lɑ.mæt/→[sæ.lɑ.mæt], *[sæ.lu.mæt] 'healthiness'	/sæ.lɑ.mæt-i/→[sæ.lɑ.mæ.ti], *[sæ.lu.mæ.ti] 'health' healthiness-noun marker

This finding is particularly striking, with interesting implications for word structure and its interaction with phonology.

5.7. Homophonous affixes

In this section I discuss two pairs of homophonous suffixes. The first pair of homophonous affixes is the 3rd person singular marker *-ef* and the noun-marking blocking affix *-ef*, and the second pair of affixes is the causative marker *-an* and the adjective/adverb marker *-an*.

The nominalizer *-ef* affix attaches to present tense verb stems. As explained earlier in this chapter, whenever this suffix attaches to a verb stem which contains an *aN* sequence, the pre-nasal

a in the derived form does not raise. On the other hand, the homophonous affix, the third person singular marker, is an inflectional suffix which plays no role in inhibiting or triggering pre-nasal raising. Below, in the examples of the left column the *-eʃ* suffix is the derivational blocking affix and in the right column, the *-eʃ* affix is the third person singular marker:

(71) *noun marker -eʃ vs 3rd person singular -eʃ*

Blocking *-eʃ*	**3rd person singular *-eʃ***
a. /ɑrɑm-eʃ/→[ɑrɑmeʃ], *[ɑrumeʃ] 'peace'	b. /ɑrɑm-eʃ/ → [ɑrum-eʃ] as in ***arum-eʃ*** *kærdæm.* 'I alleviated her/him.'
c. /dɑn-eʃ-d͡ʒu/→[dɑneʃd͡ʒu], *[duneʃd͡ʒu] know.PRS-eʃ-seek 'university student'	d. /dɑn-eʃ/ → [dun-eʃ][29] seed-3SG.POSS 'its seed'

The other pair of homophonous affixes is the adjective/adverb marker *-an* and the causativizer *-an*. The former attaches to present tense verb stems and the latter affixes to both present and past tense stems. That said, some present tense verb stems host both the causativizer and the adjective/adverb marker, creating verbs whose causativized and adjectival/adverbial forms are identical. Here, I show that if the affixed form is actively used in TP, then raising is allowed; otherwise, pre-nasal *a* does not raise. In (72), pre-nasal *a* in the adjectival/adverbial forms of (72.a), (72.b) and (72.h) and the causativized form in (72.c) do not raise as they are formal Persian words that are not used in TP. Pre-nasal raising is observed in the rest of data since they are TP lexical items.

(72) *causativizer -an vs adjective/adverb marker -an*

PRESENT TENSE VERB STEM	CAUSATIVE FORM[30]	ADJ/ADV FORM		VERB STEM GLOSS
a) /d͡ʒuʃ + -ɑn/→	[d͡ʒuʃun]	[d͡ʒuʃɑn]	*[d͡ʒuʃsun]	boil
b) /suz + -ɑn/→	[suzun]	[suzɑn]	*[suzun]	burn

[29] Also, pre-nasal *a* in complex predicates like *dan kærdæn* 'to granulate' which are used in TP is allowed to raise: *anar-o dun-eʃ kærd-æm* 'I granulated the pomegranate.'

[30] These are the bare forms of the causativized verbs.

c)	/goriz + -ɑn/⟶	[gorizɑn] *[gorizun]	[gorizun]		escape
d)	/dæv + -ɑn/⟶	[dævun]	[dævun][31]		run
e)	/keʃ + -ɑn/⟶	[keʃun]	[keʃun][32]		drag
f)	/xænd + -ɑn/⟶	[xændun]	[xændun]		laugh
g)	/gerj + -ɑn/⟶	[gerjun]	[gerjun]		cry
h)	/ʃur + -ɑn/⟶	[ʃurun]	[ʃurɑn]	*[ʃurun]	revolt

5.8. Summary

This chapter highlights how different morphological factors interact with pre-nasal raising in TP. This investigation reveals that pre-nasal raising is not merely phonological; rather, there are morphological factors that influence pre-nasal raising. These include that raising occurs only when the target and the trigger are contained within the same morpheme, as discussed in chapter four. There are blocking affixes like the noun marker *-eʃ* (**arum-eʃ* 'peace'). Also, in this chapter it was demonstrated how morphologically inflected and derived forms of words interact with lexical factors (formality, non-Arabic loanwords and recent coinages) in terms of pre-nasal raising. From this section it can be deduced that in some morphologically derived words there are potentially several factors at play that simultaneously influence raising. Particularly striking with *m* is that raising is present in non-monosyllabic *am*-final items even if a vowel-initial suffix is added, unexpectedly given the same syllable constraint.

[31] As in *dæv-un dæv-un* 'while running'
[32] As in *kesh-un kesh-un* 'while dragging'

CHAPTER SIX

LEXICAL FACTORS

6.1. Introduction

This chapter discusses those lexical items in which blockage of pre-nasal raising is not due to phonological or morphological factors. Raising in these words is blocked because they belong to one of the four lexical categories that are introduced in this chapter. The four classes of lexical items that resist raising are proper names, non-Arabic loanwords, recent coinages and formal Persian lexical items.

The structure of this chapter is as follows. In section 6.2., I discuss proper names and their interaction with pre-nasal raising. After that, I discuss pre-nasal raising in loanwords and show that non-Arabic loanwords resist raising. In addition, I demonstrate that Arabic borrowings pattern differently from other loanwords with respect to pre-nasal raising. In section 6.4, I show that pre-nasal *a* in recent coinages resists raising. The final section addresses lexical items that are considered to be formal Persian words that are not used in TP, and hence raising is blocked.

6.2. Proper names

In this section, following Modaressi (1978), Jahangiri (1980) and Kahn and Bernstein (1981), I show that proper names containing an *aN* sequence generally resist pre-nasal raising. Pre-nasal raising in these lexical items is insensitive phonological structure; that is, even if pre-nasal *a* in a proper name does not have any phonological restrictions for raising, raising is inhibited, as shown by the words in (73):

(73) *sample of proper names with aN sequence which resist raising*

PROPER NAMES			GLOSS
/ehteram/→	[ehteram]	*[ehterum]	female name
/pærvane/→	[pærvane]	*[pærvune]	female name
/pedram/→	[pedram]	*[pedrum]	male name
/ærsælan/→	[ærsælan]	*[ærsælun]	male name
/ramsær/→	[ramsær]	*[rumsær]	city name
/ilam/→	[ilam]	*[ilum]	city name

Some lexical items in Persian are used both as common nouns and proper names. When one of these words with an *aN* sequence is used as a common noun, pre-nasal *a* raises, but when that lexical item is used as a proper name, raising is blocked. Examples in (74) support this claim:

(74) *sample of words that are used both as proper name and common noun*

COMMON NOUN		PROPER NAME			GLOSS
/baran/ →	[barun]	/baran/ →	[baran]	*[barun]	rain
/rejhan/ →	[rejhun]	/rejhan/ →	[rejhan]	*[rejhun]	basil
/færman/→	[færmun]	/færman/ →	[færman]	*[færmun]	command
/ʃejtan/ →	[ʃejtun]	/ʃejtan/ →	[ʃejtan]	*[ʃejtun]	naughty (as a common noun), Satan (as a proper name)

Contrary to the above generalization, there are some proper names with an *aN* sequence in which pre-nasal *a* raises when they are used in TP (75). I assume that one of the reasons that raising is not inhibited in these words is that they are frequently used proper names in TP.[33]

(75) *proper names whose pre-nasal [a] undergoes raising*

UNDERLYING FORM	TP SURFACE FORM	GLOSS
/iran/	[irun]	Iran
/esfæhan/	[esfæhun]	name of a city
/kordestan/	[kordestun]	name of a province
/hæmedan/	[hæmedun]	name of a city
/kaʃan/	[kaʃun]	name of a city
/kerman/	[kermun]	name of a city

[33] Unfortunately, as there is not a TP corpus available, it is impossible to measure the frequency of use.

/kermɑnʃɑh/	[kermunʃɑh]	name of a city
/xorɑsɑn/	[xorɑsun]	name of a province
/ʃemirɑn/	[ʃemirun]	name of a county in Tehran
/tehrɑn/	[tehrun]	Tehran
/ræmezɑn/	[ræmezun]	male name
/ʃæʔbɑn/	[ʃæʔbun]	male name
/qolɑm/	[qolum]	male name

Proper names generally resist raising, with raising being possible in some names that appear to be frequent.

6.3. Loanwords

In this section I discuss loanwords as a class of lexical items that resist raising. Jahangiri (1980:79) states that loanwords seem more likely to keep their original form and disfavor raising; however, he does not distinguish Arabic loanwords from non-Arabic borrowings. On the other hand, Modaressi (1978:82) claims that loanwords from non-Iranian languages other than Arabic do not favor the application of pre-nasal raising. Following Modaressi (1978), I show that Arabic and non-Arabic loanwords pattern differently. I divide this section into two subsections: pre-nasal raising in non-Arabic loanwords and pre-nasal raising in Arabic loanwords.

With loanwords, it would be helpful to know when they entered into the language. However, it is not possible to give a date when a loanword was borrowed into Persian since the dictionaries and other sources used for finding the source language of loanwords do not provide such information.

6.3.1. Non-Arabic loanwords

The main observation in this subsection is that non-Arabic loanwords which phonologically meet the criterion for raising resist raising even if they are frequently used words in TP. Examples in (76) support this claim:

(76) *sample of non-Arabic loans that contain aN sequence*

UNDERLYING FORM	TP SURFACE FORM		GLOSS	SOURCE LANGUAGE
/livɑn/ →	[livɑn]	*[livun]	glass	Russian
/estekɑn/ →	[estekɑn]	*[estekun]	teacup	Russian
/mɑnitor/ →	[mɑnitor]	*[munitor]	monitor	English
/septɑmbr/ →	[septɑmbr]	*[septumbr]	September	French
/simɑn/ →	[simɑn]	*[simun]	cement	French
/vɑn/ →	[vɑn]	*[vun]	bathtub	Russian
/kɑmjon/ →	[kɑmjon]	*[kumjon]	truck	French

When a non-Arabic loanword hosts an affix, the derived form resists raising, regardless of the frequency of that word.

