

**The Impact of the Farhangestān on Modern Persian:**  
A Case Study of Political Terms in Iranian Newspapers and Blogs

Dissertation  
zur Erlangung des Grades  
der Doktorin der Philosophie (Dr. phil)  
an der Fakultät für Geisteswissenschaften  
der Universität Hamburg

vorgelegt von  
**Asmaa Shehata**  
Hamburg, 2022

**Names of the evaluators according to §7(5) and §8 of the  
doctorate regulations for the Faculty of Humanities at the  
University of Hamburg (Dated, 7 July 2010)**

Chairperson: Prof. Dr. Ludwig Paul


Deputy Chairperson: Prof. Dr. Henning Schreiber

Third Evaluator: Prof. Dr. Stefan Heidemann

Date of Disputation: 13 July 2022

**Oath according to §7(4) of the doctorate regulations for  
the Faculty of Humanities at the University of Hamburg  
(Dated, 7 July 2010)**

I herewith declare upon oath that I wrote the present dissertation independently and that I did not use any other resources and aids than those stated.

Hamburg, 9 January 2022  \_\_\_\_\_  
Place, Date, Signature



# Acknowledgment

In the name of God most gracious, most merciful. All praise to the almighty, all-knowing, and most wise. I am grateful to God for giving me the courage and patience to endure this journey of seeking knowledge.

First and foremost, I am extremely grateful to my supervisor, Prof. Dr. Ludwig Paul, for his invaluable advice, continuous academic and organizational support, and patience during my doctoral studies. His immense knowledge and plentiful experience have encouraged me in all the time of my academic research.

My deepest and special thanks go to my husband Jens Kutscher without whose support, time, and company this work would not be the same. His knowledge and experience in the field of Middle Eastern studies and his devotion as a father to our children helped me focus on my work. Many thanks are due also to my Mum and Dad who have always given me all kind of support in life.

I would furthermore like to express my gratitude to my dear friend Nouran Khallaf from Leeds University in England for thought-provoking discussions especially on the content related to computational linguistics. Many thanks go also to Zeinab Refahi Bakhsh from the University of Gilan in Iran for offering a helping hand in obtaining whatever references, sources, and previous studies from Iran I needed and for the time she spent with me as a Persian native speaker in lengthy language discussions. Special thanks are due to Nesrin Parvisi, vice-chairwoman of the APLL terminology group, and Mehnoosh Tehrani, scientific board member and researcher in the APLL terminology group, for providing me with all their publications and for not hesitating to answer my questions and providing me with academic explanations throughout the years of study.

Finally, I would like to express my appreciation to the first one who believed in my abilities in the field of Iranian Studies and Persian Linguistics when I started my studies in Egypt, the late Professor Dr. Abdelhafiz Yaqoub Hegab, and to the Egyptian Bureau for Cultural Affairs and the Educational Mission in Berlin for their support.

# Table of Content

<b>Acknowledgment</b> .....	<b>4</b>
<b>Table of Content</b> .....	<b>5</b>
<b>List of Tables</b> .....	<b>8</b>
<b>List of Figures</b> .....	<b>12</b>
<b>List of Charts</b> .....	<b>13</b>
<b>Chapter 1: Introduction</b> .....	<b>14</b>
1.1 Purpose of the Study.....	16
1.2 Research Questions.....	17
1.3 Methodology.....	18
1.4 Overview.....	19
<b>Chapter 2: Historical Background &amp; Language Evolution</b> .....	<b>22</b>
2.1 A cursory Overview of Language Academies.....	22
2.2 Linguistic Borrowing.....	24
2.2.1 Borrowing in Linguistic Scholarship.....	24
2.2.2 Academic Perspectives on Borrowing in Persian.....	26
2.2.3 The Impact of Borrowing on the Persian Lexicon.....	27
2.2.4 History of Borrowing in New Persian.....	31
2.2.5 The Impact of Arabic and Turco-Mongolian Languages.....	32
2.2.6 The Onset of Modernity and European Influences on Persian.....	33
2.3 Language Planning and Language Policy.....	36
2.4 The Academy of Persian Language and Literature.....	38
2.4.1 Etymology and History of the Term <i>Farhangestān</i> .....	38
2.4.2 Early Attempts at Institutionalizing Language Preservation.....	39
2.4.3 The First Farhangestān (1935).....	42
2.4.4 The Second Farhangestān (1970).....	45
2.4.5 Random Efforts at Coining Foreign Terms.....	46

2.4.6	The Third Farhangestān: The Academy of Persian Language and Literature (APLL, 1990) .....	48
2.5	The Terminology Group .....	51
2.5.1	Terminology .....	51
2.5.2	<i>Vāžehgozīnī</i> (Term Formation) or <i>Eṣṭelāḥšenāsī</i> (Terminology)?.....	53
2.5.3	Goals and Tasks of the Terminology Group .....	54
2.5.4	Principles and Regulations of the Terminology Group.....	56
2.5.5	What makes Persian Persian?.....	57
2.5.6	Public Relations and Raising Awareness Activities .....	58
2.6	Permissible Persian Terms in Term Formation .....	60
<b>Chapter 3: Review of Literature .....</b>		<b>63</b>
3.1	Iranian Scholarly Disputes.....	63
3.2	Previous Studies on Persian Term Formation .....	67
3.2.1	Research on Coining Political Terms .....	67
3.2.2	Research on Coining Other Specialized Terms.....	70
3.2.3	Challenges, Recommendations, and Outlook .....	73
<b>Chapter 4: Methodology .....</b>		<b>75</b>
4.1	Introduction .....	75
4.2	Selected Terms .....	75
4.3	Corpus.....	77
4.3.1	Newspapers in the Corpus Terms_Analysis_Persian (TAP).....	77
4.3.2	Blog Posts in the Corpus TalkBank Persian.....	80
4.4	The TF-IDF Equation and the Adjusted Weight of Terms.....	83
<b>Chapter 5: Corpus Analysis .....</b>		<b>85</b>
5.1	The Concept of <i>Parliament</i> .....	90
5.1.1	Analysis of the Foreign Term <i>pārlemān</i> .....	93
5.1.2	Analysis of the Approved Term <i>mağles</i> .....	100
5.2	The Concept of <i>Candidate</i> .....	112

5.3	The Concept of <i>Opposition</i> .....	117
5.4	The Concept of <i>Nation</i> .....	118
5.5	The Concept of <i>Democracy</i> and Its Derivatives.....	127
5.5.1	Analysis of the Foreign Terms <i>demūkrāt/demūkrātīk</i> .....	129
5.5.2	Analysis of the Foreign Terms <i>mardombāvar/mardomsālārī</i> .....	130
5.6	The Concept of <i>Dictatorship</i> and Its Derivatives .....	131
5.7	The Concept of <i>Regime</i> and Its Derivatives .....	134
5.8	The Concept of <i>Bureaucracy</i> and Its Derivatives .....	142
5.9	The Concept of <i>Ideology</i> and Its Derivatives .....	144
5.10	The Concept of <i>Terrorism</i> and Its Derivatives .....	150
5.11	The Concept of <i>Populism</i> and Its Derivatives .....	155
5.12	The Concept of <i>Capitalism</i> and Its Derivatives.....	156
5.13	The Concept of <i>Reformism</i> and Its Derivatives .....	160
5.14	The Concept of <i>Doctrine</i> and Its Derivatives .....	164
5.15	The Concept of <i>Strategy</i> and Its Derivatives .....	165
5.16	The Concept of <i>Protocol</i> .....	168
<b>Chapter 6: Findings and Conclusion .....</b>		<b>175</b>
6.1	Findings: Describing the Impact of Coined Terms on Modern Persian .....	175
6.2	Conclusion: Explaining the Impact of Coined Terms on Modern Persian.....	187
6.3	Recommendations: Ensuring the Positive Impact of Coined Terms on Modern Persian .....	190
<b>Bibliography .....</b>		<b>193</b>
<b>Online Resources .....</b>		<b>211</b>
<b>Appendix .....</b>		<b>217</b>

## List of Tables

Table 1. Selected foreign terms and the equivalent Persian terms approved by the APLL * <i>Ommat</i> is the term that is compared with <i>mellat</i> even if it does not qualify as a foreign term	76
Table 2. Number of followers of major Iranian newspapers on Twitter and Telegram.....	79
Table 3. Iranian newspapers in the corpus TAP (with an individual word count of >1mn) ....	79
Table 4. Uninserted newspaper websites in the corpus TAP .....	80
Table 5. Term frequency and document frequency of <i>pārlemān</i> and its derivative forms in the two corpora.....	94
Table 6. Examples of <i>pārlemān</i> in the plural form .....	95
Table 7. Selected examples of the most frequent general collocations of the foreign term <i>pārlemān</i> in TalkBank corpus .....	95
Table 8. Examples of general collocations with <i>pārlemān</i> .....	96
Table 9. Examples of collocations with the term <i>pārlemān</i> , absolute frequency, and document frequency in the corpora.....	97
Table 10. Examples of country names collocations with <i>pārlemān</i> .....	98
Table 11. Examples of the use of proper nouns in collocation with the term <i>pārlemān</i> .....	99
Table 12. Examples of the foreign term <i>pārlemān</i> in use to indicate the Iranian parliament	100
Table 13. Term frequency and document frequency of <i>mağles</i> and <i>mağlesī</i> in both corpora	100
Table 14. High-frequency compound terms with <i>mağles</i> .....	102
Table 15. Collocations of the term <i>mağles</i> in the corpora, term frequency and document frequency.....	103
Table 16. Examples of the most frequent compound terms of the term <i>mağles</i> with country names, absolute frequency, and document frequency .....	106



Table 17. Examples of <i>mağles</i> in context from the corpora.....	110
Table 18. Term frequency and document frequency of the terms in the semantic field of <i>candidate</i> .....	112
Table 19. High-frequency collocations for the approved term <i>nāmzad</i> (term frequency and document frequency).....	113
Table 20. High-frequency collocations for the approved term <i>kāndīdā</i> (term frequency and document frequency).....	114
Table 21. Examples of both borrowed and foreign terms used in the same text.....	115
Table 22. Term frequency and document frequency of the terms in the semantic field of <i>opposition</i> .....	118
Table 23. Term frequency and document frequency of the approved term <i>mellat</i> and the non-approved term <i>ommat</i> .....	122
Table 24. The most frequent collocations (compound terms) of the approved term <i>mellat</i> in both corpora.....	123
Table 25. The most frequent collocations (compound terms) of <i>mardom</i> in both corpora....	124
Table 26. The most frequent collocations (compound terms) of <i>ommat</i> in both corpora .....	125
Table 27. Examples of <i>mellat</i> , <i>ommat</i> , and <i>mardom</i> in the two corpora .....	126
Table 28. Approved terms that contain <i>sālārī</i> or <i>bāvar</i> often as a replacement for the English suffixes <i>-cracy</i> or <i>-archy</i> respectively .....	128
Table 29. Term frequency and document frequency for <i>demūkrāt</i> , <i>demkūkrāsī</i> , and <i>demūkrātīk</i> .....	129
Table 30. Term frequency and document frequency for <i>mardombāvar</i> and <i>mardomsālārī</i> ..	130
Table 31. Term frequency and document frequency of the terms in the semantic field of <i>dictatorship</i> .....	133

Table 32. Term frequency and document frequency of the approved terms <i>reżīm</i> and <i>nezām</i> .....	138
Table 33. Collocations of the approved foreign term <i>reżīm</i> .....	138
Table 34. Collocations of the approved Persian term <i>nezām</i> .....	140
Table 35. Term frequency and document frequency of the terms in the semantic field of <i>bureaucracy</i> .....	144
Table 36. Examples of <i>īde'ūlūżī</i> and <i>marām</i> in the corpora .....	147
Table 37. Term frequency and document frequency of the terms in the semantic field of <i>ideology</i> .....	149
Table 38. Term frequency and document frequency of the terms in the semantic field of <i>terrorism</i> .....	153
Table 39. Term frequency and document frequency of the terms in the semantic field of <i>populism</i> .....	156
Table 40. Examples of both <i>kāpītālīsm</i> and <i>sarmāyedārī</i> in the corpora.....	158
Table 41. Term frequency and document frequency of the terms in the semantic field of <i>capitalism</i> .....	159
Table 42. Frequency and document frequency of the foreign terms and their approved Persian equivalents in the semantic field of <i>reform</i> .....	162
Table 43. Examples of <i>eşlāḥ</i> and <i>reform</i> from the corpora.....	163
Table 44. Term frequency and document frequency of the terms in the semantic field of <i>doctrine</i> .....	164
Table 45. Term frequency and document frequency of the terms in the semantic field of <i>strategy</i> .....	167
Table 46. Examples of the foreign term <i>esterāteżī</i> and the approved term <i>rāhbord</i> in context from TalkBank and TAP .....	168

Table 47. Term frequency and document frequency of the terms in the semantic field of <i>protocol</i> .....	170
Table 48. Examples of the foreign term <i>protokol</i> in the corpora .....	172
Table 49. Examples of the approved term <i>tašrīfāt</i> in both corpora.....	173
Table 50. Weight ratios of the examined term pairs in TalkBank and TAP as well as the type of borrowing of the approved Persian equivalent term .....	188

## List of Figures

Figure 1. Distribution of online newspapers in the corpus TAP sorted by term frequency.....	80
Figure 2. Visualization of the frequency of collocations compared between <i>pārlemān</i> and <i>mağles</i> in TalkBank.....	110
Figure 3. Visualization of the frequency of collocations compared between <i>kāndīdā</i> and <i>nāmzad</i> in TalkBank.....	115
Figure 4. Visualization of the frequency of collocations compared between <i>īde 'ulūžī</i> and <i>marām</i> in TalkBank.....	148

## List of Charts

- Chart 1. Comparison of adjusted weight values of foreign terms (green) and their Persian equivalents (red) in TalkBank ..... 177
- Chart 2. Comparison of adjusted weight values of foreign terms (green) and their Persian equivalents (red) in TAP ..... 179
- Chart 3. Comparison of adjusted weight values of the terms *pārlemān* (green) and *mağles* (red = absolute weight, yellow = approximate weight) in both corpora..... 181
- Chart 4. Comparison of adjusted weight values of the terms *kāndīdā* (green) and *nāmzad* (red) and approximate weight values of *kāndīdā* (gray) and *nāmzad* (yellow) in both corpora ..... 182
- Chart 5. Comparison of adjusted weight values of the terms *ommat* (green), *mellat* (red), and *mardom* (gray) in both corpora ..... 183
- Chart 6. Comparison of weight values of the terms *režīm* (green) and *nezām* (red) as well as of approximate weight values of the terms *režīm* (gray) and *nezām* (yellow) in both corpora.. 184
- Chart 7. Comparison of adjusted weight values of the terms *īde'ūlūzī* (green) and *marām* (red = absolute weight, yellow = approximate weight) as well as of the terms *īde'ūlūzīkī* (blue) and *marāmī* (purple) in both corpora ..... 185
- Chart 8. Comparison of adjusted weight values of the terms *prūtokol* (green) and *tašrīfāt* (red = absolute weight, yellow = approximate weight) in both corpora ..... 186

# Chapter 1

## Introduction

Language and identity are two of the most important and personal elements that shape our lives. In fact, language itself can be considered a living being as it lives on the tongue of its speakers. It develops, changes, and evolves according to the developments and life changes that occur to its speakers and their place on earth. All these changes are imprinted on the individual languages. With the development of life and its different societal systems – and consequently for every human group, be it tribes, civilizations, or the current form of countries – each of them has a language in which its members, its citizens deal with one another and with those belonging to other groups: foreigners who do not speak the same language. God Almighty said in the Quran: “O mankind, we created you from a male and a female and made of you peoples and tribes so that you may get to know one another” (sura *al-Huġūrāt* 49:13).<sup>1</sup>

As a result of this communication, exchange, and overlap between peoples – whether for political, social, commercial, or other reasons – languages overlap as well and are affected by each other. The more the speakers of those languages interact, the greater the languages will interfere. This linguistic phenomenon is known as *language interference* meaning “those instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language” (Weinreich, 1968, p. 1). Since every language has its own lexicon, when such linguistic interference occurs, words move between different languages – a phenomenon known in linguistics as *linguistic borrowing*.

As for the Persian language, it has been subject to many linguistic frictions and its words have overlapped with the words of other languages, either for political, social, or other reasons. One of the most important influences on the Persian language – and its influence remains to this day – was the Arab-Islamic conquest of Iran between the seventh and ninth centuries of the Common Era. From a linguistic perspective, this resulted in the emergence of many anti-Arab tendencies and language purification movements which include also Arabophobic ideas. Turco-Mongolic languages left a strong mark on the Persian language, too, since a lot of Turkic words were brought into Persian after the Mongol invasion of Iran in the 13<sup>th</sup> century. In the late 18<sup>th</sup> and early 19<sup>th</sup> centuries – especially with the onset of modernity in Europe and the period of

<sup>1</sup> يَا أَيُّهَا النَّاسُ إِنَّا خَلَقْنَاكُمْ مِنْ ذَكَرٍ وَأُنْثَىٰ وَجَعَلْنَاكُمْ شُعُوبًا وَقَبَائِلَ لِتَعَارَفُوا

Enlightenment –, the influence of European languages began to appear in many oriental languages of which Persian is one (Mo‘īn M. , 1386 [2007]a., pp. 32-37).

In the wake of these developments, the role and importance of modern language planning emerges. The concept of language planning could be summed up in the following definition by Fishman as “the organized pursuit of solutions to language problems, typically at the national level” (Fishman, 1973, p. 24) and the definition given by Tauli according to which “language planning is the methodical activity of regulating and improving existing languages or creating new common regional, national or international languages” (Tauli, 1968, p. 56). Karam, finally, suggests that language planning “refer[s] to an activity which *attempts* to solve a language problem usually on a national scale, and which focuses on either language form or language use or both” (Karam, 1974, p. 105). In this sense, the term-formation process or the coining of new terms (neologisms) is meant to fix linguistic problems and involves a centralized institution charged with this process.

According to article 15 of the Constitution of the Islamic Republic of Iran, the Persian language is the only official language of the country. It is the only language allowed to be used in documents, official correspondence, and textbooks. The Arabic language is also mentioned in the Iranian constitution (article 16) as the religious language in which the Quran and Islamic knowledge were revealed – and Arabic is also the language that has had a lasting influence on Persian literature. Therefore, the constitution acknowledges the inclusion of Arabic in the basic education curricula as an essential part of primary education until the end of secondary school (The Constitution of the Islamic Republic of Iran, 1989). As for the English language as the universal language of communication, while not enjoying any official status under Iranian constitutional law, it is also found in the educational stages (Aḥmadīpūr, 1390[2011], p. 13).

The Academy of Persian Language and Literature (APLL) or Farhangestān is the official institution that is responsible for language planning in Iran. It is concerned with the preservation of the originality and continuity of this language as well as with its promotion. The APLL has existed since before the Iranian Revolution of 1979. What is nowadays known as the first phase was initiated in 1935 in the period of Reza Shah Pahlavi following the Turkish example. It is known as *Farhangestān-e āvval*, that is, the First Academy or the First Farhangestān. Purifying the Persian language and coining new Persian terms was the main task of the First Academy, which also included various research activities at that time such as the publication of the monthly journal *Nāme-ye Farhangestān* as well as of a book with the newly coined terms. The First Farhangestān ceased its activities in 1954 after the abdication of Reza Shah Pahlavi.

The *Farhangestān-e zabān-e Īrān* or Second Academy was effectively founded in 1970 when Mohammad Reza Shah Pahlavi declared the establishment of *Bonyad-e šāhenšāhī-ye Farhangestānhā-ye Īrān* (Imperial Foundation of Iranian Academies). The main objectives of the Second Academy were the same as those of the First Academy, which was mainly the preservation of the Persian language and its purification from all foreign terms. Among others, this Academy published the periodical *Pažūheš-nāme* (*Research Journal*). It came to its end in 1978 with the beginning of the demonstrations that ended the Shah's rule and brought about the Islamic Republic (Jazayeri, 1999).

In 1990, the Supreme Council of Cultural Revolution (*Šūrā-ye 'ālī-ye enqelāb-e farhangī*) declared the establishment of the Academy of Persian Language and Literature (*Farhangestān-e zabān-e va adab-e fārsī*) or Third Farhangestān as the official governmental institution that is mainly concerned with preserving, strengthening, and disseminating the Persian language. The current Academy consists of 14 research units tasked with subjects such as the Grammar of Persian Language and Lexicography, Iranian Languages and Dialects, Contemporary Literature, the Literature of the Islamic Revolution, Language and Computer, and last but not least Terminology. One of the main responsibilities of the terminology group (*gorūh-e vāžehgozīnī*) is to coin and approve neologisms to replace foreign terms that entered the language, which is the main concern of the present study.

Language, as stated by the prominent Iranian thinker, author, translator, linguist, and intellectual Dāryūš Āšūrī, is not just a complex set or string of words following each other. Rather, it is the attachment between them that forms the texture of speech (Āšūrī, 1392 [2014], p. 242). Therefore, the process of forming a term in the target language should ensure that the new term retains the original meaning and concept of the foreign term in a way that is harmonized with sentences and speech (Şafavī K. , 1388 [2009], p. 40). The lack of harmony between approved neologisms and different contexts may adversely affect the prevalence, success, and practical use of those words (Monšī-zādeh & Sharif, 1384 [2005], p. 20). To look deeper into the process of term formation and how neologisms are – or are not – embedded in the target language is at the core of the present study.

## **1.1 Purpose of the Study**

The objective of this thesis is to examine the impact of the Academy of Persian Language and Literature (Third Farhangestān) as the official institution in charge of Persian language planning in Iran on modern Persian by measuring the extent of the use of political neologisms proposed



by the Farhangestān in newspapers and blog posts. Since 2004, the APLL has published 16 dictionaries introducing newly coined terms to Persian and recommending them for replacing non-Persian loanwords on an annual basis. These dictionaries are organized by topics, two of which are Political Science and International Relations. It is these two topics that are in the focus of this thesis and form the point of departure for the analysis of the use of the proposed and approved neologisms in modern Persian texts.

Thanks to this case study of political neologisms in modern Persian, this thesis aims furthermore to contribute to the understanding of a hitherto understudied or even neglected institution in Western academia. In Iran, on the other hand, there have been several studies and papers on the APLL, for example, with regard to the new terms approved by the Academy and criticism thereof.

It is crucial to study and trace the impact of the Academy on modern Persian through the application of the new political terms in newspapers and blogs to gain a better insight into language politics on the one hand and into language formation or language evolution on the other. The extent of the use of the selected new terms and their success in a significant part of language use – and not only in the dictionaries published by the APLL – will be ascertained. In addition, a look at the strategies of term formation used by the Academy will support the study on the impact of political neologisms in Persian.

## **1.2 Research Questions**

The following questions will serve as the basis for the study of the APLL and the evaluation of its organization, its scope, and its impact. They can be grouped into three categories.

First comes a quantitative or statistical assessment: How many of the proposed and approved neologisms in the categories of Political Science and International Relations are actually used in newspapers and blogs? How frequently and how heavily are they used? And how does the use of approved Persian equivalents compare to that of the foreign terms they were designed to replace?

A qualitative linguistic reading of the neologisms and of the process of their development completes this overview: How does the APLL replace existing terms? Which types of borrowing can be observed in the Persian equivalents? How can the impact of the new terms be determined by linguistic, semantic, and other factors?

Finally, a contextual or social reading of the neologisms will complement the findings from a perspective of language politics: Why does the APLL initiate a term-changing process? Why does it want to change specific terms? Is there a difference in using the neologisms between different types of publications? For example, do newspapers make faster and/or more frequent use of these terms after their introduction than blogs?

Altogether, this research examines the impact of the terms coined by the Farhangestān on the Persian language used in newspapers and blogs. Based on the set of quantitative, qualitative, and contextual questions raised above, this study looks at the sources or type of publication, the type of linguistic borrowing, and the time that has passed since the formal introduction of the neologisms in question to ask whether they have been successful or not.

### **1.3 Methodology**

This study is based methodologically on a corpus analysis of the online archives of major Iranian newspapers and blogs. To be precise, two different and distinct corpora with altogether more than 500 million words serve as primary sources for the quantitative and qualitative analysis. One corpus contains the online archives of numerous Iranian newspapers while the other corpus contains countless blog posts. The two corpora hence represent a wide variety of language in use. Their sheer size helps obtain meaningful results to assess the impact of the neologisms approved by the APLL on modern Persian over time.

The first corpus is a self-created database called *Terms\_Analysis\_Persian*, which will be referred to as TAP in this thesis. It covers the largest and most widely distributed Iranian newspapers with more than 30 million words in more than 22,000 documents. It was built during the time of this research starting from 2018 till 2021 by using the online tool Sketch Engine ([www.sketchengine.eu](http://www.sketchengine.eu)) just like the second corpus.

The second corpus used in this study is *TalkBank Persian (blog posts)* or simply *TalkBank*, which was created from Persian blogs collected by Shlomo Argamon's research group at the Illinois Institute of Technology (IIT) in cooperation with Brian MacWhinney's group at Carnegie Mellon. It contains more than 474 million words in more than one million documents. This is helpful in the context of the present thesis because it sheds light on the use of terms in daily writings thus juxtaposing them with newspapers and their formal editing processes.

In order to be able to quantify the use of the selected neologisms and the foreign terms on which they are based in the corpora, this thesis will make use of the term frequency/inverse document frequency (TF-IDF) equation proposed and refined by various linguists in recent years. This helps calculate the weight of a given term and thus evaluate the extent of the impact – that is, the success or failure – of the approved neologisms vis-à-vis the loanwords.

## **1.4 Overview**

This thesis is divided into four chapters in addition to this Introduction (chapter 1) and the Findings and Conclusion (chapter 6). Chapter 2 moves along three main axes. Starting, first, with a discussion of the historical background of language academies in general, the chapter then goes on in its second section to focus on the issue of linguistic borrowing. In doing so, it answers the question of how linguistic borrowing affects the nature of languages in general and sheds light on the eventful history of linguistic borrowing in Persian in particular. Moreover, it looks at how this led to the importance of the existence of language planning and the centralization of language policy, which formed a strong basis for the formation of language academies seeking to preserve the originality of their respective language.

The third axis of chapter 2 scrutinizes the Academy of Persian Language and Literature (APLL) as well as its goals in all its stages. Moreover, it takes a closer look at the APLL terminology group and the framework within which the terminology group acts to form neologisms. As the unit concerned with studying the foreign terms that entered Persian, it works on finding authentic linguistic alternatives in the language to replace those terms and on keeping the same meaning with the aim of gaining their acceptance by native speakers. The chapter concludes with a glance at the guidelines of Persian term formation created and followed by the APLL.

Chapter 3 covers selected previous studies and efforts that deal with evaluating the impact of the terms coined and approved by the Academy of Persian Language and Literature on modern Persian. Many of them focus on the fields of Political Science and International Relations. But there have also been studies in other fields such as Medicine or Computer Science and Information Technology whose research is of value to the present study.

The fundamental focus of this study is covered in chapters 4 and 5. Chapter 4 concisely presents the methodology used in this study. The analysis is methodologically based (1) on a selection of foreign and approved terms in the field of politics that will be examined in the study, (2) on an extensive corpus analysis, and (3) on the weight calculation for all examined

terms. After presenting the selected terms, this chapter introduces and contextualizes the two corpora TAP and TalkBank through which the terms are evaluated. It also presents the so-called TF-IDF (term frequency/inverse document frequency) equation used to calculate the weight of the terms in both corpora, which will help later in comparing the approved equivalent terms to the foreign terms they were meant to replace. This is crucial to answering the leading question of this study about the impact of the terms formed and approved on modern Persian. In other words: Thanks to the weight calculation it is possible to measure the extent to which the Academy succeeded in making the neologisms it coined acceptable in newspapers and blogs, that is, among journalists and other proficient language users.

Chapter 5, finally, deals with the analysis of the selected terms. Each approved Persian term is presented alongside the foreign term it was meant to replace. All sections of chapter 5 follow the same pattern to facilitate comparisons. In the beginning of each section, the foreign terms and their Persian equivalents are presented separately. The terms are explained etymologically and, where possible, the term history is summarized and the date of their first use or entry in the Persian language is recorded. To be able to trace the terms, their history, and their evolution, some of the oldest Persian dictionaries and Iranian encyclopedias are consulted. Depending on their density, the terms are further contextualized, for example, by presenting derivatives (e.g., *demūkrāt* with *demūkrāsī*), related terms approved by the APLL, and compound terms formed with the terms in question.

The second part of each section in chapter 5 highlights the quantitative and statistical properties of each term in both corpora, that is essentially the term frequency and the document frequency. This is necessary to calculate their weight and put them in relation with each other. Some terms such as *mağles* (parliament) or *nāmzad* (candidate) have more than one meaning. In these cases, the historical-etymological discussion and the quantitative analysis are supported by a collocation analysis (context analysis) that allows for a more accurate weight calculation thanks to the use of approximated frequencies. Furthermore, where ambiguous, the terms in question will be contextualized by providing example texts from the corpora to make a better understanding of their everyday use possible.

Each section of chapter 5 concludes with a discussion of the results and implications of the weight calculation. This ascertains the actual use of each Persianized term in newspapers and blogs and allows for conclusions regarding its success or lack thereof vis-à-vis the foreign term it came to replace.

Taken together, the analysis of each term according to its type of source or publication, the determination of its type of borrowing, as well as the consideration of the time that has passed since the (formal) introduction of that term reveal and put in perspective the impact of the Third Farhangestān on modern Persian.

## Chapter 2

# Historical Background & Language Evolution

### 2.1 A cursory Overview of Language Academies

“A language that is not continuously updated is left, over time, with a marginal presence in formal and prestigious communication, and consequently, plays little or no part in international communication” (Cabr , Bagot, & Chelo, 2012, p. 2). In light of the openness that accompanies the modern world and the large number of inventions and technological developments, new terms and words often appear on the scene in most domains. Naturally, these new terms are in the language of the country where they appear. And since languages are an open space for the transfer of words from one language to another, these neologisms will remain strange and foreign to all other languages.

Language overlap is not a new thing. The history of languages is full of language interference, linguistic invasion, word spread, and transfer from one environment to another. This linguistic transfer includes also various linguistic phenomena in phonology, morphology, grammar, etc. Here, the decision remains to the native speakers: either to confront and fight the new foreign terms or to adapt to them. Creating new terms that are appropriate to the language is considered one of the most important ways to keep the native language and the identity of it. Shohamy (Shohamy, 2006, p. xv) identified the power of language through its being used as “a form of control, by imposing the use of certain languages in certain ways”.

One way to counter or control the influx of new terms into a language is language planning. Language planning as defined by Joan & Jernudd (1971, p. xvi) “is a deliberate change in language by an organization set up to bring about the change”. The aim of language planning is to transform a language or adjust linguistic activities for many reasons, the most important of which is the preservation of the language identity of a nation. Creating new terms as well as standardizing grammatical, phonological, and morphological forms of non-standard forms are some of those linguistic activities. In the definition of Kaplan & Baldauf Jr. (1997, p. 3, as quoted in Nekvapil (2011, p. 871)), “language planning is a body of ideas, laws, regulations (language policy), change rules, beliefs, and practice intended to achieve a planned change (or to stop change from happening) in the language use in one or more communities”.

Hence, language academies appeared as responsible institutions to preserve native languages. One prominent and classical example of language planning, in the context of a

nation-building process, is the centuries old Académie Française in France, founded in the 17<sup>th</sup> century in 1634 (Académie Française) and known today mostly for its strict approach to anglicisms, that is, any English terms making their way into the French language.

In the 19<sup>th</sup> and 20<sup>th</sup> centuries, with the rise of nationalism and the nation state inspired by the French Revolution,<sup>2</sup> there were also other European national movements concerned with language planning, not least with the dissolution of the multi-ethnic, multi-national Ottoman and Austro-Hungarian empires. One of these examples of language planning and policy took place in newly founded Czechoslovakia in the 1920s and 1930s, the participants of which were the linguists of the Prague Linguistic School (Nekvapil, 2011).

Examples of such efforts could also be found outside Europe. The first academy of the Arabic language was established in Damascus in 1919. In 1932, the Academy of the Arabic Language in Cairo was established by a decree of King Fouad I, with half of the members being Egyptians and the other half a selection of international members such as orientalist or non-Egyptian Arabs where the only selection criterion was competence. The main objectives of the academy were to prepare linguistic dictionaries, to discuss language matters, to approve scientific and linguistic terms, and to implement Arab heritage and cultural activities. In 1974, the Academy of Arabic Language in Baghdad was established. The three academies united in 1971 to form the general union of the academy of the Arabic language to coordinate efforts to promote the language and maintain its integrity while being updated with modern life topics (Academy of The Arabic Language, 2021).

Similarly, the Academy of the Hebrew Language was founded in 1953, five years after the creation of the State of Israel, with the aim to gather and research Hebrew language vocabulary of all periods and levels, to research the forms of Hebrew and its history, and to guide the natural development of the language according to the needs and possibilities in all fields, in vocabulary, grammar, writing, and transliteration. Modern Hebrew (Ivrit) was a special case in terms of language formation because it was not the native language of people speaking it and it existed only in religious texts (The Academy of the Hebrew Language, 2021).

Language academies as defined by Edward (2009, p. 257) are “learned institutions, found in most countries or national regions, charged particularly with the definition, the protection, the purity and the enhancement of the national language”. These academies have

---

<sup>2</sup> For the rise of nationalism and the “invention” of the nation state see Anderson (1983).

several tasks, starting from selecting foreign language terms, determining rules and techniques by which equivalent terms will be formed, and distributing the neologisms among native speakers in addition to many other tasks. As stated by Shohamy (2006, p. 66), although language academies have authority to create and renew foreign terms that do not exist in the target language and the ability of giving them national flavor, there is no guarantee that they will succeed to insert those equivalent terms in the texts and daily dialogues of users.

However, some countries are still trying to codify the use of foreign terms and to maintain their language and national identity. Iran has also established one of the most active and resourceful language academies of the world in terms of structure, organization, level of expertise, and current affairs. It was established three times in different periods throughout modern Iranian history. The most recent one dates to eleven years after the Iranian Revolution of 1979 as The Academy of Persian Language and Literature (APLL). This chapter presents the linguistic and historical background of founding the APLL in all phases as well as the establishment of the APLL terminology group that is responsible for identifying loanwords and possible Persian equivalents.

## **2.2 Linguistic Borrowing**

### **2.2.1 Borrowing in Linguistic Scholarship**

Language is an organism. It has the same features as humans, rich or poor, spreading or receding, concentrated in one place or dispersed, alive or dead. It is in constant change from the outside and inside. It is a question of lifestyle, technology, culture, entertainment, science, etc. As language is influenced by the civilization of the people speaking it, it also intertwines and overlaps with other languages as they interact directly or indirectly with each other. Unsurprisingly, there are many types of influence and impact between the languages of the world. This may be due to neighboring countries, cultural relations between nations, trade relations, wars, invasions, and occupations.

Linguistic borrowing or language borrowing has been an interest in various fields of linguistics for some time since the set of loanwords that is imported into the vocabulary of any language is one of the most easily recognizable results of intercultural contact and communication. Different scholars have defined language borrowing in different ways and according to different parameters. Most of the linguistic studies agree that language borrowing is a common result of language communication and contact situations intensified synchronously with increasing communications and relations in human society. Chaika (1982) organized



linguistic borrowing according to three levels: lexicon, morpheme, and syntax. Other studies concentrate on phonological borrowing such as Campbell (1999, pp. 57-78), who believes that terms are usually remodeled to fit the phonological and morphological structure of the borrowing language. She also states that not only terms can be borrowed, but sounds, phonological features, and syntactic constructions, too.

Bloomfield categorizes borrowing into three types: (1) dialect borrowing where the borrowed features come from within the same speaking area, (2) intimate borrowing which occurs when two languages are spoken in what is topographically and politically a single community and can be considered a kind of borrowing from an upper language to a lower one, and, finally, (3) cultural borrowing where the borrowed features come from a different language (Bloomfield, 1933, pp. 444-495). This latter type of borrowing is at the core of this thesis.

Going into more detail then, borrowing has also been formally classified by Haugen (1950) based upon the degree of modification of lexical units of the source language or the target language. He also categorizes lexical borrowing into three types by providing the following examples. Loanwords are the first type in which both form and meaning are copied completely such as the French term *capsule* which is used precisely in the same way in Persian: *kapsōl* (capsule, کپسول). Loanblends form the second type in which terms consist of a copied part and a native part such as *ček kardan* (to check, چک کردن), which is borrowed from English. The last type is loanshifts where only the meaning is copied such as *vīrāstār* (editor, ویراستار) from *vīrāstan* (to edit).<sup>3</sup>

Another classification for language borrowing (Hochett, 1958) is also very similar to the one by Haugen where loan translation is added to loanwords, loanblends, and loanshifts. In this distinction, a loan translation or calque occurs when the native language uses an item-for-item native version of the original. The prominent Iranian linguist Moḥammad Reżā Bāṭenī<sup>4</sup> uses a similar categorization as Haugen and Hochett but details it even further as will be seen in chapter 2.2.3.

---

<sup>3</sup> For more details about lexical borrowing in New Persian see, for example, Paraskiewicz (2015).

<sup>4</sup> Mohammad-Reza Bateni as his name is written in English was a prominent Iranian linguist, translator, author, and professor from Tehran University. He authored numerous books, dictionaries, and articles on linguistics and term formation. The English-Persian dictionary *Farhang Moaser Pooya*, that was published in two parts, is one of his most important works (Neḡād & Rošan, 2008). Bateni passed away in 2021 (Zendeḥ bād Moḥammad Reżā Bāṭenī, n.d.).

The phenomenon of borrowing and lending terms and mutual language influences takes place both in terms of vocabulary and grammar. The forms of borrowed terms usually differ from their original forms where they fall under the influence of the phonetic system of the borrowing language, that is, the new or target language, for example, by adding the plural form of the target language to borrowed nouns or conjugating borrowed verbs according to the grammatical rules of the target language.

### **2.2.2 Academic Perspectives on Borrowing in Persian**

According to principles and regulations that have been set forth by the terminology group of the current APLL, the definition of borrowing is “obtaining a word from a foreign language (usually European) with the same concept or purpose as in the source language” (Terminology group, 1388 [2009]b, p. 12). This phenomenon is affected by the phonetic structure of the language as much as it is affected by its grammatical structure (Hoseynī, 1380 [2001-2002], pp. 125-126).

Linguistic borrowing is either direct borrowing, in which the linguistic elements are transmitted through a direct communal contact, or it is indirect when the linguistic elements are transferred from one community to another through other languages and societies in the middle. Direct borrowing occurs in Persian terms such as *ġarrāhī* (original language: Arabic, surgery, جراحی), *dāšbūrd* (English, dashboard, داشبورد), and *otūbān* (German, highway, اتوبان). One example of indirect borrowing in Persian is *almās* (diamond, الماس) that comes originally from the Greek term *mas* and then arrived in Arabic as *al-almās* before reaching Persian as *almās*. Another example is *šāmpū* (shampoo, شامپو) that entered Persian through English from an Indian origin (Şafavī k. , 1374 [1995-1996], p. 102).

The Iranian linguist and the former head of the Linguistics Society of Iran Vida Shaghghi considers importing lexical elements from other languages and dialects one of the ways to expand a language. Speakers of any language resort to borrowing linguistic elements to meet their needs. She categorizes terms of a language into native and non-native elements based on the fact that languages borrow from each other. She states that the non-native terms could either remain the same in the borrowing language as they preserve the same meaning and form or be changed according to the phonetic, orthographic, semantic, and grammatical rules of the target language (Shaghagi, 1391[2012-2013], pp. 126-127).

Shaghghi divides the borrowing phenomena into direct borrowing – a type of borrowing that is categorized by Bateni as loanword (see 2.2.3) – and non-direct borrowing. As mentioned above, the directly borrowed terms are those that come directly from one language to another language and change after entering the borrowing language to conform to the grammatical rules of that language such as *ṣarāḥat* (Arabic: صراحة, honesty, صراحت), *fūtbāl* (English: football, فوتبال), *eṣṭabl* (Latin: stabula, stables, اصطبل), and *estekān* (Russian: стакан, teacup, استکان) (Shaghghi, 1391[2012-2013], p. 128). The non-direct borrowing, on the other hand, applies to the terms that come to a language and enter a new language through a third language such as the Greek word *logārītm* (لگاریتم) that entered Persian through French (Shaghghi, 1391[2012-2013], p. 129).