(77) *sample of derived forms of non-Arabic loans that contain aN sequence*

UNDERLYING FORM	TP SURFACE FORM		GLOSS	SOURCE LANGUAGE
/bɑnk + -i/ → bank-adj marker	[bɑnki]	*[bunki]	related to bank	French
/ɑnɑliz + -gær/→ analyze-noun marker	[ɑnɑlizgær]	*[unɑlizgær]	analyzer	French
/kɑmvɑ + -i/→ yarn-adj marker	[kɑmvɑʔi]	*[kumvɑʔi]	yarn-like	English
/ʃɑns + -i/→ chance-adj marker	[ʃɑnsi]	*[ʃunsi]	by chance	French
/ɑlmɑn + -i/→ Germany-adj marker	[ɑlmɑni]	*[ɑlmuni]	German	French
/pɑndul + -vɑr/→ pedulum-like	[pɑndulvɑr]	*[pundulvɑr]	pendulum-like	French[34]

Thus, loanwords from languages other than Arabic are resistant to raising.

[34] Source languages of these borrowings are extracted from Dehkhoda Online Dictionary (http://www.jasjoo.com/books/wordbook/dehkhoda/) and Moein Online Dictionary (http://www.jasjoo.com/books/wordbook/fa/).

6.3.2. Arabic loanwords

Unlike non-Arabic loans, Arabic loanwords can exhibit raising, and are sensitive to phonological restrictions on pre-nasal raising. For instance, as in native Persian words, in an Arabic loanword pre-nasal *a* does not raise if in that word an *an* sequence is followed by *u* (78). Moreover, in monosyllabic words with an *am* sequence raising is blocked (79), and in multisyllabic words with an *am* sequence raising is inhibited if *a* and *m* are not in the same syllable (80). In Arabic loanwords that do not have any of these phonological restrictions and are used in TP, pre-nasal *a* raises (81). In this sense, Arabic loanwords are more nativized than other loanwords.

(78) *Arabic loans with [anu] sequence*

UNDERLYING FORM	SURFACE FORM		GLOSS
/qɑnun/	[qɑnun]	*[qunun]	law
/kɑnun/	[kɑnun]	*[kunun]	focus

(79) *monosyllabic Arabic loans with [am] sequence*

UNDERLYING FORM	SURFACE FORM		GLOSS
/ɑm/	[ɑm]	*[um]	common, public
/tɑm/	[tɑm]	*[tum]	complete, full
/lɑm/	[lɑm]	*[lum]	a letter in Arabic alphabet

(80) *multisyllabic Arabic loans with [am] sequence*

UNDERLYING FORM	SURFACE FORM		GLOSS
/kɑmel/ →	[kɑ.mel]	*[ku.mel]	complete
/ʃɑmel/ →	[ʃɑ.mel]	*[ʃu.mel]	include
/edɑme/ →	[edɑ.me]	*[edu.me]	continuation
/hɑmele/ →	[hɑ.me.le]	*[hu.me.le]	pregnant
/d͡ʒɑmeʔe/ →	[d͡ʒɑ.me.ʔe]	*[d͡ʒu.me.ʔe]	society
/d͡ʒɑmed/ →	[d͡ʒɑ.med]	*[d͡ʒu.med]	solid
/moʔɑmele/ →	[mo.ʔɑ.me.le]	*[mo.ʔu.me.le]	trade
/zɑmen/ →	[zɑ.men]	*[zu.men]	guarantor
/ælɑmæt/ →	[æ.lɑ.mæt]	*[æ.lu.mæt]	sign

(81) *Arabic loans which undergo pre-nasal raising*

UNDERLYING FORM	SURFACE FORM	GLOSS
/tæmɑm/ →	[tæmum]	end
/hæmɑm/ →	[hæmum]	bathroom
/æzɑn/ →	[æzun]	call for prayer
/dævɑm/ →	[dævum]	durability
/dokkɑn/ →	[dokkun]	shop
/hæzijɑn/ →	[hæzijun]	delirium
/hærɑm/ →	[hærum]	haram
/hejvɑn/ →	[hejvun]	animal
/kætɑn/ →	[kætun]	cotton
/mejdɑn/ →	[mejdun]	square
/mizɑn/ →	[mizun]	amount
/mosælmɑn/ →	[mosælmun]	Muslim
/zæʔferɑn/ →	[zæʔferun]	saffron
/qejtɑn/ →	[qejtun]	braid
/rejhɑn/ →	[rejhun]	basil
/ʃæʔbɑn/ →	[ʃæʔbun]	month name
/ræmezɑn/ →	[ræmezun]	Ramadan

While European loanwords are recent borrowings which entered Persian in the 19th century (Kłagisz 2013:39 and Deyhime 2000), Arabic loanwords were borrowed into Persian after the Arab conquest of Persia (Iran) ca 651 AD (Morony 1986). Arabic loans are extensively used in Persian. Assuming this, we see that the older loanwords are treated like native words when phonological conditions are met. It has often been observed that loanwords adapt to the phonology of a language over time. For instance, Kemmer (2017) argues that conventionalization is a gradual process in which a word progressively permeates a larger and larger speech community. She posits that the longer a loanword has been frequently used in the borrowing language, the more it resembles the native words of the language in terms of its phonology.

This type of stratification of loanwords based on their level of assimilation to the phonology of the borrowing language is in line with Itô and Mester's (1999) Core-Periphery model of the lexicon. They present their model based on loanwords in Japanese and propose a model of

loanword adaptation in which lexical items are divided into different strata, each with its own set of constraints. The core stratum of this model is native lexical items called Yamato in Japanese. The other Japanese strata, moving outward from the core, are Sino-Japanese, Assimilated foreign words, and finally Unassimilated foreign lexical items. In this stratification, the core lexical items fulfill markedness constraints of the borrowing language; hence, the more we move outward from the core, the more violations of these markedness constraints are found, until we encounter lexical items at the periphery that only satisfy small subset of the core constraints. A deep study of the loanword adaptation in Persian is a topic of research which is beyond the scope of this study.

Nevertheless, it is interesting to note that as Arabic loanwords used in TP satisfy the markedness constraints on pre-nasal raising, it suggests that these loanwords are more adapted to phonology of TP in comparison to other loanwords; hence, Arabic loanwords are in the closest stratum to core of Persian lexicon which is native Persian. Non-Arabic loanwords are in a peripheral stratum, as they violate the pre-nasal raising markedness constraint of Tehrani Persian phonology; that is, non-Arabic loanwords are less nativized than Arabic loanwords. As a result, it can be concluded that phonological stratification of the lexicon based on the markedness constraints of a borrowing language shows that loanwords do not nativize in a random or haphazard way; rather, they are nativized through a structured pattern and process.[35]

However, not all Arabic loanwords used in TP undergo raising. For instance, Modaressi (1978:81-82) mentions that Arabic broken plural forms disfavor raising. Broken plurals do not get the regular plural marking suffixes; they pattern with internal vowel change (Ryding 2005:145). Examples in (82) are Arabic broken plurals that are used in TP but raising is blocked although there are not any phonological restrictions.

[35] Recall that I restrict my discussion to the TP lexicon; only those Arabic loanwords undergo pre-nasal raising that are used in TP. Arabic loans that are categorized as formal lexical items resist raising.

(82) *sample of Arabic broken plurals*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/ænʔam/ →	[ænʔam]	*[ænʔum]	tip
/sæham/ →	[sæham]	*[sæhum]	equities, stock
/æhkam/ →	[æhkam]	*[æhkum]	commands
/ævam/ →	[ævam]	*[ævum]	common people
/æqvam/ →	[æqvam]	*[æqvum]	relatives
/æqsam/ →	[æqsam]	*[æqsum]	types

In addition to these broken plurals, Arabic quadriliteral (83) and triliteral infinitives (84) that are used in TP resist raising despite their frequency of use. (83) shows quadriliteral roots of the forms $eC_1C_2eC_3aC_4$ where the *a* may not be expected to undergo raising.

(83) *sample of Arabic quadriliteral infinitives*[36]

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/ezdeham/ →	[ezdeham]	*[ezdehum]	confluence
/etminan/ →	[etminan]	*[etminun]	assurance

(84.a) to (84.e) have the $eC_1C_2aC_3$[37] infinitival pattern. Examples in (84.f) to (84.h) have the $eC_1teC_2aC_3$ pattern. Forms from (84.i) to (4.k) have the $esteC_1C_2aC_3$ pattern. Again, raising is blocked.

(84) *sample of Arabic triliteral infinitives*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
a) /eʔdam/ →	[eʔdam]	*[eʔdum]	execution
b) /eʔlam/ →	[eʔlam]	*[eʔlum]	announcement
c) /eslam/	[eslam]	*[eslum]	Islam
d) /elham/ →	[elham]	*[elhum]	revelation
e) /eʔzam/ →	[eʔzam]	*[eʔzum]	dispatching
f) /emtehan/ →	[emtehan]	*[emtehun]	examination
g) /enteqam/ →	[enteqam]	*[entequm]	revenge
h) /ehteram/ →	[ehteram]	*[ehterum]	respect
i) /estexdam/ →	[estexdam]	*[estexdum]	employment
j) /estehkam/ →	[estehkam]	*[estehkum]	solidity
k) /esteʔlam/ →	[esteʔlam]	*[esteʔlum]	inquiry

[36] Words in (82) and (83) might be formal for some speakers.

[37] *iman* 'faith' is in this category, but due to certain morphophonological alternations which are beyond the scope of this study, its infinitival form differs from other infinitives of this category.

Kahn and Bernstein (1981) suggest that raising in *emtehan* (84.f) is blocked since it is related to education (see the literature review chapter); however, based on the above generalization, I posit that pre-nasal *a* in *emtehan* does not raise since the word is a triliteral Arabic infinitive.

In addition to these triliteral and quadrilateral infinitives, Arabic loanwords with *C₁æC₂æC₃an* (85) and *C₁oC₂C₃an* (86) patterns disfavor raising.

(85) *Arabic loans with "C₁æC₂æC₃an" pattern which resist raising*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/særætɑn/ →	[særætɑn]	*[særætun]	cancer
/xæfæqɑn/ →	[xæfæqɑn]	*[xæfæqun]	suffocation
/zæræbɑn/ →	[zæræbɑn]	*[zæræbun]	pulse
/hæjæd͡ʒɑn/ →	[hæjæd͡ʒɑn]	*[hæjæd͡ʒun]	excitement
/nævæsɑn/ →	[nævæsɑn]	*[nævæsun]	oscillation

d͡ʒæræjan 'flow' and *tæwæfan* 'storm' which are pronounced as *d͡ʒærjan* and *tufan* respectively in Persian, have the same *C₁æC₂æC₃an* pattern, and thus pre-nasal *a* does not raise.

(86) *Arabic loans with "C₁oC₂C₃an" pattern which resist raising*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/vod͡ʒdɑn/	[vod͡ʒdɑn]	*[vod͡ʒdun]	conscience
/d͡ʒobrɑn/	[d͡ʒobrɑn]	*[d͡ʒobrun]	compensation
/bohrɑn/	[bohrɑn]	*[bohrun]	crisis
/omrɑn/	[omrɑn]	*[omrun]	construction, a university field of study
/onvɑn/	[onvɑn]	*[onvun]	title

Unlike Rohany (2102) who claims that raising in *onvan* is simply an exception to raising, based on the abovementioned generalization, I posit that blockage of raising in this word follows from its *C₁oC₂C₃an* root-and-pattern morphology.

6.4. Recent coinages

Jahangiri (1980:79) mentions that new items which are recently "made" by the Iranian Academy of Language to replace Arabic loanwords are likely to keep their original form. Following his claim, I show that pre-nasal *a* in recently coined words does not raise.

Recent coinages in Persian are usually created through affixation, and there are affixes that contain *aN* sequence. *aN* raising in these affixes is blocked even though pre-nasal *a* in the same affix raises in derived words which are not recent coinages. As shown in (87), pre-nasal *a* in the adjectival/adverbial marker *-ane* and the noun marker *-ban* does not raise since the derived words are newly coined. But in (88) raising in the same *-ane* and *-ban* suffixes is not blocked as the words are not recent coinages.

(87) *blockage of raising in -ane or -ban suffixes that are used in recently coined words*

UNDERLYING FORM	SURFACE FORM		GLOSS
/rɑj + -ɑne/ → think-noun marker	[rɑjɑne]	*[rɑjune]	computer
/jɑr + -ɑne/ → companion-noun marker	[jɑrɑne]	*[jɑrune]	subsidy
/poʃti + -bɑn/ → back-agentive	[poʃtibɑn]	*[poʃtibun]	school students' consultant
/pɑrk + -bɑn/ → park-agentive	[pɑrkbɑn]	*[pɑrkbun]	parking lot guard

(88) *pre-nasal raising in words with -ane or -ban suffixes*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/mærd+ -ɑne/ → man-adj marker	[mærdune]	manly
/zæn + -ɑne / → woman-adj marker	[zænune]	womanly
/bɑq + -bɑn/ → garden-agentive	[bɑqbun]	gardener
/bɑd + -bɑn/ → wind-agentive	[bɑdbun]	sail

Some recent coinages are revitalized from classic Persian. As shown in (89), parallel to other newly coined words, raising in these words is blocked although they are of high frequency in TP.

(89) *sample of revitalized classic Persian words used in TP*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/dæb-estɑn/ → ?-place	[dæbestɑn]	*[dæbestun]	primary school
/dæbir-estɑn/ → instructor-place	[dæbirestɑn]	*[dæbirestun]	high school
/kudæk-estɑn/ → child-place	[kudækestɑn]	*[kudækestun]	kindergarten
/bimɑr-estɑn/ → patient-place	[bimɑrestɑn]	*[bimɑrestun]	hospital
/deʒ-bɑn/ → fortress-agentive	[deʒbɑn]	*[deʒbun]	military police

The last point on this topic is that there are newly coined words that are proper names which are used for names of some nations. Raising is blocked in these words due to two lexical blocking factors; they are both proper names and recently coined words. (90) is a list of some of these words:

(90) *recently coined proper names*

UNDERLYING FORM	SURFACE FORM		GLOSS
/mæd͡ʒɑr-estɑn/	[mæd͡ʒɑrestɑn]	*[mæd͡ʒɑrestun]	Hungary
/ærmæn-estɑn/	[ærmænestɑn]	*[ærmænestun]	Armenia
/gord͡ʒ-estɑn/	[gord͡ʒestɑn]	*[gord͡ʒestun]	Georgia
/engel-estɑn/	[engelestɑn]	*[engelestun]	England
/pɑk-estɑn/	[pɑkestɑn]	*[pɑkestun]	Pakistan
/bolqɑr-estɑn/	[bolqɑrestɑn]	*[bolqɑrestun]	Bulgaria

6.5. Formal Persian words

In the Persian lexicon there are lexical items that are highly formal which are used in formal contexts like news, literary texts, poems and philosophical writings. This category of Persian words does not exist in the TP lexicon, but as Modaressi (1978:84) observes, the majority of these words have TP equivalents and synonyms. In this study I use the cover term “formal Persian

words" for this class of lexical items that are not used in TP. A list of formal Persian words with their TP equivalents is shown in (91):

(91) *list of FP words with their TP equivalents*

FORMAL PERSIAN WITH *aN* SEQUENCE			TP EQUIVALENT	GLOSS
1. /rɑjgɑn/ →	[rɑjgɑn]	*[rɑjgun]	[moft]	free of charges
2. /dehqɑn/ →	[dehqɑn]	*[dehqun]	[keʃɑværz]	peasant
3. /ærmæqɑn/ →	[ærmæqɑn]	*[ærmæqun]	[hedje], [kɑdo]	gift
4. /bɑng/ →	[bɑng]	*[bung]	[dɑd]	shout
5. /ɑʃjɑn/ →	[ɑʃjɑn]	*[ɑʃjun]	/lane/ →[lune]	nest
6. /bonjɑn/ →	[bonjɑn]	*[bonjun]	[pɑje], [æsɑs]	base
7. /doʃnɑm/ →	[doʃnɑm]	*[doʃnum]	[fohʃ]	swearword
8. /ehɑnæt/ →	[ehɑnæt]	*[ehunæt]	[tohin]	insult
9. /fæqɑn/ →	[fæqɑn]	*[fæqun]	[dɑd]	shout
10. /hed͡ʒrɑn/ →	[hed͡ʒrɑn]	*[hed͡ʒrun]	[duri]	separation
11. /herɱɑn/ →	[herɱɑn]	*[hermun]	[nɑ-omidi]	hopelessness
12. /d͡ʒɑneb/ →	[d͡ʒɑneb]	*[d͡ʒuneb]	[sæmt]	direction
13. /ketmɑn/ →	[ketmɑn]	*[ketmun]	[puʃɑndæn]	to hide
14. /mækɑn/ →	[mækɑn]	*[mækun]	[d͡ʒɑ]	place
15. /legɑm/ →	[legɑm]	*[legum]	[æfsɑr]	harness
16. /pejmɑn/ →	[pejmɑn]	*[pejmun]	[qol]	promise
17. /pejkɑn/ →	[pejkɑn]	*[pejkun]	[tir]	arrow
18. /zæmɑn/ →	[zæmɑn]	*[zæmun]	[væqt]	time
19. /rismɑn/ →	[rismɑn]	*[rismun]	[bænd], [tænɑb]	rope
20. /ʃɑdmɑn/ →	[ʃɑdmɑn]	*[ʃɑdmun]	[xoʃ hɑl]	happy
21. /ʃæbɑn/ →	[ʃæbɑn]	*[ʃæbun]	[t͡ʃupɑn] → [t͡ʃupun]	shepherd
22. /sijɑnæt/ →	[sijɑnæt]	*[sijunæt]	[negæhdɑri]	vigilance
23. /tæʔɑm/ →	[tæʔɑm]	*[tæʔum]	[qæzɑ]	food
24. /tɑzjɑne/ →	[tɑzjɑne]	*[tɑzjune]	[ʃællɑq]	whip
25. /zemɑm/ →	[zemɑm]	*[zemum]	[æfsɑr]	harness
26. /æjɑn/ →	[æjɑn]	*[æjun]	[mæʔlum]	visible
27. /zijɑn/ →	[zijɑn]	*[zijun]	[zærær]	loss
28. /kærɑn/ →	[kærɑn]	*[kærun]	[hæd]	limit
29. /mɑnænd/ →	[mɑnænd]	*[munænd]	[mesl]	similarity
30. /nijɑm/ →	[nijɑm]	*[nijum]	[qælɑf]	scabbard
31. /færd͡ʒɑm/ →	[færd͡ʒɑm]	*[færd͡ʒum]	[ɑxær]	end
32. /bɑmdɑd/ →	[bɑmdɑd]	*[bumdɑd]	[sobh], [sæhær]	dawn
33. /bɑstɑn/ →	[bɑstɑn]	*[bɑstun]	[qædim]	days of old
34. /bɑzærgɑn/ →	[bɑzærgɑn]	*[bɑzærgun]	[tɑd͡ʒer]	merchant
35. /bigɑne/ →	[bigɑne]	*[bigune]	[qæribe]	stranger
36. /dɑnɑ/ →	[dɑnɑ]	*[dunɑ]	[ɑqel]	wise

37. /pɑjɑn/ →	[pɑjɑn]	*[pɑjun]	[ɑxær]	end
38. /hengɑm/ →	[hengɑm]	*[hengum]	[væqt]	time
39. /ɑrmɑn/ →	[ɑrmɑn]	*[ɑrmun]	[hædæf]	goal

There are three points that should be highlighted. First, some of the formal Persian words are Arabic borrowings (e.g. (91.10), (91.11) and (91.12)), and some are native Persian words (e.g. (91.15) and (91.16)); given this, some of the TP equivalents of these formal words are native Persian words and some are Arabic loanwords; i.e., there are formal words that are native Persian lexical items, but their TP equivalents are Arabic loanwords. This suggests that Arabic loanwords in TP are so nativized in the TP lexicon that they are categorized as core TP lexical items, and they can be used as equivalents for formal lexical items which are native Persian. The data from (91.28) to (91.39) support this claim as in these words the formal words are native Persian lexical items that are not used in TP whereas their TP equivalents are Arabic loans. I know of no such cases where non-Arabic loanwords are the TP equivalent of a Persian formal lexical item. This observation adds support to the analysis that Arabic loanwords should be separated from non-Arabic borrowings since Arabic borrowings in comparison to other loanwords are more adapted to the TP lexicon.