The Iranian linguist Yahya Modarresi<sup>5</sup> categorizes linguistic borrowing into three types considering language borrowing as a general process that occurs in different forms in different categories in the language. Based on that, language borrowing is divided into lexical, phonological, and grammatical types (Modarresi Y. , 1368 [1989-1990], p. 64). Another Iranian linguist, Azita Afshari, considers the borrowing of terms an inescapable feature that happens in all languages with Persian being no exception (Afshari, 1386 [2019], p. 88).

As stated in the *Encyclopedia of Linguistics*, the lexicon of Persian has been highly influenced by borrowing from Arabic, but not only. It is estimated that 65 percent of the vocabulary used in modern novels and short stories is of Arabic origin. This number increases when considering conservative and formal writings. The influence of modern European languages such as French, and most recently English, is also evident in the lexicon of this language (Strazny, 2011, p. 325). For example, the French terms that also appeared in Persian in conjunction with the European cultural heydays are, among others, *restūrān* (رستوران), *kāfeh* (کافه), *māšīn* (ماشین), *telefon* (تلفن), *fīzīk* (فیزیک), *film* (فیلم), and *prūžeh* (پروژه) (Bāṭenī, 1388 [2009/2010], p. 91).

### **2.2.3 The Impact of Borrowing on the Persian Lexicon**

Clearly, Persian is full of loanwords and borrowed terms whatever their precise definition. Abolhassan Najafi, the Iranian linguist and translator and an associated member of the APLL

---

<sup>5</sup> Yahia Modarresi is an Iranian linguist and a professor at the Tehran-based Institute for Humanities and Cultural Studies, a research institute affiliated with the Ministry of Higher Education and active in a number of fields, among them literature and linguistics. He is also the editor-in-chief of *Language and Linguistics*, the Journal of the Iranian Linguistics Association (Editorial Board, n.d.).

since 1990, though emphasizing the importance of replacing foreign terms, acknowledges that the existence of foreign and imported terms is not a major risk. He argues that if there were a risk of borrowing terms for the survival of the Persian language, in its thousand-year relationship with Arabic and after borrowing tens of thousands of Arabic words it would have been destroyed (Nağafī, 1361 [1983]).

To the contrary, the prominent Iranian linguist Ahmed Samiei Gilani,<sup>6</sup> who has divided all the foreign terms that entered Persian over the years into two parts, argues that “a great amount of those words has mostly infiltrated [Persian] through trade and imported goods while the other part such as scientific, cultural, and artistic terms came through either translated works or students who studied abroad. If the flow of foreign vocabulary were slow and gradual, we might not feel much danger. We had at least a chance to digest them. But now that their numbers are on the rise, their entry must be restrained” (Samī‘ī Gīlānī, 1375 [1996], p. 2).

According to Mohammad-Reza Bateni, language has three main elements: a phonological element, a grammatical element, and vocabulary. The first one is the system which aims to define and connect the vocal elements that make up the words. The second one is the system that connects the meaningful elements of the language. Vocabulary, finally, is the set of terms available to the native speakers to be able to communicate. While he considers all three elements to be susceptible to change, the third element, vocabulary, is the most vulnerable to social changes. Even more than Haugen and Hotchett, he categorizes borrowing in languages into seven types of which he considers lexical borrowing as the most common type among all of them (Bātenī, 1388 [2009/2010], pp. 85 - 102). To some extent, Bateni’s categorization is also reflected by other linguists as will be seen.

**Loanword** is the first type referring to a term or a phrase that moves from the exporting language to the receiving one, considering the grammatical rules of the receiving language. Generally, a loanword is a lexical term that has been borrowed from another language (Campbell, 1999, p. 58). Such is the case with *kūdetā* (کودتا) in Persian from the French word *coup d’état* and, also from French, *carte postale* that entered Persian as *kārt-e postāl* (کارت پستال) (Bātenī, 1388 [2009/2010], p. 91). The word *computer* is another famous example of this type of lexical borrowing: *kāmpyūtar* (کامپیوتر). Further examples include also political terms such as *sūsyālīsm* (socialism, سوسیالیسم) and *komūnīsm* (communism, کمونیسم).

---

<sup>6</sup> Ahmed Samiei Gilani is a prominent Iranian linguist, author, translator, and an associate member of the APLL.

**Loanshift** or – as Modarresi calls it – “semantic expansion” is the second type, “caring about the meaning, not the term itself” (Modarresī Y. , 1368 [1989-1990], p. 273). This means that a term in the exporting language is given a new semantic meaning when it moves to the receiving language and expands that meaning. This type has been described also by Şafavī as “semantic generalization” (تعميم معنایی) (Şafavī k. , 1374 [1995-1996]).

Batani argues, for example, that the term *daftar* (دفتر), which used to hold the meaning of notebook, has expanded to mean *office* or the *room where people work* and that the meaning of the term *mağles* (مجلس), which originally meant sitting area or a place to sit, has expanded in Persian to mean *parliament* (Bāṭenī, 1388 [2009/2010], p. 96). However, as will be seen below, especially in the case of *mağles*, there is some scholarly dispute on how to define certain terms.

**Loan translation** is the third type of linguistic borrowing. It concentrates on translating the term or the phrase from the donor language literally to the recipient language taking into account the rules and grammar of the receiving language. Batani believes that this way of term formation is easy and effective, but he also mentions that translators should pay attention to the fact that excessive or negligent use of this method could lead to the emergence of unacceptable and unfamiliar terms (Bāṭenī, 1388 [2009/2010], p. 97) – a challenge that might account for some of the difficulties of accepting some of the political neologisms in question in this thesis. For instance, *noqṭeh-ye naẓar* (نقطه نظر) is the loan translation of *point of view* and *dūčarḩeh* (دوچرخه) is the loan translation of *bicycle*. Shaghagi defines it as “interpreting an imported phenomenon or entity and creating a word based on its purpose or extending the meaning of the word” (Shaghagi, 1391[2012-2013], p. 129).

**Loan rendition** is the type that does not translate the foreign term but explains it and finds a new local Persian term to explain it. Such is, for example, the case with *dānešgāh* (دانشگاه) that goes back to the Latin term *universitas* that means *whole* or *everything* and *čorāğ-e rāhnamā* (چراغ راهنما) instead of *indicator* (Bāṭenī, 1388 [2009/2010], p. 98).

**Loan translation and rendition** is a combined type of the two previous features which means that part of the imported term will be translated, and the other part will be rendered. Batani, by way of example, refers to *havāpeymā* (هوایپما) being used instead of the French term *aéroplane* (Bāṭenī, 1388 [2009/2010], p. 99). While it is true that *havā* is the Persian translation of the Greek prefix *aer(o)-*, both of which means air, the morpheme *peymā* is derived from the Persian verb *peymūdan*, which means *to cover (a distance)*, *to peregrinate*, *to ride* (vajehyab). Depending on the etymological explanation, the suffix *-plan* (feminine *-plane*) is a loan

rendition or translation. If we accept *-plane* to come from French where it means *level, flat, or even* (CNRTL, *Aéroplane*, 2012), then *peymā* is really a loan rendition which makes the full word *havāpeymā* a combined loan translation and rendition as mentioned by Bateni. However, there is a second etymological explanation which claims that *-plane* has its root in the Greek *planasthai* (to wander) (Webster, *aeroplane*, 2021), (Webster, *Floor*, 2021). In this case, we look at a pure loan translation because *planasthai* (to wander) translates straight into Persian *peymūdan* (to peregrinate).

**Loan blending** is the borrowing type where a term contains a part of the original language and a part of Persian such as *bombārān* (bombardment, بمباران) that consists of *bombe* as originally French and *bārān* as the Persian part. Another example is *filmbardārī* (filming, فیلم برداری), which is made of *film* and *bardārī* (Bāṭenī, 1388 [2009/2010], p. 99).

**Loan exchange**, finally, or what is known as “familiar foreign language” is the type in which one foreign term is replaced by a (more) well-known imported foreign term (Bāṭenī, 1388 [2009/2010], p. 99), for example, *āčār* (pickles, آچار) and *aks* (picture, عكس) (Shaghagi, 1391[2012-2013], p. 130). Unfortunately, the sources do not mention which less well-known terms these accepted loan exchange equivalents replaced.

As mentioned, this is just one way of categorizing and defining terms that enter Persian (or any other language for that matter). When looking at *mağles*, for example, Bateni’s assumption to consider it a loanshift is certainly correct when looking at the term from the narrow Persian inside perspective given that the term historically entered Persian from Arabic and has been part of Persian for hundreds of years. Shaghagi categorizes it as a loan exchange though where the foreign (French) term *parlement* is replaced with the foreign (Arabic) term *mağles* (Shaghagi, 1391[2012-2013], p. 130). However, the APLL does not consider Arabic terms as foreign terms if they have been part of Persian for centuries. As the equivalent Persian term of *pārlemān* approved by the APLL, *mağles* could therefore be considered a loan rendition because it is more of an explanation of the foreign term and certainly not its translation. In Arabic, it also carries the meaning of *council* and occurs in many compound words that refer to national parliaments such as *mağlis al-nawwāb* or *mağlis al-šūrā*. The term *mağles* will be discussed in more detail as an example of a political approved term later in chapter 5.1.

Kūrūš Şafavī mentions an additional linguistic type of borrowing which he calls “analogous equivalence” (Şafavī, 2005, p. 52). It occurs when a term in the target language adopts a linguistic phenomenon such as a term-formation pattern from another language. There

are three different subtypes that can be identified in Persian. First, the original language supplies both the term *and* the term-formation pattern both of which enter the target language. This is the case with *ḥammāl* (carrier/porter, حَمَّال) and *raqqās* (dancer, رَقَّاص), which have neither changed their meaning nor their spelling while transitioning from Arabic to Persian. Second, the target language accepts this term-formation pattern for an existing term in its lexicon such as *kaffās* (shoemaker, كَفَّاش) from *kafṣ* (كَفَش). Third, the new term in the target language comes from two different original languages. In this case, one of the original languages supplies the stem or root and the other original language supplies the term-formation pattern to form the new term in the target language. A good example is the Persian term *qannād* (confectioner, قَنَاد), whose root comes from Turkish *qand* (قند) and whose term-formation pattern follows the abovementioned *fa‘āl* pattern from Arabic.

Another type of borrowing that Bateni mentions in his book occurs by attaching the first letter of a few words together, that is, creating acronyms as is often the case in European languages. One example would be *sāvāk* (SAVAK, ساواک) as a shortcut for *sāzmān-e āmneyyat va eṭṭelā‘āt-e kešvar* (national organization for security and intelligence, سازمان امنيت واطلاعات, کشور). Proper nouns are usually either translated from European languages or are used as they are but in Persian letters (“Persianization”). For example, the proper nouns *sāzmān-e melal-e mottahad* (United Nations, سازمان ملل متحد) and *šerkathā-ye ‘āmel* (operating companies, شرکت های عامل) are translations while *asūšited pres* (Associated Press, اسوشيندپرس) is an English term written in Persian letters (Bāṭenī, 1388 [2009/2010], p. 100).

#### 2.2.4 History of Borrowing in New Persian

Although the Persian language has tolerated a deep-rooted transformation, attempts to retrieve the original terms and replace the non-Persian ones still exist and are strongly advocated. Iran being a great ancient civilization, its history has many aspects of these influences. Since the seventh century CE, Iran has witnessed radical developments and transformations that have influenced the Persian language that is being used today.

Three types of foreign influences have left their prominent mark on New Persian: Arabic, Turco-Mongol, and western languages (Sadeghi A. A., 2003, p. 49). Ghasemi and Setari divide historically the stages of passing foreign terms into Persian into three periods. First, the period before the Arab invasion and the extinction of the Sassanids up until the seventh century CE, which is considered the period least affected by foreign languages. Second, the period after the Arab invasion and until the Iranian Constitutional Revolution in the first decade

of the 20<sup>th</sup> century, which is divided itself into three phases: first the Arab-Islamic invasion, second the Mongol invasion and Turkish domination, and third the Qajar dynasty and its efforts to modernize which boil down to Iran's relations with Russia and France at the time. The third period started with the Constitutional Revolution and continues till today (Qāsemī & Setarī, 1389[2010]).

### 2.2.5 The Impact of Arabic and Turco-Mongolian Languages

Beginning in the seventh century and continuing for two centuries, especially after the Arab conquest of Iran when Arabic began to influence scientific and philosophical Persian texts – texts that were already written in Middle Persian (Sadeghi A. A., 2001) –, a huge number of scientific, philosophical, and cultural terms vanished and Arabic terms started to be used in spoken language in addition to the written texts. This was, for example, documented when Iranian scholars began to write in Persian again, trying to revive it, but their Persian texts were mixed with a variety of scientific Arabic terms with no attempt at coining new Persian ones instead.

The Arab invasion of Persian lands is considered by some as an attack loaded with a completely different religion, culture, and language. Iranians were in a kind of psychosis and stupor that lasted for two hundred years before they could resume life (Meskoob, 1385 [2006-2007], p. 5).

Iranians have described this period of Arab domination quite suitably with Arabic words as *Dū qarn-e sokūt* or *Two Centuries of Silence* to quote the title of a 1336[1957] book by Zarrinkoub or, in Persian, as “two hundred years in the consternation of this collapse” (“devīst sāli dar buht īn āvār”) (Meskoob, 1385 [2006-2007], p. 5). This reading may reflect the assessment of Iranian historians of what they also call the era of exclusion in the Umayyad period until the Harun al-Rashid era in the seventh and eighth centuries CE. In these two centuries, Arabic became the official language used in Iran – thus leading to a virtual silencing of Persian as a living language. Since then, Arabic terms have started to be dominating and merging with the Persian ones.

The language of classical Persian writers and poets of the eleventh to 15<sup>th</sup> centuries – strongly influenced by Arabic – set a standard of literary usage of Persian which has remained to this day (Paul, 2010, p. 80). The only preliminary attempt that the period of Arab dominance witnessed at reusing the old Persian terms or reformulating the new ones to adapt with Persian



had been made by Ibn Sina/Avicenna (c. 980-1037) in his book *Encyclopedia of Alai/Dānešnāmeḥ-ye 'Alā'ī* (دانشنامهء علائی).

Much in the same way, the Mongol invasion brought a lot of Turkic terms into Persian in the 13<sup>th</sup> century (Sadeghi A. A., 2001, p. 20). But after the Mongols were driven out of Persia, the influence of Mongolian-Turkic terms came to an end (Sadeghi A. A., 2003, pp. 50-51).

Persian is written with the Arabic alphabet or, as described by Sadeghi (Sadeghi A. A., 2003), the Perso-Arabic alphabet. The Arabic language did not only invade the Persian dictionary; it also affected the Persian grammar. The effect of Arabic grammar lasted from the 16<sup>th</sup> to the 18<sup>th</sup> centuries to take the shape that is witnessed in New Persian. Some of the borrowed grammatical elements from Arabic still used in New Persian are the Arabic plural forms with the suffix *-āt* (ات) such as *edārāt* (offices, ادارات), with the suffix *-īn* (ین) such as *mo'tādīn* (addicts, معتادین), and with the suffix *-ūn* (ون) such as *enqelābeyyūn* (revolutionaries, انقلابیون), as well as the broken Arabic plural form such as *aqvām* (people, اقوام) (Gīvī & Anvarī, 1390 [2011-2012], pp. 70-73).

“None of the books that were written in the thirteenth century, although some of them were useful and beneficial, had a pure Persian style. They were rather an imitation and translation of Arabic books” (Jalaluddin Homaei as quoted by (Bātenī, 1385[2007]b, p. 23). The first grammarian who composed a pure Persian grammar book away from translation and copying from Arabic would turn out to be Habib Esfahani (1835-1893) in 1872 (Yazıcı, 2002). He was also the first to use the word *dastūr* (grammar, دستور) instead of the Arabic word *qavā'ed* (قواعد) as he called the book *Dastūr-e soḥan* (دستور سخن, *A Grammar of Spoken Persian*) (Bātenī, 1385[2007]b, p. 24).

## 2.2.6 The Onset of Modernity and European Influences on Persian

There are many factors that played an important role in the shaping and reforming of languages. Modernity is one of them. “Languages are the pioneers of modernity” with modernity being the name given to a new concept of life and humanity that is shaped by scientific knowledge, technical and industrial organization, along with new political and social phenomena and their cultural value. All of this leads to the formation of modern languages. The relationship between languages and modernity is therefore a direct one. Hence, the correct understanding of modernity and the search for the depth of its achievements without understanding what has

happened to languages is not possible. By exploring the horizons of all aspects of human life, they have turned themselves into open languages (Ášūrī, 1386[2007/2008], pp. 89-92).

Since the late 18<sup>th</sup> and early 19<sup>th</sup> centuries and thus the beginning of historical modernity, the same development as with the impact of Arabic and Turco-Mongolian languages has occurred again, but this time with Europe when an ever-increasing number of loanwords began to appear in the daily written and spoken language as a reaction to the new relationship between Iran and the western countries. At that time, the western countries were the main source of civilization, culture, and science (Sadeghi A. A., 2003, p. 51).

Tehran became the gathering place for the elites where all scholars, thinkers, artists, writers, and poets arrived. Consequently, linguistic sophistication in conjunction with social growth and urban development were the results to be expected by this move (Sarlı, 1387 [2008-2009], pp. 311-313).

For the Iranians to keep pace with Western European civilization, they established *Dār al-Fonūn* in 1851, which was the first European-style school of higher education to be established in Iran. Through the establishment of this school, many new terms entered Persian. *Dār al-Fonūn* was not the last school to be founded in that vein with many schools following suit. As a natural next step, it became necessary to write and print new textbooks that accompanied this development and served studying in these new schools. Of course, those books were full of new foreign terms arriving in the Persian language as a natural consequence of all this cultural mixing, which led to a deep impact on and reform of the Persian language (Sarlı, 1387 [2008-2009], pp. 313-314).

The acceptance of Arabic words in Persian has passed through several stages since the beginning of the 19<sup>th</sup> century when the idea of language purification began to flourish with the likes of the Qajar prince Jalal al-Din Mirza Qajar (c. 1832-1871) and those who supported him in completely rejecting Arabic words and viewing their presence in Persian as a threat to Persian nationalism. Others dealt with it differently arguing that the presence of Arabic words in the Persian language gives strength, vitality, and a special character from Perso-Arabic loan blending. This group believed in the need to get rid of European loanwords only. With regard to Arabic, its adherents reasoned that Iranians – after using the Arabic language for a period of nearly 1,300 years in their daily language – had adopted Arabic words as an integral component of their own language. The scholar Moḥammad Qazvīnī, for example, himself not a language

purist, stated that what makes the Persian language lose its splendor is not the use of Arabic words but the use of European ones (Kia, 1998, p. 18f.).

One of the early attempts of purifying the Persian language from imported terms – especially from Arabic – goes back to Jalal al-Din Mirza Qajar when he started to write his book *Nāmeḥ-ye Khosravān (The Book of Kings)*,<sup>7</sup> modelled after the work of Europeans, on the history of Iran from the pre-Islamic era to the 19<sup>th</sup> century in pure simple Persian clear of Arabic words. He wrote this book in an attempt to save Persian from extinction at the hands of Arabs and replaced the Arabic terms with ancient Persian words (Kia, 1998, pp. 11- 12) reminding people of how great Persia was in addition to reviving the forgotten Persian language. The beliefs of Mirza were, among others, opposed by ‘Abd al-Raḥīm Ṭālebūf Tabrīzī in his book *Masālek Al-Moḥsenīn*.<sup>8</sup> As a conservative language nationalist, who did not believe in language purification, Ṭālebūf argued that Iranians should retain Arabic vocabulary and use and incorporate in their language the modern European terms.

Jalal al-Din Mirza’s approach can be explained with a look to social theory according to which the world of social institutions and practice has two distinct domains: the material, outside domain of modern science and technology and the spiritual, inner domain of cultural identity.<sup>9</sup> In the former, 19<sup>th</sup> century Europe proved its progress, development, and superiority over much of the rest of the world – including Iran. With regard to the latter, Persian nationalists divided history into the pre-Islamic period, which they considered the source of the Iranian spiritual domain, and the Islamic period, which they called “the dark centuries” when Iranians started to lose their identity. Since the nationalists believed that the Persian language maintained the sovereignty of the people and their national identity since pre-Islamic times, they consequently believed in the necessity of purifying and purging the language from foreign terms. This is in line with nationalist tendencies to transform history, culture, and language into ideological tools – in the Persian case with the goal to create a new homogeneous, non-Islamic, Persian national identity (Kia, 1998, pp. 12-18).

Different from Jalal al-Din Mirza, who advocated for a complete eradication of Arabic terms in modern Persian and a return to a language that had not been spoken in over a thousand years, and different also from Ṭālebūf, who favored instead progress and modernity by

---

<sup>7</sup> Published as Mirza, Jalal al-Din (1871). *Nāmeḥ-ye Khosravān*. 3 vols. Tehran.

<sup>8</sup> Full citation is not given in Kia’s article, but one recent issue would be Tabrīzī (1386 [2008]).

<sup>9</sup> Mehrdad Kia (Kia, 1998, pp. 9-10) applied the model by Partha Chatterjee (1993) in dividing the world of social institutions and practice into two domains.

embracing foreign terms – be they Arabic or European – to advance Persian in modern times, the reformer Mīrzā Kermānī proposed a third way. He suggested that instead of going back to the dead language of pre-Islamic times, Persian could be modernized by looking at all the living dialects and languages all over the country and by making use of genuinely Persian terms in actual use that could eventually replace foreign terms, thus preserving the Persian character of the language while at the same time adapting it to the needs of modern scientific, technological, and other discourses (Kia, 1998, p. 14).

The continuation of Iranian-European relations that had begun with the Safavid era led to many Iranians traveling to Europe for education or learning new languages well into the Pahlavi era. That was accompanied by the translation of many literary and scientific works and had a deep effect on two levels. The first effect was on Iranian cultural thoughts and the second one was a deep and profound effect on the Persian language itself. In addition to the opening up to Western European influences, Reza Shah Pahlavi sent groups of military envoys to France, Italy, and Germany for training, which contributed to this rise (Abrahamian, 2008, pp. 63-96).

Moreover, the reform movement started by the Shah spread to the arts, newspapers, and education and eventually to writers who tried to modernize written Persian in line with the spoken one (Paul, 2010, p. 80). Most of the loanwords – scientific, technological, daily life terms – that entered Persian came from French as French was the second language of education at that time, among them *rādyū* (radio, رادیو), *īdeh* (idea, ایده), *fīzīk* (physics, فیزیک), *māshīn* (car, ماشین), *mersī* (thank you, مریسی), *dīplom* (diploma, دیپلم), and so on. Sadeghi (Sadeghi A. A., 2003, p. 51) notes, however, that the first loanwords of this type had come from the Russian language as can be seen in *eskenās* (banknote, اسکناس). Persian has also borrowed a lot of scientific English words such as *kāmpyūtar* (computer, کامپیوتر) and *seryāl* (TV series, سریال). The tendency for English terms to enter Persian spoken and written has continued to this day since the Second World War.

### 2.3 Language Planning and Language Policy

Language plays a pivotal role in many parts of the world as a factor determining national identity (Paul, 1999, p. 208). It is the means of societies and peoples for communication and understanding among each other.

The expression *language planning* was mentioned for the first time by Einar Haugen in 1959 in his article about “planning for standard language in modern Norway”. According to

Haugen, “[p]lanning implies an attempt to guide a development of a language in a direction desired by the planners” (Haugen, 1959, p. 109).

Bateni sums up the main tasks of language planning in four points: first, revival of the native language as an official language. Second, modernizing the national language to meet the scientific and technical needs. Third, cleaning the language from foreign terms. Fourth, adopting consistent orthography (Bāṭenī, 1385[2007]a, p. 66).

The expression *language policy* was promoted for the first time by Fishman in 1970 as a part of applied linguistics (Fishman, 1970). In a narrow sense, language or linguistic policy can be defined as the sum of the conscious choices made in the field of relations between language and social life, specifically between language and life in the homeland, and language planning can be considered as the search for the necessary means to implement a language policy (Calvet, 1998, p. 114). Linguistic policy aims to influence the form of language on three levels defined by Calvet, namely writing (orthography), lexis, and dialects. The first level concerns creating new orthography or replacing an existing one or even changing the alphabet itself. The lexicon level concerns creating neologisms to replace the foreign terms in the language. The last level is concerned with promoting a national language among all the dialects (Calvet, 1998, p. 115). Spolsky and Shohamy (1999, pp. 31-32) consider language policy to be the effort to modify or influence the practice of each language.

In Iran, the seed for a renewed, convincing, and perhaps more coherent language policy was planted in the new 1979 Constitution of the Islamic Republic of Iran, amended in 1989. Article 15 stipulates that Persian is the “official and common language and script of the people of Iran”. Laws and other official texts as well as textbooks must be published in Persian. In addition, the constitution allows for the use of regional or tribal languages in media and education thus acknowledging the multiethnic character of the state. Arabic enjoys a special status as specified in article 16 (The Constitution of the Islamic Republic of Iran, 1989). These language provisions in the 1979/1989 Constitution stand in contrast to the 1906 Constitution and its later amendments where no such reference as to the status of Persian had been made (Irans 1906 Constitution, 1906) although the Shah’s Farhangestān had made the use of certain terms it coined compulsory by law (Paul, 1999).

In practice, in the absence of any formal laws or regulations before the establishment of the Third Farhangestān, Persian language policy was largely ideational during much of the 1980s. In a way, the political system was still in the process of consolidation and political

priorities had a different focus at the time such as the Iran-Iraq War. Nevertheless, language-related policies and activities took hold. Place names with reference to the Shah's regime were changed and Arabic as the language of the Quran and of Islam was made compulsory in primary and secondary education within a few years after the Revolution. Also, public figures such as the prominent politicians and leaders Rafsanjani and Khamenei mentioned the Persian language in some of their speeches (Paul, 1999). All these efforts led to a more consolidated and serious language policy and paved the way for the reestablishment of the Farhangestān in 1990.

## 2.4 The Academy of Persian Language and Literature

The Academy of Persian Language and Literature APLL – also known as the Third Academy – is considered the main official Iranian association of Persian language planning. “No one can deny the change and can say that the Persian language has not changed throughout the ages or ignore this change and ask the speakers to speak and write in the way of the past” (Najafi, 1390 [2011], p. 5). But to some extent, it is possible to reduce the huge change taking place in the language. That is exactly one of the essential tasks of the APLL. This chapter will sketch the history of the development of language planning institutions in Iran before going into more detail on the Third Academy.

### 2.4.1 Etymology and History of the Term *Farhangestān*

The term *Farhangestān* (academy, فرهنگستان) is widely accepted to be composed of three parts: the prefix *far* (فر) or *frā* (before, ago, فرا) going back to the Achaemenid empire and the Avesta,<sup>10</sup> the word *hang* (intention, purpose, هنگ), and the suffix *stān* (ستان) (Pourdavoud, 1322 [1943-1944], pp. 690-691). According to a comment in one of the oldest Persian dictionaries, *Borhān-e qāṭe* (برهان قاطع), the middle component *hang* comes from the Avestan stem *thang* that has the same meaning as the Latin stem *educat-* and can be translated as *education* (“bema'nā amūzeš va ta'līm va tarbeyat”). In Middle Persian (Pahlavi),<sup>11</sup> the compound *farhangestān* then came to mean *maktab* (office) or *madrāse* (school). Only later did this term acquire the new meaning of *academy* (ākādemī, آکادمی) (Tabrīzī M. Ḥ., 1342 [1963]c., pp. 1480-

---

<sup>10</sup> The Avesta is the holy book of the pre-Islamic religion of Iran which is known as Zoroastrianism. This book contains all the rules and teachings of the prophet Zarathushtra. The current version of the Avesta is what was written, compiled, and standardized during the Sassanid era from the third to the seventh centuries CE (Şafavī, 1386 [2007], pp. 60-62).

<sup>11</sup> The Pahlavi language or what is also known as Middle Persian is a Sasanian middle Iranian language or Zoroastrian. It is a subdivision of the Indo-Iranian branch of the Indo-European language family. It was used as the official language from the end of the Achaemenid dynasty to the seventh century CE with the rise of Islam (Hoseynī, 1380 [2001-2002], pp. 264-265).

1481). The word *farhangestān* is nowadays considered equivalent to the term *academy* and has been used in Persian at least since the establishment of the First Farhangestān in 1314š/1935 (Sa'adat, 1393 [2014-2015], p. 47).

There is another opinion which considers the term *Farhangestān* as consisting of two parts: the word *frahang* (sic!, فرهنگ), which meant *education* or *knowledge* in Middle Persian (Pahlavi) before it came to be associated more with *culture* in modern Persian, and the suffix *-stān* (ستان) which means *place* or *base* in the language that was spoken in the period of the Achaemenid empire and in the Avesta (Dehḥodā, 1377 [1998]a., pp. 919-920).

The current Academy of Persian Language and Literature itself strongly supports the use of the term *farhangestān* consistent with and in the tradition of its use for the so-called First Farhangestān from 1935 onwards (see chapter 2.4.3). Moreover, “its name is equivalent to a foreign word used in some countries for institutions such as the Academy of Persian Language and Literature, such as the Academy of French Language in France” (Parvīzī N. , 2021). Given the choice between *Farhangestān* and *Ākādemī*, it would indeed be incongruous with its mission of finding Persian equivalents to foreign terms if the Farhangestān had adopted the foreign term *Ākādemī* as its official name.

In the first dictionary of neologisms published by the First Farhangestān in 1319š/1941, the term *farhang* was approved as replacing the Arabic term *ma'āref* (Farhangestān-e Īrān, 1319[1941], p. 92). Therefore, it seems that the First Farhangestān drew on the meaning of *farhang* as *knowledge* when creating this name for the institution.

## 2.4.2 Early Attempts at Institutionalizing Language Preservation

The 19<sup>th</sup> century saw lively scholarly and nationalist discussions on language preservation. It took a few more decades after these debates, though, before serious efforts at institutionalizing language preservation were made.

The idea of establishing an official institution called *Farhangestān* emerged when a group of Iranian scholars felt the need to purify their native language from foreign influence. They perceived their language to be in danger from being polluted with questionable thoughts under the pretext of reform and refinement and through introducing elements unfamiliar to Persian. Ever since, they intended to select these foreign terms and to find an original Persian alternative to them (Gilani, 1374[1995], p. 136).

A 1903 news piece in the *Irān Solṭānī* newspaper refers to “the academy house” *mağles-e ākādemī* (مجلس آکادمی). The article mentions a monthly meeting held to discuss the new foreign terms in Persian and to find neologisms to replace the foreign ones. The members reported directly to the press ministry. This news article also points to monthly meetings of both “the authorship house” and “the translation house” to decide on the imported terms into Persian in that era, such as the French words *otūmobīl* (car, اتومبیل) and *ártikel* (article, آر تکل) (Sa‘adat, 1393 [2014-2015], pp. 47- 48).

Before there were organized attempts by specialists to coin new terms in the Persian language, there had been a few random attempts by non-specialists for different reasons. One of those efforts was initiated by the ministry of war in 1303š/1924 by a group of non-specialists in linguistics. Their group was formed by order of Reza Shah to revise Persian words, investigate the new terminology, and form new terms (Sarlı, 1387 [2008-2009], p. 316, and Badreh-ey, 1355[1976/1977], pp. 9-13). They were very successful in creating equivalents to the foreign terms that were common at that time, and despite the criticisms leveled by linguists for coining equivalents without a linguistic basis, these words still remain. Examples are *havāsang* (French: aéro-mètre, هواسنگ), *vābaste-ye nezāmī* (French: attaché militaire, وابستهء نظامی), *forūd-gāh* (French: aéro-drome, فرودگاه), *havānavard* (French: aéro-naute, هوانورد), *ḥalabān* (French: pilote, خلبان), and *havāpymā* (French: avion, هواپیما) (Şeddīq, 1322 [1943], pp. 2-3). The same committee replaced some Arabic terms with Persian ones, too, such as replacing *vaṭan* (وطن), which means homeland, with *mīhan* (میهن) and *mamlakat* (مملکت), which means kingdom or country, to be *kešvar* (کشور).

In 1307š/1928 or 1311š/1932,<sup>12</sup> another effort was made with the founding of *dār-ale-mo‘alemīn ‘ālī* (the school of teacher training, دارالمعلمین عالی) and a society called The Association of Coining Scientific Terms (Anğoman-e važ‘-e loğāt va eştelāhāt-e ‘elmī). This institution was concerned with coining new terms in several topics such as natural sciences, mathematics, physics and chemistry, literature, and philosophy. It was more specialized than the previous one and used a specific approach in coining new Persian terms. First, a group of students worked on choosing terms and present them to a main board for either approval or rejection. Then the approved terms were distributed to the specific groups of scholars, specialists, and writers to learn their opinions, and finally came the publication of the approved new terms. The Association laid out four principles to be followed in the process of coining the

---

<sup>12</sup> Different years are given depending on the source, for 1928 (Rūstā‘ī, 1377 [1998-1999], p. 76) and for 1932 (Şeddīq, 1322 [1943], p. 3).



neologisms: first, observing the grammar of Persian. Second, observing simplicity and conciseness. Third, keeping the international foreign terms where a Persian equivalent could not be found or would not give the proper meaning. And finally, maintaining common expressions except for the incorrect ones and replacing those with appropriate ones instead (Şeddīq, 1322 [1943], pp. 3-5, and Rūstā'ī, 1377 [1998-1999], pp. 76-77).

This Association managed to continue working until 1940 and during these years coined around 350 neologisms per year and around 45 scientific terms. That resulted in around 3,000 new terms, 400 of which entered the elementary and high school textbooks. The Association continued its work even after the founding of Farhangestān-e Iran. The following terms are examples of the approved ones by that Association: *payvasteh* (French: continu, پیوسته, continued), *tarāvoš* (French: osmose, تراوش), and *āmīḥteh* (French: mélange, آمیخته) (Şeddīq, 1322 [1943], pp. 3-5).

In 1313š/1935, specifically five months before the creation of the First Academy, the Iranian Medical Academy (Académie de Médecine, *Farhangestān-e ṭebbī-ye Īrān*, فرهنگستان طبى ايران) was established by Iranian scientists and physicians in cooperation with several European academies, which was the first time to use the term *Farhangestān* for an association instead of “academy”. Its purpose was to look at new medical terminology and its equivalent in Persian, to translate medical books, and to publish medical papers (Keyā & Gol-Golāb, 1355[1976-1977], p. 7).

Aside from finding Persian equivalents to the foreign medical terms, the Iranian Medical Academy had other goals including translating and authoring medical books in Persian, preparing dictionaries for medical terms and medicine, publishing medical studies in Persian, and cooperating with the international medicine academies (*Tārīḥī az farhangestān-e Īrān* (History of the Iranian Academy), 1322[1943], pp. 6-7).

The main goal in all those attempts was not meant from the beginning to only create new terms and look after the foreign ones, but the surrounding circumstances turned the original goal into a new direction. It was thought to be dangerous that the idea of association may be used to serve specific political interests (Gilani, 1374[1995], p. 137). Hence developed the nucleus of the First Farhangestān.

### 2.4.3 The First Farhangestān (1935)

The Iranian Academy *Farhangestān-e Īrān* (فرهنگستان ایران) was founded in 1935, along the lines of the Turkish language association (Türk Dil Kurumu), which had been founded by Mustafa Kemal Atatürk in 1932 to simplify and purify Turkish of Persian and Arabic terms. The First Academy was subsequently the first official organization tasked with language reform in Iran (Perry, 1985, p. 295). The nucleus of establishing the First Academy came after Reza Shah Pahlavi returned from Turkey in 1934 and ordered the ministry of war to form a committee for coining Persian equivalents to the military terms. Knowing about the support of the Shah for language reform, then-Prime Minister Mohammad Ali Foroughi<sup>13</sup> proposed that a community of scholars should be assigned to research foreign terms in Persian. Then the Shah declared the founding of this first authoritative Persian language academy (Jazayeri, 1999).

The statute of this Farhangestān was formulated in 16<sup>th</sup> parts (Gilani, 1374[1995], p. 138). The first clause of the charter stipulated that the Iranian Academy was founded for the maintenance, development, and promotion of the Persian language. With regard to the tasks of the Academy, the second paragraph stipulated that the Academy should delete incongruous foreign terms from the language and coin Persian terms and expressions for every branch of life, using Persian roots and terms as far as possible and not Arabic or Turkish ones (Sadeghi A. A., 2001, p. 23). This was in line with the opinion of Foroughi, who could be described as a language conservative rather than a purist and is a reminiscence of Kermanī's and Tālebūf's arguments from the previous century. Foroughi argued that only recent European terms entering Persian needed to be replaced while ancient Arabic terms should remain unchanged (Paul, 2010, p. 81).

The criteria of creating new terms were detailed in the tenth rule of the charter, which stated that the equivalent term must meet several conditions. In the first place, it should be common, familiar, certain, and valid in addition to having a definite Persian root and a correct derivation. If this requirement did not apply to the proposed term, then an Arabic term could be proposed or a compound term of Arabic and Persian. If the right alternative could not be found,

---

<sup>13</sup> Mohammad Ali Foroughi, who was also known as *Dokā al-Molk*, was a teacher, statesman, scholar, and man of letters. He was the first prime minister under Reza Shah Pahlavi. He also initiated the idea of creating the Farhangestān in 1314Š/1935 and convinced Reza Shah Pahlavi to establish an official organization to plan the Persian language. Foroughi then became its first president (Afshar & Fakhreddin, 1999). Foroughi believed that the Persian language suffers from two things, namely deficiency and defect. As for deficiency, it means the absence of linguistic alternatives to new scientific terms. Defect means that the language is mixed with many foreign words, especially Arabic (Kāfi, 1371[1992], p. 41).

then it would be possible to choose a term from the abandoned Persian terms with the same meaning or a close one. In the last place came the option of keeping the international foreign borrowed term to be used in its original language (Gilani, 1374[1995], p. 139). These criteria actually resemble Bateni's categorization of borrowing with the second option being a mix of loan translation, loan rendition, loan blending, and loan exchange. The third option evokes the idea of creating loanshifts while the fourth option refers to the retention of loanwords.

In addition to the committee of the Academy that was responsible for replacing the foreign terms with Persian ones, the Academy included seven more committees as follows (Sadeghi A. A., 2001, p. 23): committee for general administrative terminology, committee for legal terminology, committee for scientific terminology, committee for Persian grammar, committee for compiling a Persian dictionary, committee for the guidance of public thoughts and collecting local hymns, tunes, and expressions, and committee for studying medical terms.

Some researchers consider the terminology group of the First Farhangestān to be successful as it took a moderate way away from Persian intolerance against Arabs and conforming to the nature of languages and their development (Kāfī, 1371[1992]). Nevertheless, the committee that was formed to work on preparing dictionaries also did some research into grammar and geographical terms and republished old books (Gilani, 1374[1995], p. 138).

The Academy established its own publishing house and printed the periodical *Nāmeḥ-ye Farhangestān* for five years in addition to many books such as *Vāžehhā-ye nū* (*Neologisms*, واژه های نو), which has seen several editions since its first publication in 1319š/1941 and is a good example of how the First Academy accomplished its work. Evaluating *Vāžehhā-ye nū*, the poet Vaḥīd Dastgard criticized in his last published article that “all the words mentioned in this book are wrong and in vain from the beginning to its end” (Dastgard, 1327 [1948-1949], p. 43). He expressed sharp criticism because he felt that the terms approved by the Academy did not carry the same meanings as, for example, those used by Ferdowsi. It is clear that, as a professional user of language, he was dissatisfied with the term-formation process and, more so, with its results.

Nonetheless, others have acknowledged the achievements of the First Academy. It was able to take the initiative of term formation and word selection away from non-qualified individuals and institutions and reduce the excessive purification of the language. “The main result of the efforts of Farhangestān-e Iran is the standardization and approval of the scientific

terms which are used in the textbooks of primary and middle schools” (Sarlı, 1387 [2008-2009], p. 319, and Sadeghi, 1370 [1991]).

Regarding the First Academy, Bateni praised its success in using living terms in the process of approving Persian equivalents as opposed to dead terms. Among the criteria of choosing equivalents were that the terms sound harmonious and the letters of one term do not repel each other. This is considered one of the reasons for the survival of the approved terms so far. As an example, Bateni mentioned *answering machine*. When this product entered the Iranian market, the merchants started to use *monšī-ye telefonī* (literally “telephone secretary”, منشی تلفنی). But later the First Academy succeeded to replace it with *payāmgīr* (literally “message receiver”, پیامگیر) (Nežād, 1388 [2010], pp. 142-144).

An in-depth case study of some 70 neologisms approved by the First Farhangestān shows that the rate of success 70 years after the coining of these new terms stands at about 90 percent. According to this research, the success of an approved term rests mostly on semantics and not so much on morphology or etymology. Along these lines, newly coined terms from the fields of military and (political) administration can generally be considered more successful than others hinting at the suspicion that extralinguistic, socio-political factors – such as the circumstances for creating the First Farhangestān – play an important role in determining the success of neologisms (Paul, 2010, p. 99).

In 1319š/1941, the terminology group succeeded to approve 1,700 terms and another 2,000 terms in 1320š/1942. In order to know more about the accomplishments of that version of the Academy, some examples of the common terms that were approved and are still in use up to date include *forūdghāh* (airport, فرودگاه), *govāhī-nāmeḥ* (certificate, گواهی نامه), and *mardomšenāsī* (anthropology, مردم شناسی). In 1322š/1944, the Academy stopped working on creating new terms and instead turned to intensifying its interest in literature and research (Gilani, 1374[1995], pp. 140-141). As the Academy was mainly influenced by the political situation of the time, after the abdication of Reza Shah Pahlavi in 1941 and the death of Foroughi in 1942, it lost its main supporters, decreased its activities, and came officially to end in 1954.

#### 2.4.4 The Second Farhangestān (1970)

The Iranian Language Academy, *Farhangestān-e zabān-e Īrān* or Second Farhangestān, was established by a decree of Mohammed Reza Pahlavi with the statute of “the imperial foundation of the Iranian academy” (Sharifi, 1391[2013], p. 1091). It had set two goals to achieve (Keyā & Gol-Golāb, 1355[1976-1977], p. 20):

1. Keeping the Persian language on its long-standing cultural base and preparing it to meet the various scientific, technological, and cultural needs of the country.
2. Researching, studying, and investigating all current and previous Iranian languages and dialects, especially for further identification and promotion of the Persian language.

The activities of the Academy began with four research groups: The department of term selection and coining, the department of the study of Persian vocabulary, the department of the study of Persian grammar and writing system, and the department of ancient, middle, and modern Iranian languages and dialects (Sa‘adat, 1393 [2014-2015], p. 50). The tasks of the department of term selection and coining were the main tasks of this Academy.

The total number of terms approved by the Second Academy are variously given with up to 6,650 (Şaffār Moqaddam, 1374 [1996], p. 136). Modarresi reports that the Second Academy coined and approved 2,000 terms but published only 200 of them (Modarresī Y. , 1371[1992], p. 187). Şaffār Moqaddam counts 1,470 Persian equivalents among 1,515 foreign terms coined and approved by the Second Academy (Moqaddam, 1374 [1996], p. 134).

Given its structure and organization, the Second Farhangestān is considered by some as more disciplined than the First Farhangestān (Mansouri, 1376 [1997], p. 22). The criteria of choosing and structuring the equivalent terms in this Academy were stricter than in the first one. The new terms had to be originally Persian and of the common and daily used terms – during that time – in the first place, such as *pāyghāh* (base, پایگاه). In the case of the absence of such a condition came the accessibility of using a term from textbooks which was not very common during that time, such as *nūf* (*şedā*) (sound, نوف صدا). Then came the choice of using a term from Old or Middle Persian or its dialects, such as *barzīdan* (manipulate, برزیدن), which was taken from Middle Persian. The last option was to create a new term which observed the consistency of the language and its grammatical and phonological principles, such as *dast-e nāmeḥ* (manual, دست نامه) (Şaffār Moqaddam, 1374 [1996], pp. 136-138).

However, this is not undisputed. For example, Kāfī criticizes the Second Academy’s arbitrary and monopolistic decisions and its attention to the thoughts and opinions of non-experts (Kāfī, 1363 [1984-1985], p. 20). Samiei Gilani agrees with Kāfī’s opinion about the failure of the Second Farhangestān and its approved terms. While testifying to the success of implementing linguistic research in coining new terms, he holds that the approved terms were not accepted by native speakers of Persian. This is likely due to their anti-Arabic approach and thus constitutes a practical echo of Jalal al-Din Mirza’s radical purist approach from the mid-19<sup>th</sup> century. The majority in the Academy excluded and replaced any Arabic terms, however well-rooted in Persian, and replaced them with non-familiar terms such as *hamīstād* (system, همیستاد) or *pardīz* (park, پردیز) (Gilani, 1374[1995], pp. 142-143). In contemporary Persian, the former is usually referred to as *nezām* (نظام) whereas the latter is nowadays *pārk* (پارک), *bāġ-e mellī* (باغ ملی), or *gardešgāh* (گرددشگاه).

On the Second Academy, Bateni seconds that, unlike the First Academy, it was not very successful in choosing equivalents that the public would like to use. He also attributes the reason for this to the intolerance and strong adherence to using abandoned terms. As an example, he mentions the term *zeddeḥamleh* (counterattack, ضدحمله). The Academy replaced it with *pātak*. Although *pātak* was used in the media during the eight years of the Iran-Iraq War (1980-1988) even after the dissolution of the Second Academy, people still use *zeddeḥamleh* (Nežād, 1388 [2010], p. 143).

Jazayeri claims that the absence of long-standing traditions in linguistic research and the lack of a consistent methodology for coining scientific terms were problems that both the Second Academy as well as the first one faced. And while the use of the approved terms by the First Academy was mandatory, the use of the approved terms by the second one was optional (Jazayeri, 1999). Whatever its perceived failure or success, it was eventually shut down by the victory of the Iranian Revolution in 1979.

#### 2.4.5 Random Efforts at Coining Foreign Terms

The writer, diplomat, and three-times Prime Minister Mohammad Ali Foroughi wrote in his letter to the Farhangestān in 1315š/1936 “that language is the other side of culture and preserving and maintaining the Persian language and Iranian culture is tantamount to preserving Iranian nationalism” (Sarlı, 1387 [2008-2009], p. 315, and Foroughi, 1354 [1975-1976], p. 2). Therefore, the belief in the idea that the support of Iranian nationalism depended heavily on the support, preservation, and restoration of Persian itself was one of the main reasons of

establishing and reforming the Farhangestān many times. Foroughi advocated for the Persian language as a kind of preservation of Iranian nationalism and an independent Iranian national identity.

Some researchers either individually in their studies or in scientific groups worked on coining new terms or replacing foreign ones. Sadeghi sorts them as individual attempts and group circles respectively. One of these circles was the society of coining terms and expressions that had to be used in the Persian encyclopedia *Dā`erat ol-ma`āref-e fārsī* (دائرةالمعارف فارسی). The same encyclopedia, in 1959, published a leaflet including 700 geographical, geological, and meteorological terms (Sadeghi A. A., 2001, p. 24).<sup>14</sup>

The *Dictionary of Human Sciences (Farhang-e `olūm-e ensānī, فرهنگ علوم انسانی)* is considered one of the individual efforts for coining new Persian terms. This dictionary includes 12,000 coined Persian neologisms that were translated from English (Ášūrī, 1374[1995]). Its author Ášūrī made another contribution to the field in 1975 when he published it in *The Dictionary of Humanities*, which is an English-Persian and French-Persian dictionary that includes terms in the fields of philosophy and humanities (Ášūrī, 1394[2015]).

In the field of humanities, Ariyanpur has also coined around 2,000 philosophical Persian terms from English in the course of translating Mohammed Iqbal's book *The Development of Metaphysics in Persia* (Sadeghi A. A., 2001, p. 24).

Zarnikhi in his paper (Zarnikhi, 2010) mentions other attempts related to Persian terminology, among them *vāžegān* (واژگان), a set of dictionaries that were published by the Iranian University Press in the period from 1959 to 1980 and contained 22,960 English physics terms and 30,220 Persian equivalents (Mansouri, 1988).<sup>15</sup>

Abolhassan Najafi authored a book that was prepared in the form of a dictionary in which he dealt with correcting the common linguistic Persian mistakes in written and spoken language. In addition, he discussed the morphological and grammatical issues and mistakes that mostly appeared in association with the loanshifts in Persian with the aim of stabilizing the integrity of the Persian language and preserving nationalism (Najafi, 1390 [2011], p. 4).

---

<sup>14</sup> Sadeghi attributes the first attempt of coining Persian and replacing Arabic terms to Avicenna in his encyclopedia *Dānešnāme-ye `Alā`ī*. Nevertheless, this attempt was not so successful and was not followed by any scholar for a long time to come (Sadeghi A. A., 2001, p. 19).

<sup>15</sup> For more information about that collection of dictionaries see Mansouri (1996).

In his book, Najafi defines the following criteria to confirm the correctness or the incorrectness of the terms or the phrases. First, he mentions the ancient language based on the most authentic works left over the past 1,400 years, secondly contemporary spoken language, and thirdly contemporary written language (Najafi, 1390 [2011], p. 5). According to Najafi, a term is considered correct if it meets all three conditions such as *peyrūzmand* (victor, پیروزمند) and *peydāyeš* (appearance, پیدایش). The terms are also considered correct but abrogated if they occur in the first case only even if it is obsolete, for example *fāleġ* (فالج) instead of *falaġ* (paralysis, فلج) and *be tark goftan* (به ترک گفتن) instead of *tark goftan* (abandon, ترک گفتن). The terms that meet the second and third conditions are also considered right such as *adam/ādamī* (human, آدمی/آدم) and *šabāhat* (similarity, شباهت).

#### **2.4.6 The Third Farhangestān: The Academy of Persian Language and Literature (APLL, 1990)**

After the Iranian Revolution, starting from 1366š/1987, four institutions called “Farhangestān” were established: The Academy of Science, the Academy of Medical Sciences, the Academy of Persian Language and Literature, and, finally, the Academy of Art (Sa‘adat, 1393 [2014-2015], p. 51). Although there was no official existence of the Academy of Persian Language and Literature before 1987, there had been some linguistic activities running through *markaz-e našr-e dānešgāhī* (مرکز نشر دانشگاهی) until the late 1980s (Sarī, 1387 [2008-2009], S. 321).

The Academy of Persian Language and Literature is the official institution that is responsible for preserving, strengthening, and disseminating the Persian language in line with article 15 of the Constitution of 1979/1989, which makes Persian the official language of Iran. These were basically the main goals of the First and the Second Farhangestān, too. Therefore, the Third Academy was also named Farhangestān (Sadeghi, 1375 [1996], p. 149).

In 1999, the Iranian council of ministers prohibited the use of foreign names, titles, and terms. Exceptions to this rule are only Arabic terms and some foreign terms from other languages that were common in the Persian language from ancient times. According to article 4 of the decree of the council of ministers of Iran, Arabic terms and phrases that come from Islamic texts, teachings, and culture that are consistent with the context of the standard Persian language are not considered foreign words (Council of Ministers, 1999).

In order to implement this decree, the Persian Language Preservation Council (*šūrā-ye ḥefẓ va tarvīġ-e zabān-e fārsī*) was established, which works in cooperation with the APLL in



identifying foreign elements, then submitting them to the Academy to find Persian equivalents, and finally trying to promote them (article 10 of the decree of the council of ministers of Iran). Each one of the ministry of cultural and Islamic guidance, the Islamic Republic of Iran Broadcasting IRIB, the APLL, and other relevant organizations are compelled to provide appropriate social and cultural grounds for the implementation of these regulations through the mass media (article 18) (Council of Ministers, 1999).

These legal provisions help the Academy implement its choices for approved terms. For example, it is compulsory in all school textbooks from elementary to high school to insert the terms approved by the Academy. According to the vice-chairwoman of the APLL terminology group Nesrīn Parvīzī, this work has almost been completed as of October 2021. However, in tertiary education such as in universities, the state does not provide specific textbooks. Instead, university professors and other lecturers are free to choose the books they deem relevant to their students – and “most of the teachers are reluctant” to use the terms approved by the APLL or books that contain them. As a result, spreading these neologisms in universities will still take longer (Parvīzī, 2021).

It is also mandatory for all public-sector institutions and associations that receive funding from the government to use the terms approved by the Academy – at least in written documents and texts. This includes, for example, the laws, minutes, and other documents drafted by the Iranian parliament (*Mağles-e šūrā-ye eslāmī*), which the parliamentary administration forwards to the APLL terminology group for reviewing and checking. In addition, private-sector companies that seek official registration with the state are also obliged to use approved language in accordance with a 1375 (1996/1997) law on prohibiting the use of foreign names, titles, or terms (Parvīzī N. , 2021).

The statute of the Academy of Persian Language and Literature was drafted at the sessions number 208 and 209 of the Council of the Islamic Revolution, approved by the Supreme Council of the Cultural Revolution (*šūrā-ye ‘ālī-ye enqelāb-e farhangī*, شورای عالی انقلاب فرهنگی), and submitted to the government. So, in 1990, the Academy was officially established.

According to the third article of the Academy’s statute, the President of the Republic is the nominal head of the Academy. But the operational business and activities are run by the president of the Farhangestān who is appointed by a decree of the President of the Republic and approved by the Supreme Council of the Cultural Revolution for a four-year term. This office

was held alternately by Ḥasan Ḥabībī<sup>16</sup> (1991-1995 and 2004-2008) as well as Gholam Haddad Adel (1995-2004 and since 2008 to date).

Its goals are set out in the statute of the Academy as follows (Haddad-Adel, 1382 [2003-2004]):

1. Maintaining the strength and originality of the Persian language, the second language of the Islamic world, as one of the pillars of Iranian national identity conveying knowledge and Islamic culture.
2. Developing a clear and refined language to clarify the scientific and literary thoughts and creating continuity between the past, present, and future generations.
3. Promoting the Persian language and literature and extending its domain inside and outside Iran.
4. Expanding the Persian language to fit the circumstances of the time and for the preservation of human life and the development of science and technology.

In 1992, the Academy of Persian Language and Literature, in order to fulfill its duties and to respond to the needs of the Persian language, began with the formation of four groups: terminology, lexicography, Persian grammar and orthography, as well as languages and Iranian dialects. From its inception until 2012, the Academy gradually increased the number of working groups to become 14. Adding to the previously existing four groups they are as follows: teaching Persian language and literature, Islamic Revolution literature, comparative literature, contemporary literature, foreign studies, spell checking, encyclopedia of literary research, language and computer, Encyclopedia of Persian Language and Literature in the Subcontinent, as well as Encyclopedia of Persian Language and Literature in Asia (The Academy of Persian Language and Literature, 2019).

Aside from the tasks assigned by the APLL to its works, it organizes and participates in other cultural and scientific activities such as organizing several educational classes (e.g., teaching term formation, learning other languages like German or Khwarezmian, teaching computational skills, organizing conferences for Persian students and academic researchers, and holding seminars about terminology and Persian dialects) as well as sending and receiving students in cultural exchange programs. Moreover, the Academy is a member of the

---

<sup>16</sup> Ḥassan Ḥabībī was the first vice-president of the Islamic Republic of Iran (1989-2001).

international information center for terminology of UNESCO (Info-term), which is headquartered in Vienna (Baradaran & Akbarpour, 1395[2017], p. 5).

In addition to fighting the incursion of foreign terms into the Persian language, a couple of new problems appeared on the scene doubling the problems that Persian faces especially with modern technological life and the spread of computer devices and the Internet and of academic, educational, or scientific publications. Part of these problems is the Persian writing system (orthography). Since compound words in Persian could be written either separately or connected using “pseudo-space”, spelling causes many problems in discourse understanding (see chapter 3.1 and the introduction of chapter 5). The APLL has since 1372š (2013/2014) considered it one of its important issues to study and solve. Its target was to compile and edit the Persian script system and standardize the writing system (Kashefi, Mitra, & Kanani, 1389 [2010], p. 2).

The APLL published a booklet that compiles rules and regulations of the Persian writing system to be able to establish consistency of the Persian orthography. Especially the spread of the Internet and the large number of its users may lead to misusing the way of writing the language in the long run – a phenomenon that is also visible in the blog posts that serve as a primary source in the present study (see chapter 5). Ensuring correct spelling and codifying orthography is therefore an essential task for the APLL as the governmental institution entrusted with preserving the Persian language and everything related to it. The Academy took responsibility for each detail related to the Persian language starting from the writing system to speech and pronunciation problems, approving a standard language to be used in spoken, written, academic, and scientific parts of life, and ratifying specific terms selected by competent committees to replace the unwanted foreign terms used in Persian and preparing dictionaries with the selected terms in all fields (cf. The Academy of Persian Language and Literature, 1391[2012-2013]).

## 2.5 The Terminology Group

The following section will highlight the main principles and regulations that apply to the current terminology group (*gorūh-e vāzehgozīnī*) in the APLL.

### 2.5.1 Terminology

The term *term* has always been a subject of controversy by terminology schools and so many interpretations have been given to express it. It was defined in the *Glossary of Terms Used in*

*Terminology* as a “lexical unit consisting of one or more than one word which represents a concept inside a domain” (De Bessé, Nkwenti-Azeh , & Sager, 1997, p. 152).

*Terminology* – the “study of terms” – is a multi-meaning term that works in a lot of contexts. The communicative theory of terminology introduces *terminology* as an interdisciplinary field that draws its theoretical foundations from linguistics, cognitive science, and social sciences (Cabré M. T., 1999, p. 25).

Cabré argues that *terminology* is a discipline in and of itself. It is based on the study of multi-dimensional terminological units that are at the same time “units of knowledge, units of language, and units of communication”. A term in this sense is more than just a word. A terminological unit depends on concept, term (definition), and situation and can consequently be approached through any of these three entry points. It differs from a merely lexical unit regarding semantics and pragmatism, but not regarding its phonological, morphological, or syntactic features. Moreover, a terminological unit according to Cabré is embedded in the larger framework of communication and thus depends on text and context (Cabré M. T., 2000).

Another perspective is that *terminology* is used in three different senses: the principles and conceptual concepts that determine the field of term analysis, strategies and rules that are used in the terminological activities, and a set of terms in specialized fields (Bešārat, 1397[2016], S. 216).

The terminology group of the Third Academy launched its formal activities in 1995. Term formation (*vāžehgozīnī*, واژه گزینی) is a deliberate process of finding or creating a specific term in order to express a particular meaning and concept in one of the fields of science, technology, professions, and art. More than one term could also be found for the same concept for historical or other reasons (Terminology group, 1388 [2009]b., p. 20). Since the elimination of foreign terms in the Persian language was one of the most important elements of establishing the Academy in its three stages, this unit has been one of the most important working groups formed from the First Academy till the third one.

Since the main task of this group is to create and find suitable, appropriate, and compatible alternatives to foreign terms, it was necessary to set a strong basis and measures to answer the following question: On which basis will the new terms be chosen?

Eventually, the Academy council, after a lengthy discussion and detailed research taking into account the history of at least 60 years of term formation in Iran, approved a set of

principles and rules of term formation at its 58<sup>th</sup> meeting which took place on October 3, 1994, with further rules being introduced and approved in the 111<sup>th</sup> session of the council held on June 23, 1997. These principles and regulations were released in the form of nine clauses. They have been criticized in many studies, most importantly in a study entitled *Barrasī-ye vāžehhā-ye 'omūmī-ye mošavvab-e Farhangestān-e zabān-e va adab-e fārsī* (بررسی واژه های عمومی مصوب (فرهنگستان زبان و ادب فارسی) (Monšī-zādeh & Sharif, 1384 [2005]). In 2005, the principles and regulations of term formation were discussed again and some amendments were made by a group of specialists (Terminology group, 1388 [2009]b., p. 7).

### 2.5.2 *Vāžehgozīnī* (Term Formation) or *Eṣṭelāḥšenāsī* (Terminology)?

*Vāžehgozīnī* (term formation, واژه گزینی) is the term used by the Academy of Persian Language and Literature to indicate the process of finding or creating equivalents to foreign terms in Persian (Terminology group, 1388 [2009]b., p. 19). There was a dispute over using *eṣṭelāḥšenāsī* (اصطلاح شناسی) or *vāžehgozīnī* for that process as many researchers preferred *eṣṭelāḥšenāsī* as a translation of *terminology* (Bešārat, 1397[2016], p. 216).

In the Principles and Regulations of Terminology (Terminology group, 1388 [2009]b., p. 17), the definition of the term *vāžeh* (واژه) is given in two forms as follows:

1. “A term that consists of one or more meaningful components and grammatically belongs to one of the following categories: noun, pronoun, adjective, adverb, verb, number, preposition, conjunction, or sound. Such a term could be simple or derivative or a compound or derivative compound” (own translation).
2. “A term that at least has a specific definition in one of the scientific, technical, professional, and artistic fields. Such a term could be a word, a phrase or an abbreviated form” (own translation).

According to the second definition, the APLL considered *vāžeh* as a translation of both *word* and *term* that covers both meanings (Terminology group, 1388 [2009]b., p. 18). By contrast, Zarnikhi (Zarnikhi, 2010) opines that *eṣṭelāḥšenāsī* is the best equivalent for *terminology*. He argues that *eṣṭelāḥ* is an equivalent to *term* and *-šenāsī* for *-logy*. He insists that *term* and *word* are explicitly not synonyms, but distinct.

In order to clarify the main meaning and use of the term *vāžehgozīnī*, Besharat Fathi in his study (Bešārat, 1397[2016], pp. 205-210) researched that definition in the original text of

the principles and regulations of terminology by the Farhangestān. He found that this term holds two meanings. The first is corresponding to standardization which is related to some terminological, organizational, and systematic activities such as coordinating, approving, and releasing new terms. The second meaning is particularly related to the process of forming, creating, and selecting new terms. It is this second concept that could be compared to the English term “term formation”.

### 2.5.3 Goals and Tasks of the Terminology Group

The main objectives of the terminology group are to strengthen and expand the Persian language as well as to meet the growing cultural, scientific, and technical needs in addition to coordinating term formation, word structuring, and finding equivalents for the foreign language terms that have entered Persian. The main tasks can be summarized as follows (Terminology group, 1388 [2009]b., pp. 25-29):

1. Planning and managing the terminology group
2. Organizing the situation of the foreign non-original terms that entered Persian and finding equivalents for those terms
3. Contributing to the standards and concepts of various sciences and arts of the Persian language

In practice, the term-formation process follows specific steps between different units within the APLL. The group’s work usually begins with specialized committees studying the foreign terms extracted from publications (e.g., magazines, newspapers, textbooks, and similar word collections) that contain terms which are either new to their fields or new to Persian or occur frequently in language use. Then they create term sheets to describe in detail the selected terms or “candidate terms”, that is, terms which are seriously considered to enter the Persian lexicon. These sheets contain information, among others, on the etymology, term history, definitions, and existing equivalents (even in languages other than Persian if and where relevant). In addition, the terms are divided into categories to be studied and researched in accordance with the principles and regulations of the terminology group. In order to decide which term will be in which specific category, the council of the group holds a weekly meeting with its members. The group prepares the categorization according to the following points (Parvīzī, 1385 [2006-2007], p. 10, and Akbari, 2020, p. 35):

1. Terms that do not need an equivalent and are considered Persian;
2. Terms that need an equivalent;
3. Terms that need an equivalent, but the foreign terms need to be used alongside the Persian equivalents for some time to gradually take their place;
4. The specialized terms that need specialists and experts to take the decision.

On this basis, a book titled *Rāhnamāyī-ye ta'yīn-e ovlaviyyat dar vāzehgozīnī* (*Manual of Setting Priorities in Term Formation*, راهنمایی تعیین اولویت در واژه‌گزینی) was published beside the principles and regulations of term formation in order to set rules for selecting foreign terms and their categories.

After thorough deliberation, the abovementioned term sheets together with the proposed Persian equivalents are forwarded to so-called Harmonization Councils. It is the responsibility of these Harmonization Councils to make sure that the proposed terms are consistent with existing terms before the Technical Committees take over. These Committees are composed of two members – the vice-chair of the terminology group and another linguist – who review the proposed terms linguistically and might return them to the specialized committees for modification. At the end of the term-formation process, the ten to 15 members of a terminology council discuss the pros and contras of all propositions and, if they agree, give their approval for the new term to be published (Akbari, 2020, p. 35).

In 1997, the specialized terminology groups started their activities and cooperation with the Academy. In 2005, to increase the participation of institutions, associations, and scientific centers in the field of term formation, the collaboration with scientific communities began with the support of the scientific committees of the ministry of health and medical education. The number of scientific centers and institutions that have cooperated with the Farhangestān has increased to more than 70 (Parvīzī, 1385 [2006-2007], pp. 11-12).

In order to ensure the dissemination of the terms approved by the Farhangestān to the artists, authors, scientists, and eventually to all people, the terminology group of the Farhangestān has cooperated with the Iranian National Standards Organization (Sāzmān-e mellī-ye estāndārd-e Īrān).<sup>17</sup>

---

<sup>17</sup> The activities of the Iranian National Standards Organization are defined in five areas: standardization, conformity assessment, metrology, qualification, and standard research activities. The organization is responsible for developing and promoting standards and for developing new national standards or updating existing standards (Sāzmān-e mellī-ye estāndārd-e Īrān (The Institute of Standards and Industrial Research of Iran), 2019).

### 2.5.4 Principles and Regulations of the Terminology Group

The terminology group of the Academy of Persian Language and Literature has defined a set of principles regarding the process of creating neologisms instead of the foreign terms that entered the language. This set of principles was published by the Academy in a booklet called *Oṣūl va zavābeṭ-e vāzehgozīnī: hamrāh bā šarḥ va tūzīḥāt* (*The Principles and Regulations of Term Formation: With Explanations and Illustrative Examples*) in 2009 (Terminology group, 1388 [2009]b., pp. 43-58). This section will highlight some of these principles:

1. Persian grammar must be observed in the process of term formation.
2. The selected term must be suitable for the corresponding term-formation process, such as derivations and combinations.
3. Observing the phonetic elements of Persian and loanwords in the process of term formation, the pronunciation of new terms should be in accordance with the phonemes and syllabic rules in Persian.
4. The Persian writing system rules should be considered in the process of term formation.
5. Existing terms in the Persian language – whatever their origins – can be used in forming new Persian terms from foreign terms provided that they can be prefixed or suffixed according to Persian rules.
6. In case of homonymy<sup>18</sup> and polysemy<sup>19</sup> in both source language and Persian, the terminology group introduced a set of rules to follow:
  - a. When a foreign term carries more than one concept (homonymy), equivalents to each of these concepts must be found in Persian.
  - b. Given a foreign term in a specific field, only one equivalent in Persian should be selected unless that foreign term in the same field has more than one meaning (synonymy). In order to avoid confusion in Persian, it is admissible to find or create more than one equivalent to express different concepts of that foreign term separately.
  - c. When one concept is carried by more than one foreign term (polysemy in the source language), it is better to find only one equivalent in Persian.

---

<sup>18</sup> Homonymy: words which are written in the same way and/or sound alike, but which have different meanings altogether (Richards, Schmidt, Kendricks, & Kim, 2002, p. 241).

<sup>19</sup> Polysemy: (of a word) having two or more closely related meanings in the same semantic field (Richards, Schmidt, Kendricks, & Kim, 2002, p. 406).



- d. It is admissible to use only one Persian equivalent for different foreign terms in many domains (homonymy in Persian).

### 2.5.5 What makes Persian Persian?

The Persian language, like any ancient language that has its own literature and civilization, has acquired its words and vocabulary from multiple sources. Persian vocabs that are used in contemporary Persian in Iran, by the definition of lexicologists, came mainly from the following languages: Pahlavi, Parthian, Hindi, Arabic, Turkish, Mongolian, and European languages, especially French, English, and Russian (Ṭabāṭabā'ī, 1386 [2007-2008], p. 118).

According to the linguistic theories, it is possible to claim that the purely Persian terms are those commonly used in Persian today. But it should be noted that the number of common terms in each language over a given period is not large compared to its local lexical storage. The famous Iranian linguist 'Alā' al-Dīn Ṭabāṭabā'ī suggested the following criteria to identify authentic Persian terms: all the terms taken from Iranian languages, Hindi, Arabic, Turkish, and Mongolian and still used in contemporary Persian among people (Ṭabāṭabā'ī, 1386 [2007-2008], pp. 119-121).

The term *Farsi* itself was disputed by the Academy of Persian Language and Literature. *Farsi* is Persian's Iranian endonym that has gained popularity in English and several other languages in recent decades, primarily due to the rapid influx of Iranian migration after the Islamic Revolution (Akbari, 2020).

In 1992, the Academy of Persian Language and Literature highly recommended that the application of the term *Persian* is to be encouraged in international communication. The Academy argued that *Persian* has been applied in scientific, cultural, and political documents for centuries and has semantic, cultural, and historical associations. If one used *Farsi* instead, one would make the “grave mistake” of ignoring this long history. Some within the APLL emphasized that it had been a big mistake to change the name of the country from Persia to Iran in 1935 when the Persian government asked for countries with which it had diplomatic relations to call Persia Iran (Akbari, 2020, p. 26, and Yarshater, 1989).

### 2.5.6 Public Relations and Raising Awareness Activities

The terminology group actively publishes and advertises the approved terms on a variety of platforms.<sup>20</sup> Since 2004, it has published 16 dictionaries under the title *Farhang-e vāžehhā-ye mošavvab-e farhangestān* (*A Collection of Phrases Approved by the Academy*) containing over 60,000 approved Persian equivalents altogether or roughly 4,000 terms every year. The approved terms published in these dictionaries can then still be recalled or revised within a three-year period from publication. Each volume is processed through the office of the president, which forwards copies of each volume to all governmental institutions so as to encourage implementation. As mentioned before, the Farhangestān itself does not have executive power to enforce the application of the terms it approves (Akbari, 2020, S. 35). Instead, it relies on a mix of certain legal provisions that are meant to ensure their use (see chapter 2.4.6), on their linguistic power of persuasion, and on target-oriented public relations and raising awareness work.

Aside from its general dictionaries, the terminology group has released several specialized dictionaries addressing disciplines such as Chemistry (*Hezār vāžeh-ye šīmī, A Thousand Terms of Chemistry*, 1392 [2013-2014]), Engineering (*Hezār vāžeh-ye ‘olūm-e mohandesī, A Thousand Terms of Engineering*, 3 volumes, 1394-1395 [2015-2017]), and Linguistics (*Hezār vāžeh-ye zabānšenāsī, A Thousand Terms of Humanities*, 1397 [2018-2019]).<sup>21</sup> More recently, it announced its plans to prepare a dictionary with special attention to the Covid-19 pandemic and related medical-epidemiological terminology (The Academy of Persian Language and Literature, 2020). Other publications on a variety of terminology-related subjects complete the terminology group’s publishing activities.

These dictionaries are not only printed and distributed as books, but their content is also available online. The Academy has its own website<sup>22</sup> where one can search for foreign or approved terms including their definitions. Furthermore, the Academy cooperates with the online platform Vajehyab.<sup>23</sup> This private Persian translation association offers several services

---

<sup>20</sup> For an abridged version of this chapter and the information compiled herein cf. Shehata (2021).

<sup>21</sup> For a complete list of specialized dictionaries see the Farhangestān website at *Ātār-e montašer šode dar gorūh-e vāžehgozīnī*, 1398 [2019].

<sup>22</sup> <https://apll.ir/>

<sup>23</sup> <https://www.vajehyab.com/>

to its users including access to online Persian dictionaries. Another cooperation partner is the online search engine Vajehyar<sup>24</sup> that offers also analytical tools (cf. Parvīzī, 2021).

The terminology group furthermore cooperates with the media to foster the spread of the approved terms making use of the wide reach of traditional media such as newspapers, radio, and TV stations. The flow of feedback to the terminology group is ensured by the APLL communications office, which facilitates the further development and evolution of the research criteria on the one hand and the direct and indirect dissemination of linguistic content on the other.

This cooperation, however, proves difficult sometimes as pointed out by the vice-chairwoman of the terminology group Nesrīn Parvīzī, who stated that the media would use the Farhangestān's proximity to and affiliation with the government as an excuse to not use the approved neologisms. She believes that the Academy itself is not free from fault because it had not done everything in its power to publicize and advertise the approved terms in the past. However, in recent years, the APLL set up a (social) media unit specifically tasked with advertising the new terms. This special department has managed to reach out to a new audience interested in Farhangestān activities and to engage them in an open exchange on term formation (Parvīzī N. , 2021).

The (social) media unit of the APLL has been successful in accessing and tapping various social media channels where the account of the terminology group usually features under the name of *Cheshmocheragh* (چشم چراغ). Its first social media channel was created on Telegram<sup>25</sup> in July 2016 and had more than 2,500 subscribers in January 2022. On Telegram, the terminology group invites the public to join its researchers in discussing terms that are still in the process of being approved and in engaging in other ways with Persian language and literature by sharing texts, photos, voice notes, etc.

In 2019, the terminology group created accounts also on Instagram, Facebook, and Twitter. The first post on the terminology group's Instagram profile<sup>26</sup> was published in February 2019. It is also the profile with the largest following reaching almost 6,800 followers by January 2022. At the same time, there had been more than 1,150 posts meaning that its social media team had published on average more than one post per day. The terminology group is also very

---

<sup>24</sup> <https://vajehyar.inoor.ir/>

<sup>25</sup> <https://t.me/cheshmocheragh>

<sup>26</sup> [https://www.instagram.com/\\_cheshmocheragh/](https://www.instagram.com/_cheshmocheragh/)

active on Facebook,<sup>27</sup> which it joined in June 2019. The *Cheshmocheragh* group on Facebook had more than 1,800 members in January 2022 and presented an average of two posts daily. On Twitter<sup>28</sup>, finally, the terminology group has been active since August 2019 and attracted more than 2,200 followers by January 2022. The number of tweets varies greatly throughout the weeks with breaks of a few days at some times and up to four tweets daily at other times.

The fifth and last social media platform where the terminology group maintains an account is Soroush Plus. It is an Iranian produced messenger app that is only accessible to subscribers and can only be downloaded with an Iranian mobile phone number.<sup>29</sup>

## 2.6 Permissible Persian Terms in Term Formation

The council of the Academy defined a set of terms that may be used in the formation of the new equivalent terms instead of the loan words. They belong to a set of three categories as follows (Terminology group, 1388 [2009]b., p. 29):

1. Old terms used in the Persian language regardless of their origins
2. Terms that belong to the living languages and dialects of Persian
3. Terms that belong to languages of Middle and Old Persian

The first category can be further divided into:

- a. All originally Persian terms that are documented in authentic dictionaries
- b. All the originally Arabic terms that are either still in use in Persian today or at least the terms which have been used in authentic and valuable poetry and prose books, especially till the end of the fifth Islamic century (corresponding roughly to the early twelfth century CE). The acceptance of Arabic words<sup>30</sup> rooted in Persian as an integral part of the language proves the hypothesis of many current studies such as that by Ariyae (2019), which states that the more frequently a loanword is used in the borrowing language and the more it resembles that language's native words, the more it is rooted in the borrowing language and becomes an integral part of it. For example, such terms contain no marker that prevent them from being Persianized.

---

<sup>27</sup> <https://m.facebook.com/vazhe.gozini.56>

<sup>28</sup> <https://twitter.com/persiantermino1>

<sup>29</sup> For more on Soroush Plus see <https://hi.splus.ir/>.

<sup>30</sup> For a detailed historical discussion see chapter 2.2.6.

- c. All the originally Indian, Turkish, Greek, and Mongolian books and all similar ideas that are still in use either in current Persian or in use in prestigious specialized texts
- d. All the European terms which abide by the following conditions:
  - Terms are still common among Persian speakers today
  - Terms are simple in their original source language, flexible enough to fulfill the rules of term formation, and accept all the morphological combinations made on them
  - Terms of which the experts of the Academy admit with evidence that it is not important to find alternatives in Persian so the terms could be used as in their original language

As for the second category, the Persian language has already benefited from many elements of Iranian languages during its development. Until today, it is still under the influence of borrowing from and lending to living Iranian dialects and languages. The living languages and dialects meant in this category are Kurdish, Balochi, Gilani, and Mazanderani, dialects such as Sivandi in the region of Fars, Zoroastrian in Yazd and Kerman, as well as other Persian dialects such as Shirazi, Kermani, Khorasani, and Kashani. In the Farhangestān, the use of terms from these languages and dialects is mostly recommended with a view to replacing scientific and specialized terms.

Among these terms used by the Farhangestān is *hūšāb*, which is used in the Nishapuri dialect and refers to a type of juice. But the terminology group has used it to refer to *compote*. Also, the Farhangestān coined the word *zāmeḥ*, which is taken from the term *zāmād* in Kurdish and refers to the English term *groom*. But the terminology group in the Farhangestān shortened it and added the suffix *-h* to refer to the English term *sperm* or *spermatozoid* (Terminology group, 1388 [2009]b.).

Finally, the third category looks at the history of Persian. According to the distinction made by Iranian linguists, the Persian language can be divided into three stages. They are as follows: the old era, the middle era, and the modern era. The old era started from the ancient times and lasted till the end of the Achaemenid empire. The languages used in the old era were Old Persian and the language of the Avesta. Old Persian was the official language of the Achaemenid Empire while Avestan was the language of the religious writings (Şafavī k. , *Ášenā-ye bā tāriḩ-e zabānhā-ye īranī* (A Short History of Iranian Languages), 1386 [2007], p. 60).

The middle era lasted from the beginning of the Parthian Empire till the end of the Sasanian Empire and the appearance of Islam. The languages in this era were Parthian,

Akkadian, Middle Persian, Sogdian, Bactrian, Khwarezmian, Saka, and Old Ossetic. Linguists categorize languages in this epoch into two parts: eastern languages such as Sogdian, Khwarezmian, Bactrian, and Saka, as well as western languages. The western Iranian languages of the middle era were known as Pahlavi and themselves distinguished as either Sasanian Pahlavi or Parthian Pahlavi (Lazard, 1384 [2006], pp. 70 - 73).

The modern era starts with the appearance of Islam and continues to this day. New Persian has many sections such as Dari, Pashto, Tati, Balochi, and so on. But among all these languages appearing in this time, New Persian has prevailed and become the official language of Iran (Vafaei, 1390 (2011/2012), p. 7).

## Chapter 3

### Review of Literature

#### 3.1 Iranian Scholarly Disputes

Languages can be considered to be living organisms exposed to what living beings are exposed to. Words of these languages travel from one language to another and their shape changes depending on the place in which they are used. As discussed in chapter 2, the history of languages is replete with the borrowing of words between each other. There is no living language that is safe from influencing or being influenced by another language.

Generally, one can claim that the more foreign words a language has, the more attempts are made to get rid of them and replace them with words original to the receiving language accepted by native speakers. By the same token, the more borrowed words in a language, the more likely an institution such as the Academy of Persian Language and Literature in Iran will be found to assume the responsibility of defending that language and replacing such neologisms with native equivalents. Whether in the fields of term formation or creating neologisms and equivalents to loanwords, whether evaluating approved terms or measuring their success, this topic has always been a subject of study, concern, and scrutiny by linguists and other scholars throughout history – not least in Iran.

Because the history of language is full of linguistic borrowing, linguists have dedicated countless studies to this issue. A great number of researchers has already discussed the equivalents coined by the APLL and the measure of their success. It has been covered, discussed, and researched by many Iranian linguists such as Ṭabāṭabā'ī, Bateni, and Khanlari, who will be introduced in this chapter.

Beginning with the First Academy in 1935 up to the third one that is active nowadays, all Persian language academies have been and continue to be responsible for the process of finding equivalents to foreign terms and publishing them. The First, Second, and Third Farhangestān are all the subject of various studies criticizing, evaluating, and assessing the success or failure of the coined terms they approved. Numerous articles, papers, and academic studies have dealt with everything related to the language academies in Iran be it about their methods, their objectives, or the morphological process of coining neologisms and equivalents. The situation of loanwords in Persian and the discussion of the importance of receiving these

words and whether it is necessary to replace them or accept this phenomenon as a natural matter of linguistic interference is also at the center of such articles, papers, and studies.