Second, words like *modam* ‘frequently’, *d͡zæhan* ‘world’, *ruz-ane* ‘daily’ are used in TP, but they are not highly frequent words; in addition, they have more frequently used TP equivalents, *hæmæsh*, *donja* and *hær ruz*, respectively. Assuming that pre-nasal raising is an ongoing process, it might be expected that in the future pre-nasal *a* in these words will undergo raising.

Third, the formal Persian words in (91) are bare forms, without any affixes, but there are words that are categorized as formal Persian words as a result of affixation and hence raising in them is blocked. For the discussion of morphologically derived formal Persian words see the morphology chapter.

6.6. Interface of blocking factors

The last point in this chapter concerns words in which raising is blocked due to more than one factor. First, I discuss those words whose pre-nasal *a* does not raise due to two lexical factors, and then I address the interface of phonological and lexical blocking factors.

6.6.1. Two lexical blocking factors in one word

As shown above, formal Persian words with an *aN* sequence disfavor raising; similarly, pre-nasal *a* in Arabic broken plural loanwords that are used in TP does not raise. There are formal Persian words which are Arabic broken plurals. Raising in these words is inhibited by two lexical blocking factors, as shown in (92).

(92) *Formal words that are Arabic broken plurals*

UNDERLYING FORM	SURFACE FORM		GLOSS
/ælhɑn/ →	[ælhɑn]	*[ælhun]	voices, songs
/æsnɑm/ →	[æsnɑm]	*[æsnum]	idols
/ozzɑm/ →	[ozzɑm]	*[ozzum]	high rank people
/owzɑn/ →	[owzɑn]	*[owzun]	rhythms
/hokkɑm/ →	[hokkɑm]	*[hokkum]	governors
/æd͡ʒsɑm/ →	[æd͡ʒsɑm]	*[æd͡ʒsum]	masses
/æhʃɑm/ →	[æhʃɑm]	*[æhʃum]	cattle
/xoddɑm/ →	[xoddɑm]	*[xoddum]	servants

6.6.2. Phonological and lexical factors interface in words with *am* sequence

In this subsection I focus on multisyllabic words with an *am* sequence in which raising is blocked due to more than one factor. In (93) pre-nasal *a* does not raise since the word is a formal Persian lexical item; furthermore, *a* and *m* are not in the same syllable. (94) is a list of non-Arabic loanwords whose *a* and *m* are not in the same syllable. (95) illustrates Arabic broken plurals in which *a* and *m* are not in the same syllable.

(93) *formal words with [am] sequence whose [a] and [m] are not in the same syllable*

UNDERLYING FORM	SURFACE FORM		GLOSS
/ɑ.miq/ →	[ɑ.miq]	*[u.miq]	mix
/ɑ.ʃɑ.mi.dæn/ →	[ɑ.ʃɑ.mi.dæn]	*[ɑ.ʃu.mi.dæn]	to drink
/qɑ.mæt/ →	[qɑ.mæt]	*[qu.mæt]	height
/ge.rɑ.mi/ →	[ge.rɑ.mi]	*[ge.ru.mi]	respectable
/ke.rɑ.mæt/ →	[ke.rɑ.mæt]	*[ke.ru.mæt]	benevolence
/le.ʔɑ.mæt/ →	[le.ʔɑ.mæt]	*[le.ʔu.mæt]	meanness

(94) *non-Arabic loans with [am] sequence whose [a] and [m] are not in the same syllable*

UNDERLYING FORM	SURFACE FORM		GLOSS	SOURCE LANGUAGE
/ɑ.mɑ.tor/ →	[ɑ.mɑ.tor]	*[u.mɑ.tor]	amateur	French
/de.rɑ.mɑ.tik/ →	[de.rɑ.mɑ.tik]	*[de.ru.mɑ.tik]	dramatic	English
/mɑ.mɑn/ →	[mɑ.mɑn]	*[mu.mɑn][38]	mom	French

(95) *Arabic broken plurals with [am] sequence whose [a] and [m] are not in the same syllable*

UNDERLYING FORM	SURFACE FORM		GLOSS	SOURCE LANGUAGE
/æ.vɑ.mel / →	[æ.vɑ.mel]	*[æ.vu.mel]	factors	Arabic
/mæ.d͡ʒɑ.meʔ/ →	[mæ.d͡ʒɑ.meʔ]	*[mæ.d͡ʒu.meʔ]	associations	Arabic
/æ.sɑ.mi/ →	[æ.sɑ.mi]	*[æ.su.mi]	names	Arabic
/mæ.zɑ.min/ →	[mæ.zɑ.min]	*[mæ.zu.min]	contents	Arabic
/æ.rɑ.mæ.ne/ →	[æ.rɑ.mæ.ne]	*[æ.ru.mæ.ne]	Armenians	Arabic

6.7. Summary

This chapter shows that there are five categories of lexical items in the Persian lexicon that disfavor raising: proper names, non-Arabic loanwords, Arabic broken plurals and infinitives, recent coinages, and formal Persian lexical items.

Given the interaction of lexical and phonological blocking factors, I should add that an important general result is that there are words whose pre-nasal *a* does not have any phonological restriction for raising but raising in them is blocked due to morphological and lexical blocking

[38] *[mɑ.mun] and *[mu.mun] are not acceptable.

factor(s). For instance, pre-nasal *a* in *mehran* ‘male name’ does not raise as the word is a proper name. Raising in *antibijutik* ‘antibiotics’ is inhibited as it is a non-Arabic loanword and pre-nasal *a* in *sazman* ‘organization’ does not raise as the word is a recent coinage. Hence, for pre-nasal raising it is required that the *aN* sequence meet the phonological requirements, but this does not guarantee raising, as there might be other non-phonological factors at play which block raising in a lexical item.

CHAPTER SEVEN

INTERACTION OF BLOCKING FACTORS

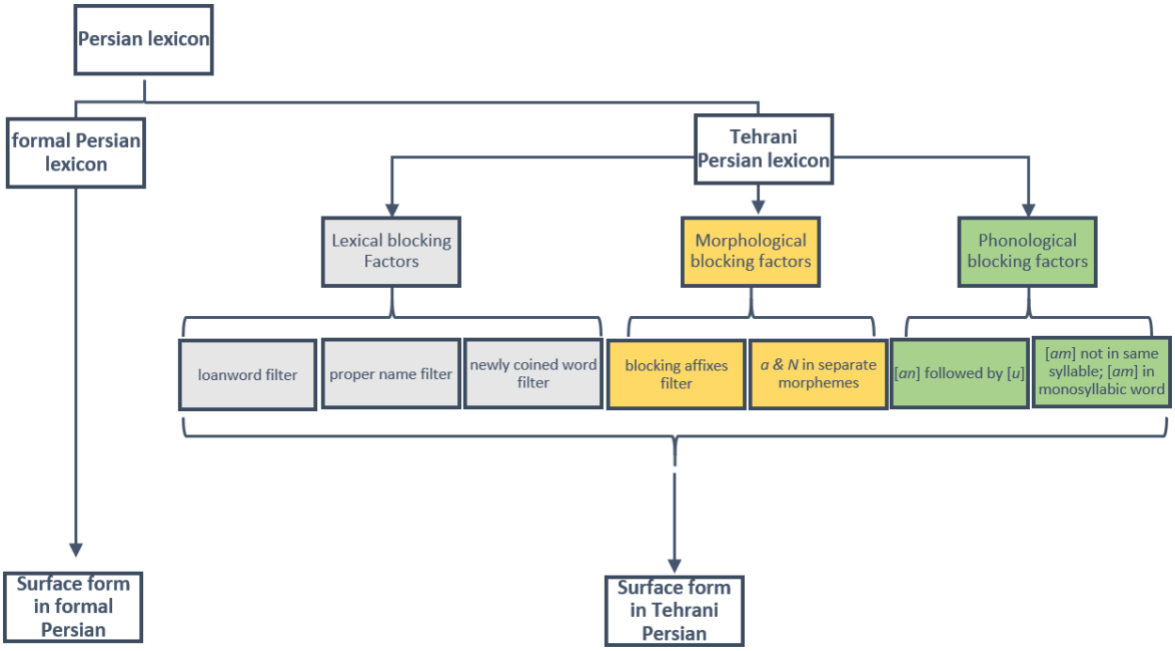
7.1. Introduction

In this study I have proposed that there are three categories of blockers of *a* to *u* raising in TP; phonological (chapter four), morphological (chapter five) and lexical (chapter six). Each category contains subcategories that are discussed in the related chapter. This chapter investigates TP lexical items in which there is more than one blocking factor, aiming to examine if there is any interaction between them, and presents an analysis of the strength of the blocking factors by looking at the percentage of the lexicon that individual and intersecting blocking factors account for.