“Poverty of words in Persian is nothing but poverty that comes from cultural poverty.” It is this provocative statement with which Sīrūs ‘Alī Nežād started his 1987 interview<sup>31</sup> with the linguist, Iranian literary scholar, author, researcher, and politician Parviz Natel Khanlari<sup>32</sup> explaining the reason for the increase of foreign terms in Persian, especially in the field of technology and science, that, in his opinion, was due to the flourishing of Western society, while at the same time asking about what Iranians could do to preserve Persian as a language for national identity. Khanlari argued that “the creative society has all the right to create and make new words and terms. When we import these [technical and scientific] phenomena, we import their words with them” (Nežād, 1388 (2010), p. 43).

Asked about his opinion on the impact of Persianization efforts, Khanlari replied that as much as he was trying to find more than one alternative to the foreign terms in the dictionaries he authorized, he had come to believe that not all the terms – especially the scientific ones – had to be Persianized. He also considered it a waste of time and effort. While saying so he used foreign terms from French to make his point: *dīksyūnerhā* (dictionaries, دیکسیونرها) and *āmātūrīm* (amateurs, آماتوریم) (Nežād, 1388 (2010), pp. 44, 49-50).

However, concerning the obstacles of term formation in Persian, ‘Alā’ al-Dīn Ṭabāṭabā’ī divided them into extralingual obstacles (“mavāne’-e berūnzabānī”) and intralingual obstacles (“mavāne’-e derūnzabānī”). The extralingual barriers are related to the culture of the Persian-speaking community in the face of diverse orientations in society. On the one hand, there are the “Persianists”<sup>33</sup>, who believe that all Latin and Arabic religious, scientific, and philosophic terms must be replaced by Persian ones. On the other hand, there are the “Arabicists”, who believe in borrowing from Arabic. Then came the “Latinists” (“lātīngarāyī”), who are considered among the biggest obstacles to Persian term formation because they believe that

<sup>31</sup> This interview took place when Khanlari was released from prison after having been arrested during the Islamic Revolution. The interview was directed by the famed Iranian journalist Sīrūs ‘Alī Nežād, editor of the Iranian publications *Adīneh*, *Donyā-ye soḥan*, and *Zamān*.

<sup>32</sup> Parviz Natel Khanlari is a prominent linguist and scholar of Persian language and literature, translator, and a professor in Tehran University. He is the founding editor of the important Iranian periodical *Soḥan*. He was in 1964 the secretary general of Bonyad-e Farhang-e Īrān (the Iranian Cultural Foundation). This Foundation was set up to preserve the heritage of the Iranian nation. He also served as secretary general of the Academy of Literature and Arts (Farhangestān-e adab va honar-e Īrān) (Āḍarang & EIr, 2016).

<sup>33</sup> This term is an own translation of the Persian term *fārsīgarāyān* (فارسی گرایان) mentioned by Ṭabāṭabā’ī. Likewise, the term *Arabicists* is a translation of *‘arabīgarāyān* (عربی گرایان). Both translations are in analogy with other *-isms*, which the APLL usually translates as *-garāyī* (identifying its followers with the suffix *-garāyān*).



specialized loanwords – especially from Latin, English, or French – should remain in their original language instead of being Persianized (Ṭabāṭabā'ī, 1376 [1997/1998], pp. 32-34).

In this regard, the Iranian linguist Mohammad Reza Bateni presents three different views. The first view believes that foreign words including the Arabic ones must be totally eliminated from Persian and the language must be pure – the Persianist doctrine so-to-speak. But in his opinion, this view cannot be implemented as there are no other available words in modern Persian to refer to the same meaning. In this case, the Academy has to search in old Persian or Pahlavi to find alternatives and then find a way to make people use them in their daily conversations (Bāṭenī, 1388 [2009/2010], pp. 67-68).

The second view argues that there is no completely pure language and Arabic words help enrich the Persian language. The focus is rather on the syntax and the linguistic structure of the languages. Proponents of this view believe that this would require from all Persian learners to learn Arabic grammar before learning Persian. One such example could be plural forms in Persian. It would, according to this view, not be enough to learn about the suffixes -*hā* and -*ān*, but learners would have to learn about the Arabic irregular plural rule (Bāṭenī, 1388 [2009/2010], pp. 67 - 68).

Finally, the third view, which Bateni follows himself, is concerned more with the morpheme. This opinion agrees with keeping the simple Arabic words without using their morphological derivations (Bāṭenī, 1388 [2009/2010], p. 67). The second and third view reflect the Arabicist doctrine according to Ṭabāṭabā'ī.

Bateni believes that the situation is somehow different when speaking about European words in Persian. Despite the multiplicity of European loanwords in the daily spoken language such as *māšīn* (car, ماشین), *ūtūbūs* (bus, اوتوبوس), or *telefon* (phone, تلفن), it is easy to find Persian equivalents to these words. However, it is not necessary to find equivalents to all the specialized terms in countless fields. Bateni advises the Academy to focus more on controlling translation from European languages as he considers this a direct way of influencing Persian (Bāṭenī, 1388 [2009/2010], p. 70).

Bateni, in dealing with Persian language planning and reforming, states that the APLL in Iran as the official body responsible for language planning faces many challenges, among them the orthographic system and its ambiguities. These challenges are inherent to the writing system whether the language is written in books or online and extend into the proper use of the Persian alphabet. This can be explained with the large number of linguistic structures and the

different ways of writing them, which may lead to differences in meaning in addition to aesthetic differences. One of these challenges is the type of spacing that is used not only between words in a sentence or a phrase, but even within words. These types of spacing are known as *fāṣele-ye yek harfī* (فاصلهء یک حرفی) which is the regularly used space among the terms of the sentences and as *nīm fāṣele* (نیم فاصله) (semi space) which is the space used to connect the morphemes of the compound term. This issue is one of the most common problems of writing Persian and will be discussed in more detail in chapter 5 because it also affects any corpus analysis.

The study by Ranjbar et al. compiles and discusses common mistakes and ambiguities in the Persian writing system. It investigates the conformity of authors and their writing behavior with the guidelines of the Academy of Persian Language and Literature in terms of writing terms continuously, semi-spacely, or separately (Ranjbar, Abbaspour, Sotudeh, & Moloodi, 1398 (2019)).

Another common issue is that some people get confused about writing a letter like *alef* with or without *madda* (أ or ا) and do not adhere to the correct spelling of words. Similarly, nunation (*tanwīn*) and gemination (*tašdīd*) – two phonetic features in Arabic words – might cause issues when written in Persian words. Therefore, since the Persian alphabet, according to article 15 of the Iranian Constitution, is the official script in the Islamic Republic of Iran and considering that the Farhangestān is the official entity concerned with preserving the Persian language and writing system, the Academy issued specific rules and guidelines for writing the language including a pamphlet containing these rules (The Academy of Persian Language and Literature, 1391 [2013/2014]).

As for the characteristics of the Arabic language that may be found in Persian, Bateni mentions the pronunciation of Arabic loanwords in Persian, many of which have a different pronunciation in Persian although some Iranian scholars and authors believe that Arabic loanwords should be pronounced as they were in Arabic. He suggests that the Academy should advise on the right Persian pronunciation considering the pronunciation that corresponds to the characteristics of Persian regardless of the Arabic pronunciation (Bāṭenī, 1388 [2009/2010], pp. 66-68).

## 3.2 Previous Studies on Persian Term Formation

Since this study examines the impact of the APLL considering the political terms, this chapter presents a set of previous works that examine the issue of how much the Academy of Persian Language and Literature succeeded – or failed to succeed – in choosing and coining new terms by presenting them and their exposure to people. The first part of this review of literature presents previous studies conducted on similar political and international relations terms as the present thesis while the second part covers similar studies on measuring the success of the approved terms from non-political fields.

Several studies have dealt with evaluating or measuring the spread of the terms approved by the APLL in all fields among native speakers or a certain category of people such as journalists, translators, or university students. As the present study investigates the case of political terms, this chapter will initially shed light on previous studies that deal with the same semantic field. This specialized overview will be followed by a more general survey on studies about other semantic fields such as IT or general terms.

Interestingly, the APLL itself encourages young scholars to prepare their Master theses in the field of coining new terms and evaluating the impact of the currently approved terms so that it can benefit from their results and findings in the future. For this purpose, the APLL established its own Research Institute for Terminology Studies (*Pažūheškadeh-ye moṭālā‘āt-e vāžehgozīnī*) in 2012 (Parvīzī, 2021, cf. *Pažūheškadeh-ye moṭālā‘āt-e vāžehgozīnī* (Research Institute for Terminology Studies), n.d.). However, none of the Master theses on term formation and the impact of approved neologisms that are discussed below were part of this APLL research program.

### 3.2.1 Research on Coining Political Terms

Mona Valinezhad, in her Master dissertation submitted in 2016, investigates whether technical translators use the equivalents suggested by the APLL in their translations of political content. She uses a questionnaire including eleven foreign terms and their equivalents suggested by the APLL. The questionnaire was filled by 20 technical translators and one political expert. The political terms selected by the author are mentioned in the first volume of approved terms published by the APLL, which contains 84 approved political terms. However, Valinezhad excludes the political *isms* in her own study thus limiting the remaining terms to eleven.

According to the findings of Valinezhad, the terminology group of the APLL was not successful enough to achieve its main goal of spreading the approved terms among the Iranian

public in general and translators in particular (Valinezhad, 2016). However, the sample in this study – eleven terms evaluated by 20 translators and one political expert – is small although the results point in the same direction as those of larger studies.

In her 2009 Master thesis, Parisa Keshavarz examines the rate of using the political terms suggested by the APLL in the press. Taking into consideration the type of the publication – either a newspaper or a magazine –, the gender of the author, and the context, Keshavarz finds that the rate of using the APLL-coined terms is considerably limited. Only 44 terms out of a total of 196 coined terms were in actual use based on the analysis of 110 newspaper articles. Moreover, it is concluded that none of the variables “gender of the author”, “type of publication”, or “context” were effective on the rate of the usage of the coined terms (Keshavarz, 1388[2009]). While the number of examined terms is almost 18 times higher than in Valinezhad’s later study, the analysis of 110 newspaper articles still has limited reach to evaluate the success or failure of approved terms in general.

Another Master student, Azimeh Sattari (Sattari, 1388 [2009]), researches the application rate of APLL-coined terms compared to that of borrowed terms based on the content of Iranian radio programs. The researched terms were selected from the first, second, and fourth published volumes of the approved terms, and the statistical analysis is based on 103 programs. In addition to testing whether the results would be different according to the type of program (news programs or talk shows), it takes also into account whether the program is recorded or broadcast live. Unlike Valinezhad and Keshavarz, Sattari concludes that the political neologisms approved by the APLL – whether in news programs or talk shows – were, by comparison, in greater use than the borrowed terms. In much the same way, the frequency of the new approved terms was higher in the recorded programs than the live ones. This suggests that the approved Persian equivalents are used more likely in spoken language that is scripted as opposed to unprepared spoken language.

However, the previous studies of Sattari and Keshavarz do not directly compare the foreign terms and their Persian equivalents unlike the present study, but they rather compare the success of the coined terms considering their borrowing patterns and the types of borrowing the Academy followed in coining the approved terms. Their leading question is whether the borrowing pattern and the type of borrowing play a role in the failure or the success of the extent of use of the terms.

Akbar Hesabi et al. (Hesabi, Amirian, & Nazari, 2017) examine in their study the impact of 60 political terms approved by the APLL using the Hamshahri online corpus<sup>34</sup> and distributing a questionnaire that includes 60 political terms to 90 Master students. The descriptive results of that study note that two thirds of English Translation and English Literature students comply with the political terms coined by the APLL while fewer than half of Political Science students accept those new terms. On top of this, most of the overall participants of the survey are not even aware of the approved political terms suggesting that implementation strategies need to be intensified if the APLL hopes to spread them in the news. However, at the time when this study was completed, the corpus was already considerably out of date to research, express, and judge the success of the chosen terms approved by the APLL.

Aside from the studies that deal with political terms, there are numerous studies about the Persian equivalents approved by the APLL in general. One of them is a survey conducted by Sa'īd Ketābī et al. Similar to the method of the Keshavarz study mentioned above, the gender of participants is the main factor to evaluate the acceptance and the transparency of the general APLL-approved terms in their study. The research design comprises 60 university students (29 female and 31 male) to answer a questionnaire of 50 sentences. Each of these sentences presents one term approved by the APLL. The results of this 2010 study show that the success and transparency of the approved terms were slightly higher for female students compared to male students (Ketābī, 'Akīlī, & Abūlḥasanī, 1389 [2010]).

Another survey was conducted by Mahdi Aghili in his Master thesis. It investigates the transparency and acceptability of the neologisms coined by the Academy of Persian Language and Literature compared to the borrowed terms depending on the factor of the educational level of the participants that answered the survey. The variables that Aghili defines for the selected participants are their knowledge of English and their educational degree (Bachelor or Master students). He picked 30 Master students and 30 Bachelor students, half of each majored in English. They all answered a 50-item questionnaire that includes 50 coined terms.

The coinages were selected from the 1997 and 1998 compilations of general approved terms<sup>35</sup> published by the Academy of Persian Language and Literature. Both volumes contain 262 new terms among which the researcher determined 120 to begin with. As a preliminary step, the participants were then asked to select the final 50 terms. The final questionnaire itself

---

<sup>34</sup> The Hamshahri corpus was developed by the Database researcher group of the faculty of engineering at the University of Tehran. The corpus focuses on the Iranian online newspaper *Hamshahri* in the years from 1996 to 2002.

<sup>35</sup> Terminology group (1376 [1997]) and Terminology group (1377 [1998]).

was made up of two forms: the first one to measure the transparency – and thus the easy understanding – of the terms approved by the Academy and the second to evaluate the acceptance of the respective terms.

Based on the results, Aghili concludes that the foreign terms that already exist among Iranian Persian speakers are more accepted than the equivalent neologisms suggested by the APLL. In addition, he also finds that while the field of study of the participants was no indicator of how transparent and acceptable term creations by the Farhangestān were, the level of education turns out to be an effective variable: the more educated people are, the more likely they are aware of the approved Persian equivalents (Aghili, 2010).

Finally, Tāherah Aḥmadīpūr in her doctoral dissertation investigates the most important non-linguistic factors in accepting the terms approved by the APLL among Persian speakers in Iran. In order to do so, she follows the model introduced by Robert L. Cooper which states that language planning is a part of social planning effecting social change. In addition to gender, age, and education level, she therefore looks at several variables such as exposure to life outside Iran and, if applicable, number of years spent outside Iran, foreign language proficiency, daily media usage, and main job tasks (translating, writing, teaching, or none of the above). The research design includes 450 participants, 150 each from the three Iranian cities of Tehran, Esfahan, and Kerman. These locations were chosen because, according to Aḥmadīpūr, they reflect at the same time the highest exposure to internationality (Tehran) and the least impact of foreigners on Iranians (Esfahan and Kerman). In essence, although there is no clear pattern for the success or failure of the proposed Persian equivalents, Aḥmadīpūr concludes that the Academy could increase the acceptance of neologisms by making a greater effort at raising awareness about them (Aḥmadīpūr, 1389 [2010]).

### **3.2.2 Research on Coining Other Specialized Terms**

Measuring the success of the APLL in spreading the approved terms and creating a standard dictionary is not only limited to literary and linguistic fields, but there are many studies discussing the extent of the APLL's approved terms in various fields such as in medicine, physics, and information technology. Abdul Amir Hazbavi examines the success of the terms approved by the APLL in the field of computer and information technology in 16 books that were translated into Persian. In his descriptive study, 140 approved terms are introduced to be examined and compared to the foreign terms that exist in those books in order to determine the extent to which the translators used the approved terms or not. The results of Hazbavi's study

show that the translators tend to avoid the approved terms in their translations as they use only 43 approved terms while the remaining almost 100 terms bring together a variety of methods from simple transcription to calques. According to these findings, the APLL failed to standardize the Persian terminology in the field of computer and information technology (Hazbavi, 2012).

Much in the same way, Šaymā Mo'menī and Mehernūš Faḥḥārzādeh investigate in their study the dominant process followed by the terminology group to coin new terms in the field of computer and information technology. To do so, they select 302 English IT terms in addition to their Persian equivalents approved by the APLL. Then, they pick seven books in the field of computer science and information technology as a database to study and investigate the approved terms. This study finds that every English term in this field has an equivalent approved by the APLL. Therefore, the Academy has accomplished its task of term formation. However, with regard to the actual use of these approved terms, the authors conclude that the APLL is not very successful and needs to work on this matter (Mo'menī & Faḥḥārzādeh, 1395 [2016]).

Some studies deal also with medical terms such as the study by Farzaneh Bahadori and Samad Mirza Suzani, who investigate the most widely used terms approved by the APLL in the dental field. They examine the perceived adequacy, awareness among native speakers, and actual use of the approved Persian equivalents. The study is based on a questionnaire distributed among 69 male and female students at the school of dentistry of Shiraz University. Roughly half of the students were in the fourth semester and the other half in the eighth semester. The findings of this study claim that the higher the educational level of students, the more they find the approved equivalents adequate, the more they are aware of them, and the more they apply them in their language. In other words, the use of approved terms among the eighth semester students was higher than among the fourth semester students (Bahadori & Suzani, 2017).

Ġazāl Mohāġerīzāde criticizes the fifth published volume of the terms approved by the APLL and takes a closer look at individual term-formation groups in various fields. The main focus of her article is on the critique of terms that could be smoother, more fluent, and easier to understand and on terms for which more appropriate equivalents have already been made but have not been used. She claims that the most important problem facing the Academy is its lack of consideration for being up to date based on the assumption that the terms selected by the terminology group are not compatible with the current era. Therefore, Mohāġerīzāde argues that the spreading of such terms to colloquial language might create a cultural rift (“gosestegī-ye farhangī”). In her study, she investigates 122 terms. 41 of them are related to the field of

astrology while only two are political terms. The remaining terms are distributed among other fields. As for her evaluation of the term-formation groups of the APLL, she praises the results of the linguistic department because she believes that the neologisms they form or even choose are suitable and fit the language and the grammar whereas she considers the terms coined by the astronomy department as the worst. She criticizes that the department does not use all the rich terms found in the book *Al-tafhīm* by Abū-Rayḥān Bayrūnī instead of trying to coin neologisms. She believes that before the department coins new terms, it should dig deep into ancient linguistic treasures. On the one hand, the members of the department could save themselves the trouble of coining new words and, on the other hand, they could thusly highlight what old Iranians achieved in the various arts and sciences (Mohāgerīzāde, 1388 [2009]).

In her book, Fatemeh Akbari deals with the English abbreviations in Persian as well as the official terminology plan that the terminology group of the APLL follows to replace these English abbreviations with others that suit the nature of the Persian language. The main objective of this study is to understand the plan adopted by the terminology group, its applicability from a morphological perspective, and the compatibility of the abbreviations as a way to form terms in the Persian language. The study is based on the terms approved by the Academy during the period 1997-2013. The findings can be summed up in two points. First, not all English abbreviated forms obtain a Persian equivalent. Second, due to the irregular construction of the coined Persian abbreviation forms, most of them are difficult to decode (Akbari, 2020).

Omīd Maḡd in his study investigates the nine principles that the terminology group follows to coin new terms based on his selection of 40 approved terms. Those selected terms do not represent a certain field, though. He finds that the terms that the Academy approved are not easy to understand. According to Maḡd, the APLL also exaggerates with literal translations, which makes the terms lose their meanings and tastes from the original foreign term. Finally, his study also concludes that the approved terms lack adherence to Persian grammar (Maḡd, 1388[2009]).

There are furthermore a lot of studies to investigate the success of the terms approved by the First Farhangestān, among them two studies by ‘Alī Kāfī. His first study is on the approved mathematical terms revealing that the First Academy was not successful enough in coining neologisms to reflect mathematical concepts. Kāfī claims that the main reason for their failure was the primary concern of the Farhangestān not to use Persian words of Arabic origin rather than its concern for finding an equivalent (Kāfī, 1386 [1989]). In the second study, Kāfī



investigates the success of the approved terms of the First Farhangestān in all the fields such as military, phytology, zoology, administrative and nation, banking and commercial, political, medical, judicial, and legal, etc. The results of his research vary according to the field of words. Despite the success of the political, commercial, administrative, and judicial terms in addition to all terms related to government-affiliated institutions, Kāfī determines that geographical, engineering, meteorology, mathematical, and medical terms did not meet with much success. However, the approved terms that are related especially to the scientific disciplines of botany and zoology were successful and accepted. He hypothesizes that the reason for their success may be due to the fact that these terms had no previous Persian or Arabic equivalent (Kāfī, 1371[1992]).

### **3.2.3 Challenges, Recommendations, and Outlook**

The previously reviewed studies on the Farhangestān and its term-formation processes highlight the challenges that the Academy is still facing in the 21<sup>st</sup> century. Most of them point out that the success of the approved terms and thus the impact of the Farhangestān on modern Persian is limited. These results are largely based on surveys with comparatively small samples, but they are nonetheless indicative of some of the problems that the Academy should address in the eyes of the researchers.

Most of them include recommendations for the Academy to intensify its public relations work and raise awareness for its activities in general and the approved terms in particular. Greater awareness generally increases the probability of putting the approved terms in use. But awareness is not everything. If one assumes that the purpose of coining new terms is their application in practice, it should be at the core of the term-formation process to coin terms that are easy to understand and use. Here, some studies find that the Academy could improve its linguistic principles. Problematic issues seem to be overly literal loan translations and an out-of-sync approach to the needs and requirements of the modern world.

There are a few instances where the terms coined and approved by the Farhangestān did enjoy some degree of success, though. Several studies suggest, for example, that higher education correlates with greater awareness and a higher likelihood of well-educated people to use the approved terms. Another prominent example is the case where political neologisms are frequently used in scripted radio news programs as opposed to live broadcasts. Although the type of publication is not always a decisive factor as could be seen from another comparison between magazines and newspapers, here the type of publication is an effective factor in

determining the impact of the Academy on modern Persian. It remains to be seen how this plays out between newspapers and blogs.

## Chapter 4 Methodology

### 4.1 Introduction

“Language is not a complex set of words together, but a pair of relations that create the contexture of speech” (Ášūrī, 1392 [2014], p. 242). In this sense, the research in this study is based on a quantitative and qualitative corpus analysis where not only foreign political terms entering Persian and their approved Persian equivalents as such are compared, but where also the context and relations of these terms in the texts and documents where they appear are taken into account.

In this chapter, the methodology of the study will be presented and discussed. It is divided into three distinct sections. First, the foreign terms and their approved Persian equivalents selected for this research will be presented. Then, the two corpora used for the study and the characteristics of each one will be introduced. Finally, the idea of the TF-IDF weight of the term as opposed to the pure term frequency will be explained.

### 4.2 Selected Terms

Between 2004 and 2021, the terminology group in the Academy of Persian Language and Literature published 16 volumes of approved terms (Terminology group, 1382-1398 [2004-2021]) in addition to a consolidated version of the terms from 1997 to 2007 (Terminology Group, 1376-1385[1997-2007]). There is a total of 1,066 approved political terms mentioned in the published volumes – among them a significant number of compound terms. Not all the volumes contain approved political terms. But the political terms are distributed over most of the volumes.<sup>36</sup>

Table 1 shows a list of the terms that are discussed in this thesis. Some of them belong to the same narrow semantic field. More specifically, the list presents the pairs of foreign terms and corresponding approved Persian equivalents. They can roughly be divided into four sets of terms that belong together. The first set comprises terms related to parliament (*parliament, candidate, opposition*). The second set consists of the *nation* as an overarching entity and four

---

<sup>36</sup> This information was sent to the author by the administration office of the APLL, headed by Mehnoosh Tehrani, the head of the communication broadcasting office of the Academy. An updated Microsoft Excel sheet of early 2021 includes all the approved terms and all the related details.

political systems: *democracy* and *dictatorship* as two qualifying political systems as well as the concepts of *regime* and *bureaucracy* as two descriptive or functional political systems. The third set of related terms deals with *ideology* in general and *isms* in particular. Finally, the last set of terms covers the three terms *doctrine*, *strategy*, and *protocol*, which play a more functional role in the political field. In other words, the third and fourth sets could be summarized as policy-oriented while the first two sets are rather polity-oriented.

Foreign term		Approved equivalent term		English translation
<b>Polity-Oriented Terms</b>				
<b>(1) Parliament</b>				
<i>pārlemān</i>	پارلمان	<i>mağles</i>	مجلس	Parliament
<i>kāndīdā</i>	کاندیدا	<i>nāmzad</i>	نامزد	Candidate
<i>opīzīsyūn</i>	اپوزیسیون	<i>gorūh-e moḥālef</i>	گروه مخالف	Opposition
<b>(2) Political Systems</b>				
<i>ommat*</i>	امت	<i>mellat</i>	ملت	Nation
<i>demūkrāsī</i>	دموکراسی	<i>mardomsālārī</i>	مردمسالاری	Democracy
<i>demūkrātīk</i>	دموکراتیک	<i>mardomsālārān</i>	مردمسالاران	Democratic
<i>demūkrāt</i>	دموکرات	<i>mardombāvar</i>	مردمباور	Democrat
<i>dīktātur</i>	دیکتاتور	<i>mostabed</i>	مستبد	Dictator
<i>dīktātūrī</i>	دیکتاتوری	<i>estebdād</i>	استبداد	Dictatorship
<i>režīm</i>	رژیم	<i>nezām</i>	نظام	Regime
<i>būrūkrāsī</i>	بوروکراسی	<i>dīvānsālārī</i>	دیوانسالاری	Bureaucracy
<i>būrūkrāt</i>	بوروکرات	<i>dīvānī</i>	دیوانی	Bureaucrat
<b>Policy-Oriented Terms</b>				
<b>(1) Ideology</b>				
<i>īde`ulūzī</i>	ایدئولوژی	<i>marām</i>	مرام	Ideology
<i>īde`ulūzīk/ī</i>	ایدئولوژیک/ی	<i>marāmī</i>	مرامی	Ideologic
<i>terūrīsm</i>	تروریسم	<i>vaḥšatafkanī</i>	وحشت افکنی	Terrorism
<i>terūrīst</i>	تروریست	<i>vaḥšatafkan</i>	وحشت افکن	Terrorist
<i>pūpūlīstī</i>	پوپولیستی	<i>'avāmgarāyī</i>	عوامگرایی	Populist
<i>kāpītālīsm</i>	کاپیتالیسم	<i>sarmāyedārī</i>	سرمایه داری	Capitalism
<i>kāpītālīst</i>	کاپیتالیست	<i>sarmāyedār</i>	سرمایه دار	Capitalist
<i>reform</i>	رفرم	<i>eṣlāḥ</i>	اصلاح	Reform
<i>reformīst</i>	رفرمیست	<i>eṣlāḥ-e ṭalab</i>	اصلاح طلب	Reformist
<i>reformīsm</i>	رفرمیسم	<i>eṣlāḥ-e ṭalabī</i>	اصلاح طلبی	Reformism
<b>(2) Functionality</b>				
<i>doktorīn</i>	دکترین	<i>rahnāme</i>	رهنامه	Doctrine
<i>esterātežī</i>	استراتژی	<i>rāhbord</i>	راهبرد	Strategy
<i>esterātežīk</i>	استراتژیک	<i>rāhbordī</i>	راهبردی	Strategic
<i>protokol</i>	پروتکل	<i>tašrīfāt</i>	تشریفات	Protocol

Table 1. Selected foreign terms and the equivalent Persian terms approved by the APLL \* *Ommat* is the term that is compared with *mellat* even if it does not qualify as a foreign term

### 4.3 Corpus

This study is methodologically based on a corpus analysis of the online archives of major Iranian newspapers and blogs.<sup>37</sup> The selection of the research corpora is one of the most important factors to get the closest results to the actual standing of the selected foreign and approved terms among native Persian speakers. To achieve that, two corpora are used in this study. Each one represents the broad variety of different categories.

The first corpus is a self-created database called *Terms\_Analysis\_Persian* (TAP). It was built during the time of this research starting from 2018 till early 2021 by using the online tool Sketch Engine ([www.sketchengine.eu](http://www.sketchengine.eu)) just like the second corpus. The second corpus is called *TalkBank Persian (blog posts)* and based mostly on Iranian online blogs. Unlike TAP, TalkBank goes back to a research initiative of a group of scholars in the United States. Thus, Iranian newspapers and blogs serve as the primary sources of this research.

Sketch Engine is an online platform that provides large high-quality word databases, lexical data, wordlists, and lexicons in many languages such as English, Arabic, French, Spanish, German, Persian, and other languages as well. Concerning Persian, this platform provides concordances (examples of use in context), N-grams (multiword expressions), a one-click dictionary, a word list (frequency list), and key words (terminology extraction). This platform supports researchers with building a corpus or corpora and analyzing the data and gives also statistical results.

#### 4.3.1 Newspapers in the Corpus *Terms\_Analysis\_Persian* (TAP)

Although official and reliable numbers are hard to come by with, the newspapers in the corpus *Terms\_Analysis\_Persian* (TAP) can be counted among the largest and most widely distributed Iranian newspapers. The selection of the newspapers represents a wide variety of political orientations, some of which belong to the free newspapers sector while others are sponsored by the government. They cover the whole spectrum of political orientation from hardline or conservative papers such as *Kayhan* (controlled by the Supreme Leader) to more moderate titles such as *Hamshahri* (published by the Tehran Mayor's office) to reformist-leaning papers such as *Aftab-e Yazd* (affiliated with circles around former President Mohammad Khatami) (Khiabany, 2010, p. 84 and Mahtafar, Khiabany, & Niknejad, 2009).

---

<sup>37</sup> For an earlier discussion of the corpora see Shehata (2021).

Until the late 2000s and early 2010s, the largest of them sold daily copies of several hundred thousand. Although such figures should be viewed with caution, more recent numbers suggest a decline of daily circulation – some of it possibly due to economic constraints related to the sanction’s regime against Iran. As of August 2018, *Hamshahri* still sold around 180,000 copies per day while *Jam-e Jam* was estimated at 70,000 to 80,000 daily copies. Most other papers – among them *Aftab-e Yazd*, *Shargh*, *Arman*, and *Etemad* – sold less than 7,000 copies per day (Khiabany, 2010, p. 85), (Payvand, 2018). Of course, this is a glimpse at the print circulation only.

A look at the number of online followers, for example on Twitter and Telegram (see table 2 below), reveals yet different results and a dynamic development. On Twitter, *Shargh* has more than 400,000 followers with an increase of almost 20,000 between November 2020 and January 2022. ISNA gained the largest number of new followers in 14 months with an increase of more than 40,000 or 20 percent whereas *Hamshahri* almost doubled the number of its followers on Twitter albeit on a low level and still tailing all other major newspapers. *Shargh* also has the largest combined following on Twitter and Telegram with more than 450,000 followers. *Jam-e Jam Daily* attracts over 160,000 followers on Telegram and ranks first on that platform despite a drop of more than 40,000 followers between November 2020 and January 2022. *Iran* and again *Hamshahri* managed to almost double the number of their followers on Telegram in the same period.

*Jam-e Jam* is the only newspaper that does reasonably well among both print and social media audiences. For other newspapers, results are reversed. While *Shargh* and *Etemad* gather a significant number of online readers, *Hamshahri* is not yet very popular among social media users although its growth rates are steep.

Twitter				Telegram			
Account	Followers (11/2020)	Followers (01/2022)	Active since	Account	Followers (11/2020)	Followers (01/2022)	Active since
@SharghDaily	422,836	442,226	August 2013	SharghDaily	11,929	12,431	2015
@isna_farsi	207,060	248,772	July 2015	isna94	82,787	84,705	2015
@EtemadOnline	173,971	185,926	April 2014	etemadonline	N/A	67,462	2017
@IranNewspaper	110,602	124,543	June 2016	irannewspaper	19,359	43,551	2016
@vatanemrooz	50,794	58,077	February 2014	vatanemrooz	3,278	3,054	2015
@jamejamCPI	48,551	54,860	October 2013	jamejamdaily	204,482	160,415	2019

@KayhanNewsFa	34,825	37,727	December 2013	kayhannews_1	1,567	1,584	2018
@hamshahrinews	14,899	26,739	September 2017	hamshahrinews	14,020	27,583	2016
@aftabeyazd_ir	Account blocked for violating Twitter's terms and conditions			Aftabeyazd_ir	2,528	2,428	2016

Table 2. Number of followers of major Iranian newspapers on Twitter and Telegram

In order to collect all the documents from the newspapers and perform the terms analysis, Sketch Engine was used to build a corpus that contains 30,991,286 words and 34,929,443 tokens<sup>38</sup> in 22,274 documents. The corpus is up to date as of June 25, 2021. The following table 3 shows selected Iranian newspapers that were used in building the corpus, including the number of words in each of them.

Newspaper	Number of words
Payām-e mā	5,477,218
Farhaḥtīgān	3,742,567
Jomleh	3,737,256
Jām-e Jam	3,256,729
Kayhān	1,805,583
Vaṭan-e emrūz	1,700,235
Āftāb-e yazd	1,531,872
Fardā	1,348,921
Ḥemāyat online	1,253,746
Īrān	1,140,675
ISNA (Iranian Students' News Agency)	1,076,706
Hamshahri online	1,001,044

Table 3. Iranian newspapers in the corpus TAP (with an individual word count of &gt;1mn)

The following figure highlights the space occupied by each newspaper in the corpus. it visualizes that only four newspapers – *Payām-e mā*, *Farhaḥtīgān*, *Jomleh*, and *Jām-e Jam* – make up more than 50 percent of the whole corpus. Each of them individually accounts for more than ten percent of the words saved in TAP. Another sizable portion consists of small newspapers with a 13-percent share in the corpus. All other prominent newspapers such as *Kayhan*, *Aftab-e Yazd*, *Iran*, and *Hamshahri* represent between three and six percent each.

<sup>38</sup> A token is any part of the corpus that can be written such as a word or a punctuation mark.

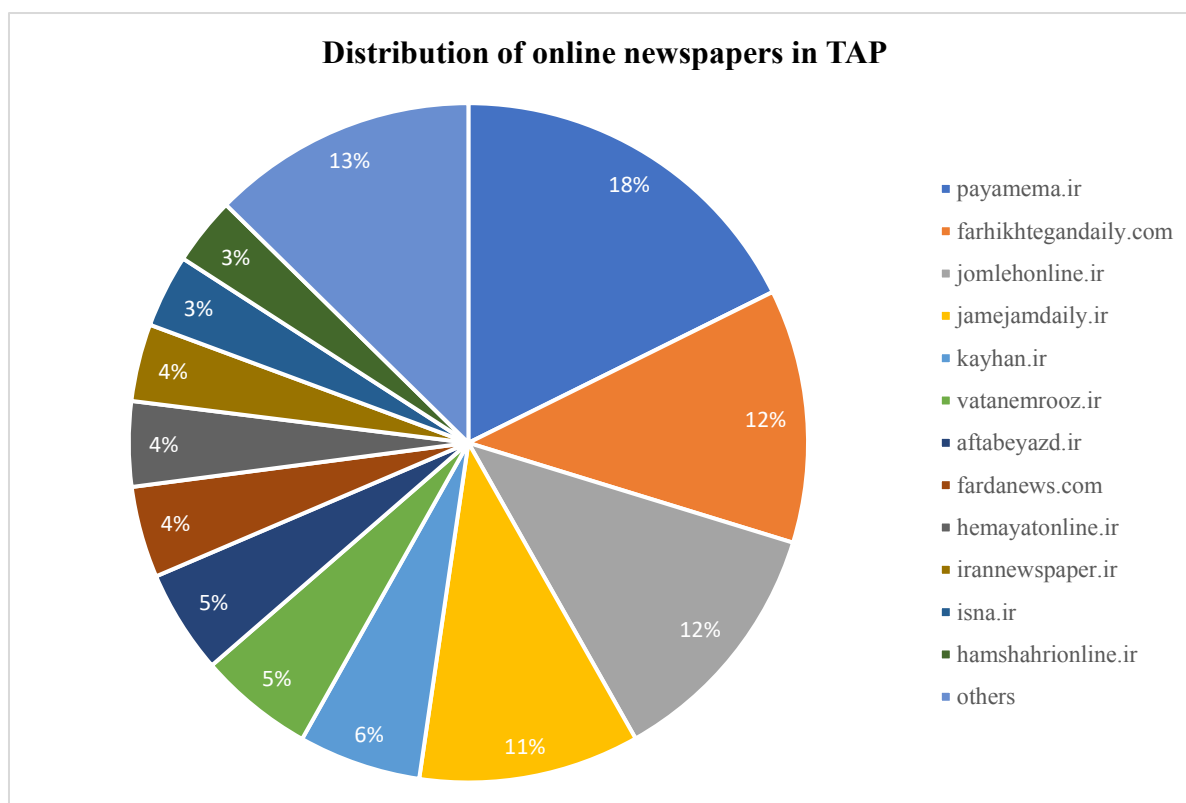


Figure 1. Distribution of online newspapers in the corpus TAP sorted by term frequency

Sketch Engine has a few limitations that prevent some of the newspapers websites from being included in the corpus. The html format is one of these limitations as Sketch Engine cannot, for example, browse webpages that are built using pdf files. Neither can it download protected webpages such as those behind paywalls or with password access. Very short, isolated paragraphs might also be ignored. The following table 4 contains the newspapers not accepted by the corpus because of these technological reasons. As can be seen, the number of words saved for each of them is significantly lower than that of other newspapers and lower than what would be expected by a newspaper.

Newspaper website	Number of words
<a href="https://www.asianews.ir/">https://www.asianews.ir/</a>	12,683
<a href="http://khorasannews.com/">http://khorasannews.com/</a>	26,011
<a href="http://www.eslahat.news/">http://www.eslahat.news/</a>	3,830

Table 4. Uninserted newspaper websites in the corpus TAP

### 4.3.2 Blog Posts in the Corpus TalkBank Persian

The second corpus that was used in this study is TalkBank Persian created from Persian blogs collected by Shlomo Argamon's research group at the Illinois Institute of Technology (IIT) in



cooperation with Brian MacWhinney's group at Carnegie Mellon.<sup>39</sup> It is not clear from Sketch Engine when TalkBank was created and last updated. A cursory overview of blogs saved in the corpus reveals references to the second term of President Hassan Rouhani. Hence, it contains texts at least up to the second half of 2017.

The general openness of the Internet and the spread of computers and mobile phones in the hands of most people has led to the emergence of new types of media and press since the 1990s. Web logs are one of these types. The term *web log* has been shortened to *blog* and has come to mean a more or less regularly published online journal by a single author or a collective of writers on any topics of their choice. Thanks to the variety of topics and authors, but also thanks to the not always formal language, Megerdoomian believes that the texts in Persian blogs contain elements commonly found in colloquial speech and informal conversations (Megerdoomian, 2010, p. 21). This highlights the importance of using a blogs corpus in the present study to reflect the actual use of terms among blog writers. Considering the absolute number of bloggers, the Iranian blogosphere is one of the most active ones and the Persian language ranked ninth in the world among Internet users in the mid-2000s (Megerdoomian & Hadjarian, 2010).

The first blog in Iran<sup>40</sup> goes back to Salman Jariri, who created <http://www.globalpersian.com/salman/weblog.html> in 2001 and has been active on that website since. He was later followed by Hossein Derakhshan, who not only wrote his own blog, but also released a construction guide and workshop videos in Persian for the Iranian youth so that they could create their own blogs to express their own opinions and thoughts. He aimed for a hundred Iranian blogs in a year. But what happened was not expected as within two years the number of new Iranian blogs increased to hundreds of thousands (Khiabany & Sreberny, 2007b., pp. 564-565).

Khiabany and Sreberny believe that the availability of software in addition to the expansion of Internet access and usage as well as the existence of a large number of technologically well-educated young Iranians should be considered some of the factors for the rapid growth of blogs in Iran (Khiabany & Sreberny, 2007b., pp. 564-565). Shirazi states that “thousands of Iranian youths, men and women, are active in today's blogging and are among

---

<sup>39</sup> For more information on the TalkBank project itself see its website at <https://talkbank.org/>.

<sup>40</sup> For further information about the situation of blogging in Iran see the following studies: Johnston (2008), Khiabany & Sreberny (2007a.), Khiabany & Sreberny (2007b.), and Khiabany (2007).

the world's leading bloggers". He believes that the reason behind this rapid spread of blogging in Iran and the explosion of text messaging is nothing but a reflection of the fact that Iranian youth are in search of change in the Iranian traditional life towards gender equality, free speech, the respect of human rights, and access to an unrestrained source of information (Shirazi, 2008, S. 302). This development can be observed in parallel with the rising numbers of newspaper followers in social media channels such as Twitter and Telegram as described in the previous chapter.

The blog-based corpus TalkBank Persian is freely and readily available at Sketch Engine. It contains 549,165,952 tokens and 474,773,547 words in 1,092,892 documents. Its configuration is slightly different from that of Terms\_Analysis\_Persian in the way the results are presented and accessible to researchers. When the selected term is searched the result shows the frequency of the term in the whole corpus (number of hits), the number of hits per million tokens, the percentage compared to the whole corpus, and the distribution of hits in the corpus, as well as the frequency of documents where the term is mentioned.

The available tools for this corpus are a concordance, which gives examples of use in context, a wordlist, N-grams – that is the frequency list of multi-word units –, a thesaurus, which is the function that shows synonyms and similar words for every word, as well as keywords, which show the terminology extraction of one-word and multi-word units. This corpus also shows the Persian collocations and word combinations by grammatical relations and compares collocations of two words.

TalkBank Persian was built by using a huge and various number of blogs. This is helpful in the context of the present study because it sheds light on the real use of words in daily writings and not only in formally edited newspapers. Consequently, this reflects the impact of the terms approved by the APLL in daily writings.

#### 4.4 The TF-IDF Equation and the Adjusted Weight of Terms

To be able to measure the effectiveness of the terms approved by the APLL in contrast to the foreign ones there are two ways. Firstly, and obviously, the frequency of the specified term in the corpus and that of its equivalent could be counted and compared. However, this method might not be very accurate to determine the power – or weight – of the tested term in newspapers and blogs because the term could be used many times in only a few texts thus suggesting strong use when, in reality, it would be poorly distributed. In other words: A simple word count might give the impression of a high frequency of a term that might not be spread far and used strongly in the whole corpus.

Secondly, the weight of the term can be measured by using the document number of the whole corpus, the term frequency [tf], and the document frequency [df]. This equation is known by the name TF-IDF where IDF stands for the inverse document frequency. This method assesses the value of a term in the corpus. In order to calculate the term frequency of the term in the document, the following equation will be used:

$$TF(i, j) = \frac{\text{Term } i \text{ frequency in document } j}{\text{Total words in document } j}$$

When T(i) is the term and j is the document, then the inverse document frequency will be calculated with the following equation:

$$IDF(i) = \log_2 \left( \frac{\text{total documents}}{\text{documents with term } i} \right)$$

The weight W of the chosen term i – be it foreign or its approved Persian equivalent – will be measured by the following mathematical formula:

$$W_{i,j} = tf_{i,j} \times \log_2 \left( \frac{N}{df_i} \right)$$

In this equation,  $tf_{i,j}$  is the number of occurrences of i in j, N is the total number of documents in the corpus, and  $df_i$  is the number of documents that contain the term i.

The formula TF-IDF is used here to weigh the importance of a term in specific texts, corpora, or society. TF-IDF is a mathematical test determining how important and effective a term is to a subject in a set of documents. It is achieved by combining two metrics: first, how

many times a term appears in a text or a document, and second, the term's reciprocal frequency in a text across a series of documents.<sup>41</sup>

TF-IDF helps to understand the distinctive frequency of a term in a text when compared to other texts. The TF-IDF score is obtained by dividing the term frequency by the number of texts where that term occurs. Thus, it stands in contrast to a pure term count or a term frequency relative to the text length and provides a more in-depth understanding of the weight of the terms compared to other terms. The TF-IDF score circumvents the impression that a term is important or relevant just because it occurs frequently, which may be due to the fact that a specific author prefers to use that term while others might choose different terms to express the same idea (Ramos, 2003).

In the present study, the term weight is very often extremely light as will be seen. For some terms, the weight does not exceed the eighth position after the decimal point. This is due to the political nature of the terms in question which, with a few exceptions, are not very common words in everyday language. Consequently, in order to simplify reading and avoid values such as 0.000000123, all weight values will be multiplied by the factor one billion (1,000,000,000), which in the example would produce a weight of 123. This also facilitates comparing all values later in a logarithmic chart, which becomes necessary because some of the weight values are very far apart from each other and could not be displayed in the same chart without losing its informative value. Multiplying all weight values by one billion will ensure that they are all larger than 1 thus making it possible to visualize them on a logarithmic scale (where values smaller than 1 cannot be projected). At the same time, the ratio and relation between all weight values will remain the same.

---

<sup>41</sup> For more information about this formula see Cheng, Yang, Zhao & Gao (2018), Beel, Langer, & Gipp (2017), Rajaraman & Ullman (2011), and Aizawa (2003).

## Chapter 5

# Corpus Analysis

This chapter presents the analysis of the foreign political terms compared to the Persian equivalents approved by the APLL for those terms and their use in online newspapers and blogs with a special focus on their individual weight. Generally, the following three types of term pairs could be expected:

1. Foreign terms that are in wide use and have been either approved as such or continue to dominate despite an approved Persian equivalent term
2. Persian terms that are in use in parallel with foreign terms
3. Persian equivalents that are distinct from the foreign terms and have successfully replaced them

But it is conceivable that there are special cases where this rough categorization into three groups might not be sufficient to explain the success of Persian equivalents or the continued use of foreign terms. This is, for example, possible when one of the two terms – foreign or Persian – is a proper noun that is used in Persian. In addition, as outlined in chapter two, some foreign terms might carry more than one meaning. Therefore, while the use or success as a political term could be limited another meaning of the term might be very widespread. The quantitative corpus analysis in this chapter will take these possibilities into account and be complemented by a text and context analysis as well as a qualitative reading of selected texts to determine the respective meanings of the terms in question.

Methodologically, regarding the context analysis, the terms will be divided into two groups. The first group presents the terms that hold only one connotation or meaning, which is largely unequivocal. In this case, the corpus will be searched for the terms in order to determine their frequency as well as the document frequency which represents the number of documents in which those terms appear. Using the TF-IDF equation outlined in chapter four, the weight of each term will be calculated to evaluate them and quantify the successful terms.

The second group is ambiguous by comparison as each of the terms in this group carries more than one meaning and many connotations. In this case, to begin with, collocations will be identified to help contextualize these terms. A collocation is a sequence or combination of words that occur together more often than would be expected by chance. A collocation could be any striking accumulation of two or more words that more frequently than others occur

together. Such cluster words do not have to be adjacent to each other; they could also be one or several words apart. In this step, collocations are used to determine the meaning or connotation of the selected term in question. In a collocation table, such a term is the node while the words associated with it are called collocates. A special type of collocation would be a compound term that is linked through *ezāfe*, for example, *mağles-e šūrā* where *mağles* is the node and *šūrā* is the collocate.

This chapter looks at selected term pairs of political loanwords and their approved Persian equivalents in three dimensions. Most importantly, thanks to the corpus analysis, it is possible to calculate the weight of each term in blogs and newspapers. This means that in the end each term – foreign or Persianized – will have a weight value that can be compared and from which conclusions can be drawn as to the impact of each term on the language used in blogs (TalkBank) and newspapers (TAP). Since some terms have more than one connotation, the complementary collocation analysis helps approximate the weight value of a term in its political meaning. The analysis of the terms and their weight values against the background of the type of their source is the first factor that will be examined.

This analysis is supported by looking at two auxiliary factors that might influence the success of a given term: the type of borrowing and the time that has passed since its introduction. The first factor is based on Bateni's categorization of linguistic borrowing (see chapter 2.2.3). In the following, the term pair analyses will shed light on the etymology and term history of each term in Persian. This analysis facilitates the categorization of each Persianized term according to its type of borrowing.

The second factor – that is, the time factor – is based on the point in history when the Persianized term was formally approved by the Academy. This can easily be determined thanks to the volume or year when the Academy published the dictionary of approved terms in which that term is first mentioned. However, some terms have a longer history of their own before formal approval by the Academy. Therefore, a reading of Persian dictionaries will complete this look at the time factor.

It is also necessary before beginning the terms analysis to hint at limitations in this part of the study that come with the Persian writing system or Persian orthography. In the context of the present study, especially the spacing system in writing Persian is a factor that needs to be considered. There are two types of spacing in Persian. The first one – which is known as *fāṣele-ye yek harfī* (فاصلهء یک حرفی) – is a regular space that is used between words in a sentence or a

phrase. The second type of spaces – which is known as *nīmḥāsele* (semi space, نیمفاصله) – is usually placed between the parts of a compound term or compound verb. It maintains the shape of the letters in the final-letter shape even in the middle of the word when it is expected to take the middle-letter shape (The Academy of Persian Language and Literature, 1391 [2013/2014], p. 10). Such is the case, for instance, in the following terms that are part of this study: دیوانسالاری (*dīvānsālārī*, bureaucracy), وحشتافکنی (*vaḥšatafkanī*, terrorism), سرمایه‌داری (*sarmāyedārī*, capitalism), and اصلاح‌طلبی (*eṣlāḥṭalabī*, reformism). Some of the most frequent semi-space words are conjugated verbs such as می‌کنم (*mīkonam*, I do), می‌خورم (*mīḥūram*, I eat), and می‌بینم (*mībīnam*, I see).

Not all authors and writers apply these rules and distinguish between the types of spaces in Persian, particularly in the blogs corpus TalkBank, which is based on daily posts by people who are not necessarily authors applying official orthography or working under editorial oversight (see chapters 2.4.6 and 3.1). A term like دیوانسالاری could also be found written as دیوان سالاری or دیوانسالاری, which would change the results of the search. Therefore, in such cases, all possible forms of spelling and findings will be added cumulatively to the final results, as this study concentrates on the use of the terms regardless of their orthographic shape.

Bloggers are in good company, though, when it comes to the lack of knowledge or awareness of the official spelling rules. According to a study on the compliance of authors and users with the APLL rules on writing words continuously, with semi-spaces, or separately, only 23.25 percent of writers follow the APLL guidelines. More than three fourths of all users ignore these rules. What is more, not all authors were even aware of how to insert the semi-space on the keyboard; only 62.2 percent of them know how to do so. The study is based on a survey of two questionnaires each of which consists of twelve items of selected rules among the APLL approved guidelines that was distributed among postgraduate students at Shiraz University as well as a content analysis of the three Persian scientific databases Magiran,<sup>42</sup> Regional Information Center for Science and Technology RICeST<sup>43</sup>, and SID<sup>44</sup> (Ranjbar, Abbaspour, Sotudeh, & Moloodi, 1398 [2019]).

<sup>42</sup> Magiran is a database of Iranian publications. It was launched in 1380 (2000/2001) by the private sector as a platform for introducing and presenting the scientific and specialized press of the country in the nascent information industry of the country and a useful tool for researchers (Darbāreh-e Magiran, 2021).

<sup>43</sup> The RICeST is a science and technology center affiliated to the World Academy of Science and Iran's Ministry of Science: <https://ricest.ac.ir/>.

<sup>44</sup> It is a scientific information database launched by Jihad University and supported by the Iran Center for e-Commerce Development and the Iranian Ministry of Industry, Mine and Trade: <https://www.sid.ir/fa/journal/>.

While the corpora serve as primary sources for the term analysis, the lexical definitions of the terms as well as their etymological and historical origins can be found in dictionaries and encyclopedias that serve as primary sources for the discussion. Three of the most authoritative such resources in Persian are the general dictionaries and encyclopedias of ‘Alī Akbar Dehḥodā, Moḥammad Mo‘īn, and Ḥasan ‘Amīd.

*Loḡatnāme-ye Dehḥodā* was first published in 1939 and has gone through several phases in numerous volumes before being concluded in 2006. Today it also extends to a vast online resource (Dehḥodā, 1377 [1998]). The idea for *Loḡatnāme-ye Dehḥodā* was born at the end of the First World War when ‘Alī Akbar Dehḥodā returned to Tehran and began studying Persian and Arabic texts in the fields of literature and science including, among others, astronomy, jurisprudence, mathematics, medicine, theology, and wisdom. In the 1940s, the Iranian Ministry of Education decided to publish *Loḡatnāme-ye Dehḥodā* based on about four million notes collected by him and his associates. By 1946, the first volume of 486 pages all with the letter Alef was published. After the Second World War, publication continued with funding from the Iranian parliament. One of Dehḥodā’s many assistants would be Moḥammad Mo‘īn, who, in time, would publish his own dictionary (Dehḥodā, 1337[1958], p. 397).

*Farhang-e Mo‘īn* has been compiled in six volumes in numerous editions. It was first published after Mo‘īn’s death in 1972 by Ğa‘far Šahīdī. One of the most important features that distinguishes it is that it contains most of the words found in Persian – whatever their origin, whether they are originally Persian words or foreign words of Arabic or European, Turkish, Mughal, or Indian descent (Mo‘īn, 1386 [2007]a., p. 76).

*Farhang-e ‘Amīd* is a dictionary in three volumes. It includes Persian words and literary and scientific terms including those of Arabic, European, and Turkish origins. It was first published with 936 pages in 1957 and revised and updated with 1,114 pages in 1964. Before *Farhang-e ‘Amīd*, Ḥasan ‘Amīd had already published *Farhang-e Nū* between 1930 and 1955.

Aside from these general dictionaries and encyclopedias, an important source for specifically political terms are the various dictionaries published by the Iranian linguist Dāryūš Āšūrī. He authored two dictionaries that focus on political terms and definitions in addition to numerous other political, philosophical, and literary works.

To check the availability of the equivalents for the terms under study, a list of the terms approved by the Academy which were published in 16 volumes was crosschecked. The list was



sent to the author by Mehnoosh Tehrani, the head of the communication broadcasting office of the Academy of Persian Language and Literature. The volumes are cited as (Terminology group, 1382-1398 [2004-2021]).

Two important online resources for tracing the origins of the foreign terms in European languages are the website and database of the French Centre National de Ressources Textuelles et Lexicales ([www.cnrtl.fr](http://www.cnrtl.fr)) and the English-language online dictionary published by Merriam-Webster ([www.merriam-webster.com](http://www.merriam-webster.com)). These two resources prove very useful with regard to etymology, particularly because many of the foreign political terms entering Persian have French origins and are therefore recorded in the CNRTL database.

Another primary source for tracing the foreign words that are in use in contemporary Persian are old Persian dictionaries. As will be seen, this research supports the assumption that all the loanwords in question were added to the Persian lexicon not only after the French Revolution, which is the origin of many political terms, but even well into the 20<sup>th</sup> century.

For this purpose, it is necessary to probe such works as *Borhān-e qāte* (برهان قاطع), which was compiled before 1062/1651 by Moḥammad Ḥosseyṅ Ḥalaf Tabrīzī alias Borhān (Tabrīzī, 1342 [1963]a., p. 79). None of the selected foreign terms under scrutiny in this study can be found in any of the five volumes of *Borhān-e qāte*. Likewise, in *Farhang-e Ānnanderāğ* authored by Moḥammad Pādešāh in 1888, none of the loanwords are mentioned in any of the volumes of the dictionary (Pādešāh, 1363 [1984]).

*Farhang-e Nafīsī*, to give one more recent example, was authored by ‘Alī Akbar Nafīsī known as Nāẓem Alatebā’. He authored his dictionary until his death in 1924. However, the books were printed and published for the first time in the decade after his death in 1939 by Sa’īd Nafīsī. Although *Farhang-e Nafīsī* was published about 50 years after *Farhang-e Ānnanderāğ*, it did not mention any of the loanwords under study, either (Nafīsī, 1343 [1964]).

By contrast, some of the examined Persian terms of Arabic origin can be found in *Farhang-e Nafīsī*. These include, for example, *dīvān*, *mağles*, and *marām* where specifically the term *marām* is indexed with a different meaning from *ideology* (see chapter 5.9) (Nafīsī, 1355[1977]b., p. 3242). Even though it could be expected that originally Arabic terms are listed in the older Persian dictionaries, too, this is not the case for lemmas or lexemes such as *dīvān*, *eşlāh*, *mağles*, *marām*, *mellat*, *mostabed*, and *ommat*. None of them can be found between the 13<sup>th</sup> and 16<sup>th</sup> centuries CE (Ġārūtī (1356 [1978]), Ṭūsī (1336 [1958]), Naḥğavānī (1355 [1977]),

and Eṣfahānī (1337 [1959])). In the following, each term analysis will take a closer look at the word history and etymology of the terms in question.

## 5.1 The Concept of *Parliament*

The loanword *pārlemān* (parliament, پارلمان) was imported to the Persian dictionary from the French language. It is mentioned in many of the Persian dictionaries, such as in *Loḡatnāme-ye Dehḡodā*, *Farhang-e ‘Amīd*, and *Farhang-e Mo ‘īn*. Dehḡodā gives a very in-depth explanation of the term, dating back to the French Revolution, while others only provide a short lexical definition of the term. Dehḡodā relates the term *pārlemān* to the French supreme court of justice before 1791, the association of dignitaries in the former French royal court, as well as the national assembly emerging during the French Revolution. He also refers to the English House of Commons with its seat in the Houses of Parliament. Dehḡodā thus expands the purely lexical definition of the term by mentioning the history of the term (Dehḡodā, 1339[1960], pp. 31-32).

The English term *parliament* is linguistically a French import (*parlement*). It can be traced to the twelfth century with its first occurrence around the year 1100 where it literally meant a “conversation” between people (from French: *parler*, to speak). By around 1165, it began to refer to a deliberative or consultative assembly and within another one hundred years specifically acquired the meaning of the judiciary authorities in the Kingdom of France (CNRTL, *Parlement*, subst. masc., 2012).

The idea of the term – referring to a consultative or legislative body – goes historically back to antiquity with such institutions, for example, in ancient Greece and Rome and then again during the late European Middle Ages in Spain, England, and France. Dāryūš Āšūrī considers the earliest example of a parliament-like institution to be the Icelandic Althing (Āšūrī, 1376[1997], p. 77).

Dehḡodā restricts himself in his definition of the modern concept of parliamentarism to the feudal French Middle Ages and the abovementioned judiciary *parlements*, which evolved out of the King’s Council (*Conseil du Roi*) under the rule of King Louis IX in late 13<sup>th</sup> century France. In addition to their judiciary role, the *parlements* also enjoyed jurisdiction over matters of taxation, education, and politics including, according to Dehḡodā, such inferred rights as to impeach the government (“ḡaq-e estīzāḡ az dūlat va ḡokūmat”) (Dehḡodā, 1339[1960], p. 32) although it remains unclear what exactly he means by that.

The first and most prominent such court was the Parlement de Paris. Throughout the centuries, the number of *parlements* grew and, by the time of the French Revolution, there were thirteen *parlements* all over France. In 1770, the former president of the Parlement de Paris and then-Chancellor of France René Nicolas de Maupeou<sup>45</sup> put forth his ideas of reforming the *parlements* and establishing the royal judiciary on a new basis. But his reforms met with strong opposition and did not ultimately succeed (Dehḥodā, 1339[1960], p. 32).

Ironically, while Maupeou was determined to support the King and royal authority against the increasingly powerful and assertive *parlement*, his forceful actions may have led to the accelerated downfall of the Ancien Régime and the success of the French Revolution two decades later. Even more so, his idea that the *parlement* was to be reformed in a way that saw the separation of its judicial from its political functions met with the Revolutionaries' approval (Chisholm, 1911) – one step in the development of the *parlement* into modern-day parliaments. However, altogether, Dehḥodā sticks to a narrow linguistic definition of parliament while largely ignoring the political-historical context, which would suggest that, for example, in France, the Estates General rather than the *parlements* held the functions of and developed into the French parliament (*Assemblée Nationale*).

Similarly, Moḥammad Mo'īn refers with the term *pārlemān* to three institutions: first, the association of dignitaries in the former French court, second, the French supreme court of justice before 1791, and third, the council of deputies in the constitutional and republican eras, which had the task of legislating in the country (Mo'īn M. , 1386 [2007]a., p. 655).

The term *maḡles* (parliament, مجلس) is the term approved by the APLL to be used instead of *pārlemān* as an equivalent to *parliament*. It was approved in the first volume of the collection of the approved terms published by the APLL in 2004. The APLL defined it as “the committee of representatives who carry the legislative duty in the country” (Terminology group, 1383 [2004]). The term *maḡles* was approved by the Academy two times. Aside from the aforementioned approval in the field of Political Science and International Relations, it was also approved in the field of Performing Arts to express the English term *tableau vivant* (Terminology group, 1388 [2009]a.).

The Academy has also approved several derivative forms of *maḡles* since 2004. *Maḡlesmeḥvarī* (مجلس محوری) is the approved term by the APLL in the second volume in 2006

---

<sup>45</sup> For a short biographical overview of Maupeou see Chisholm (1911).

as an equivalent to *pārlemāndārī* (پارلماندارى) as stated by (Ášūrī, 1376[1997], p. 77), *pārlemāntārīsm* (پارلمانتاریسم) as stated by (‘Amīd, 1365[1986]a., p. 496), and *pārlemāntārīzm* (پارلمانتاریزم) as mentioned by Ášūrī (Ášūrī, 1364 [1985], p. 54) – all referring to the English term *parliamentarism*. In addition to *mağles*, the approved term contains the morpheme *meħvarī*, which refers to the *pivot* or *axis* of an issue. *Mağlesmeħvarī* therefore suggests that the role of parliament is central to the political system thus making it a loan rendition in Bateni’s system of categorizing foreign terms. The APLL defined it as a constitutional system in which both the legislative and executive powers have their origins in parliament (Terminology group, 1384 [2005]).

In the seventh volume of the collection of the approved terms published by the APLL in 2010, *tahemağlesī* (backbencher, ته‌مجلسی) was approved to express an ordinary member of the British parliament. Backbenchers usually sit in the back row – hence their name – and have no outstanding position in the ruling government or opposition party (Terminology group, 1389 [2010]). The morpheme *tah* is not a loan translation according to Bateni’s categorization because *tah* translates as *bottom*, not as *back*. Instead, one could speak of a loan rendition thus making the whole approved term *tahemağlesī* a combined loan translation and rendition.

*Nezām-e mağlesī* (نظام مجلسی) is another approved term by the APLL which contains the morphemes *nezām* (system, order, regime) and *mağles* (parliament) to form a compound term that represents the democratic parliamentary system, in which the head of the winning party in the parliamentary elections typically obtains executive power and the legislative and executive authorities are not totally divided (Terminology group, 1394 [2016]).

Although the Academy approved the term *mağles* formally only in 2004, it had been used long before. For example, the 1906 constitution of Iran is entirely concerned with the so-called National Consultative Assembly (*mağles-e šūrā-ye mellī*, articles 1ff.) and the Senate (*mağles-e senā*, articles 43ff.) as two new parliamentary chambers to represent the Iranian people vis-à-vis the Shah (Irans 1906 Constitution, 1906) and (Foundation for Iranian Studies, n.d.). It is likely that the Persian constitutionalists adopted the term *mağles* from previous developments in the neighboring Ottoman Empire where the term *mağles* can be traced at least to the 1876 constitution. In its article 42, the Ottoman Constitution of 1876 had established a new legislative and consultative body called the General Assembly (*mağles-e ‘omūmī*), whose name may have served as a model for the Persian constitution thirty years later (The Ottoman Constitution (23 December 1876), n.d.) and (Dostūr (Law, part 4), 1295 [1878]).

*Mağles* is a word coming to Persian (and Ottoman Turkish) from Arabic. It essentially means a place to sit with the prefix *ma-* designating a place and the root consonants *ğ-l-s* (*ğalasa*, to sit). It also refers to a group of people that is designated for consideration of the issues and matters entrusted to them. The word *mağles* has produced many compound words such as Egypt's former lower house of parliament *mağlis al-ša'b* (lit. Council of the People) and the British House of Commons, which translates into Arabic as *mağlis al-'umūm* (Mustafa & El Zayyat, 1960, p. 130). In addition, the Arabic term *mağlis al-šūrā* is used in many countries to refer to one chamber of parliament such as in Egypt before the revolution of 2011 or a consultative assembly that advises the monarch such as in Saudi Arabia.

Given its rich history in Persian and its Arabic lineage, *mağles* can be considered a loanshift according to Bateni's system of categorization. Here, the original meaning of a *place to sit* in the exporting language has shifted and expanded to mean a *place or setting of deliberations*, a *council*, in both the exporting and the receiving language.

The leading question is: Does the term *mağles* feature in Iranian publications as a generic term for any parliament in any country in the world? Or do the authors rather refer to specific parliaments – either in Iran only or in other countries – that use the term *mağles* already for their parliaments as a proper noun?

### 5.1.1 Analysis of the Foreign Term *pārlemān*

It is first necessary to take a closer look at the use of the foreign term *pārlemān* in the two corpora. This helps determine how it is used compared to the approved term *mağles*. In TalkBank, the term *pārlemān* as a stem produces 83,832 results among more than 470 million words whereas in TAP only 1,958 results are found among 30 million words, which is proportionally significantly less than could be expected from the TalkBank count. Consequently, it seems that in newspapers the term *pārlemān* is not as widespread as in blogs. However, when compared to the number of documents in which the term *pārlemān* occurs in both corpora, the difference is not so pronounced. In both corpora, each text mentions the term approximately two times (see below table 5).

These numbers include the term in both nominative and genitive cases. In the following section, the grammatical forms in which the term *pārlemān* appeared are presented in order to list its collocations and whether these collocations show side by side or in the same sentence. As mentioned, the extent of the term's frequency will be given as well, that is, the document

frequency in which the term appears without repeat. This will later facilitate the calculation and comparison of the weight of the foreign term *pārlemān* in specific contexts with the weight of the approved term *mağles*. The below table shows the main forms of *pārlemān* in TalkBank.

Term Forms	Frequency		Document Frequency	
	TalkBank	TAP	TalkBank	TAP
<i>pārlemān</i> پارلمان	83,832	1,958	42,007	861
<i>pārlemānī</i> پارلمانی	57,354	507	31,397	278
<i>pārlemānhā</i> پارلمان‌ها	475	32	395	27

Table 5. Term frequency and document frequency of *pārlemān* and its derivative forms in the two corpora

It is interesting to note that *pārlemān* occurs in the plural form *pārlemānhā* more than 800 times across the two corpora. This could indicate that it is used generically although there is also a chance that the plural form just happens to refer to two or more specific parliaments. Nevertheless, this suggests that the foreign term *pārlemān* is preferred by some authors compared to the approved equivalent one. Table 6 provides several examples of the use of the plural form *pārlemānhā* from the corpus.

<b>Example 1</b>	در سراسر دنیا تنها 4/11 درصد از نمایندگان انتخابی پارلمان‌ها زن بودند که نسبت به سال 1997، 4/2 درصد افزایش یافته است.
Transliteration	Dar sarāsar-e donyā tanhā 4/11 daršad az namāyandegān-e entehābī-ye pārlemānhā zan būdand ke nesbat be sāl-e 1997, 4/2 daršad afzāyeš yāfteh ast.
Translation	Worldwide, only 11.4 percent of elected parliamentarians were women, an increase of 2.4 percent from 1997.
<b>Example 2</b>	دیپلماسی پارلمان‌ها در دهه گذشته نقش بسیار جدی در بخش دیپلماسی داشته‌اند.
Transliteration	Dīplomāsī-ye pārlemānhā dar dahe-ye gozašte naqš-e besyār-e ġeddī dar baḡš-e dīplomāsī dāštehānd.
Translation	Diplomats of parliaments have played a very important role in the field of diplomacy over the past decade.
<b>Example 3</b>	رئیس مجلس شورای اسلامی ادامه داد: البته پارلمان‌ها کار اجرایی ندارند اما می‌توانند در تصویب مقررات برای دولت‌ها نقش داشته باشند و پیرامون قطع رابطه دولت‌ها با رژیم صهیونیستی از وزن خود استفاده کنند و نیز برای راه انداختن کاروان‌های جدیدی برای کمک‌رسانی به غزه فعالیت داشته باشند.
Transliteration	Rayīs-e mağles-e šūrā-ye eslāmī edāmeḥ dād: albatteh pārlemānhā kār eġrāyī nadārand ammā mītavānand dar tašvīb-e moqararāt barā-ye dūlathā naqš dāsteh bāšand va payrāmūn-e qat‘-e rābeteh-ye dūlathā bā režīm-e šehyūnīstī az vazn-e ḡūd estefādeh konand va nīz barā-ye rāh andāḡtan-e kāravānhā-ye ġadīdī barā-ye komakrasānī be ġazze fa‘āleyat dāsteh bāšand.
Translation	The speaker of the Islamic Consultative Assembly continued: Certainly, parliaments do not have executive work, however, they can play a role in approving regulations for governments and using their clout to sever ties between governments and the Zionist regime and also work to launch new convoys to aid Gaza.

<b>Example 4</b>	همکارانی که با فضای حاکم بر پارلمان‌ها آشنا هستند ، به خوبی می‌دانند که در همه جای دنیا و از جمله ایران حضور در پارلمان و گفتگوی رو در رو با نمایندگان تا چه اندازه بر تصمیمات آنها موثر است.
Transliteration	Hamkārānī ke bā faẓā-ye ḥākem bar pārlemānhā ašnā hastand, be ḥūbī mīdānand ke dar hame-ye ġāy-e donyā va az ġomle-ye Īrān ḥoẓūr dar pārlemān va gofteġū-ye rū dar rū bā namāyandegān tā ĉe andāzeh bar tašmīmāt anhā mo'at̄ter ast.
Translation	The colleagues that are familiar with the atmosphere in parliaments are well aware of the extent to which their presence in parliament and life dialogue with representatives around the world, including Iran, influences their decisions.

Table 6. Examples of *pārlemān* in the plural form

*Pārlemān* is connected to numerous collocations in the corpus. These collocations serve to determine the context of the term and the type of content that frequently occurs with the term. The results of the search for frequent collocations with *pārlemān* are divided into two tables. First, table 7 shows collocations of a general nature with words that either precede or follow the search term *pārlemān* along with their frequency in the corpus and the document frequency (followed by selected context examples in table 8). Second, results for *pārlemān* in collocation with specific countries or place names are collected separately in table 9.

Term collocations	Frequency		Document frequency	
	TalkBank	TAP	TalkBank	TAP
<i>mo'āven-e pārlemānī</i> معاون پارلمانی	45,740	172	25,008	86
<i>pārlemānī-ye vezārat</i> پارلمانی وزارت	31,715	4	16,330	4
<i>mo'āven-e pārlemānī-ye vezārat</i> معاون پارلمانی وزارت	31,653	3	16,279	3
<i>rayīs-e pārlemān</i> رییس پارلمان	3,069	11	1,657	8
<i>namāyandegān-e pārlemān</i> نمایندگان پارلمان	1,117	37	863	31
<i>namāyande-ye pārlemān</i> نماینده پارلمان	868	54	671	41
<i>a'żā-ye pārlemān</i> اعضای پارلمان	767	15	663	14
<i>'oẓv-e pārlemān</i> عضو پارلمان	289	12	249	10
<i>vokalā-ye pārlemān</i> وکلای پارلمان	594	0	316	0

Table 7. Selected examples of the most frequent general collocations of the foreign term *pārlemān* in TalkBank corpus

Quantitatively, what stands out in these tables is the frequency of the collocations of the term *pārlemān* as a foreign term compared to its document frequency. As can be seen from tables 5 and 7 above, the average term count per document is between 1.0 and 2.0 with the exception of *pārlemān* in TAP where it is counted 2.27 times on average (see table 5). This means, if an author uses the term *pārlemān*, it is used overall only between one or two times in the text. One explanation could be that the texts are comparatively short – news agency messages perhaps – where the term appears once in the headlines and once in the text body itself. This stands in sharp contrast to the term *maġles* as will be seen in the next chapter.

Qualitatively, what is interesting about the data in table 7 is that the following terms (*'oẓv*, *namāyande*, *vakīl*) come as synonyms to imply the meaning of a member of parliament,

a deputy, a people's representative. The APLL approved *namāyande* for this purpose in the first volume of the approved terms in 2004 (Terminology group, 1383 [2004]). The following table presents various examples of the general collocations.

<b>Example 1</b>	معاون پارلمانی وزارت علوم، نیز در پی تصویب کلیات طرح حذف کنکور، در مجلس گفت: موفقیت طرح کنکور به توانمندی آموزش و پرورش در اجرای طرح بستگی دارد.
Transliteration	Mo'āven-e pārlemānī-ye vezārat 'olūm, nīz darpī tašvīb-e koleyāt-e tarḥ-e ḥazf-e konkūr, dar maḡles goft: movfaqeyat-e tarḥ-e konkūr be tavānmandī-ye āmūzeš va parvareš dar egrāy-e tarḥ bastagī dārad.
Translation	Following the approval of the general plan for eliminating the entrance exam in parliament/in the Majles, the parliamentary deputy of the Ministry of Science said: the success of the entrance exam plan depends on the ability of education to implement the plan.
<b>Example 2</b>	معاون پارلمانی رییس‌جمهور، با بیان اینکه برخی خواص بر مواضع قبلی و اشتباه خود اصرار می‌ورزند، گفت: آنان حتی از به کار بردن کلمه فتنه‌اگراه دارند.
Transliteration	Mo'āven-e pārlemānī-ye rayīs-e ḡomhūr, bā bayān-e īnke barḥī ḥavāš bar mavāze'-e qablī va eštabāh-e ḥūd ešrār mīvarzand, goft: anān ḥatā az be kār bordan-e kaleme-ye fitne ekrāh dārand.
Translation	The parliamentary vice president, stating that some properties insist on their previous and wrong positions, said: They are even reluctant to use the word sedition.
<b>Example 3</b>	رییس پارلمان به طور موقت دولت را در دست گرفت و بعد انتخابات جدید برگزار شد.
Transliteration	Rayīs-e pārlemān be tūr-e mo'aaqqat-e dūlat rā dar dast gereft va ba 'd-e entehābāt-e ḡadīd bargozār šod.
Translation	The president of the parliament temporarily took over the government and then new elections were held.
<b>Example 4</b>	دومین مسالهای که می‌خواهیم پارلمان به آن رسیدگی کند کاهش امتیازهای ویژه اعضای پارلمان است.
Transliteration	Dovvomīn-e mas'āle-ye ke mīḥāhīm pārlemān be ān resīdgī konad kāheš-e emteyāzhā-ye vīže-ye a'zā-ye pārlemān ast.
Translation	The second issue we want parliament to discuss is the reduction of special privileges for members of parliament.

Table 8. Examples of general collocations with *pārlemān*

In example 1 in table 8, *maḡles* and *pārlemān* are used in the same sentence. Both terms are mentioned only 62 times together in TalkBank carrying the same meaning and implementing the concept of parliament. In the example above, *maḡles* refers to the Iranian parliament so that both translations “said in parliament” and “said in the Majles” would be valid. But taking into consideration the size of the corpus and the number of times that both terms are shown in the same text, 62 times does not constitute a significant frequency.

In order to measure the usage of *pārlemān* in the corpus as a foreign term in Persian which is still in use by bloggers and journalists, it is necessary to investigate the other elements that are attached to it in the articles in the corpus. *Pārlemān* appears with a second type of collocations in the corpus, that is, numerous nations and political entities, which will be clarified in table 9.



Collocations		Absolute frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>Orūpā</i>	اروپا	2,748	115	1,899	35
<i>Ītāleyā</i>	ایتالیا	2,313	3	1,205	3
<i>‘Arāg</i>	عراق	1,037	322	581	125
<i>Berūtāniyā</i>	بریتانیا	407	23	364	16
<i>Engelīs</i>	انگلیس	423	36	397	21
<i>Meṣr</i>	مصر	365	16	182	6
<i>Koveyt</i>	کویت	270	24	161	10
<i>Lobnān</i>	لبنان	260	130	148	75
<i>Yūnān</i>	یونان	116	3	87	3
<i>Holand</i>	هلند	113	3	68	3
<i>Esrāyīl</i>	اسرائیل	107	0	91	0
<i>al-ša‘b<sup>46</sup></i>	الشعب	81	0	48	0

Table 9. Examples of collocations with the term *pārlemān*, absolute frequency, and document frequency in the corpora

As can be seen from table 9 above, *pārlemān* comes to denote the parliamentary assembly of several countries and the European Union in general. The coming table 10 gives examples from the context in the corpora of some of these “country collocations”. In addition, the first example in the table proves once more that the use of both *pārlemān* and *mağles* changes according to the name of the country to which it is linked. While Iran’s parliament is referred to by the author as *mağles*, bilateral parliamentary cooperation is described with the adjective *pārlemānī*.

<b>Example 1</b>	رئیس مجلس ایران در این دیدار با بیان اینکه مجلس یازدهم به ابتکار کمیسیون عالی مشترک پارلمانی ایران و روسیه شتاب بیشتری خواهد بخشید، تأکید کرد: مجلس آمادگی دارد همکاریهای پارلمانی گستردها با روسیه داشته باشد.
Transliteration	Ra’īs-e mağles-e Īrān dar īn dīdār bā bayān-e īnke mağles-e yāzdahom be ebtakār-e komīsyūn-e ‘āl-ye moštarak-e pārlemānī-ye Īrān va Rūsye šetāb-e bīštarī ḥāhad baḥšīd, ta’kīd kard: mağles-e āmādegī dārad hamkārīhā-ye pārlemānī-ye gostardehā bā Rūsye dāšte bāšad.
Translation	The Speaker of the Iranian Parliament in this meeting, stating that the 11 <sup>th</sup> Parliament will accelerate the initiative of the Joint Parliamentary High Commission of Iran and Russia, stressed: The Parliament is ready to have extensive parliamentary cooperation with Russia.
<b>Example 2</b>	یک هیئت از پارلمان ایتالیا از منطقه کردستان فدرال دیدن کرد.
Transliteration	Yek hay’at az pārlemān-e Ītālyā az mentaqe-ye Kordestān-e federāl dīdan kard.
Translation	A delegation from the Italian Parliament was seen in the federal Kurdistan region.

<sup>46</sup> *Pārlemān-e al-ša‘b* (پارلمان الشعب) in this context refers to the Egyptian parliament. It seems to be strange that the author chose the term *pārlemān* instead of the term *mağles*, which is already the proper name used in Arabic for the Egyptian parliament (formerly mağlis al-ša‘b, مجلس الشعب) and which is the approved – and one may add common – term that is used in Persian to refer to parliament.

<b>Example 3</b>	یکی از سرکرده‌های گروهک تروریستی منافقین که در رم سرگرم دیدار با نمایندگان پارلمان ایتالیا است، بار دیگر با تصمیم حاکمیتی دولت عراق برای اخراج اعضای فرقه رجوی و بستن اردوگاه عراق جدید (اشرف) مخالفت کرد.
Transliteration	Yekī az sarkardehā-ye grūhek-e terūrīstī-ye monāfeqīn ke dar rom saregarm-e dīdār bā namāyandegān-e pārlemān-e Ītāleyā ast, bār dīgar bā tašmīm-e ḥakemeyatī-ye dūlat-e ‘Araq barāye eḥrāg-e a’zā-ye firqe-ye raḡavī va bastan ordūgāh-e ‘Arāq-e ḡadīd (ašraf) moḥālefāt kard.
Translation	One of the members of the hypocritical terrorist group, in his meeting with one of the members of the Italian parliament in Rome, opposed once again the Iraqi government’s decision to expel members of the Rajavi sect and close the new Iraqi camp (Ashraf).
<b>Example 4</b>	پس از آنکه در سال 1986 اکثریت پارلمان فرانسه به احزاب مدنی و غیردولتی آن کشور تعلق گرفت، بلافاصله شاهد یک ائتلاف بودیم و بدین ترتیب میتران سوسیالیست از شیراک گلیست دعوت به تشکیل کابینه کرد.
Transliteration	Pas az ānke dar sāl-e 1986 akṭareyyat-e pārlemān-e Frānse be aḥzāb-e madanī va ḡayre-dūlatī in kešvar ta‘aloq gereft, belāfāšele šāhed-e yek ītelāf būdīm va bedīn tartīb Mīterān-e sūsyālīst az Šīrāk-e golīst da‘vat be taškīl-e kābīne kard.
Translation	In 1986, after the majority of the French parliament members joined the country’s civil and non-governmental parties, we immediately witnessed forming a coalition, and thus the Socialist Mitterrand invited the Gaullist Chirac to form a government.
<b>Example 5</b>	حسین سبحانی‌نیا ، در واکنش به تصویب طرح تعطیلی سفارت ایران در منامه از سوی پارلمان بحرین، گفت: در صورتی که بحرین ، این طرح را اجرایی کند ، بدون شک عکس‌العمل حداقلی ایران ، مقابله به مثل خواهد بود.
Transliteration	Ḥoseyn Sobḥānīnyā, dar vākoneš be tašvīb-e tarḥ-e ta‘īlī-ye sefārat-e Īrān dar Manāme az sūye pārlemān-e Baḥrayn goft: dar šūrātī ke Baḥrayn īn tarḥ rā eḡrāyī konad, bedūn šak ‘aks el-‘amal-e ḥadāqālī Īrān, moqābale be meṭl ḥāhad būd.
Translation	Hossein Sobhaninia, in response to the Bahraini parliament’s approval of the plan to close the Iranian embassy in Manama, said: If Bahrain implements this plan, there is no doubt that the minimal reaction of Iran will be reciprocal action

Table 10. Examples of country names collocations with *pārlemān*

The term *pārlemān* is also mentioned in the TalkBank corpus 1,511 times as *ḥabarnegār-e pārlemānī* (parliamentary reporter, خبرنگار پارلمانی) such as *ḥabarnegār-e pārlemānī-ye ḥabargozārī-ye fāres* (خبرنگار پارلمانی خبرگزاری فارس), *ḥabarnegār-e pārlemānī-ye ḥabargozārī-ye dānešgūyān-e Īrān (ISNA)* (خبرنگار پارلمانی خبرگزاری دانشجویان ایران), *ḥabarnegār-e pārlemānī-ye ĪRNĀ* (خبرنگار پارلمانی ایرنا), and *ḥabarnegār-e pārlemānī-ye Mehr* (خبرنگار مهر پارلمانی). *Pārlemānnyūz* (پارلمان‌نیوز) is also one of the findings in the corpus that refers in these contexts to the Islamic Consultative Assembly’s news agency. The following examples show selected sentences of the concordance of these compound terms (proper nouns) in the corpus.