7.2. A schematic model

As discussed in the introductory chapter and elsewhere, I assume that the Persian lexicon consists of two main sub-lexicons, the formal Persian (FP) lexicon and the TP lexicon. When a word is part of the FP lexicon, there is no relationship between that word and pre-nasal raising since pre-nasal raising occurs only in TP, not in formal Persian; in other words, being part of the formal Persian lexicon is itself a blocking factor. (96) shows the relationship between these two sub-lexicons with pre-nasal raising and its blocking factors:

(96) *a schematic model of the Persian lexicon*



The Persian lexicon is divided into FP and TP, as discussed above. Pre-nasal raising and the factors that block it are relevant to the TP lexicon, as shown in (96). The model shows that FP words, being in a separate part of the Persian lexicon, systematically block raising even when its conditions are met; for instance, for the data in (97) there are not any blocking factors to disallow raising; however, raising does not occur. This suggests that when a word is part of the FP lexicon, none of the blocking factors is considered for that word, it simply does not undergo raising. Only those lexical items that are in the Tehrani Persian lexicon are subject to raising.

(97) *sample of FP words which do not have any blockers but resist raising*

UNDERLYING FORM	SURFACE FORM		GLOSS
/æf.gɑ.ne/	[æf.gɑ.ne]	*[æf.gu.ne]	aborted child
/ær.mæ.qɑn/	[ær.mæ.qɑn]	*[ær.mæ.qun]	gift
/ɑ.rɑm.gɑh/	[ɑ.rɑm.gɑh]	*[ɑ.rum.gɑh]	cemetery
/bɑng/	[bɑng]	*[bung]	shout
/fæ.qɑn/	[fæ.qɑn]	*[fæ.qun]	shout
/bɑs.tɑn/	[bɑs.tɑn]	*[bɑs.tun]	ancient times
/hen.gɑm/	[hen.gɑm]	*[hen.gum]	time
/le.gɑm/	[le.gɑm]	*[le.gum]	harness
/mæ.ʃɑm/	[mæ.ʃɑm]	*[mæ.ʃum]	smell
/ni.jɑm/	[ni.jɑm]	*[ni.jum]	scabbard
/pɑ.jɑn/	[pɑ.jɑn]	*[pɑ.jun]	end
/rɑj.gɑn/	[rɑj.gɑn]	*[rɑj.gun]	free of charges
/se.nɑn/	[se.nɑn]	*[se.nun]	arrowhead, spearhead
/tæ.ʔɑm/	[tæ.ʔɑm]	*[tæ.ʔum]	food
/mo.ʔɑ.ne.de/	[mo.ʔɑ.ne.de]	*[mo.ʔu.ne.de]	animosity
/qi.jɑm/	[qi.jɑm]	*[qi.jum]	uprising
/mɑ.nænd/	[mɑ.nænd]	*[mu.nænd]	akin, similar
/feq.dɑn/	[feq.dɑn]	*[feq.dun]	deprivation
/xɑm.jɑ.ze/	[xɑm.jɑ.ze]	*[xɑm.ju.ze]	yawning

Second, as shown in the model, on the path of a word to enter from TP lexicon to the surface form, there are lexical, phonological and morphological blocking factors that influence pre-nasal raising. One possibility is that there are no relevant blocking factors, and the pre-nasal *a* successfully passes the blocking filters, as shown in (98).

(98) *sample of TP words which undergo raising*

UNDERLYING FORM	SURFACE FORM IN TP	GLOSS
/ɑ.vi.zɑn/	[ɑ.vi.zun]	hanging
/ɑn/	[un]	that
/bɑ.dɑm/	[bɑ.dum]	almond
/ko.dɑm/	[ko.dum]	which

Another possibility is that one or two of the blocking filters disallow pre-nasal raising and the word surfaces with *aN*, as shown in (99).

(99) *sample of TP words which do not undergo raising*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS	REASON FOR NOT RAISING
1. /ɑ.sɑn.sor/	[ɑ.sɑn.sor]	*[ɑ.sun.sor]	elevator	non-Arabic loanword
2. /sæ.rɑ.ne/	[sæ.rɑ.ne]	*[sæ.ru.ne]	per capita	recently coined word
3. /zɑ.nu/	[zɑ.nu]	*[zu.nu]	knee	*an* preceding *u*
4. /dɑ.mæn/	[dɑ.mæn]	*[du.mæn]	skirt	*a* and *m* not in same syllable

7.3. Intersection of two blocking factors in a word

This section addresses those words in the TP lexicon in which pre-nasal raising is inhibited by more than one blocking factor. First I discuss the interface of morphology and phonology, and then the interaction of morphological and lexical factors. Finally, I highlight the interface of phonological and lexical factors.

7.3.1. Morphology-phonology interface

As discussed in the morphology chapter, if *a* and its following nasal consonant are not in the same morpheme raising is blocked (100.a). In the phonology chapter it is argued that raising in multisyllabic words that contain an *am* sequence is inhibited if *a* and *m* are not in the same syllable (100.b). In the same chapter it is shown that raising in monosyllabic words with an *am* sequence is blocked as well (100.c).

(100) *sample of TP words with different morphophonological blocking factors*

a. /bæ.rɑ-m/ for-me	→	[bæ.rɑm]	*[bæ.rum]	'for me'	different morphemes
b. /dɑ.mæn/	→	[dɑ.mæn]	*[du.mæn]	'skirt'	different syllables
c. /dɑm/	→	[dɑm]	*[dum]	'trap'	monosyllable

Using the blocking factors of (100.a) and (100.c), in this section I list those words in which raising is prohibited due to both phonological and morphological factors. (101) is a list of

monosyllabic words containing an *am* sequence whose *a* and *m* are not in the same morpheme, and raising is blocked.

(101) *two intersecting blockers: phonological (single syllable) & morphological (different morphemes)*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/bɑ-m/ with-me	[bɑm]	*[bum]	with me
/d͡ʒɑ-m/ place-my	[d͡ʒɑm]	*[d͡ʒum]	my place
/mɑ-m/ we-too	[mɑm]	*[mum]	we too
/pɑ-m/ foot-my	[pɑm]	*[pum]	my foot

Similarly, (102) is a list of multisyllabic words with an *am* sequence whose *a* and *m* are not in the same syllable nor are they in the same morpheme.

(102) *two intersecting blockers: phonological (syllable boundary) & morphological (different morphemes)*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/bɑ-mæqz/ with-core	[bɑ.mæqz]	*[bu.mæqz]	meaningful
/bɑ-mæze/ with-taste	[bɑ.mæ.ze]	*[bu.mæ.ze]	tasty, cute
/bɑ-mæʔni/ with-meaning	[bɑ.mæʔ.ni]	*[bu.mæʔ.ni]	meaningful
/bɑ-mæʔrefæt/ with-gallantry	[bɑ.mæʔ.re.fæt]	*[bu.mæʔ.re.fæt]	gallant
/bɑ-mohæbæt/ with-kindness	[bɑ.mo.hæ.bæt]	*[bu.mo.hæ.bæt]	kind
/bɑ-morovæt/ with-compassion	[bɑ.mo.ro.væt]	*[bu.mo.ro.væt]	chivalrous
/lɑ-mæzhæb/ not-religion	[lɑ.mæz.hæb]	*[lu.mæz.hæb]	heretic
/nɑ-mærd/ not-man	[nɑ.mærd]	*[nu.mærd]	scoundrel
/nɑ-mæfhum/ not-meaning	[nɑ.mæf.hum]	*[nu.mæf.hum]	unclear

/nɑ-monɑseb/ not-suitable	[nɑ.mo.nɑ.seb]	*[nu.mo.nɑ.seb]	unsuitable
/nɑ-monæzæm/ not-organized	[nɑ.mo.næ.zæm]	*[nu.mo.næ.zæm]	disorganized
/nɑ-mæʔqul/ not-rational	[nɑ.mæʔ.qul]	*[nu.mæʔ.qul]	irrational
/nɑ-mowzun/ not-rhythmic	[nɑ.mow.zun]	*[nu.mow.zun]	unrhythmic
/nɑ-motenɑseb/ not-proportionate	[nɑ.mo.te.nɑ.seb]	*[nu.mo.te.nɑ.seb]	disproportionate
/nɑ-mosɑʔed/ not-appropriate	[nɑ.mo.sɑ.ʔed]	*[nu.mo.sɑ.ʔed]	inappropriate
/nɑ-morætæb/ not-organized	[nɑ.mo.ræ.tæb]	*[nu.mo.ræ.tæb]	disorganized

These examples support the main observation of this chapter: if two blocking factors occur in the same word, the result is the same; pre-nasal raising is blocked which suggests that the intersection of two blockers in the same lexical item does not affect each other's function; each blocker separately inhibits raising.

7.3.2. Phonology-lexical factors interface

In this section words are given in which raising is blocked due to two factors, one phonological and the other lexical. This section is about the intersection of different blockers in words with an *am* sequence only. The reason is that there are not any interactions between the phonological blocker of words with an *an* sequence and other lexical blocking factors.

7.3.2.1. Monosyllabic words with an *am* and lexical factors

The only monosyllabic word with an *am* sequence in which there is an interface of blockers is *sam*. This word is a male proper name and as mentioned above, it is a monosyllabic word with an *am* sequence.

7.3.2.2. Multisyllabic words with an *am* sequence and lexical factors

In this subsection I discuss those words in which phonological and lexical blocking factors intersect. The phonological factor is common: *a* and *m* are not in the same syllable. There are four different lexical blockers. Hence, as the phonological blocking factor is constant, the words are categorized based on these four factors: (103) is a list of proper names, (104) includes recent coinages, in (105) the words are non-Arabic borrowings, and (106) is a list of triliteral Arabic infinitives. In all cases, raising fails to apply.

(103)[39] *two intersecting blockers: phonological (syllable boundary) & lexical (proper name)*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/ɑ.me.ne/	[ɑ.me.ne]	*[u.me.ne]	female name
/ɑ.mu/	[ɑ.mu]	*[u.mu]	name of a river
/ɑ.mol/	[ɑ.mol]	*[u.mol]	city name
/ɑ.mer/	[ɑ.mer]	*[u.mer]	male name
/fæ.rɑ.mærz/	[fæ.rɑ.mærz]	*[fæ.ru.mærz]	male name
/hæ.xɑ.mæ.neʃ/	[hæ.xɑ.mæ.neʃ]	*[hæ.xu.mæ.neʃ]	dynasty name
/hɑ.med/	[hɑ.med]	*[hu.med]	male name
/hen.gɑ.me/	[hen.gɑ.me]	*[hen.gu.me]	female name
/d͡ʒɑ.mi/	[d͡ʒɑ.mi]	*[d͡ʒu.mi]	poet's name
/ne.zɑ.mi/	[ne.zɑ.mi]	*[ne.zu.mi]	poet's name
/rɑ.min/	[rɑ.min]	*[ru.min]	male name
/sɑ.mi/	[sɑ.mi]	*[su.mi]	Semitic
/sjɑ.mæk/	[sjɑ.mæk]	*[sju.mæk]	male name
/væ.rɑ.min/	[væ.rɑ.min]	*[væ.ru.min]	city name
/vɑ.meq/	[vɑ.meq]	*[vu.meq]	male name
/ke.rɑ.mæt/	[ke.rɑ.mæt]	*[ke.ru.mæt]	male name

(104) *two intersecting blockers: phonological (syllable boundary) & lexical (recent coinage)*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/bæxʃ nɑ.me/ distribution letter	[bæxʃ nɑ.me]	*[bæxʃ nu.me]	edict
/bɑr nɑ.me/	[bɑr nɑ.me]	*[bɑr nu.me]	bill of loading

[39] Names of remote villages and cities in Iran and names of books are not listed in (102).

load letter			
/go.zær nɑ.me/ transit letter	[go.zær nɑ.me]	*[go.zær nu.me]	passport
/pɑ.sox nɑ.me/ answer letter	[pɑ.sox nɑ.me]	*[pɑ.sox nu.me]	answer sheet
/bæ.sɑ.mæd/	[bæ.sa.mæd]	*[bæ.su.mæd	frequency
/sɑ.mɑ.ne/	[sɑ.mɑ.ne]	*[su.mɑ.ne][40]	system

(105) *two intersecting blockers: phonological (syllable boundary) & lexical (non-Arabic loans)*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/de.rɑ.mɑ.tik/	[de.rɑ.mɑ.tik]	*[de.ru.mɑ.tik]	dramatic
/ɑ.mɑ.tor/	[ɑ.mɑ.tor]	*[u.mɑ.tor]	amateur
/ɑ.muk.si.si.lin/	[ɑ.muk.si.si.lin]	*[u.muk.si.si.lin]	name of a medicine
/ɑ.mu.ni.jɑk/	[ɑ.mu.ni.jɑk]	*[u.mu.ni.jɑk]	ammoniac
/his.tɑ.min/	[his.tɑ.min]	*[his.tu.min]	histamine
/mɑ.mɑn/	[mɑ.mɑn]	*[mu.mɑn][41]	mom

(106) *two intersecting blockers: phonological (syllable boundary) & lexical (triliteral Arabic infinitive)*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/e.dɑ.me/	[e.dɑ.me]	*[edume]	continuation
/e.qɑ.me/	[e.qɑ.me]	*[equme]	raising (for prayer)
/mo.ʔɑ.me.le/	[mo.ʔɑ.me.le]	*[moʔumele]	trade
/tæ.ʔɑ.mol/	[tæ.ʔɑ.mol]	*[tæ.ʔu.mol]	discussion, reaction

7.3.3. Interaction of two lexical factors

This section addresses those words in which raising is blocked due to two lexical blocking factors.

(107) is a list of proper names that are non-Arabic loanwords as well.

[40] *sɑ.mu.ne* and *su.mu.ne* are not acceptable either.
[41] *mamun* and *mumun* are not acceptable either.

(107)[42] *two intersecting lexical blockers: proper name and non-Arabic loanword*

UNDERLYING FORM	SURFACE FORM IN TP		GLOSS
/kɑ.nɑ.dɑ/	[kɑ.nɑ.dɑ]	*[ku.nɑ.dɑ]	Canada
/mi.lɑn/	[mi.lɑn]	*[mi.lun]	Milan
/fe.rɑnk/	[fe.rɑnk]	*[fe.runk]	Frank (male name)
/ʒɑn.vi.je/	[ʒɑn.vi.je]	*[ʒun.vi.je]	January
/de.sɑmbr/	[de.sɑmbr]	*[de.sumbr]	December
/no.vɑmbr/	[no.vɑmbr]	*[no.vumbr]	November

7.4. The hierarchy of blockers based on the degree of restrictiveness

In this section I categorize the blocking factors in a hierarchical order based on their degree of restrictiveness. That is, the more a blocking factor prevents raising, the higher its ranking is in the hierarchy of blockers. For instance, the morpheme boundary blocker is a top-ranking factor since there is never pre-nasal raising in words whose *a* and *N* are in two separate morphemes. Likewise, pre-nasal raising is never found in non-Arabic loanwords or in recent coinages. This means that they are top-ranking factors. However, there are blocking factors that encounter exceptions, and having exceptions means that they are not as restrictive as the abovementioned factors, so they are not top-ranking blockers.

7.4.1. Quantified information about the database

Before categorizing each blocker, it is necessary to provide quantified information about the database. As shown in (108), the database includes 1780 words with an *aN* sequence. 33.39% of this number are formal Persian (FP) lexical items. These words are excluded from the following calculations since they are insensitive to pre-nasal raising and hence insensitive to the blocking factors.

[42] This list is a sample of loanwords that are proper names. For further data in this regard, see the database.

(108) *quantified description of the database*

LEXICAL ITEMS	NUMBER	PERCENTAGE
Total number of TP & FP words	1780	100%
Tehrani Persian words	1176	66.06%
Formal Persian words	604	33.93%
TP & FP words with *am*	498	27.97%
TP & FP Words with *an*	1257	70.61%
TP & FP Words with *an* & *am*	25	1.40%

7.4.2. Quantification of the TP database (FP words excluded)

(109) shows that the phonological, morphological and lexical factors, altogether, block pre-nasal raising in 71.25% of the TP lexical items.

(109) *quantified description of TP database*

TP LEXICAL ITEMS	NUMBER	PERCENTAGE
Total[43]	1176	100%
aN > **uN*	838	71.25%
aN > *uN*	338	28.74%

In (110) the second cell of the third column shows what percentage of words with an *am* sequence out of all of the words with an *am* sequence in the TP database resist raising. The second cell of the fourth column shows what percentage these resisting words in the whole TP database occupy; for instance, 90.64% of the words with an *am* sequence do not undergo raising. This (252 words out of the total of 278) is 21.42% of the TP database.

The third column in (110)-(112) shows the same type and proportion of information about the words in which raising is not inhibited.

(110) *quantified description of TP words with [am] sequence*

TP WORDS WITH *an* SEQUENCE	NUMBER	PERCENTAGE	% IN TP WORDS
total	278	100%	23.63%
am > **um*	252	90.64%	21.42%
am > *um*	26	9.35%	2.21%

[43] Formal Persian words are excluded.

(111) *quantified description of TP words with [an] sequence*

TP WORDS WITH *an* SEQUENCE	NUMBER	% IN *an* WORDS	% IN TP WORDS
total	883	100%	75.08%
an > **un*	572	64.77%	48.63%
an > *un*	311	35.22%	26.44%

(112) *quantified description of TP words with [an]&[am] sequence*

TP WORDS WITH *am* & *an SEQUENCES*	NUMBER	% IN *am* & *an* WORDS	% IN TP WORDS
total	15	100	1.27%
no raising	14	93.33%	1.19%
am & *an* > *am* & *un*[44]	1	6.66%	0.08%

7.4.3. Quantification of individual blocking factors

Each table from (113) to (123) shows the number and percentage of a group of words in which raising is blocked by a common factor. These tables also show the number and percentage of counterexamples to each factor.

The structure of these tables is the same as above. For example, the second cell of the third column in (113) shows that 97.4% of proper names resist raising; this is 41.41% of the whole TP database. In these tables the third row gives information about the exceptions. For instance, the third row of (113) shows that 13 proper names or 2.6% of all proper names, undergo raising. This number represents 1.10% of the TP lexicon. When the variable in that cell is 0%, it means that that blocking factor does not have counterexamples; in other words, that blocker is completely restrictive and blocks raising in 100% of the words to which it could apply.

(113) *quantified description of proper names in TP*

PROPER NAMES IN TP	NUMBER	% IN PROPER NAMES	% IN TP WORDS
Total	500	100%	42.51%
aN > **uN*	487	97.4%	41.41%
aN > *uN*	13	2.6%	1.10%

[44] From among words with both an *am* and an *an* sequence only *an* sequence in *amjane* 'vulgar' undergoes raising.

(114) *quantified description of Arabic broken plurals in TP*

BROKEN PLURALS IN TP	NUMBER	% IN BROKEN PLURALS	% IN TP WORDS
Total	8	100%	0.6%
*aN > *uN*	8	100%	0.6%
aN > uN	0	0%	0%

(115) *quantified description of Arabic infinitives in TP*

ARABIC INFINITIVES IN TP	NUMBER	% IN ARABIC INF.	% IN TP WORDS
Total	31	100%	2.63%
*aN > *uN*	31	100%	2.63%
aN > uN	0	0%	0%

(116) *quantified description of morphologically derived words in TP*

MORPHOLOGICALLY DERIVED WORDS IN TP[45]	NUMBER	% IN MD	% IN TP WORDS
Total	70	100%	5.95%
*aN > *uN*	70	100%	5.95%
aN > uN	0	0%	0%

(117) *quantified description of words with [a] and [N] in separate morphemes*

[ɑ-N] IN TP	NUMBER	% IN [ɑ-N]	% IN TP WORDS
Total	64	100%	5.44%
*aN > *uN*	64	100%	5.44%
aN > uN	0	0%	0%

(118) *quantified description of recent coinages in TP*

RECENT COINAGES IN TP	NUMBER	% IN RECENT COINAGES	% IN TP WORDS
Total	38	100%	3.23%
*aN > *uN*	36	94.73%	3.06%
aN > uN	2	5.26%	0.17%

(119) *quantified description of words with [anu] sequence in TP*

*anu > *unu* IN TP	NUMBER	% IN *anu > *unu*	% IN TP WORDS
Total	6	100%	0.51%
*anu > *unu*	6	100%	0.51%
anu > unu	0	0%	0%

[45] Morphologically derived words are words that don't have any morphophonological/lexical blockers, but don't raise as they are derived from words in which raising is blocked.