<b>Example 1</b>	با این حال شنیده‌های خبرنگار پارلمانی « وطن امروز » حاکی از رصلاحت شدن 32 تا 40 نفر از نمایندگان فعلی مجلس شورای اسلامی است
Transliteration	Bā īn ḥāl šenīdehā-ye ḥabarnegār-e pārlemānī “Vaṭan-e Emrūz” ḥākī az radšalāḥeyyat šodan-e 32 tā 40 nafar az namāyandegān fe‘lī-ye maḡles-e šūrā-ye eslāmī ast.
Translation	However, a parliamentary reporter of <i>vaṭan-e Emrūz</i> presented hearsay of the disqualification of 32 to 40 members of the current parliament.

<b>Example 2</b>	به گزارش خبرنگار پارلمانی خبرگزاری فارس، غلامعلی حداد عادل مشاور عالی مقام معظم رهبری و رییس کمیسیون فرهنگی مجلس شورای اسلامی امسال سخنران مراسم 12 بهمن است
Transliteration	Be gozāreš-e ḥabarnegār-e pārlemānī-ye ḥabargozārī-ye Fārs, Golām ‘Alī Hadād ‘Ādel mošāver-e ‘ālī-ye maqām-e mo‘azzam-e rahbarī va rayīs-e komīsyūn-e farhangī-ye maḡles-e šūrā-ye eslāmī emsāl soḥanrān-e marāsem-e 12 bahman ast.
Translation	According to the parliamentary correspondent of Fars news agency, Gholam-Ali Haddad Adel, the high Advisor to the Supreme Leader and the Chairman of the Cultural Commission of the Islamic Consultative Assembly, will be the speaker of the ceremony of 12 Bahman this year.
<b>Example 3</b>	علی مطهری در گفتگو با خبرنگار پارلمانی مهر، با اعلام این خبر اظهار داشت: لیست منتقدان دولت دهم در دو یا سه روز آینده منتشر خواهد شد
Transliteration	‘Alī Moṭaharī dar goftegū bā ḥabarnegār-e pārlemānī-ye Mehr, bā e ‘lām-e īn ḥabar ezhār dāšt: līst-e montaqedān-e dūlat-e dahom dar dū yā seh rūz-e āyande montašer ḥāhad šod.
Translation	Ali Motahari, in an interview with Mehr parliamentary reporter, stated that a list of government critics will be released in the next two or three days.
<b>Example 4</b>	محمد کریم شهبزاد روز چهارشنبه در حاشیه جلسه علنی مجلس شورای اسلامی، در گفتگو با خبرنگار پارلمانی ایرنا در واکنش به تذکر علی مطهری گفت: مردم به کمبصیرتی این افراد آگاهی دارند
Transliteration	Moḥamad Karīm Šahrazād rūz-e čahāršanbe dar ḥāšī-ye ḡalase-ye ‘alanī-ye maḡles-e šūrā-ye eslāmī, dar goftegū bā ḥabarnegār-e pārlemānī Īrnā dar vākoneš be tazzakor-e ‘Alī Moṭaharī goft: mardom be kambašīratī-ye īn afrād agāhī dārand.
Translation	Mohammed Karim Shahrazad announced, on the sideline of the public session of the parliament of Iran, in an interview with IRNA parliamentary correspondent in response to Ali Motahari’s remark that: “People are aware of the insight of these individuals.”

Table 11. Examples of the use of proper nouns in collocation with the term *pārlemān*

Finally, the term *pārlemān* exists in the corpus followed by *Īrān* referring to the Iranian Parliament (پارلمان ایران). This compound term produces only 185 hits. Considering the size of the corpus, this number of hits is negligible. A few examples are nevertheless shown in the following table 12.

<b>Example 1</b>	اینها دلایلی است که پارلمان ایران – مجلس شورا و سنا – را بر آن داشت که به اتفاق آرا رای به ملی کردن صنعت نفت بدهد.
Transliteration	Īnhā dalāyelī ast ke pārlemān-e Īrān – maḡles-e šūrā va senā – rā bar ān dāšt ke be etefāq-e arā rā’y be mellī kardan šan‘at-e naft bedahad.
Translation	These are the reasons that prompted the Iranian parliament – the consultative assembly and senate – to vote unanimously to nationalize the oil industry.
<b>Example 2</b>	مجلس ششم در طول تاریخ پارلمان ایران اگر نخواهیم بگوییم بی نظیر، باید تاکید کنیم کم نظیر بوده است.
Transliteration	Maḡles-e šešom dar tūl-e tāriḡ-e pārlemān-e Īrān agar naḥāhīm begūyīm bīnazīr, bāyad tākīd konīm kamnazīr būde-ast.
Translation	The sixth Majles throughout the history of the Iranian parliament, if we don’t admit that it is unique, we have to emphasize that it has been rare.
<b>Example 3</b>	وی به تشکیل کمیته حمایت از انتفاضه فلسطین در مجلس اشاره کرد و آن را مصداق توجه خاص پارلمان ایران به مساله فلسطین خواند.
Transliteration	Vey be taškīl-e komīte-ye ḥemāyat az entefāze-ye Feleštīn dar maḡles ešāre kard va ān rā mešdāq-e tavaḡoh-e ḥāš-e pārlemān-e Īrān be mas’āle-ye Feleštīn ḥānad.
Translation	He referred to the formation of a committee in support of the Palestinian Intifada in the parliament and called it an example of the Iranian parliament’s special attention to the Palestinian issue.

<b>Example 4</b>	پارلمان ایران قانونی را تصویب کرد که روابط دیپلماتیک با انگلیس را به سطح کاردار کاهش می‌دهد و اخراج سفیر انگلیس را در مهلتی دو هفته‌ای درخواست می‌کند.
Transliteration	Pārlemān-e Īrān-e qanūnī rā taṣvīb kard ke ravābeṭ-e dīplomātīk bā Enḡlīs rā be saṭḥ-e kārdār kāheš mīdahad, va eḥrāḡ-e safīr-e Enḡlīs rā dar mohlatī-ye dū hafte dar ḥāst mīkonad.
Translation	The Iranian parliament has passed a law reducing the diplomatic relations with Britain to a lower level and called for the British ambassador to be expelled within two weeks.
<b>Example 5</b>	در چهارشنبه آخر سال 1390 محمود احمدی نژاد برای اولین بار در پارلمان ایران بعنوان مسؤل اول اجرایی کشور، مخاطب سوال تعدادی از نمایندگان مردم قرار گرفت.
Transliteration	Dar čahāršanbe āḡar-e sāl-e 1390, Maḥmūd Aḡmadī Nežād barāye avvalīn bār dar pārlemān-e Īrān be 'onvān-e mas'ūl-e avval-e eḡrāyī-ye kešvar, moḡḥāṭeb-e so'āl-e te'dādī az namāyandegān-e mardom qarār gereft.
Translation	On the last Wednesday of 1390, Mahmoud Ahmadinejad was addressed by a number of people's representatives in the Iranian parliament for the first time as the country's first executive leader.

Table 12. Examples of the foreign term *pārlemān* in use to indicate the Iranian parliament

The weight of the term will be calculated using the TF-IDF equation (see chapter 4.4) as explained before in order to compare it later with the weight of the approved term *maḡles*. *Pārlemān* as a foreign term is seen 83,832 times in 42,007 documents in TalkBank which makes a TF of 1.76573E-07 (i.e., 83,832 divided by 474,773,547) and the inverse document frequency 1.41525558 (i.e.,  $\log(1,092,892/42,007)$ ). Therefore, the weight of that term in the corpus is 2.49895E-07 (i.e., TF multiplied by IDF), which is adjusted to 250 (i.e., multiplied by one billion) for the purpose of this study. The weight of *pārlemān* in TAP, on the other hand, is significantly higher with a value of 171,874. Clearly, its use in newspapers is more pronounced than in blogs.

### 5.1.2 Analysis of the Approved Term *maḡles*

The term *maḡles* produces 480,465 hits in TalkBank in addition to 34,166 hits in TAP. However, one of the most important reasons for this increased occurrence is the average repetition of the term in the same text more than three times in TalkBank and even more than five times in TAP with the same meaning. The number of hits for the term without repetition, that is the document frequency, is 141,879 in TalkBank and 6,659 in TAP. The main forms of the term in addition to its frequency and the document frequency are given in the following table 13.

Term forms	Frequency		Document frequency	
	TalkBank	TAP	TalkBank	TAP
<i>maḡles</i> مجلس	480,465	34,166	141,879	6,659
<i>maḡlesī</i> مجلسی	10,590	444	7,259	310

Table 13. Term frequency and document frequency of *maḡles* and *maḡlesī* in both corpora

What is complex about the term *mağles* is that it covers a wide range of meanings and connotations in Persian and is not only restricted to the meaning of *parliament* as is the case with *pārlemān*. This makes the process of identifying the elements that carry the meaning of *parliament* difficult. However, this is the prerequisite for comparing it later with the weight of the foreign term *pārlemān* and so that the more common term with regard to this meaning could be determined. Therefore, in this section, the use of the term *mağles* will be analyzed quantitatively and qualitatively in general and with the meaning of *parliament* in particular – whether it be the Iranian Islamic Consultative Assembly or otherwise, in addition to other Iranian institutions with the name *mağles* such as *mağles-e ħobregān* (Assembly of Experts).

As mentioned, *mağles* could mean meeting place, association, forum, and assembly for consultation, negotiation, and conversation. It also refers to an assembly for discussing the general affairs of the people and consulting what is beneficial for them. Furthermore, it means a meeting place of hospitality or a place for people to party. It could in this sense also carry the meaning of dancing party, mourning assembly, preaching ceremony, ritual gathering, sermon assembly, or board of grievances (Dehğodā, 1352[1974], pp. 461-464).

For there to be a fair comparison between the foreign term and the approved one, both of them must be compared with each other in the same semantic field meaning that both terms must denote a legislative body. Therefore, the official legislative bodies in Iran after the Constitutional Revolution in 1906 and in the current Islamic Republic of Iran since 1979 will be introduced.

The first parliament in Iran was formed after the Constitutional Revolution by a decree of Muzaffar al-Dīn Shah. Unlike today's parliament of Iran, the first parliament was bicameral: Alongside the establishment of the national consultative assembly (*mağles-e šūrā-ye mellī*), a second chamber was created called senate (*mağles-e senā*) (Dehğodā, 1352[1974], pp. 464-465).

From the issuance of the Constitutional Decree in 1906 to the Islamic Revolution in 1979, the national consultative assembly convened for a total of 24 terms although the country was at times left without an assembly such as in times of turmoil. After the Iranian Revolution, the legislature formed again with the name of the parliament remaining *mağles-e šūrā-ye mellī* (National Consultative Assembly) according to the constitution of the Islamic Republic of Iran. However, the members of the first parliament renamed it *mağles-e šūrā-ye eslāmī* (مجلس شورای اسلامی).

اسلامی, Islamic Consultative Assembly). This name was formally adopted in the revision of the constitution in 1989 and remains the official name of Iran's unicameral parliament today.<sup>47</sup>

Even if not part of a compound word, the term *mağles* itself might still refer to a parliament despite its other meanings. For example, within one news article or corpus document, after introducing the full compound word for the first time, the writer might choose to mention it in a short version as *mağles* only. The use of the TF-IDF equation turns out to be very helpful here because it adjusts the weight of the (repeated) term by the number of its appearances within the same text or document.

In order to identify the use of the term *mağles*, the corpus is searched for high-frequency compound terms that relate to it. This leads to the results of the following proper nouns: *mağles-e ħobregān* (assembly of experts, مجلس خبرگان), *mağles-e šūrā* (مجلس شورا), *mağles-e šūrā-ye eslāmī* (مجلس شورای اسلامی), *mağles-e šūrā-ye mellī* (مجلس شورای ملی), and *mağles-e senā* (مجلس سنا). The below table illustrates the most frequent compound words with *mağles* in both corpora.

	Frequency		Document frequency	
	TalkBank	TAP	TalkBank	TAP
<i>mağles-e šūrā</i>	51,324	5,155	28,715	2,853
<i>mağles-e šūrā-ye eslāmī</i>	48,494	4,713	27,217	2,673
<i>mağles-e šūrā-ye negahbān</i>	24,385	2,827	10,398	765
<i>mağles-e ħobregān</i>	13,293	360	6,348	220
<i>mağles-e šūrā-ye mellī</i>	2,875	98	2,094	56

Table 14. High-frequency compound terms with *mağles*

In order to determine the context in which the term *mağles* occurs most, table 15 below shows the high-frequency collocations with *mağles* (nouns or verbs).

<sup>47</sup> For further information about the national assembly and the senate of Iran from 1906 until the Islamic Revolution see Nāteqiyān Far & Nāteqiyān Farī (1389 [2010]).

Collocations	Frequency (number of hits)		Document frequency	
	TalkBank	TAP	TalkBank	TAP
<i>namāyande</i> نماینده	136,536	5,078	59,731	2,297
<i>dūlat</i> دولت	82,551	5,654	41,480	1,840
<i>rayīs</i> رئیس	74,336	1,305	29,478	605
<i>šūrā-ye eslāmī</i> شورای اسلامی	63,855	4,713	27,493	2,853
<i>šūrā</i> شورا	51,860	5,155	28,950	3,129
<i>enteḥābāt</i> انتخابات	44,237	5,489	17,631	2,421
<i>qānūn</i> قانون	42,799	2,932	16,775	1,115
<i>‘oḡv</i> عضو	39,277	2,330	17,955	1,022
<i>komīsyūn</i> کمیسیون	39,126	3,522	15,445	1,272
<i>vezārat</i> وزارت	31,836	1,305	11,697	634
<i>reyāsat</i> ریاست	27,855	1,256	18,113	516
<i>taṣvīb</i> تصویب	27,395	2,704	13,258	1,603
<i>rahbarī</i> رهبری	22,462	699	12,647	366
<i>ḡalase</i> جلسه	21,872	1,568	10,981	646
<i>e ‘temād</i> اعتماد	19,546	373	14,856	203
<i>būdḡe</i> بودجه	19,358	1,849	9,290	555
<i>dūre</i> دوره	17,632	1,643	8,360	712
<i>lāyḡe</i> لایحه	17,533	1,054	8,267	358
<i>ḡobre(gān)</i> خبرگان	15,900	623	6,652	256
<i>mo ‘araft</i> معرفی	15,196	427	9,724	202
<i>kardan/šodan</i> کردن/شدن				
<i>hay ‘āt</i> هیات	12,499	934	5,548	283
<i>maṭraḡ</i> مطرح	11,955	1,235	6,613	667
<i>frākseyūn</i> فراكسيون	11,450	608	4,555	267
<i>hay ‘at</i> هیئت	10,737	515	3,904	194
<i>dastūr</i> دستور	7,935	725	4,520	388
<i>nāyeb</i> نایب	6,666	300	3,193	161
<i>oṣūlḡarā</i> اصولگرا	5,725	178	2,818	99
<i>āyīnnāme</i> آیین نامه	2,989	120	1,154	45
<i>šūrā-ye mellī</i> شورای ملی	2,358	137	1,250	65

Table 15. Collocations of the term *maḡles* in the corpora, term frequency and document frequency

The collocations in the previous table can be categorized into four distinct groups. The first group covers the names given to different institutions, whereas the second group contains collocations that describe functional units *within* those institutions. Such functional units refer to individuals, groups, and organizational bodies. The third group is concerned with procedural aspects of the parliamentary system. Finally, although comprising only two collocations, it is worth distinguishing the fourth group from the previous three groups because they are of an antagonistic nature.

The majority of the collocations in the first group belongs to the constitutional framework of the Islamic Republic of Iran with terms such as *šūrā* and *šūrā-ye eslāmī* with a combined total of more than 125,000 hits in the corpora. The collocations *rahbarī* and *ḡobre(gān)* rank second and third with a total of more than 23,000 and 16,000 hits respectively. Of the roughly 51,000 collocations with *maḡles-e šūrā* in TalkBank 48,494 are identical with

*mağles-e šūrā-ye eslāmī*. Similarly, of the roughly 5,100 collocations with *mağles-e šūrā* in TAP 4,713 are identical with *mağles-e šūrā-ye eslāmī*. This leaves almost 3,000 hits in both corpora where *mağles-e šūrā* is not used with the collocation *eslāmī*. While it is likely that many or even most of them still refer to the Iranian parliament by way of an abbreviated use of the term that is synonymous with *mağles-e šūrā-ye eslāmī*, it is conceivable that *mağles-e šūrā* means legislative and/or consultative assemblies in other countries.

Several countries have parliament-like institutions that are called *mağles-e šūrā* by name. This is, for example, the case in Bahrain (*mağles-e šūrā-ye Baḥrayn*), Oman (*mağles-e šūrā-ye ‘Omān*), Qatar (*mağles-e šūrā-ye Qaṭar*), and Saudi Arabia (*mağles-e šūrā-ye ‘Arabestān/ ‘Arabestān-e so ‘ūdī*). However, the use of these collocations is extremely limited with only eight, four, four, and 17 hits respectively.

Among this first group, one also finds the collocation *šūrā-ye mellī* which does not necessarily refer to the (former) Iranian parliament. Even though *mağles-e šūrā-ye mellī* was the name given to Iran’s parliament from 1906 till it was abolished by the Islamic Revolution in 1979, it does not mean that the term is no longer in use. To the contrary, this is, for example, the case in Afghanistan where the – now *de facto* defunct – national parliament was called *šūrā-ye mellī dūlat-e ġomhūrī-ye eslāmī-ye ‘Afgānestān* (شورای ملی دولت جمهوری اسلامی افغانستان) before the Taliban retook control of the country in August 2021 (National Assembly of IR of the Afghanistan, n.d.)

In example 13 of table 17, it can be seen that the text uses *mağles-e šūrā-ye mellī* referring most likely to the two houses of the Afghan parliament although there is a chance that the two “parliaments” that it refers to might mean the U.S. Congress and the Afghan parliament. Still, this is an obvious example of the use of *mağles* outside the Iranian context. If referring to the Afghan parliament, the writers applied with *mağles* a Persian term to the institution that – in its native Dari – is not referred to as *mağles*, but instead as *šūrā-ye mellī dūlat-e ġomhūrī-ye eslāmī-ye Afgānestān* (شورای ملی دولت جمهوری اسلامی افغانستان) (National Assembly of IR of the Afghanistan, n.d.) – not to speak of the possibility that the text might include the U.S. Congress. The number of hits for *mağles-e Afgānestān* in the two corpora is 787. The compound word *mağles-e Āmrīkā* occurs even 1,481 times in the two corpora (see table 16).

By the way, the compound word *mağles-e moḥtaram* (مجلس محترم) is also used to refer to the Iranian *mağles-e šūrā -ye eslāmī*. It produces 413 results in the corpus.



The second group of collocations can be characterized as functional on different levels of parliamentary representation. Here one finds the individual member of parliament such as *namāyande* (member) and *nāyeb* (vice-) with more than 141,000 and close to 7,000 hits respectively, *rayīs* (president) and *reyāsat* (presidency) with more than 75,000 and 29,000 hits respectively, as well as *‘ozv* (member) with more than 41,000 hits.

Regarding *rayīs*, a more nuanced analysis is necessary. Of the 75,000 hits for *rayīs*, roughly 23,000 occur as *rayīs-e mağles* thus clearly pointing at the speaker of parliament. The compound terms *rayīs-e mağles-e vozarā* and *rayīs-e mağles-e vazīrān*, which would lead away from the legislative to the executive branch, occurs only three times, all of them in TalkBank. Only about 15,000 occur as part of the compound word *rayīs-e ġomhūrī*. These hits have to be added to the fourth group below. The remaining 37,000 hits cannot easily be subsumed in one or the other group.

Furthermore, the second group contains organizational bodies such as *komīsyūn* (committee) with more than 42,000 hits and *frākseyūn* (parliamentary group) with more than 12,000 hits in addition to *hay’āt/hay’at* (board, commission, panel) with more than 11,000 hits each. Finally, it contains the collocate *oşūlgarā* (fundamentalist) with almost 6,000 hits featuring as an ideological subgroup of *frākseyūn*.

The third group deals with procedural aspects of parliaments. It contains, of course, those responsibilities that are in the original domain of parliaments, that is, law making and drafting the state budget. These two responsibilities occur in the collocations of *qānūn* (law), *lāyhe* (act, bill), and *būdġe* (budget), which yield more than 45,000, 18,000, and 21,000 results respectively. Moreover, collocations regarding the legislative period can also be found in this third group. These are, for instance, *enteġbāt* (elections), *dūre* (period), and *ġalase* (session). *Enteġbāt* alone occurs almost 50,000 times in the corpus followed by *ġalase* with over 23,000 and *dūre* with more than 19,000 hits. Finally, the third group also covers parliamentary tasks as reflected in the words *ra’y e ‘temād* (*gereftan*) (cast a vote of confidence), *mo ‘arafī* (*kardan*) (announce, present, introduce), and *maṭraḥ* (*kardan*) (propose), as well as in the word *taşvīb* (ratification, confirmation, sanction). These collocates appear almost 20,000 times, more than 15,500 times, more than 13,000 times, and roughly 30,000 times respectively.

Collocations in the fourth and last group, can be summarized as belonging to the executive branch. In this sense, they are in an antagonistic relationship with the legislative branch of parliament. The three words in this group are *dūlat* (government), *vezārat* (ministry),

and *rahbarī* (pertaining to the Supreme Leadership Authority) with over 88,000 hits, over 33,000 hits, and over 23,000 hits respectively. As mentioned above in the first group, at least 15,000 hits for *rayīs-e ġomhūrī* would also have to be included in the fourth group.

When it comes to indicating the parliaments of other nations referred to by the name of *maġles*, the following table 16 presents the frequency and the document frequency of the term's country collocations.

Collocations	Absolute frequency		Document frequency	
	TalkBank	TAP	TalkBank	TAP
<i>Turkiye</i> ترکیه	5,166	34	4,695	26
<i>Ámríkā</i> آمریکا	1,243	238	920	166
<i>‘Arāq</i> عراق	981	51	637	29
<i>Afġānestān</i> افغانستان	776	11	545	7
<i>Frānse</i> فرانسه	320	9	306	9

Table 16. Examples of the most frequent compound terms of the term *maġles* with country names, absolute frequency, and document frequency

As the previous table 16 reveals, *maġles* is used 5,200 times in collocation with *Turkiye* to represent the Grand National Assembly of Turkey. In Persian, the Turkish parliament is formally referred to as *maġles-e mellī-ye kabīr-e Turkiye* or just briefly as *maġles-e Turkiye* (cf. e.g. (TRT, 2021)), which is a direct translation from the name given to the assembly in Turkish itself: Türkiye Büyük Millet Meclisi (Türkiye Büyük Millet Meclisi (The Grand National Assembly of Turkey), n.d.). This formal full version of the name of the Turkish parliament appears only five times in the two corpora, all of them in TalkBank.

The second most results for country collocations of *maġles* relate to the United States. *Maġles* is used to refer to both houses of the U.S. Congress: the U.S. House of Representatives and the U.S. Senate, where the frequency of the first, *maġles-e namāyandegān-e Ámríkā*, reaches 748 and the second, *maġles-e senā-ye ‘amríkā*, reaches 1,129 hits. *Senā-ye Ámríkā* alone has a frequency of 2,557 in the corpora which indicates that *maġles* is not an integral part of this compound. The fact that the upper chamber of the first Iranian parliament was called *maġles-e senā* probably helps to make the connection between *maġles* and any parliamentary body called Senate anywhere in the world thus also supporting the attempt of the Farhangestān to establish *maġles* as the Persian equivalent to *pārlemān*. The compound term *maġles-e senā* is used as well to refer to the senate or upper house of the bicameral parliaments of the following countries: Afghanistan, Belgium, Canada, France, Italy, Kazakhstan, Pakistan, and the Netherlands.

The term *mağles* is also used with Iraq hinting at two very different political bodies. The first body simply means the Iraqi parliament calling it *mağles-e 'Arāq* with 256 hits. In Iraq, however, there is also a political party called Islamic Supreme Council of Iraq (*mağles-e a 'lā-ye eslāmī-ye 'Arāq*) with 756 hits in the corpora. From an Iranian perspective, this party might be given due credit because it was created by Shiite Iraqi exiles in Iran in the early 1980s with the help of the new Iranian government and with the aim to overthrow the regime of Saddam Hussein finally entering Iraqi politics as a political party after his downfall in 2003 (Supreme Council for Islamic Revolution in Iraq (SCIRI), 2015).

In conclusion, the approved term *mağles* as a general term is the most frequent term of those examined in this study and has 480,465 hits in the corpus in 141,879 documents. The TF is 0.001011988 and the inverse document frequency IDF is 0.886659128. Consequently, the adjusted weight of *mağles* is 897,288 in TalkBank, which amounts to one of the highest weight values for any of the selected terms in this thesis. After making the same calculation for TAP, the adjusted weight of *mağles* in the newspapers-based corpus is 666,804.

But since not all the connotations of this term refer to Parliament, it is not accurate to compare its weight to the weight of the foreign term *pārlemān*. Therefore, in order to approximate the real weight of the term, *pārlemān* will be compared to the compound words *mağles-e šūrā* and *mağles-e ĥobregān*, that reflect the current Iranian political system, as well as the compound words *mağles-e senā* and *mağles-e namāyandegān* to reflect the use of the term outside the Iranian context. In addition, other forms of *mağles* with the meaning of *parliament* that are used to calculate the adjusted weight of *mağles* are *mağles-e moħtaram*, which is exclusively used for the Iranian parliament, as well as the country collocations from table 16 above. The cumulative adjusted weight of *mağles*, where *mağles* definitely has the meaning of *parliament*, would therefore be 233,333 in TalkBank and 251,503 in TAP. Of course, this approximation excludes mentions of the term *mağles* where it might mean *parliament* but does not occur in collocation with any of the above signifiers.

The following table presents various examples with the term *mağles* from both corpora. The examples were selected to reflect the variety of how the term *mağles* is used in different contexts. It is used with Iran, Afghanistan, Iraq, and sometimes with European countries. However, the dominant theme links to Iran.

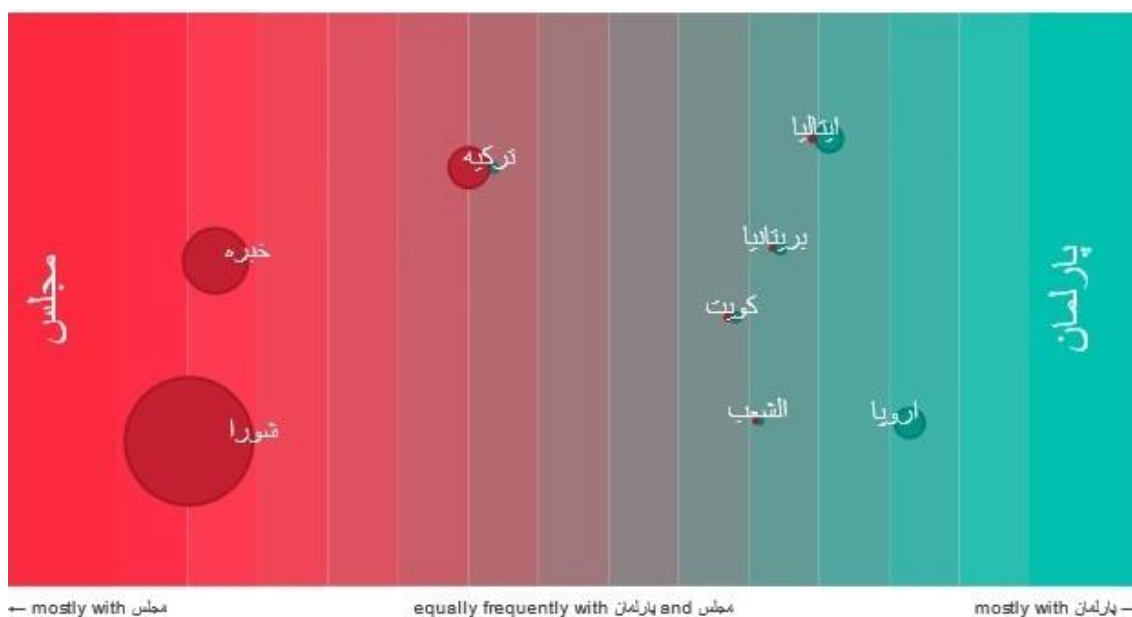
<b>Example 1</b>	لوايح محاکمه وزرای سابق و انحلال ساواک ، در جلسه مجلس شورا به تصویب رسید.
Transliteration	Lavāyḥ-e mohākeme-ye vozarā-ye sābeq va enḥelāl-e SAVAK, dar ḡalase-ye maḡles-e šūrā be taṣvīb rašid.
Translation	The court regulations of the former ministers and the dissolution of SĀVĀK were approved in the session of the Consultative Assembly.
<b>Example 2</b>	نایب رئیس مجلس شورای اسلامی اظهار داشت: نگاه مجلس این است که تربیون ان نباید در عرصه انتخابات دهمین دوره ریاست جمهوری به عرصه دفاع یا تخطیه نامزدها تبدیل شود.
Transliteration	Nāyeb-e rayīs-e maḡles-e šūrā-ye eslāmī ezhār dāšt: negāh-e maḡles īn ast ke: trībūn-e ān nabāyad dar ‘arše-ye enteḥābāt-e dahomīn-e dūre-ye reyāsāt-e ḡomhūrī be ‘arše-ye defā’ yā taḥṭīḥ-e nāmzadhā tabdīl šavad.
Translation	The Vice-President of the Islamic Consultative Assembly said: The opinion of the Assembly is that its dais should not become an arena for defense or mistreatment of candidates in the arena of the tenth presidential elections.
<b>Example 3</b>	ایت الله هاشمی رفسنجانی رئیس مجلس خبرگان رهبری به خیل عظیم راهپیمایان روز جهانی قدس پیوست.
Transliteration	Āyātollāh Hāšemī Rafsanjānī, rayīs-e maḡles-e ḥobregān-e rahbarī be ḥīl-e ‘azīm-e rāhpīmāyān-e rūz-e ḡahānī-ye qods payvast.
Translation	Ayatollah Hashemi Rafsanjani, chairman of the Assembly of Experts, joined the large number of marchers on international Quds Day.
<b>Example 4</b>	شنیده‌ها حکایت از تأیید صلاحیت آقای جلالیان کاندیدای هشتمین دوره انتخابات مجلس شورای اسلامی در حوزه انتخابیه کنگان دیر و جم دارد.
Transliteration	Šenīdehā-ye ḥekāyat az tāyīd-e ṣalāḥeyyat-e āqā-ye Galālyān kāndīdā-ye haštomīn-e dūre-ye enteḥābāt-e maḡles-e šūrā-ye eslāmī dar ḥūze-ye enteḥābīḥ-e Kangān Dīr va Ğam dārad.
Translation	There are rumors that Mr. Jalalian has been confirmed as the candidate for the eighth term of the Islamic Consultative Assembly in the Kangan Dir and Jam constituencies.
<b>Example 5</b>	عضو فراکسیون اکثریت مجلس از تلاش جدی دولت و مجلس برای رفع موانع در رشد و توسعه اقتصادی مناطق محروم خیر داد و گفت: مجلس و دولت برای کمک به شکوفایی اقتصادی مناطق مختلف کشور به‌ویژه شهرهای مرزی و محروم عزم جدی دارند.
Transliteration	‘Ozv-e frakseyūn-e aktāreyat-e maḡles az talāš-e ḡeddī-e dūlat va maḡles baray raf‘-e mavāne’ dar rušd va tūse‘eh-e eqtešādī-ye manāteq-e maḥrūm ḥabar dād va goft: maḡles va dūlat baray komak be šekūfāyī-ye eqtešādī-ye manāteq-e moḥtalef-e kešvar be vīže šāhrehā-ye marzī va maḥrūm ‘azm ḡaddī dārand.
Translation	The member of the majority faction of the parliament announced the serious efforts of the government and the parliament to remove obstacles to the economic growth and the development of deprived areas and said: The parliament and the government are determined to help the economic prosperity of different regions of the country, especially border and deprived cities.
<b>Example 6</b>	خبرگزاری فارس: رهبر معظم انقلاب انتخاب ایت‌الله مهدوی کنی به ریاست مجلس خبرگان را انتخابی به‌جا و به‌حق خواندند و افزودند: ایشان از ابتدای انقلاب تا امروز در همه عرصه‌ها از جمله روحانیت سیاست مسایل جاری کشور حوزه و دانشگاه و دیگر عرصه‌ها، شخصیت برجسته و بارزی بوده‌اند.
Transliteration	Ḥabargozārī Fars: rahbar-e mo‘azam-e enqelāb, enteḥāb-e Āyātollāh Maḥdāvī konī be reyāsāt-e maḡles-e ḥobregān rā enteḥābī be-ḡā va be-ḥaq ḥāndand va afzūdand: īšān az ebtedā-ye enqelāb tā emrūz dar hame-ye ‘aršehā az ḡomleh-e rūḥāneyat-e seyāsāt-e masāyel-e ḡārī-ye kešvar-e ḥūzeh va dānešgāh va digār-e ‘aršehā, šaḥšīyat-e bargaste va bārezī būdeh-ānd.
Translation	Fars News Agency: The Supreme Leader of the Revolution called the election of Ayatollah Mehdevi Koni as the President of the Assembly of Experts a right choice and added: From the beginning of the revolution until today, he has been a prominent and noble figure in all fields, including the clergy, the current affairs of the country, the seminary, the university, and other fields.

<b>Example 7</b>	نماینده مردم کرمان در <b>مجلس خبرگان</b> رهبری گفت: باید مواظب فرصت‌طلبان بود تا گوهر انقلاب را به تاراج نبرند.
Transliteration	Namāyandeh-e mardom-e Kermān dar maḡles-e ḥobregān-e rahbarī goft: bāyad movāzeb-e forṣattalabān būd tā gūhar-e enqelāb rā be tārāḡ nabarand.
Translation	The representative of the people of Kerman in the Council of Leading Experts said: We must beware of the opportunists so that they do not plunder the essence of the revolution.
<b>Example 8</b>	عضو کمیسیون اقتصادی <b>مجلس</b> گفت: کمیسیون اقتصادی مجلس، ازاد شدن ورود کالای همراه مسافر در مناطق ازاد اروند، جلفا و انزلی و همچنین تشکیل 32 منطقه ازاد و ویژه اقتصادی جدید را در کشور به تصویب رساند.
Transliteration	‘Ozv-e komīsyūn-e eqtešādī-ye maḡles goft: komīsyūn-e eqtešādī-ye maḡles, āzād šodan-e vorūd-e kālā-ye hamrāh-e mosāfer dar manāteq-e āzād Arvand, Ğolfā va Anzalī va hamčonīn-e taškīl-e 32 manteqeh-e āzād va veže-ye eqtešādī-ye ḡadīd rā dar kešvar be tašvīb rasīd.
Translation	The member of the economic commission of the parliament said: the economic commission of the parliament approved the free entry of goods with passengers in Arvand, Jolfa and Anzali free zones and also the formation of 32 new free and special economic zones in the country.
<b>Example 9</b>	عضو کمیسیون صنایع و معادن مجلس گفت: رییس جمهور قصد دارد، دو نماینده از مجلس شورای اسلامی را به عنوان وزرای پیشنهادی نفت و صنایع و معادن به <b>مجلس معرفی کند</b> .
Transliteration	‘Ozv-e komīsyūn-e šanāye‘ va ma‘āden-e maḡles-e goft: dū namāyande az maḡles-e šūrā-ye eslāmī rā be ‘onvān-e vazarā-ye pišnahādī-ye naft va šanāye‘ va ma‘āden be maḡles mo‘arefī konad.
Translation	A member of the Industries and Mines Committee of the parliament said: The president intends to introduce two representatives of the Islamic Consultative Assembly as the proposed ministers of oil as well as industries and mines to the parliament.
<b>Example 10</b>	<b>مجلس نمایندگان</b> پارلمان فرانسه لایحه جنجالی ممنوعیت پوشیدن روبنده در اماکن عمومی را تصویب کرده است.
Transliteration	Maḡles-e namāyandegān-e parlemān-e Frānse lāyḡe-ye ḡanḡālī-ye mamnū‘eyat-e pūšīdan-e rūbande dar amāken-e ‘omūmī rā tašvīb karde-ast.
Translation	The French parliament has passed a controversial bill banning the wearing of headscarves in public places.
<b>Example 11</b>	<b>جلسه</b> کمیته تربیت بدنی کمیسیون فرهنگی <b>مجلس</b> شورای اسلامی به منظور بررسی علل حذف تیم ملی فوتبال امید با حضور علی کفاشیان رییس فدراسیون فوتبال ایران و دبیر کل کمیته ملی المپیک ایران برگزار شد.
Transliteration	Ġalase-ye komīteh-e tarbeyat-e badanī-ye komīsyūn-e farhangī-ye maḡles-e šūrā-ye eslāmī be manzūr-e barrasī-ye ‘alal-e ḡazf-e tīm-e mellī-e fūtbāl omīd bā ḡozūr-e ‘Alī Kafāšyān, rayīs-e federāsyūn-e fūtbāl Īrān va dabīr-e kol-e komeyte-ye mellī-ye olampīk-e Īrān bargozār šod.
Translation	The meeting of the Physical Education Committee of the Cultural Commission of the Islamic Consultative Assembly to investigate the reasons for the elimination of the national football team, was held in the presence of Ali Kafashian, the President of the Iranian Football Federation and Secretary General of the National Olympic Committee of Iran.
<b>Example 12</b>	پیرو یکشنبه جنجالی؛ شکایت نمایندگان از هیات ریسه مجلس تعدادی از نمایندگان مجلس از هیات ریسه به کمیسیون اصل 90 شکایت کردند.
Transliteration	Pīrav-e yek šanbe-ye ḡanḡālī, šekāyat-e namāyandegān az hay‘āt-e rayīs-e maḡles te‘dādī az namāyandegān-e maḡles az hay‘āt rayīs-e be komīsyūn-e āšl-e 90 šekāyat kardand.
Translation	Following a controversial Saturday, Complaints of MPs from the Presidium of the Parliament. Several MPs complained about the Presidium to the Article 90 Commission.