(120) *quantified description of words with blocking affixes in TP*

BLOCKING AFFIX (*-eʃ* & *-ani*) IN TP	NUMBER	% IN *-eʃ* & *-ani*	% IN TP WORDS
Total	13	100%	1.10%
aN > **uN*	13	100%	1.10%
aN > *uN*	0	0%	0%

(121) *quantified description of non-Arabic loans in TP*

LOANWORDS IN TP	NUMBER	% IN LOANWORDS	% IN TP WORDS
Total	108	100%	9.18%
aN > **uN*	108	100%	9.18%
aN > *uN*	0	0%	0%

(122) *quantified description of monosyllabic words with [am] sequence*

TP MONOSYLLABIC WITH *am*	NUMBER	% IN TP MONOSYLLABIC WITH *am*	% IN TP WORDS
TOTAL	16	100%	1.36%
am > **um*	15	93.75%	1.27%
am > *um*	1	6.25%	0.08%

(123) *quantified description of multisyllabic words with [a] and [m] in separate syllables*

a.m > **u.m* IN TP	NUMBER	% IN *a.m* > **u.m*	% IN TP WORDS
Total	164	100%	13.94%
a.m > **u.m*	157	95.73%	13.35%
a.m > *u.m*	7	4.26%	0.59%

This quantified description shows that the phonological blockers of words with an *am* sequence are less restrictive than morphological and lexical factors.

7.4.4. Quantification of intersecting blockers

In the following tables, first I show the interaction of morphology with phonology; then the phonology-lexical factors interface and finally the interface of two lexical factors in the same word. Tables from (124) to (131) show that when two blockers intersect in a word, without any exception, raising fails to occur. Thus, there is not any hierarchy of restrictiveness between one class of intersecting blockers with other classes.

7.4.4.1. Morpho-phonology interface

(124) *quantified description of two blockers: morphological (morpheme boundary) and phonological (monosyllabic word with [am] sequence)*

a-N & am IN MONOSYLLABIC WORD	NUMBER	% IN TP WORDS
Total	4	0.34%
*aN > *uN*	4	0.34%
*aN > *uN*	0	0 %

(125) *quantified description of two blockers: morphological (morpheme boundary) and phonological (syllable boundary)*

a-N & [a.m]	NUMBER	% IN TP WORDS
Total	38	3.23%
*aN > *uN*	38	3.23%
*aN > *uN*	0	0%

7.4.4.2. Phonology-lexical factors interface

(126) *quantified description of two blockers: lexical (proper name) and phonological (syllable boundary)*

PROPER NAME *& [a.m]*	NUMBER	% IN TP WORDS
Total	47	3.99%
*a.m > *u.m*	47	3.99%
a.m > u.m	0	0%

(127) *quantified description of two blockers: lexical (Arabic infinitive) and phonological (syllable boundary)*

ARABIC INFINITIVE *& [a.m]*	NUMBER	% IN TP WORDS
Total	4	0.34%
*a.m > *u.m*	4	0.34%
a.m > u.m	0	0%

(128) *quantified description of two blockers: lexical (recent coinage) and phonological (syllable boundary)*

RECENT COINAGE *& [a.m]*	NUMBER	% IN TP WORDS
Total	7	0.59%
*a.m > *u.m*	7	0.59%
a.m > u.m	0	0%

(129) *quantified description of two blockers: lexical (non-Arabic loans) and phonological (syllable boundary)*

LOANWORDS & *[a.m]*	NUMBER	% IN TP WORDS
Total	6	0.51%
a.m > **u.m*	7	0.51%
a.m > *u.m*	0	0%

7.4.4.3. Two lexical factors interface

(130) *quantified description of two lexical blockers: proper name and non-Arabic loans*

PROPER NAME & LOANWORD	NUMBER	% IN TP WORDS
Total	17	1.44%
aN > **uN*	17	1.44%
aN > **uN*	0	0%

(131) *quantified description of two lexical blockers: proper name and recent coinage*

PROPER NAME & RECENT COINAGE	NUMBER	% IN TP WORDS
Total	5	0.42%
aN > **uN*	5	0.42%
aN > **uN*	0	0%

According to the number of exceptions that each blocker encounters, in (132) I show the hierarchical order of the blocking factors. The proper name blocker with 2.6% of exceptions ranks right after the top-ranking blockers. The syllable boundary blocker in multisyllabic words with an *am* sequence, with 4.26% of exceptions, is in the third place. And the least restrictive blocker with 6.25% of exceptional cases is the blocker of monosyllabic words which contain an *am* sequence.

(132) *hierarchy of blockers based on degree of restrictiveness*

HIERARCHY	NAME OF BLOCKING FACTORS	PERCENTAGE OF EXCEPTIONS
1	recent coinage blocker	0%
	non-Arabic loanwords blocker	0%
	blocking affix	0%
	morpheme boundary blocker	0%
	dissimilating *a* in *anu* sequence	0%
	interface of 2 blockers in the same word	0%
2	proper names	2.6%
3	*a* and *m* in separate syllables	4.26%
4	monosyllabic words with *am* sequence	6.25%

7.5. Patternless exceptions

The analysis presented in this study encounters 27 patternless exceptions. These exceptions can be classified into two categories. One is words whose pre-nasal *a* should undergo raising but doesn't (133). The other consists of multisyllabic words with an *am* sequence in which raising should not occur since *a* and *m* are not in the same syllable; nevertheless, pre-nasal *a* raises to *u*, shown in (134). (135) shows the percentage these irregularities occupy in the TP lexicon.

(133) *list of words which should undergo raising but don't*

UNDERLYING FORM	SURFACE FORM		GLOSS
/ælʔan/ →	[ælʔan]	*[ælʔun]	now
/ænd͡ʒam/ →	[ænd͡ʒam]	*[ænd͡ʒum]	do (something)
/buran[46]/ →	[buran]	*[burun]	thunderstorm
/bambul/ →	[bambul]	*[bumbul]	deceit
/dang/ →	[dang] or [dong]	*[dung]	sixth of a property
/xijanæt/ →	[xijanæt]	*[xijunæt]	betrayal
/xandan/ →	[xandan]	*[xundan[47]]	household
/sowhan/ →	[sowhan]	*[sowhun]	file, rasp
/sælam/ →	[sælam]	*[sælum]	hello
/qaneʔ/ →	[qaneʔ]	*[quneʔ]	convinced
/gambu/ →	[gambu]	*[gumbu]	fat
/ostovane/ →	[ostovane]	*[ostovune]	cylinder

[46] As in *bad o buran* 'wind and thunderstorm'
[47] *xandun* and *xundun* are not acceptable either

/mɑneʔ/ →	[mɑneʔ]	*[muneʔ]	obstacle
/mæsɑne/ →	[mæsɑne]	*[mæsune]	bladder
/mærɑm/ →	[mærɑm]	*[mærum]	creed
/mæqɑm/ →	[mæqɑm]	*[mæqum]	rank
/kowhɑn/ →	[kowhɑn]	*[kowhun]	hump
/dɑstɑn/ →	[dɑstɑn]	*[dɑstun]	story
/mæd͡ʒɑni/ →	[mæd͡ʒɑni]	*[mæd͡ʒuni]	free of charges

(134) *list of words which shouldn't undergo raising but do*

UNDERLYING FORM	SURFACE FORM	GLOSS
/dɑ.mɑd/ →	[du.mɑd]	groom
/ɑ.mæ.d-æn[48]/ → came-INF	[u.mæ.dæn]	to come
/bær ɑ.mæ.d-e-gi/ → ahead came-adj marker-noun marker	[bær u.mæ.de.gi]	bulge
/bɑr ɑ.mæ.d-æn[49]/ → growth came-INF	[bɑr u.mæ.dæn]	to grow
/fo.rud ɑ.mæ.d-æn/ → descent came-INF	[fo.rud u.mæ.dæn]	to land (for plane)
/vær ɑ.mæ.d-æn/ side came-INF	[vær u.mæ.dæn]	to rise (for paste)
/piʃ ɑ.mæd/ → before came	[piʃ u.mæd]	incident
/bɑm[50]/ →	[bum]	roof

Note that not all morphologically derived forms of past tense stem *amæd* 'came' undergo raising. For instance, pre-nasal *a* in *dæramæd* 'income' does not undergo raising.

To summarize, (135) shows the number and percentage of exceptions.

[48] Its inflections also undergo raising.

[49] As in *æz pæs bær amædæn* 'manage to do a task'

[50] Recall that monosyllabic words with an *am* sequence resist raising.

(135) *quantified description of patternless exceptions in TP*

PATTERNLESS EXCEPTIONS	NUMBER	% IN TP WORDS
Total	27	2.29%
Words that shouldn't raise but do	7	0.59%
Words that should raise but don't	20	1.70%

7.6. Summary

This chapter highlights the distinction between the FP and TP lexicons and focuses on TP lexical items in which raising is inhibited by one or more one blocking factor. I have shown the interaction of different blockers in the same word and have attempted to demonstrate that these intersecting blockers neither trigger nor nullify the function of each other.

CHAPTER EIGHT

CONCLUSION

8.1. Introduction

This study attempts to identify the factors which block pre-nasal raising in Tehrani Persian (TP). In the literature, historical (Sadeghi 2001 and Miller 2011), sociolinguistic (Modaressi 1978, Jahangiri 1980) and phonological analyses (Kahn and Bernstein 1981, Kalbasi 2001, Bakhtiari 2008, Rees 2008, and Rohany 2012) are proposed to identify the influential elements of this phenomenon. Despite the valuable endeavours and analyses, there remain unanswered questions in the literature and contradictory analyses which face counterexamples. The aim of this study is to present a comprehensive analysis which answers the questions raised in the literature.

The first step in identifying the blocking factors was to create a database of words which contain an *aN* sequence. This database includes 1782 lexical items of formal Persian (FP) and Tehrani Persian. As shown in the model chapter, 33.93% of the words in the database are FP lexical items that are excluded in the analyses presented in this study; the reason is that pre-nasal raising occurs only in TP, and FP words are insensitive to this phenomenon.

The database builds on three categories of blocking factors, phonological, morphological and lexical, which influence raising in TP. Each of these factors includes subfactors which are discussed the relevant chapters.

What distinguishes the phonological analysis in this study from the phonological analyses in the literature is that I claim that the phonological blocking factors in words with an *am* sequence differ from the blocking factors in words which contain an *an* sequence (see the phonology

chapter). In addition, I introduce the morphological domain of pre-nasal raising. It occurs only within a morpheme. While it is agreed in the Persian literature that blocking occurs when the *aN* sequence fall between morphemes, it is a surprise from a phonological perspective, as it is unexpected to find a rule operating only within morphemes.

In the morphology chapter, the effect of inflectional and derivational affixes on pre-nasal raising is discussed. The findings show that these affixes do not play a role in triggering or blocking pre-nasal raising (for the discussion of the morphologically derived FP words and the interaction of morphology with lexical factors, see the morphology chapter). Nonetheless, there are two derivational suffixes which pattern differently: the noun marker *-eʃ* and the adjective marker *-ani*. When *-eʃ* attaches to a base which contains an *aN* sequence, raising in the derived form fails to occur. And when *-ani* suffixes to a base, the pre-nasal *a* in the affix resists raising. Particularly interesting is that when a non-monosyllabic *am*-final word hosts a vowel-initial suffix, pre-nasal raising is not sensitive to the surface syllable structure of the inflected/derived form; raising applies although in the resyllabified form *a* and *m* are not in the same syllable.

Chapter six addresses the influence of lexical factors on pre-nasal raising. In this chapter I show that proper names, non-Arabic loanwords, Arabic broken plurals, Arabic infinitival forms and recently coined words do not undergo raising. An important result of this chapter is that loanwords that are borrowed into Persian fit into the Core-Periphery model proposed by Itô and Mester (1999). Based on this model, there is a degree of nativization among loanwords in a language based on the duration of the time these borrowings have been used in the language; the older a loanword, the more it adapts to the phonology of the borrowing language. Given that, since pre-nasal raising is observed in Arabic loanwords only (except infinitives and broken plurals), I claim that Arabic loans are more nativized to the phonology of Tehrani Persian than non-Arabic

loanwords. Of interest is the exceptional Arabic infinitives and broken plurals – one might expect them to pattern with the rest of the Arabic loanwords.

Chapter seven focuses on the interface of two blocking factors in the same word to investigate whether two blockers influence each other's function in triggering or nullifying pre-nasal raising. The result was that the two intersecting blockers do not have any effect on each other; they are two separate mechanisms that pattern differently from one another and these blockers by chance coincide in some words.

In the same chapter, I present a quantified description of the database and show what percentage of the TP database is occupied by which blocking factor. Some of these blockers encounter exceptions, suggesting that that blocker is not totally restrictive in preventing pre-nasal raising. According to the calculations presented in chapter seven, there is a hierarchy of restrictiveness among the blocking factors. Some blockers like the morpheme boundary are top-ranking factors since they do not have any counterexamples, but some blockers like the proper name factor have exceptions. An important point about this hierarchy is that two intersecting blockers in the same word do not have any exceptions.

In the end, this multidimensional analysis shows that blockage of pre-nasal raising in TP is systematically under the influence of different factors. The analysis predicts the raising or non-raising of pre-nasal *a*. In the literature, there are different and sometimes contradictory analyses. This study attempts to provide an account of pre-nasal raising and identify the blocking factors, revealing that much of which has been assumed to be exception in the literature is in general, systematically accounted for, with very few exceptions.

8.2. Questions for further studies

This work raises many questions for future research. Derived environment effects are one interesting question: as discussed in chapter four, it is striking that raising occurs only within a morpheme. Why should this be? Another is what it means for the lexicon to be stratified by factors such as Arabic broken plurals and infinitives. In addition, I have made appeals to frequency. In order to substantiate these claims, a corpus-based analysis of the interaction of frequency with pre-nasal raising is required. A final crucial question is the synchronic status of raising. Both Modaressi and Jahangiri identify sociolinguistic factors that are important in whether raising occurs or not. A study of within and between individuals is likely to shed light on whether raising is synchronic or not, and also provide insight into the structure of the Persian lexicon.

References

Burzio, Luigi. 2011. Derived environment effects. *The Blackwell Companion to Phonology*. Malden, MA: Blackwell.

Deyhime, Guity. 2000. "France xvi. Loanwords in Persian," *Encyclopaedia Iranica* X/2, pp. 181-184, available online at http://www.iranicaonline.org/articles/france-xvi-loan-words-in-Persian- (accessed on 30 December 2012).

Hodge, Carleton T. 1957. Some aspects of Persian style. *Language* 33.3, 355-369.

Inkelas, Sharon. 2014. *The interplay of morphology and phonology*. Vol. 8. Oxford University Press.

Itô, Junko, and Armin Mester. 1999. The phonological lexicon. Tsujimura, Natsuko, eds. *The Handbook of Japanese Linguistics*. Blackwell Publishers, Malden, Massachusetts, and Oxford, U.K. 62-100.

Jahangiri, Nader. 1980. A sociolinguistic study of Tehrani Persian. PhD Dissertation. University of London, UK.

Kahn, Margaret, and Jared Bernstein. 1981. Progress of the uN Change in Persian: dæbestan vs. tabestun. *Acta Orientalia Academiae Scientiarum Hungaricae* 35.1, 133-140.

Kahnemuyipour, Arsalan. 2003. Syntactic categories and Persian stress. *Natural Language & Linguistic Theory*, *21*(2), 333-379.

Kalbasi, Iran. 2001. Spoken and Written Persian (In Persian). *Farhang* Vol 37 & 38, 49-68.

Kang, Yoonjung. 2011. Loanword Phonology. In Marc van Oostendorp, eds. *The Blackwell Companion to Phonology: General issues and segmental phonology*. Vol. 1. John Wiley & Sons. 1—25.

Kemmer, Suzzane. 2017. Retrieved on April 6, 2018 from http://www.ruf.rice.edu/~kemmer/Words/loanwords.html.

Kłagisz, Mateusz. 2013. Hints on French loanwords in modern New Persian. *Romanica Cracoviensia* 13(1), 38-51.

Kiparsky, Paul. 1969. Linguistic Universals and Linguistic Change. In E. Bach and R. Harms, eds. *Universals in Linguistic Theory*. New York, Holt, 170-210.

Kiparsky, Paul. 1973a. Phonological representations: Abstractness, opacity, and global rules. In Osamu Fujimura (ed.) *Three dimensions of linguistic theory*, 57–86. Tokyo: Taikusha.

Kiparsky, Paul. 1993. Blocking in nonderived environments. In Hargus & Kaisse (1993), 277–313.

Labov, William. 2007. Transmission and diffusion. *Language* 83.2, 344-387.

Mahmoodi-Bakhtiari, Behrooz. 2008. Tehrani Dialect (In Persian). *The Center for Great Islamic Encyclopedia* Vol 16, 575-579.

McCarthy, John J. 2003. OT constraints are categorical. *Phonology*, 20 (1), 75-138.

Miller, Corey. 2011. A Holistic Treatment of/ān/to [un] in Persian. Proceedings of ICPhS XVII, 1386-1389.

Miller, Corey. 2013. Variation in Persian vowel systems. *Orientalia Suecana* 61, 156-169. McCarthy 2003a

Modaressi, Yahya. 1978. A sociolinguistic investigation of modern Persian. PhD Dissertation. University of Kansas, USA.

Morony, Michael G. 1986. Arab ii. Arab conquest of Iran, *Encyclopaedia Iranica*, II/2, pp. 203-210, available online at http://www.iranicaonline.org/articles/arab-ii (accessed on 30 December 2012).

Pisowicz, Andrzej. 1985. *Origins of the New and Middle Persian Phonological system*. Nakladem Uniwersytetu Jagiellonskiego

Perry, John. 2003. Persian as a Homoglossic Language. In *Iran, questions et connaissances.* Vol. 3, *Cultures et sociétés contemporaines*, ed. B. Hourcade, pp. 11–28. Studia Iranica, Cahier 27. Actes du 4e Congrès européen d'études iraniennes organisé par la Societas Iranologica Europea, Paris, 6–10 septembre 1999. Paris: Peeters.

Perry, John. Brian Spooner, and William L. Hanaway. 2012. New Persian: Expansion, Standardization, and Inclusivity. *Literacy in the Persianate World: Writing and the Social Order*, pp.70-94.

Rahbar, Elham Rohany. 2012. Aspects of Persian phonology and morpho-phonology. PhD Dissertation. University of Toronto, Canada.

Rees, Daniel A. 2008. Towards proto-Persian: An optimality theoretic historical reconstruction. PhD Dissertation. Georgetown University, USA.

Ryding, Karin. 2014. Arabic phonology. In *Arabic: A Linguistic Introduction*. Cambridge: Cambridge University Press.13-22.

Sadeghi, Ali Ashraf. 2001. Historical Issues of Persian Language (In Persian). *Sokhan* Publications, Tehran, Iran.

Windfuhr, Gernot, ed. *The Iranian Languages*. Psychology Press, 2009.

Dictionaries:

Amid, Hasan. 1965. Farhange Amid (Amid Dictionary). Tehran: Javidan publications.

Collins French Dictionary | Translations, Definitions and Pronunciations https://www.collinsdictionary.com/dictionary/english-french.

Dehkhoda Online Dictionary (http://www.jasjoo.com/books/wordbook/dehkhoda/)

Ganjoor Persian poetry database (https://ganjoor.net/)

Moein Online Dictionary (http://www.jasjoo.com/books/wordbook/)