<b>Example 13</b>	توافقنامه همکاری‌های دراز مدت همکاری‌های راهبردی امریکا با افغانستان در میان موجی از نگرانی‌های کشورهای همسایه و منطقه یک ماه پیش از سوی روسای جمهور افغانستان و امریکا در کابل امضا شد و هر دو مجلس شورای ملی این موافقتنامه را تصویب کردند.
Transliteration	Tavāfoqnāme-ye hamkārīhā-ye derāz-e moddat-e hamkāryhā-ye rāhbordī-ye Āmrīkā bā Afgānestān dar meyān-e mūḡī āz negrānīhā-ye kešvarhā-ye hamsāye va manteqeh-e yek māhe pīš az sūy-e Rūsā-ye ḡomhūr-e Afgānestān va Āmrīkā dar Kābol emzā šod va har dū maḡles-e šūrā-ye mellī īn movāfeqatnāme rā tašvīb kardand.
Translation	The US-Afghanistan Strategic Partnership Agreement with Afghanistan was signed by the Afghan and US Presidents in Kabul a month ago amid a wave of concerns from neighboring countries and the region, and both houses of parliament (or: both parliaments) ratified the agreement.

Table 17. Examples of *maḡles* in context from the corpora

The following figure 2 is a visualization that represents the frequency of collocations comparing between *maḡles* and *pārlemān* in TalkBank. This tool allows users to compare between word forms or lemmas by contrasting collocations. This function, which is called the Word Sketch, helps to compare the use of two different word forms of the same lemma via their collocates. It is used here to compare between the two terms *pārlemān* and *maḡles* in order to gain a better understanding of how both terms are used and to provide insights into the differences in use and meaning between the given terms.

Figure 2. Visualization of the frequency of collocations compared between *pārlemān* and *maḡles* in TalkBank

It is clear from figure 2 that the term *maḡles* is more frequently collocated with the terms *ḡobregān*, *šūrā*, and *Turkiye* while the term *pārlemān* appears more frequently with *Ītāliyā*, *Orūpā*, *Koveyt*, *Berītāniyā*, and *alša ‘b*. This highlights that among the authors of Iranian blogs the term *pārlemān* refers to parliaments of other countries and the term *maḡles* is strongly linked to the Iranian parliament.

So, returning to the initial question of whether the term *mağles* features in Iranian blogs and newspapers as a generic term for any parliament in any country in the world, it can be stated that *mağles* is used only within limits to denote parliaments in countries other than Iran especially when those countries do not use the term domestically. The most prominent exception is the United States Congress with its two chambers: the House of Representatives and the Senate.

Instead, the authors mostly refer with *mağles* to the Iranian parliament and occasionally to parliaments of other countries, especially when these parliaments go by the name of *mağles* in their native languages as is the case in Bahrain, Oman, Qatar, and Saudi Arabia, as well as in Turkey. In the case of Afghanistan, the term is added to the proper name of the national parliament by analogy even though it is not part of its official name.

One important indicator for the generic use of *mağles* could be the frequency of its plural form *mağleshā*. It could be argued that using the plural indicates an abstraction of the term away from the identification of one specific parliament in the singular form. However, *mağleshā* is not used outside the context of the Iranian parliament thus reflecting the overall one-sidedness of the term.

A comparison of the weight values of *pārlemān* and *mağles* shows that the foreign term is less in use than its Persian equivalent in both corpora. The case is clear in blogs where the political meaning of *mağles* has a weight of 233,333 in TalkBank while the weight of *pārlemān* barely reaches 250. But also in newspapers, the political meaning of *mağles* with 251,503 has a higher weight than *pārlemān* with just 171,874.

In conclusion, *pārlemān* as a foreign term is still in wide use in Persian blogs and newspapers to refer to any parliament in the world. By contrast, the use of *mağles* in the meaning of parliament is strongly limited to the Iranian legislative body of the same name.

## 5.2 The Concept of *Candidate*

Another French loanword in Persian is *kāndīdā* (candidat, کاندیدا). It refers to the applicant or person running in an election campaign or selection process for a certain position or job (Moʻīn M. , 1386[2007]c., p. 2874). *Kāndīdā* is originally derived from the Latin term *candidatus* or *candidus* referring to the *white gown* worn by those striving for public office in ancient Rome (CNRTL, *Candidat*, 2012).

Moʻīn in his dictionary defines the term *nāmzad* in two distinct ways although both are based on the idea of applying for a position in the widest sense. First, it indicates the above-mentioned applicant for a specific office or a job. Second, it refers to a young boy or girl who made a promise to get married, a fiancé or fiancée, and in this sense a candidate for marriage (Moʻīn M. , 1383 [2004], p. 1095). Although the terminology group of the APLL approved the term *nāmzad* as an equivalent to *kāndīdā*, it already existed years earlier and was already mentioned in dictionaries such as (ʻAmīd, 1365[1986]c., p. 1939) as a synonym for *kāndīdā*.

Dehḥodā traces the use of the term back to the eleventh-century historiography *Tārīḥ-e Bayhaqī* by Abūfazīl Bayhaqī. It covers the history of the kingdom of Masʻūd Ġaznāvī and the Ghaznavid dynasty. This book was published in 30 volumes, only five of which have survived. It was also translated into Arabic and Russian (Yūsufī, 1988).

Dehḥodā also defines the term *nāmzad* as a person who comes into possession of something after a period of time, be it a material or touchable person or thing or be it ruling a state or a country (Dehḥodā, 1341[1963]a., pp. 248-249).

The Academy published *nāmzad* as the Persian equivalent to *kāndīdā* in the second volume of approved terms. The approved term *nāmzad* consists of two parts: first, the noun *nām* that means *name*, and second, the verbal noun *zad* from *zadan* (to hit, زدن). It could therefore loosely be translated as “to forge a name” or “to make a name for oneself”, that is, to become famous or to establish oneself in a specific field of expertise. It is therefore not a loan translation of the French word *candidat*. Instead, it is an explanation of the term or a loan rendition in Bateni’s system of categorization.

Term form	Frequency		Document frequency	
	TalkBank	TAP	TalkBank	TAP
<i>kāndīdā</i> کاندیدا	39,491	3,390	17,091	1,085
<i>nāmzad</i> نامزد	60,393	5,994	26,270	1,945

Table 18. Term frequency and document frequency of the terms in the semantic field of *candidate*



At first glance, both the foreign term *kāndīdā* and its Persian equivalent *nāmzad* occur frequently in the corpora. Even the ratio of both term pairs in the two corpora is similar: *Nāmzad* occurs 1.53 times more often in TalkBank than *kāndīdā* while the ratio in TAP is 1.77. The document frequency is also similar in both corpora even if not across corpora. *Kāndīdā* occurs roughly 2.3 times per document in TalkBank and even three times per document in TAP. The same ratios apply to *nāmzad* in each corpus.

However, the numbers in the previous table should be approached with caution and not taken as accurate. The Persian term *nāmzad* allows for multiple interpretations as mentioned before. In addition to the meaning of (political) candidate or nominee, in a social context it could mean groom or fiancé(e). Therefore, what should be considered for comparison with *kāndīdā* are only the texts that use *nāmzad* as a person nominated or running for a specific position.

The two corpora are hence searched for the terms *nāmzad* and *kāndīdā* in collocation with terms that have political connotations such as *enteḥābāt* (elections) and others. Table 19 below introduces the high-frequency collocations with the approved Persian equivalent *nāmzad*.

Term collocation		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
enteḥāb(āt)	انتخاب(ات)	20,986	2,538	10,295	1,064
reyāsat	ریاست	13,176	785	7,175	411
ḡomhūrī	جمهوری	7,597	726	4,735	384
maḡles	مجلس	5,121	657	3,115	404
ḡāyeze	جایزه	1,956	192	1,318	87
Oskār	اسکار	1,611	76	786	37
ezdevāḡ	ازدواج	150	2	140	2
arūsī	عروسی	63	6	59	6

Table 19. High-frequency collocations for the approved term *nāmzad* (term frequency and document frequency)

Table 19 reveals three groups that should be distinguished when analyzing the term *nāmzad* and its meanings or connotations. First, it has the unambiguous political connotation that is at the center of this research and the comparison with *kāndīdā*. Most of the results are associated with this group, among them *enteḥābāt* with roughly 23,500 hits, *reyāsat* (presidency) with almost 14,000 hits, *ḡomhūrī* (republic) with more than 8,000 hits, and *maḡles* (parliament) with approximately 6,000 hits across the two corpora.

The second group, somewhat surprisingly, contains references to candidates for awards and prizes, especially the Academy Awards or Oscars for motion picture accomplishments in

Hollywood. The collocation *gāyeze* (prize, award) occurs more than 2,000 times and *Oskār* more than 1,600 times in the two corpora. Even if this specific context is not political, it still deals with candidates in the first sense of the word as defined by Mo‘īn. This shows that *nāmzad* is used with the meaning of *candidate* in contexts that are not only political.

It is, by the way, interesting to note that not only *nāmzad* but also *kāndīdā* is used to refer to a nominee for the Academy Awards. The results of a correlation analysis show that *Oskār* is mentioned with *kāndīdā* 90 times in TalkBank and three times in TAP whereas it is mentioned with *nāmzad* 1,611 times in TalkBank and 76 times in TAP. These results lean clearly in favor of the Persian equivalent term *nāmzad* with a user rate of 94.5 percent (1,687 times in an overall total of 1,780 occurrences).

Finally, the third group comprises mentions of *nāmzad* in the semantic field of marriage. Collocations in this group are not frequent, however, with only 152 hits for *ezdevāg* (marriage) and 69 hits for *‘arūsī* (wedding).

It is also important to highlight the most frequent collocations of the foreign term *kāndīdā* in order to customize the comparison with *nāmzad* in the political context. Table 20 below shows the most frequent collocations with the loanword *kāndīdā* in both corpora.

Term collocation		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
enteḥāb(āt)	انتخاب(ات)	13,052	1,280	7,586	624
reyāsat	ریاست	8,767	513	5,366	314
ḡomhūrī	جمهوری	4,037	352	3,022	240
maḡles	مجلس	1,944	68	1,452	48
eṣlāh	اصلاح	1,320	261	924	159
oṣūlgarā(yān)	اصولگرا(یان)	1,176	73	747	58
ḡebhe	جبهه	858	15	566	14

Table 20. High-frequency collocations for the approved term *kāndīdā* (term frequency and document frequency)

Unlike *nāmzad*, whose collocates can be categorized into three different groups, the most frequent collocates of *kāndīdā* all stem from the semantic field of politics. The most frequent collocate of *kāndīdā* is also *enteḥābāt* like in the case of *nāmzad*. *Reyāsat*, *ḡomhūrī*, and *maḡles* follow suit in the same order as with *nāmzad*. Their individual frequency in both corpora is significantly lower than for *nāmzad*, though. Aside from the four mentioned collocates that refer generally to the political system, the remaining three frequent collocates of *kāndīdā* can be identified more precisely with political parties or movements.

Remarkably, *nāmzad* and *kāndīdā* are used as synonyms in the same news articles. In the TAP corpus they occur together in 30 documents. In TalkBank they are even mentioned in

1,122 documents together. The following table shows two examples of this case from both corpora.

<b>Example 1</b>	بدلیل عملکرد وزارت ورزش در دوره اخیر، شرایط در انتخابات فدراسیون های ورزشی به گونه ای رقم می خورد که هرگاه شنیده شود نامزدی مورد حمایت وزارت خانه قرار گرفته است هیات ها به سمت کاندیدا رقیب متمایل می شوند.
Transliteration	Bedalīl-e ‘amalkard-e vezārat-e verzeš dar dūreh-e aḥīr, šarāyet dar entehābāt-e federāsyūnhā-ye verzešī be gūneh-e raqam mīḥūrad ke hargāh šenīde šavad nāmzadī-ye mūred-e ḥemāyat-e vezārat ḥāneh qarār gereftēh-ast hay’ āthā be samt kāndīdā raqīb motamāyel mīšavad.
Translation	Due to the performance of the ministry of sports in the recent period, the situation of the elections in sports federations is such that whenever it is heard that a candidate has been supported by the ministry, the delegations lean towards the rival candidate.
<b>Example 2</b>	اگر یک زمانی آقای لاریجانی بیاید و بگوید من کاندیدا و نامزد هستم، میتواند جزو افرادی باشد که در شورای وحدت بررسی میشود.
Transliteration	Agar yek zamānī āqā-ye Lārījānī beyāyīd va begūyīd man kāndīdā va nāmzad hastam, mītavānad gozv-e afrādī bāšad ke dar šūrā-ye vaḥdat barrasī mīšavad.
Translation	If one day Mr. Larijani comes and says, “I am a candidate [and a candidate]”, he can be one of the people who will be considered in the Unity Council [a political party platform attached to the fundamentalist factions in Iran].

Table 21. Examples of both borrowed and foreign terms used in the same text

The visualization in figure 3 shows the most frequent collocations that appear with the foreign term *kāndīdā* or the approved term *nāmzad* in TalkBank. The colorful circles present the frequency of each collocation with the respective term with green representing *nāmzad* and red representing *kāndīdā*.

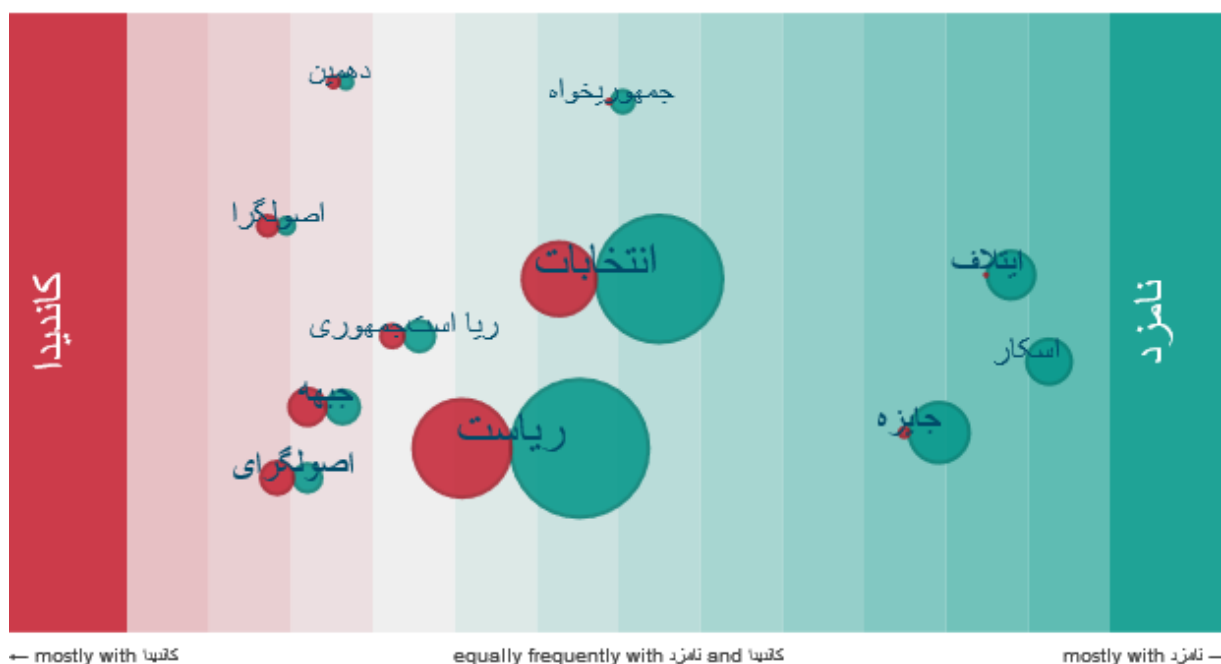


Figure 3. Visualization of the frequency of collocations compared between *kāndīdā* and *nāmzad* in TalkBank

As becomes clear from figure 3 above, the term pair *nāmzad*/*kāndīdā* occurs quite frequently with the same collocates in TalkBank. The only exception from this “rule” in the chart is the collocate *Oskār*, which only occurs with *nāmzad*. This is unusual when considering other term pairs such as *pārlemān*/*maḡles* or *īde`ūlūzī*/*marām* (see chapter 5.9 below) where collocations are always associated with one of the terms only.

Nonetheless, certain collocates are more favored by one of the two terms than the other. For example, *enteḡābāt* and *reyāsat* are very frequent collocates that are more likely to occur with *nāmzad* than with *kāndīdā*. By contrast, *ītelāf* (coalition) and *ḡāyeze* are less frequent collocates whose almost exclusive use is together with *nāmzad*. Collocates of *kāndīdā* are always shared with *nāmzad* even if their majority use is with *kāndīdā*. This is, for instance, the case with *ḡebhe* (front, as in the name of a political party) and *oṣūlgarā* and its derivatives.

In conclusion, it seems that *nāmzad* as an established Persian term is not only more frequent but also more successful in the political field than the foreign term *kāndīdā*. Comparing the overall weight of both terms supports this claim: *nāmzad* has an adjusted weight of 205,958 in TalkBank and 363,152 in TAP while *kāndīdā* has an adjusted weight of only 150,205 in TalkBank and 271,416 in TAP.

Even when only taking into account a purely political connotation of both terms, *nāmzad* is still heavier than *kāndīdā*. For this purpose, one can, for example, look at the two terms in collocation with *enteḡābāt*. The adjusted weight of *kāndīdā* in collocation with *enteḡābāt* is 59,341 in TalkBank and 126,111 in TAP. For *nāmzad*, it is 89,551 in TalkBank and 204,857 in TAP and thus higher in both corpora.

What approximates the true weight of *kāndīdā* and *nāmzad* in their political meanings even more accurately, is a weight count that does not only look at *enteḡābāt*, but also at the other frequent collocates *reyāsat*, *ḡomhūrī*, and *maḡles*. In Sketch Engine, it is possible to search the corpora in a way that excludes duplicates so that for this approximation all texts that include *kāndīdā* or *nāmzad* with one of the four collocates can be added. Doing so yields an adjusted weight of *kāndīdā* of 75,260 in TalkBank and of 146,735 in TAP as opposed to an adjusted weight of *nāmzad* of 105,732 in TalkBank and of 220,081 in TAP. By more than one standard, the Persian equivalent can therefore be considered more successful in newspapers and blogs than the foreign term.

### 5.3 The Concept of *Opposition*

The term *opūzīsyūn* is the borrowed French term that is used in Persian to denote the opposition party and to represent the political opposition to a theory or a specific policy (Mo'īn M. , 1383 [2004]). Even though *opūzīsyūn* is borrowed from French, the etymology of the term goes back to the Latin word *oppositiō* that consists of the prefix *opposi-* (from the verb *oppōnere*, to place over or against, to argue in reply) and the suffix *-tiō/-tion* (Webster, Opposition, 2021) and (CNRTL, opposition, 2012).

The broadest meaning of *opūzīsyūn* refers to the efforts of groups, parties, unions, and individuals that – whether through legislative or other means – work to attain aims that are opposite to the goals and policies of those in power, be they political, economic, or social in nature. In a narrow sense, it is the name given to a group whose existence is recognized by the government and/or by the constitution in a parliamentary system. In liberal systems of government, such as in the United Kingdom (one comes to think of Her Majesty's Most Loyal Opposition), France, and Germany, this type of opposition parties is accepted by the government and the public as a group that challenges government policies and could stand in to replace the actual government after elections or a change of coalition alignments and at the same time participates in parliamentary debates and decision-making. The first such political opposition party appeared in England in the 18<sup>th</sup> century. By contrast, in countries ruled by one party, any opposition to the ruling party and the government is considered an opposition to the public order and the constitution (cf. Āšūrī, 1376 (1997), pp. 8-9).

The terminology group in the first volume of approved terms agreed on using *gorūh-e moḥālef* to replace the foreign term *opūzīsyūn* (Terminology group , 1382 [2004]). The coined term *gorūh-e moḥālef* consists of two parts. The first part is *gorūh* (گروه). This word comes originally from the Middle Persian *grūh*. It could carry a lot of connotations, but the principal translation is *group* or *sector*. Its meaning is not limited to a group of humans; it may mean a group of animals or a group of any category ('Amīd, 1365[1986]c., p. 2049). The Academy has used this term in coining a lot of other terms in various topics and fields.

The second part of the approved term is *moḥālef*, which is originally an Arabic adjective that refers to the opposite of something. It means *against*, *anti-*, or *adverse*. It could also refer to an enemy. The same meaning is used with the same word in Persian (Dehḥodā, 1353[1974], p. 630). *Gorūh-e moḥālef* would therefore translate as *opposite group* or *adverse group*. It serves as a good example of what Bateni considers a loan translation if one accepts the lexical

meaning of *opposition* to be the opposite group in parliament or of a political idea. If one compares the etymologies of both terms, though, the Persian equivalent could also be considered a loan rendition where *moḥālef* is equivalent to *opposition* and *gorūh* explains the context.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>opūzīsyūn</i>	اپوزيسيون	11,814	213	5,986	128
<i>gorūh-e moḥālef</i>	گروه مخالف	301	8	255	8

Table 22. Term frequency and document frequency of the terms in the semantic field of *opposition*

Table 22 above presents the frequency and the document frequency of both the foreign term *opūzīsyūn* as well as the approved one *gorūh-e moḥālef* to be able to determine the success or failure factors of each. In absolute numbers, *opūzīsyūn* occurs significantly more often in each corpus than its Persian equivalent. It has a share of 97.5 percent in TalkBank and of 96.4 percent in TAP. *Opūzīsyūn* is also more frequent per document in both corpora with almost two and over 1.6 mentions per document in TalkBank and TAP respectively. By contrast, *gorūh-e moḥālef* is mentioned only between one and 1.2 times in TAP and TalkBank respectively.

The adjusted TF-IDF weight of *opūzīsyūn* is 56,272 in TalkBank and 32,245 in TAP. The Persian equivalent *gorūh-e moḥālef* records an adjusted weight of only 2,303 in TalkBank and 1,951 in TAP. Therefore, the foreign term continues to be at least 16 times heavier than the approved Persian equivalent.

This is so despite the fact that *gorūh-e moḥālef* was coined in the first volume of approved terms by the APLL. Consequently, time does not seem to be the only measure of the success of any of the approved terms. That is, older terms are not necessarily more successful over time. Instead, the fluidity and awareness of the new term play a role. If it is difficult to pronounce or if the APLL does not sufficiently promote it, it might simply remain out of use or go unnoticed.

## 5.4 The Concept of *Nation*

Throughout history, the issue of Iranian national identity has been and remains one of the most important and fundamental issues in the life of Persians and Iranians. Manifestations of nation, national self-esteem, nationalism, and all terms that relate to this topic are among the most prominent features of the Iranian people. However, there is more than one term in Persian that could refer to *nation*, some of which are in wider use and cover different concepts than others.

This analysis of the complex concept of *nation* is therefore somewhat removed from the other term analyses in this study. In addition, it is distinct from the other terms and term pairs discussed in this study in that none of the three terms in the focus of this section are of European origin. Instead, with *mellat* and *ommat* coming from Arabic and *mardom* being originally Old Persian,<sup>48</sup> they have a long history of shifting meanings in Persian.

The first Persian translation of *nation* is *mellat* (ملت). At the same time, *mellat* covers the concept of *people*, which links it to the Persian term *mardom* (people, مردم) (Paul, 1999, p. 192) as will be evidenced below. In this sense, it has come to describe an inherent Aryan feature of Iranian social and historical consciousness with its civilization extending to about 4000 BCE. The formation of the Persian empire in the period of Sassanid rule, which included a great part of the ancient world and happened to be the last Persian dynasty before the Islamic conquest in the mid-seventh century CE, is an important reference point for this idea or concept of *mellat*.

Ironically, the term *mellat* entered Persian coming from Arabic with an originally religious connotation. In Arabic, *milla* (ملة) describes a religious group or congregation. It occurs several times in the Quran, often together with the prophet Abraham/Ibrahim (*milla Ibrāhīm*, ملة ابراهيم), such as in sura *al-Nisā'* 4:125 and sura *Yūsuf* 12:37. Before its redefinition in the current era to mean *nation*, *mellat* meant *religion* (Paul, 1999, p. 193) or *religious way*. Long after the Islamic conquest, “[*m*]illat, a collective formation associated with the community of believers, was dissociated from Islam and the creator God, and anchored to the life-giving mother-nation (*mām-i vaṭān*) and the mother tongue (*zabān-i mādarī*)” (Tavakoli-Targhi, 1990, p. 78).

Nevertheless, the most influential event in Iranian history is arguably the Arab-led Islamic conquest of Persia, which lasted for two centuries and has had a major impact on Iran’s history, literature, civilization, and language. As discussed in chapter 2.2.6, the Arab-Islamic influence has been heavily attacked and rebuked many times. Nowhere can this better be observed than in the context of a national awakening or national consciousness. The main concern of such anti-Arab and anti-Arabic movements and tendencies has been to preserve the Iranian self, identity, and nationalism. Arabophobic ideas and sentiments can be noticed in a lot of the important literary works and rhetorical masterpieces in Iran throughout history. These are the works that portrayed the period of Islamic rule in Iran as a period of suffering, devastation, ignorance, and dictatorship whereas the pre-Islamic period was portrayed as a lost

---

<sup>48</sup> For the etymology of *mardom* see chapter 5.5 with the discussion of *mardombāvar* and *mardomsālārī*.

utopia with righteous rulers based on the description of Mirza Aqa Khan Kirmani that considered this “rhetorical use of history” as “necessary for uprooting the malicious tree of oppression and revitalizing the power of *millīyat* (nationalism) in the character of the Iranian people” (Tavakoli-Targhi, 1990, p. 83). All the attempts to purify the Persian language from Arabic and to establish purely Persian equivalents – along with attempts to reform the structure of the Persian language itself<sup>49</sup> – have eventually led to reconsidering the concept of *mellat* to be *mellat-e Īrān* (the – secular – Iranian nation) instead of *mellat-e šī‘ah-ye eṭnā ‘ašrī* (the – religious – Shiite nation) (Tavakoli-Targhi, 1990, pp. 91,93).

Ášūrī describes *mellat* as a large group of people united by a shared culture and knowledge. This connection creates a sense of unity to one another as well as a feeling of belonging among those who are part of the same group. The occupation of a shared geographic territory and the feeling of attachment to a certain land play an important role in the rise of this feeling. He states as well that the vital force that unites the nation together arises from the sense of belonging to one’s own history, religion, culture, and language. In the modern era, country, nationality, and government are the main elements that determine the concept of the *nation* by modern standards (Ášūrī, 1376[1997], pp. 306-308).

The national consciousness that refers to a sense of belonging to a country or a specific nation and the requirement of building a national government and possessing a power or state that is solely one’s own has a short history. In fact, the direct connection between the concepts of *nation* (*mellat*) and *state* (*dūlat*) is a matter of the modern era with a history that is linked to Europe’s intellectual, political, and social developments over the last two centuries.

The other Persian term that is used to denote the concept of *nation* is *ommat* (امت). Historically, *ommat* is an originally Arabic word, too, coming from the Arabic *umma* (أمة). *Al-umma* in Arabic describes a group of people most of whom are of one origin and share inherited characteristics, interests, and aspirations or are united by a single matter of religion, place, or time (Academy of the Arabic Language in Cairo, 2004, p. 27). *Al-umma* refers also to the community (of believers) (Ibn Manzūr, 1922, p. 27). It is first mentioned in the Quran in sura *Āl ‘Imrān* 3:104 and 3:110 where it says that among those who believe there should be a group or community (*umma*) that invites others to do good, promotes virtue, and prevents vice (3:104) and that those who believe are in fact the best community (“ḥayr umma”) that was sent to

<sup>49</sup> For more information on language purification activities in Iran see Hidayat, (1288 (1881)) quoted in Tavakoli-Targhi (1990).



mankind (3:110). The concept of *umma* is therefore significantly older than the idea of nation, nation state, and nationalism that was introduced by the French.

Dehḥodā deals with *ommat* in two ways. First, he describes it as a community to which a prophet has come and specifies it as a group of people who are followers of a prophet. Second, he defines *ommat* as a group of each type (“ṣanf”) of people (Dehḥodā, 1342[1963], pp. 160-161). These two approaches reflect the two concepts of *ommat* as a religious community on the one hand and a community of people with generally shared features on the other. By contrast, in a reversal of other scholarly definitions, Dehḥodā uses religion, creed, and sharia also to define *mellat* (Dehḥodā, p. 1059). He is therefore in one line with Nafīsī, who defined both *ommat* and *mellat* as religion (“dīn”) and sharia (“šarī‘at”) aside from assigning to *mellat* other religious definitions as well (Nafīsī, 1355[1977]b., pp. 349, 384).

However, according to the Iranian jurist and philosopher Moḥammad Ğa‘far Langrūdī, *ommat* is a synonym of *mellat* in the field of the new Iranian legal terms as both refer to the modern concept of *nation* (Langrūdī M. , 1392[2013], p. 79). In addition, Langrūdī states that the modern concept of nation (*mellat*) embraces a group of human beings who live in a specified territory and are subject to the power of a government (Langrūdī M. , 1367 [1988], p. 687).

The terminology group of the APLL approved the term *mellat* in the fifth volume of approved terms in order to express the concept of a group of people who consider themselves constituents of a society with a distinct political identity from other human beings, share common historical and cultural traditions, and live mostly in a specific land (Terminology group, 1387 [2008]). In other words: The APLL introduced *mellat* as a term to cover the concept of the English term *nation*. This is also supported by the compound terms of *mellat* coined by the APLL. Since *mellat* does not replace a specific foreign term, it should be considered a loanshift according to Bateni’s types of borrowing where the Arabic term *mellat* has been given a new semantic meaning after moving to the receiving Persian language.

This loanshift is particularly clear in a compound word such as *mellat-sāzī*. For the concept that refers to the process of strengthening centrist factors between various cultural groups in a country with the aim of creating and promoting a sense of societal nationalism rather than a sense of ethnic nationalism, the terminology group coined the compound *mellat-sāzī* in the eleventh volume (nation building, ملت‌سازی). Here, the term *mellat* is followed by the morpheme *sāzī*, which is the infinitive form (*sāz* with suffixed *ī*) of the verb *sāhtan* that means

to construct or to form. In Bateni's system of categorization, it therefore classifies as a loan translation of the English term *nation building*.

The following analysis sheds light on both *ommat* and *mellat* and the contexts in which each of the two terms are used, that is, whether they have a religious or political implication and whether the use conforms to the definition set by the APLL or otherwise. In addition, the term *mardom* (people) will be highlighted in order to compare it to *mellat* covering the concept of the *people* and the *nation*. The following table presents the frequency and document frequency of *ommat* and *mellat* in TalkBank and TAP.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>ommat</i>	امت	14,857	716	9,103	293
<i>mellat</i>	ملت	214,294	7,255	88,882	2,807

Table 23. Term frequency and document frequency of the approved term *mellat* and the non-approved term *ommat*

From the previous table it can be seen that *mellat* is by far the more frequent term in both newspapers and blogs. It occurs roughly 14 times more than *ommat* in TalkBank and roughly ten times more in TAP. The frequency of each term per document is roughly the same across the corpora with around 2.5 mentions per text except for *ommat* in TalkBank, which occurs only about 1.6 times per document.

The proper noun *sāzmān-e melal-e motaḥed* (United Nations Organization, سازمان ملل متحد) is already excluded from the total frequency count of the term *mellat* in table 23 above. By itself, *sāzmān-e melal* has 61,600 unrepeated hits in TalkBank and 1,083 unrepeated hits in TAP. This would be more than 40 percent of all hits in TalkBank and over one fourth of all hits in TAP.

Arguably, blogs and newspapers speak more about the secular nation (*mellat*) than the religiously connotated nation (*ommat*). A closer look at each term and its collocations is therefore helpful. The following table 24 gives the most frequent compound collocations of *mellat* in both corpora. Here, *mellat* comes to denote a nation in the sense of a population living in a specific geographic territory governed by a specific government in accordance with the definition made by the APLL.

Collocations		Frequency (number of hits)		Document frequency	
		TalkBank	TAP	TalkBank	TAP
mellat-e Īrān	ملت ایران	48,804	2,363	24,974	1,111
mellathā-ye manteqe	ملت‌های منطقه	3,253	69	2,054	57
mellat-e Feleštīn	ملت فلسطین	3,228	237	1,715	120
mellathā-ye Āsyā	ملت‌های آسیا	2,300	15	1,471	11
mellathā-ye ġahān	ملت‌های جهان	1,594	38	1,341	36
mellat-e Meṣr	ملت مصر	1,237	0	722	0
mellat-e 'Arāq	ملت عراق	1,038	146	754	81
mellat-e Baḥrayn	ملت بحرین	944	7	593	7

Table 24. The most frequent collocations (compound terms) of the approved term *mellat* in both corpora

Of the most frequent compound terms, *mellat* occurs exclusively in collocation with country names such as Iran, Palestine, or Egypt. The plural form *mellathā* is the node for collocates that describe other geographically defined territories such as Asia or the world at large.

The most frequent collocate is by far *Īrān*, which dominates the collocation analysis. There are various adjectives that are used with *mellat-e Īrān* in both corpora. In addition to the more than 50,000 hits of the compound *mellat-e Īrān* in the two corpora, the following compounds are also found: *mellat-e 'azīz-e Īrān* (ملت عزیز ایران), *mellat-e šarīf-e Īrān* (ملت شریف ایران), *mellat-e bozorg-e Īrān* (great nation of Iran, ملت بزرگ ایران), *mellat-e qahremān-r Īrān* (the heroic nation of Iran, ملت قهرمان ایران), *mellat-e ġayūr-e Īrān* (the gallant nation of Iran, ملت غیور ایران), *mellat-e bozorgevār-e Īrān* (the grand nation of Iran, ملت بزرگوار ایران), and *mellat-e mazlūm-e Īrān* (the oppressed nation of Iran, ملت مظلوم ایران). Altogether, they produce 6,541 non-repeated hits in TalkBank. In TAP, *mellat-e Īrān* is also mentioned with the same adjectives 438 non-repeated times. These adjectival compounds make up 20 percent of all documents in TalkBank and 28 percent of all documents in TAP. Such descriptions of the Iranian nation are therefore more common in newspapers.

As mentioned before, *mellat* could be translated as *people* in addition to *nation*. Persian has another term for *people* which is *mardom*. The next section will look at the term *mardom* in blogs and newspapers. The term *mardom* is among the highly frequent terms in both corpora. It is mentioned 781,801 times in TalkBank, of which 237,127 are without recurrence. In TAP, it has a frequency of 55,517 hits in 11,311 documents. Table 25 below shows the most frequent collocations of *mardom* in both corpora.

Collocations		Frequency (number of hits)		Document frequency	
		TalkBank	TAP	TalkBank	TAP
mardom-e Īrān	مردم ایران	44,898	1,318	27,311	912
mardom-e īn kešvar	مردم این کشور	10,194	205	8,442	182
mardom-e Āmrīkā	مردم آمریکا	9,826	396	8,298	240
mardom-e Afġānestān	مردم افغانستان	9,541	88	5,096	49
mardom-e Tehrān	مردم تهران	7,490	562	5,015	392
mardom-e manteqe	مردم منطقه	6,488	209	4,966	170
mardom-e Mešr	مردم مصر	5,544	14	2,989	5
mardom-e Baħrayn	مردم بحرین	5,430	49	2,894	24
mardom-e šahr	مردم شهر	4,533	335	3,850	284
mardom-e ġahān	مردم جهان	4,283	292	3,559	264
mardom-e ostān	مردم استان	3,851	261	2,742	210
mardom-e mosalmān	مردم مسلمان	3,535	74	2,883	65
mardom-e ‘ādī	مردم عادی	3,379	322	3,022	254
mardom-e maḡlūm	مردم مظلوم	2,367	122	1,908	95
mardom-e šarīf-e Īrān	مردم شریف ایران	2,168	27	1,874	25
mardom-e ‘azīz	مردم عزیز	2,045	323	1,844	267
mardom-e enqelābī	مردم انقلابی	1,664	77	1,445	65

Table 25. The most frequent collocations (compound terms) of *mardom* in both corpora

A look at table 25 with the most frequent collocations of *mardom* allows to distinguish four groups of compound terms. First, as in the case with *mellat*, the most frequent collocates of *mardom* include country names such as Iran, (United States of) America, Afghanistan, and Egypt. They make up six of the top eight compound terms. This group also includes a frequent compound term with a country name and an adjective (*mardom-e šarīf-e Īrān*). Second, collocates that make general reference to a country or territory such as *īn kešvar* (this country), *manteqe* (region), and *šahr* (city) can be discerned. The third group is made up solely of *mardom-e mosalmān* because it is the only one that has a religious component. Finally, some of the less frequent collocations are adjectival descriptions of the people such as ‘*ādī* (ordinary), *maḡlūm* (oppressed), and *enqelābī* (revolutionary). These compound terms are largely congruent with those of *mellat* supporting the assumption that both *mellat* and *mardom* are used synonymously and interchangeably.

Aside from these findings, the term *mellat* is not limited to a purely secular political implication, but it is also used in both corpora with the religious political connotation of nations that have an Islamic system of government (such as Iran) or that are simply predominantly Muslim. This does not necessarily mean that the term means the Muslim community of believers as a category distinct from the modern nation-state. This is, for instance, the case in the compound terms *mellathā-ye eslāmī* (Islamic nations) and *mellat-e mosalmān* (Muslim nation) with a total of 1,131 hits in 963 documents in TalkBank and 45 hits in 40 documents in TAP.

Collocations		Frequency (number of hits)		Document frequency	
		TalkBank	TAP	TalkBank	TAP
ommat-e eslāmī	امت اسلامی	6,362	288	4,140	139
ommat-e eslām	امت اسلام	3,590	132	2,572	84
ommat-e mosalmān	امت مسلمان	1,499	12	1,106	12
ommat-e Hezbollāh	امت حزب الله	641	20	514	7
ommat-e Īrān	امت ایران	315	0	300	0
ommat-e ‘arabī	امت عربی	270	0	223	0
ommat-e vāhedeh	امت واحده	210	6	161	6
ommat-e enqelābī	امت انقلابی	84	1	74	1
ommat-e bozorg-e eslām	امت بزرگ اسلامی	49	3	44	3
ommathā-ye pīšīn	امت های پیشین	57	0	44	0
ommat-e gozašteḥ	امت گذشته	56	0	37	0
ommat-e moslemeh	امت مسلمه	21	0	15	0

Table 26. The most frequent collocations (compound terms) of *ommat* in both corpora

The collocation analysis of table 26 reveals that among the over 15,000 hits of the term *ommat* in both corpora, almost 12,000 hits occur as a part of a compound term that relates to a religious – specifically Islamic – connotation such as *ommat-e eslāmī* (Islamic nation), *ommat-e eslām* (nation of Islam), and *ommat-e mosalmān* (Muslim nation). Another frequent collocation of *ommat* with more than 660 hits in both corpora is the Lebanese political party and militant group *Hezbollāh*. Only two compounds have a distinct national connotation referring to the Iranian and Arab nations with a combined total of almost 600 hits. Table 27 below gives several examples of the practical application of *ommat*, *mellat*, and *mardom* in the two corpora. It shows especially how *mellat* and *mardom* could be used interchangeably (see example 5).

<b>Example 1</b>	قائم مقام شورای هماهنگی تبلیغات اسلامی گفت: همزمان با آخرین جمعه ماه رمضان، رهبر معظم انقلاب از طریق شبکه های داخلی و بین المللی با امت اسلام سخن خواهند گفت.
Transliteration	Qā'em-e maqām-e šūrā-ye hamāhangī-ye tablīgāt-e eslāmī goft: hamzamān bā āḥarīn-e ġom'eh-ye māh-e ramazān, rahbar-e mo'azzam-e enqelāb az ṭarīq-e šabakehā-ye dāḥelī va bayn ol-melalī bā ommat-e eslām soḥan ḥāhand goft.
Translation	The deputy chairman of the Islamic propaganda coordination council said: Simultaneously with the last Friday of Ramadan, the Supreme Leader of the Revolution will speak to the Muslim ummah through local and international channels.
<b>Example 2</b>	شهردار تهران در پیامی توییتری، ضمن تشکر از ریاست پیشین مجلس، به منتخبان جدید ملت تبریک گفت.
Transliteration	Šahredār-e Tehrān dar payām-e tūyīterī, zemn-e tašakkor az reyāsat-e pīšīn-e maḡles, be montahabān-e ġadīd-e mellat tabrik goft.
Translation	In a Twitter message, the mayor of Tehran thanked the former speaker of the parliament and congratulated the new elected members of the nation.

<b>Example 3</b>	رئیس ستاد کل نیروهای مسلح با بیان اینکه انتفاضه ملت فلسطین فراتر از سرزمینهای اشغالی رفته و وارد مرحله سرنوشت سازی شده است، گفت: زمزمه پایان اسرائیل غاصب در خیابانهای تلآویو به گوش میرسد.
Transliteration	Ra'īs-e setād kol-e nīrūhā-ye mosallah bā bayān-e īnke entefāze-ye mellat-e Feleštīn farātar az sarzamīnhā-ye ešgālī rafte va vāred-e marḥale-ye sarnevešt-e sāzī šode ast, goft: zamzameh-ye pāyān-e Esrā'īl-e gāṣeb dar ḥeyābānhā-ye Talāvīv be goš mīrasad.
Translation	The chief of the general staff of the armed forces claimed that the Intifada of the Palestinian nation had gone beyond the occupied areas and entered a pivotal stage and said: "The murmur of the end of the occupier Israel can be heard in the streets of Tel Aviv."
<b>Example 4</b>	رئیس سازمان اطلاعات و شناسایی وزارت دفاع دولت نجات ملی یمن در پیام خود به وزیر اطلاعات ایران از آزدگان جهان درخواست کرده لغو محاصره ظالمانه ملت ایران توسط آمریکا را مطالبه کنند.
Transliteration	Ra'īs-e sāzmān-e eṭṭelā'ī va šenāsā-ye vezārat-e defā'-e dūlat-e neḡāt-e mellī-ye Yaman dar payām-e ḥūd be vazīr-e eṭṭelā'ī-e Īrān az āzād-gān-e ḡahān-e darḥāst kardeh laḡū-ye moḥāṣare-ye zālemāne-ye mellat-e Īrān tavassoṭ-e Āmrīkā rā motālabē konand.
Translation	In a statement to the Iranian Minister of Intelligence, the head of the Yemeni National Salvation Government's Ministry of Defense's Intelligence and Identification Organization urged the world's free people to demand the removal of the US's oppressive blockade on the nation of Iran.
<b>Example 5</b>	رئیس فراکسیون امید مجلس شورای اسلامی در نامهای به مردم ایران و موکلین خود ضمن ارائه گزارشی از عملکرد فراکسیون امید به تشریح چالشهای مهم کشور پرداخته و راهکارهایی برای برون رفت از وضعیت فعلی ارائه کرده است.
Transliteration	Ra'īs-e frākseyūn-e omīd-e maḡles-e šūrā-ye eslāmī dar namhā-ye be mardom-e Īrān va movakelīn-e ḥūd žemn-e erā'eh-e gozārešī az 'amalkard-e frakseyūn-e omīd be tašrīḥ-e čālešhā-ye mohem-e kešvar pardāḥte va rāhkārhā-ye barāye berūn raft az važ'eyyat-e fe'lī erā'eh kardeh ast.
Translation	While presenting a report on the Omid faction's performance, the president of the Omid faction of the Islamic Consultative Assembly described the country's major issues and gave remedies to overcome the current scenario in a letter to the Iranian people and her clients.

Table 27. Examples of *mellat*, *ommat*, and *mardom* in the two corpora

Calculating the weight of both terms that are used as translation to the English term *nation* shows a heavy adjusted weight of the term *mellat* of 491,876 in TalkBank and 350,734 in TAP. By contrast, *ommat* records only a weight of 65,070 in TalkBank and 88,606 in TAP. By comparison, the adjusted weight of *mardom* in TalkBank reaches 1,092,732 while in TAP it is 101,967. As expected, the rather secular notions of *mellat* and *mardom* are significantly more prominent and more dominant in newspapers and blogs alike than the religiously connotated equivalent *ommat*.

## 5.5 The Concept of *Democracy* and Its Derivatives

The loanwords *demūkrāsī* (democracy, دموکراسی), *demūkrātīk* (democratic, دموکراتیک), and *demūkrāt* (democrat, دموکرات) like so many other foreign terms arrived in the Persian lexicon via French. While the noun form *demūkrāsī* was borrowed from the French term *démocratie* (Mo‘īn M. , 1386 [2007]b., p. 1561), it goes back to the Greek word *demokratia* that consists of the morphemes *demos* (people) and *kratia* (strength, power) (Ášūrī, 1376[1997], p. 157) and (Sa‘īdeyān, 1378 [1999]b., p. 1966).

In his history of the term, Ášūrī discusses the concept of democracy looking back at its emergence in antiquity among Greeks and Romans. It was first introduced in the Greek city states such as Athens in the fifth century BCE. In ancient Greece, democracy was exercised directly where all people except for women and slaves participated in the law-making process, where leadership positions were rotated, and where court judges were likewise chosen by lot. There was no parliament and there were no representatives. The Roman Republic implemented some characteristics of democracy with the election of representatives, but democracy vanished with the arrival of the empire around the beginning of the Common Era (Ášūrī, 1376[1997], p. 157).

By contrast, Dehḥodā emphasizes the modern definition of *demūkrāt* as a person who believes in democracy. He traces the term *demūkrāt* historically to 1792 when the American Democratic-Republican Party was established under the leadership of Thomas Jefferson to support middle-income strata of society such as craftsmen, small farmers, and merchants, as well as owners of small industries and manufactories (Dehḥodā, 1351[1972], p. 245).

These loanwords have been widely used in the Persian language, which is evidenced by the fact that they are found in many Persian dictionaries such as *Loḡatnāme-ye Dehḥodā*, *Farhang-e ‘Amīd*, and *Farhang-e Mo‘īn*. The term *demūkrāsī* is also referenced in Ášūrī’s political encyclopedia, where it is used alongside the term *mardomsālārī* (Ášūrī, 1376[1997], pp. 157-162), a term that was later approved by the APLL to replace the loanword *demūkrāsī*.

The APLL approved the term *mardombāvar* as an equivalent to *demūkrāt* in the sixth volume of approved terms in 2010 (Terminology group, 1388 [2009]a.). It consists of *mardom*, which means *people*, which in turn is taken from the Middle Persian word *martōm* (human) (‘Amīd, 1365[1986]c., p. 2216) going back to the Old Persian *marta-tauxman* (Hassandoust, 1393[2014]c., p. 2592). The second part of the approved term, *bāvar*, goes back to the Middle

Persian *wābar* and the Old Persian *vāvāra* with the meaning of *believe, trust, or faith* (Hassandoust, 1393[2014]a., pp. 400-402) and (‘Amīd, 1365[1986]a., p. 371). The approved Persian term *mardombāvar* therefore translates as believer in (the) people. According to Bateni’s categorization, it qualifies as a mixed loan translation and rendition of the term *democracy* where *mardom* stands as a translation of *demo-* (people) and *bāvar* is an explanation of *-crat*.

In the same volume, the Academy approved *mardomsālārī* to replace the borrowed term *demūkrāsī*. This equivalent is also a compound term made up of two parts with the first again being *mardom* as explained before. The second part is *sālārī* dating back to the Middle Persian *sardar* and Old Persian *sara-dāra*, which means *leader, chief, head, and chieftain* (Hassandoust, 1393[2014]b., p. 1648). *Mardomsālārī* could therefore be translated as leadership of the people, which in Bateni’s system of categorizing Persianized terms, would be a loan translation with *sālārī* replacing *-cracy* meaning the power to rule.

The terminology group has used the two morphemes *bāvar* and *sālārī* for creating neologisms and replacing foreign terms in various cases and for a diverse range of topics. In the field of political and foreign relations, there are numerous neologisms compiled in the following table 28.

Equivalent terms		English translation	Volume
<i>dīvānsālārī</i>	دیوان سالاری	Bureaucracy	2
<i>ṣolḥbāvar</i>	صلح باور	Pacifist	4
<i>noḥbesālārī</i>	نخبه سالاری	Elitism	5
<i>hamgānsālārī</i>	همگان سالاری	Pantisocracy	6
<i>šāyestesālārī</i>	شایسته سالاری	Meritocracy	6
<i>zarsālārī</i>	زر سالاری	Plutocracy	6
<i>kammāyehsālārī</i>	کم مایه سالاری	Mediocracy	6
<i>oḡūlbāvar</i>	افول باور	Declinist	8
<i>dozdsālārī</i>	دزد سالاری	Kleptarchy – Kleptocracy	11
<i>yekesālārī</i>	یکه سالاری	Autocracy	12
<i>artešsālārī</i>	ارتش سالاری	Stratocracy	12
<i>meyāngānsālārī</i>	میانگان سالاری	Mesocracy	12
<i>taksālārī</i>	تک سالاری	Monocracy	12
<i>zar-mardomsālārī</i>	زر- مردم سالاری	Pluto-Democracy	12
<i>besyārsālārī</i>	بسیار سالاری	Polyarchy	12
<i>bordehdārsālārī</i>	برده دار سالاری	Slavocracy	12
<i>dūsālārī</i>	دو سالاری	Diarchy – Dyarchy	12
<i>kārgarsālārī</i>	کارگر سالاری	Ergatocracy	12
<i>ḥīšāvandsālārī</i>	خویشاوند سالاری	Nepotocracy	12
<i>zamīndārsālārī</i>	زمین دار سالاری	Plantocracy	12

Table 28. Approved terms that contain *sālārī* or *bāvar* often as a replacement for the English suffixes *-cracy* or *-archy* respectively



Clearly, the terminology group is consistent in translating the English suffixes *-cracy* and *-archy* with *sālārī*, both of which indicate the meaning of a form of government, social or political class, or theory of social organization. However, of all the terms mentioned in the previous table 28, only six are mentioned at least once in the two corpora. The term *artešsālārī* has only one hit in TalkBank, equally *taksālārī* and *hīšāvandsālārī*. The approved term *noḥbesālārī* is mentioned only five times in TalkBank. The fifth term that has any hits in the corpora among the approved terms that contain the suffix *sālārī* is *šāyestesālārī* as it is mentioned 303 times in TalkBank and 58 times in TAP. Finally, the term *dīvānsālārī* will be discussed in detail later in chapter 5.8. Being the specialized terms from political science most of them are, it is not surprising that they feature so little in the corpora that address a larger audience of newspaper and blog readers.

### 5.5.1 Analysis of the Foreign Terms *demūkrāt/demūkrātīk*

The word *democracy* has found its way as a loanword into numerous languages, for example, in Arabic (الديموقراطية), German (*Demokratie*), Romanian (*democrație*), and Turkish (*demokrasi*) to name but a random few. This foreign term is used widely as well in Persian. The following table 29 shows the frequency of *demūkrāt* and *demūkrātīk* in newspapers and blogs.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>demūkrāt</i>	دموکرات	7,621	2,922	4,308	775
<i>demūkrātīk</i>	دموکراتیک	12,197	320	6,474	196
<i>demkūkrāsī</i>	دموکراسی	34,000	1,114	15,846	458

Table 29. Term frequency and document frequency for *demūkrāt*, *demkūkrāsī*, and *demūkrātīk*

In order to give an accurate result of the success of the term, proper nouns have to be excluded. The term *demūkrātīk* appears frequently in the names of political parties such as the US Democratic Party. *Hezb-e demūkrātīk* (حزب دموکراتیک), as a proper noun for the US Democrats, is mentioned 342 times in TalkBank. Similarly, *hezb-e demūkrātīk-e Kordestān* (حزب دموکراتیک کردستان), the Democratic Party of Kurdistan, has 218 hits in the same corpus. Referring to the German “Democratic” parties (Social Democratic Party, Christian Democratic Union, Free Democratic Party), *hezb-e demūkrātīk* produces 101 hits. The Azerbaijani Democratic Faction is mentioned in Talkbank 37 times as *hezb-e demūkrātīk-e Āzarbayğān* in – what seems to be – an explanatory attempt by the authors since its original name is not *hezb* but *ferqe* (فرقه).

It is also part of several official country names. For example, the Democratic Republic of the Congo (DRC, *ğomhūrī-ye demūkrātīk-e Kongū*) and the former German Democratic Republic (GDR, *ğomhūrī-ye demūkrātīk-e Ālmān*) carry the term *democratic* in their formal names. Both of them – and several other country names – occur a few times in the corpus: the DRC produces 117 hits, the GDR 10 hits, and other countries with the official name *ğomhūrī-ye demūkrātīk* 397 hits.

### 5.5.2 Analysis of the Foreign Terms *mardombāvar/mardomsālārī*

The terminology group of the APLL has approved the Persian term *mardombāvar* as an equivalent to the foreign term *demūkrāt* and approved the term *mardomsālārī* as an equivalent to the loanword *demūkrāsī*. For *democratic peace* the terminology group approved *şolh-e mardomsālārān* (صلح مردمسالاران) and for *democratic peace theory* they approved *nazare-ye şolh-e mardomsālārān* (نظریه صلح مردمسالاران). The democratic peace theory is defined in the eighth volume of approved terms by the APLL as the theory that postulates that democratic systems are not at war with each other. The following table 30 represents the frequencies of the use of the approved terms in TalkBank and TAP.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>mardombāvar</i>	مردمباور	204	0	181	0
<i>mardomsālārī</i>	مردمسالاری	5,856	476	3,656	220

Table 30. Term frequency and document frequency for *mardombāvar* and *mardomsālārī*

In the TAP corpus, no occurrence of the approved term *mardombāvar* can be observed, and a low occurrence in TalkBank with only 204 hits. However, the approved Persian equivalent term *mardomsālārī* occurs 476 times in TAP and 5,856 times in TalkBank. Both the terms *nazare-ye şolh-e mardomsālārān* and *şolh-e mardomsālārān* produce zero hits in the two corpora.

Interestingly, a previous study on the same term pair *mardomsālārī/demūkrāsī* from 2017 reveals similar findings. There the approved term *mardomsālārī* occurs 566 times in the corpus in contrast to the foreign term *demūkrāsī*, which has 8,364 hits. The corpus used in that study represents news articles in the years from 1996 till 2007 (Hesabi, Amirian, & Nazari, 2017, p. 14). This means for the study of Hesabi et al. that *mardomsālārī* amounts to only about 6.3 percent of the combined total use (8,930 hits) of the two terms.

Even though the term *mardomsālārī* is still used significantly less than the term *demūkrāsī* in absolute terms, the abovementioned findings hint at an increased relative use of *mardomsālārī*. It occurs 6,332 times in the two corpora, while *demūkrāsī* occurs 35,114 times. This means that it has more than doubled its share since its approval by the APLL in 2010 now amounting to 15.3 percent of the combined total use (41,446) of the two terms.

Tables 29 and 30 show that the frequencies of the two equivalents approved by the APLL *mardombāvar* and *mardomsālārī* in both corpora are lower compared to those of the foreign terms *demūkrāt*, *demūkrātīk*, and *demūkrāsī*. This is especially true in the case of the borrowed term *demūkrāt* as its equivalent *mardombāvar* shows zero occurrences in TAP and a mere ten hits in TalkBank. Its weight in TAP is consequently 0 whereas in TalkBank *mardombāvar* has an adjusted weight of 1,625. The foreign term *demūkrāt* records a weight of 274,513 in TAP and of 38,594 in TalkBank. It is therefore higher than that of the approved term in both corpora.

In the same vein, the adjusted TF-IDF value of *mardomsālārī* in TalkBank is 30,534 in contrast to the adjusted weight of *demūkrāsī* of 131,672 and the adjusted weight of 57,222 for *demūkrātīk*. In TAP, on the other hand, the adjusted weight of *mardomsālārī* is 63,457 while that of *demūkrāsī* is 121,253 and that of *demūkrātīk* is 43,894.

The results for the term pairs in the semantic field of democracy are hence mixed. While it is little surprising that the approved Persian term *mardomsālārī* is heavier in newspapers than in blogs underscoring the prevalence of formally approved language in this medium, altogether the foreign terms are significantly heavier than their Persian equivalents. In TAP, *demūkrāsī* is almost twice as heavy than *mardomsālārī* and in TalkBank it dominates even at more than 4:1. The case is even more obvious with *demūkrāt*, which – due to the absence of any mention – is infinitely heavier than *mardombāvar* in TAP and still about 24 times heavier in TalkBank.

## 5.6 The Concept of *Dictatorship* and Its Derivatives

Despite the fact that the term *dīktātūr* has come to have a negative meaning in our time, its earliest use in history was not negative as it was once a position awarded to someone in times of crisis in the ancient Roman Republic. In those times, this person assumed full authority to rule the army for a specified period of, for instance, six months. Thus, while the Roman dictatorship was a legitimate institution, it was the generals who gained power through illegal means in the late era of the republic. The title of *dictator* was seized; the institution of

dictatorship was declared illegal by Sulla and Julius Caesar. From then on, the dictator was the one who ruled the government with full authority and without legal restrictions in the absence of a monarchy (Ášūrī, 1376[1997], pp. 170-171). Consequently, in the modern sense of the term, ‘Amīd defines as *dīktātūr* a ruler who leads a people or a party arbitrarily (‘Amīd, 1365[1986]b., p. 1175).

The etymology of the term goes back to the Latin word *dictātor*. It is thought to come from the verb *dictāre* that means, on the one hand, to *say repeatedly* or *speak words loudly for someone else to write down* and, on the other hand, to *issue an order*, which is probably the connotational root of *dictator*. The word *dictato(u)r* has been used in English referring to “a person granted absolute emergency power” at least since the twelfth century (Webster, 2021).

In its fourth volume of the approved terms, the terminology group chose the term *mostabed* (dictator) to replace the foreign one *dīktātūr*. In addition, *estebdādī* was endorsed as a translation of *dictatorial* while the term *estebdād* (dictatorship) was selected to be the officially approved equivalent of *dīktātūrī*. The Academy defines *estebdād* as the arbitrary and unrestrained rule of an individual or a group. The approved terms *mostabed*, *estebdādī*, and *estebdād* were imported to Persian from Arabic where *mostabed* refers to the one who makes decisions about himself and others without consulting the others. The terms have a long history in Persian and are now reused to mean *dictatorship* and its derivatives. Speaking in Bateni’s categories, *mostabed* would be a loan translation of *dīktātūr* although one could argue that *mostabed* – being an Arabic loanword to begin with – really is a loan exchange. In its guidelines, the APLL set forth, though, that Arabic terms are not considered loanwords in the same sense as European loanwords.

The terminology group also applied the term *estebdād* in a few compound words such as *estebdād-e nezāmī* (military dictatorship, استبداد نظامی) to signify autocratic military rule that is usually established following a coup. Furthermore, *estebdād-e prūletāreyā* (dictatorship of the proletariat, استبداد پرولتاریا) is defined as the rule of a society by the proletariat that wields political power in order to move towards the establishment of a state of socialism and communism.

According to Ášūrī, however, *estebdād* and *dīktātūrī* describe two distinct concepts. He translates *estebdād* in his political dictionary as *absolutism*, and he reserves *dīktātūrī* to exclusively refer to *dictatorship*. He defines *estebdād*, on the one hand, as a system characterized by the absence of traditional or legal restrictions on government power and the

extent of power used in practice (Ášūrī, 1364 [1985], p. 14). On the other hand, he defines *dīktātūrī* as a type of authority in which the ruler is not bound by the existence of laws. It is also the system in which people follow the ruler out of fear of him and not in application of the law, a system that concentrates the power in the hands of one person using terror to serve the benefits of a small group for personal gains (Ášūrī, 1364 [1985], p. 95).

Ášūrī continues to elaborate that absolutism (*estebdād*) used to exist in ancient civilizations such as in ancient Egypt, Babylon, and many more, except for the Greek and Roman civilizations where the concept of dictatorship originated. From the 16<sup>th</sup> century onwards, *estebdād* took a new form when governments and kingdoms in Europe started to revolt against papal authority. In the following two centuries, authority was then based on the ruler's right to unlimited power: absolutism. The 19<sup>th</sup> century – against the outcome of the American and French Revolutions that brought an end to absolutism – laid the foundations and led to the expansion of legitimate governments (“*hokūmathā-ye qānūnī*”) around the world in the 20<sup>th</sup> century (Ášūrī, 1376[1997], p. 25).

In several dictionaries such as *Farhang-e 'Amīd*, the terms *hūdkaṃe* (خودکامه), *hūdra'y* (خودرای), and *hūdsar* (خودسر) are listed as synonyms for the term *mostabed* ('Amīd, 1365[1986]c.). These terms will also be examined to check if they are in use to denote *dictator* or if they have other connotations. The following table 31 presents the term frequency and document frequency of all the above-mentioned terms in the semantic field of *dictatorship* combining different ways of spelling.

Term forms		Term frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>dīktātūr</i>	دیکتاتور	20,887	302	12,051	197
<i>dīktātūrī</i>	دیکتاتوری	12,215	164	7,623	108
<i>estebdād</i>	استبداد	10,664	380	6,260	188
<i>estebdādī</i>	استبدادی	3,591	70	2,520	50
<i>mostabed</i>	مستبد	3,653	101	2,686	78
<i>hūdkaṃe</i>	خودکامه	2,522	20	1,766	17
<i>hūdsar</i>	خودسر	1,617	20	1,322	20
<i>hūdra'y</i>	خودرای	177	9	139	7

Table 31. Term frequency and document frequency of the terms in the semantic field of *dictatorship*

In absolute numbers, the foreign term *dīktātūr* is clearly more frequent in the corpora than its approved equivalent *mostabed* occurring three times more in newspapers and almost six times more in blogs. When it comes to the terms *dīktātūrī* and *estebdād*, the difference is not so pronounced, but still narrowly leaning in favor of the foreign term in TalkBank. Only in the newspapers recorded in TAP does *estebdād* outnumber *dīktātūrī* by almost 70 percent vis-

à-vis 30 percent – albeit on a low level of only 544 hits among 23,000 combined hits in both corpora for both terms. The dominance of non-conforming language in blogs and of an adherence to official language in newspapers becomes evident once again.

In the study by Hesabi et al. (2017), findings show a similar result where the approved term *estebdād* (2,668 hits) is more frequent in newspapers than the loanword *dīktātūrī* (1,527) by about 63.6 percent. However, with the term pair *mostabed* and *dīktātūr* the frequency is reversed with 567 and 1,262 hits in their newspaper-based corpus respectively.

The independently proposed alternatives *hūd̄kāme* instead of *estebdād* to mean *absolutism*, *tyranny*, or *despotism*, as well as *hūdsar* and *hūdra* 'y as synonyms with the meaning of *arbitrary* and *despotic* are relatively little used in the corpora. *Hūd̄kāme* accounts for a little less than ten percent of the share between *dīktātūrī*, *estebdād*, and *hūd̄kāme* across the two corpora. As for *hūdsar* and *hūdra* 'y, the absolute numbers for both terms would even be lower as they are also used in social contexts to represent stubborn and opinionated people. They could denote the misuse of anything whether the abusive use of a drug or any other abusive act.

In TalkBank, *dīktātūr* has an adjusted weight of 86,120 while *dīktātūrī* has an adjusted weight of 55,481. In TAP, the adjusted TF-IDF score of the terms *dīktātūr* and *dīktātūrī* is 41,371 and 25,757 respectively. By contrast, the terms *mostabed*, *estebdādī*, and *estebdād* record adjusted weights of 20,078, 19,947, and 50,358 respectively in TalkBank. In TAP, the same approved Persian equivalents have adjusted TF-IDF scores of 16,961, 12,793, and 45,746 respectively. Altogether, therefore, the foreign term *dīktātūr* weighs heavier throughout the corpora than *mostabed*. For *dīktātūrī*, the result is mixed: Its weight is slightly heavier than *estebdād* in TalkBank whereas in TAP this result is reversed. The term *estebdādī*, again, is lighter than either of the foreign terms in both corpora.

## 5.7 The Concept of *Regime* and Its Derivatives

The term pair *režīm/nezām* is a special case in at least two regards. First, the APLL did not only accept the foreign term *režīm* in Persian, but also approved the Persian equivalent *nezām*. Second, both the borrowed term and the equivalent term approved by the APLL carry several different meanings in different semantic fields. The Academy approved the loanword *režīm* mostly not in a political context, but in the context of the dietary or nutrition system. It will therefore be necessary to take a closer look at both terms quantitatively and qualitatively.

According to Mo‘īn’s dictionary, the approved foreign term *režīm* means *system, procedure, or a way of ruling*. It has two distinct definitions of which only the first one carries a political meaning that refers to the type of government or the way of ruling a country, for example, a democratic system (“*režīm-e demūkrāsī*”) or dictatorial system (“*režīm-e dīktātūrī*”). It is interesting to note that Mo‘īn here uses the two terms *demūkrāsī* and *dīktātūrī* that were also replaced by the Farhangestān. The second definition refers to the food system or a diet or dietary regime that could be followed either for health reasons or to control the weight of the body (*režīm-e gāzāyī*) (Mo‘īn M. , 1383 [2004], p. 502) and (‘Amīd, 1365[1986]b., p. 1233).

Similarly, the dictionary *Dā‘erat-e el-ma‘āref-e nū* introduces several types of *regimes* only one of which refers to political systems. Under the header of *režīm-e seyāsī* one finds also *režīm-e eqtešādī*, and *režīm-e kešāverzī*. Here, *režīm-e seyāsī* is understood as a type of government that is usually called a “political system” (*sīstem-e seyāsī* (سیستم سیاسی) or *nezām-e seyāsī* (نظام سیاسی)) and can further be divided into distinct subgroups such as regimes (“*režīmehā*”) with free elections, single-party systems, and dictatorial or autocratic regimes. The economic and agricultural regimes are also included among the political regimes. The former (*režīm-e eqtešādī*) refers, for example, to a capitalist system or a planned economy whereas the latter (*režīm-e kešāverzī*) takes into account the organization, concentration, and exploitation of land ownership (Sa‘īdeyān, 1378 [1999]b., p. 2138). Sa‘īdeyān covers these types of political regimes in one paragraph while dedicating another six pages and twelve entries to various healthcare related regimes such as regimes for those with diabetes, heart diseases, high blood pressure, etc.

The Persian term *režīm* comes from the French term *régime* (Mo‘īn M. , 1383 [2004], p. 502). In French, it goes back to the 13<sup>th</sup>/14<sup>th</sup> century where it acquired both meanings almost simultaneously. The first known mention of *regimen* with the meaning of *governing* or *administering* in French occurred in 1265; by 1350, its spelling had changed to *regime*. In 1314, French saw the first mention of *regime* in the sense of *medical treatment* (CNRTL, *Regime*: subst. masc., 2012). This is an indication that both meanings of the term *regime* can be traced to old French.

While the terminology group confirmed the term *režīm* in the two mentioned fields of politics and nutrition, it is mostly applied in the field of nutrition. There are 98 compound terms such as *režīm-e čarbīzādā* (fat flush diet, رژیم چربی‌زدا), *režīm-e mīvehjārī* (fruitarian diet, رژیم

(میومخواری), *reżīm-e māye* (liquid diet, رژیم مایع), *reżīm-e porkālorī* (liquid diet, رژیم پُرکالری), and many other terms in this context, which the APLL coined based on *reżīm*.

The terminology group of the APLL accepted the term *reżīm* four times in the political field although it would usually give preference to *neẓām*. In the fifth volume of the approved terms, *reżīm-e baroftāde* (ancient regime or ancien régime, رژیم برافتاده) was listed as a system of government overthrown by revolution or great social change. In the sixth volume, *reżīm-e dastnešānde* (puppet regime, رژیم دست‌نشانده) was approved to imply a regime whose officials act in the interests of another government. The term *reżīm-e solṭānī* (sultanic regime, رژیم سلطانی) was approved in the seventh volume to indicate a system or regime in which an autocratic person rules the people through his or her bureaucracy and military power. And finally, *reżīm-e neẓāmī* (military regime, رژیم نظامی) was approved in the eleventh volume to represent a government of military rule. It is obvious that the word *reżīm* was approved with a negative connotation; at least the first three definitions of the mentioned compound words point at the government of the Shah before the Islamic Revolution.

In addition to *reżīm*, the terminology group approved the term *neẓām*, which is originally an Arabic word. The APLL has used this term aside from the approved foreign term to be able to explicitly highlight different concepts. The coined term *neẓām* is also approved in various fields, the political category being one of them. In the second volume of approved terms, *neẓām* (regime, نظام) was approved to denote an organization with a legal framework to regulate domestic and international relations. Then in the fourth volume, *estebdād-e neẓāmī* (military dictatorship, استبداد نظامی) was approved to imply a rule of military authoritarianism that comes after a coup. In the seventh volume, two compound terms were approved, the first being *neẓām-e šāyastegarā* (merit system, نظام شایسته‌گرا) to denote a system of governance that designates and promotes workers based on their capability to carry out their administrative tasks properly and impartially. The second one in said volume was *neẓām-e rīšsafīdī* (pre-state political system, نظام ریش‌سفیدی) to refer to a very rudimentary political organization headed by a chairman or council of elders. Then come the two terms that were approved in the ninth volume: *neẓāmhā-ye baynelmelal* (international regimes, نظام‌های بین‌الملل) to imply the systems that are based on a set of rules, norms, and decision-making procedures that coordinate the behavior of governments or other international actors on a particular issue, and *neẓām-e āmneyyatgostar* (security regime, نظام امنیت‌گستر) to refer to the rules and norms agreed by governments to interact with each other in order to increase mutual security.



The eleventh volume contains more than one compound word with the approved term *nezām*. To explain the temporary provision of military rule in part of or all the country, following an announcement of state of emergency, the terminology group chose *ḥokūmat-e nezāmī* (martial law, حکومت نظامی). To implement the following step of a military rule over civilian government which usually follows a civil war, the terminology group approved *ḥokūmat-e nezāmeyān* (military government, حکومت نظامیان). *Nezām-e garāyī* (militarism, نظام گرای) is an approved term to indicate the belief in and implementation of the superiority of military institutions and values above civilian institutions and values. The person who believes in militarism is *nezām garā* (militarist, نظام گرا). The adjective *nezām-e garāyāne* (militaristic, نظام گرایانه) was also approved in the same group. Next comes the military governor *farmāndār-e nezāmī* (فرماندار نظامی) followed by the type of government in which the highest authority of the country is embodied by a military person: *ḥokmrānī-ye nezāmeyān* (military rule, حکمرانی نظامیان). Finally, in the twelfth volume, one finds the parliamentary system *nezām-e maḡlesī* (نظام مجلسی).

*Nezām* could be characterized as a loan translation of *reżīm* according to Bateni's system of categorization if one accepts the term *nezām* to be an originally Persian term. The APLL defined words of Arabic origin that have long been part of the Persian dictionary as quasi-Persian. In disregard of this guideline, it could be argued that *nezām* is a loan exchange of *reżīm*, where the foreign term *reżīm* is replaced by the more well-known imported foreign term *nezām*.

Among all these approved compound terms, only three produce significant results. *Ḥokūmat-e nezāmī* scores 1,529 hits in TalkBank and 42 in TAP, *ḥokūmat-e nezāmeyān* occurs 174 times only in TalkBank, and *farmāndār-e nezāmī* yields 196 results also in TalkBank only. The terms *estebdād-e nezāmī* and *nezām-e garāyāne* occur six and four times respectively. Yet other terms show one hit in at least one of the two corpora: *nezām-e šāyastegarā*, *nezām-e garāyī*, *nezām garā*, and *ḥokmrānī-ye nezāmeyān*. Some of these compound terms, however, do not exist in either corpus and seem to play no role in practice – at least outside specialized academic contexts –, among them *nezām-e rīšsafīdī*, *nezāmhā-ye baynelmelal*, *nezām-e amneyyatgostar*, and *nezām-e maḡlesī*.

In this chapter, the approved term pair *reżīm/nezām* will be investigated in addition to the weight of each term in order to evaluate the use of both of them in the two corpora. Some of the other compound words mentioned in the previous section that relate to both approved terms will also be examined.

Term form		Term frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>reżīm</i>	رژیم	162,836	7,562	69,492	1,763
<i>nezām</i>	نظام	397,051	24,748	135,736	6,501

Table 32. Term frequency and document frequency of the approved terms *reżīm* and *nezām*

The data in the previous table 32 gives the number of hits for the terms *reżīm* and *nezām* in addition to the number of hits for the individual documents where these terms occur indicating their frequency in the corpora without repeat. What becomes clear from looking at the quantitative corpus analysis is that both *reżīm* and *nezām* occur often in the corpora. With an accumulated frequency of more than 420,000 hits, *nezām* is clearly more frequent, though, than *reżīm* with just about 170,000 hits. When looking separately at the two corpora, *nezām* occurs over three times as much as *reżīm* in TAP, but only two and a half times more in TalkBank. Assuming that the two terms are used mostly in a political context and anticipating the interpretation below that *reżīm* has a negative connotation, it could be argued that bloggers give regimes that are perceived to be negative in their opinion relatively more room than journalists and newspaper editors who might be more inclined to report about developments in positively connotated political systems (so as not to say “regimes”). It will be seen from the table of collocations below that this might be the case for the emotional topic of Israel and the Israeli-Palestinian conflict, on the one hand, and Iran, on the other, respectively.

However, these numbers do not accurately reflect the use of either term in its political context. Like for example in the case of *pārlemān* and *mağles*, the collocations of both terms *nezām* and *reżīm* will be counted with regard to the meaning of *regime* in order to control the use of both terms. The term *reżīm* comes in collocation with four distinct groups of words that will be summarized in the following. The following table 33 illustrates the collocations of *reżīm* in the TalkBank and TAP corpora.

Collocations		Frequency (number of hits)		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>šeheyūnistī</i>	صهیونیستی	60,435	4,198	34,555	893
<i>Esrāyīl</i>	اسرائیل	16,131	798	10,021	322
<i>dādgostarī</i>	دادگستری	13,815	3	13,781	3
<i>ğazāyī</i>	غذایی	9,509	433	4,521	189
<i>ašğālgar</i>	اشغالگر	3,238	207	1,923	124
<i>verzeš</i>	ورزش	2,498	34	1,092	28
<i>Qazāfī</i>	قذافی	2,242	3	1,312	3
<i>pahlavī</i>	پهلوی	2,163	81	1,532	67
<i>Sūryeh</i>	سوریه	1,264	274	1,049	130
<i>šāhenšāhī</i>	شاهنشاهی	691	12	537	12
<i>salṭanatī</i>	سلطنتی	680	1	462	1
<i>nāmašrū</i>	نامشروع	404	13	401	12

Table 33. Collocations of the approved foreign term *reżīm*

First and foremost, for *režīm*, one collocation stands out: It occurs more than 60,000 times in collocation with *šeheyūnīstī* (Zionist). Other frequent collocations also refer directly or indirectly to Israel such as *Esrāyīl*, *ašgālgar* (occupier), and *nāmašrū* (illegal), which occur more than 16,000, 3,000, and 400 times respectively.

Second, *režīm* is also used considerably with collocates that describe the former imperial regime of Iran, that is, the system of government before the Islamic Revolution. Such words are *pahlavī* with more than 2,000 hits as well as *šāhenšāhī* (imperial) and *salṭanatī* (monarchical) with roughly 700 hits each.

The third group of collocations which occur frequently with *režīm* can be summarized as generally dictatorial and authoritarian. It contains names such as *Qazāfī* – meaning Libya’s former leader Muammar Gaddafi – with more than 2,200 and *Sūryeh* – implying the Syrian regime of Bashar Al-Assad – with more than 1,200 hits.

Finally, the fourth group of collocations is made up of words that are generally not political. It includes *dādgostarī* (judicial) with more than 13,000 hits, which could refer to the judicial system in any given country (although it is frequently applied in the context of Israel) and could therefore be considered part of the system of government. In addition, the two collocations *gāzāye* with almost 10,000 hits and *verzeš* with 2,500 hits have no political implications.

A direct comparison between the sources is revealing here. Except for the collocation *šeheyūnīstī*, almost all hits occur in blogs. The number of hits in newspapers is negligible with mostly two-digit frequencies. Only the compound collocation *režīm-e šeheyūnīstī* occurs with at least four-digit results in each corpus, that is, more than 60,000 times in blogs and more than 4,000 times in newspapers. The share of *šeheyūnīstī* in TalkBank is roughly 93.5 percent vis-à-vis 6.5 percent in TAP. For *ašgālgar* it is 93.99 percent vis-à-vis 6.01 percent. In all other cases, it tilts in favor of TalkBank with at least 95 percent.

The fact that the country name *Esrāyīl* occurs a mere 798 times in TAP as a collocation of *režīm* – as opposed to more than 16,000 hits in TalkBank – might be an indication that journalists try to find other names to describe that country and its political system without mentioning it by name. It could also indicate that newspapers – being a more official news source than blogs – refrain from mentioning the name *Israel* as some sort of *damnatio memoriae*

– erasing it from public awareness – because it is not officially recognized, and its policies are negatively connotated.

Collocations		Frequency (number of hits)		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>eslāmī</i>	اسلامی	61,328	1,937	29,376	973
<i>kešvar</i>	کشور	38,473	2,199	22,075	1,216
<i>ḡomhūrī</i>	جمهوری	32,279	1,146	18,359	671
<i>Īrān</i>	ایران	30,076	1,473	25,749	876
<i>seyāsī</i>	سیاسی	21,932	331	11,888	184
<i>eqtešādī</i>	اقتصادی	10,711	989	6,293	600
<i>ámūzešī</i>	آموزشی <sup>50</sup>	8,396	341	6,322	156
<i>moqaddas</i>	مقدس	6,289	205	4,523	161
<i>bānkī</i>	بانکی	4,978	504	2,454	271
<i>pardāht</i>	پرداخت	2,884	1,867	1,961	953

Table 34. Collocations of the approved Persian term *nezām*

With regard to the term *nezām*, three distinct groups of collocations can be identified. First, the top four collocations with *nezām* are *eslāmī* (Islamic), *kešvar* (state), *ḡomhūrī* (republic), and *Īrān* (Iran) with more than 63,000 hits, more than 40,000 hits, more than 33,000 hits, and more than 31,000 hits respectively. In addition, the collocation *moqaddas* (holy, sacred) also occurs with *nezām* in an Iranian context. All of them mean the system of government in the Islamic Republic of Iran.

This positive connotation is even reflected in the name of one of the Iranian state institutions: *maḡma‘-e tašhīṣ-e mašlahat-e nezām* (expediency discernment council, مجمع تشخیص مصلحت نظام). It has a frequency of 12,032 hits in 7,473 documents in TalkBank and 655 hits in 305 documents in TAP.

Second, with reference to the banking system, the two collocations *bānkī* (banking) and *pardāht* (deposit) occur more than 5,000 and more than 4,600 times respectively. While almost all hits for *bānkī* occur in the blogs corpus, the hits are more equally divided between the two corpora for *pardāht*, with 2,800 hits in the blogs corpus and 1,800 in the newspapers corpus. Part of this second group of financial or economic terms is the collocation *eqtešādī* itself with more than 11,000 hits in both corpora.

Third, and finally, the last group contains general collocations such as *seyāsī* (political) and *ámūzešī* (educational), which more or less refer to specific types of systems or regimes at

<sup>50</sup> One of the interesting findings about this word is that it is written right according to the Persian spelling system in the newspaper-based corpus TAP as *ámūzešī* (آموزشی), while in the blogs corpus the writers are not committed to the right way as it is exclusively spelt as *amūzešī* (آموزشی).

large. The collocation *eqtešādī* from group 2 could also be more generally subsumed in group 3, leaving group 2 with banking terms only.

In conclusion, while *režīm* – not unlike the English word *regime* or the German word *Regime* – seems to have a negative connotation, the term *nezām* evokes more positive associations – always speaking from an Iranian perspective, of course. This is in line with the compound words approved by the APLL according to which compounds based on *režīm* can be considered as describing negative institutions, namely those of the Iranian imperial regime under the Shah. The connotational categorization is corroborated by the trends revealed in the collocation analysis.

A look at the absolute numbers reveals the following weight values for the two terms in question. The results are ambiguous. The approved term *nezām* carries an adjusted weight of 757,585 in TalkBank and is clearly heavier than *režīm* with an adjusted weight of 410,420. However, in TAP, *režīm* is slightly heavier with an adjusted weight of 433,642 compared to *nezām* with an adjusted weight of 426,197 although it could be argued that they are equally heavy given the ratio of 50.4 percent vis-à-vis 49.6 percent. This is interesting insofar as oftentimes newspapers favor the approved Persian equivalent over the foreign term, which is narrowly not the case for the term pair *režīm/nezām*.

It seems that the newspaper articles in the corpus deal more with what they perceive to be negative regimes. The collocate *šeheyūnistī* is the single most frequent collocate of the term pair *režīm/nezām* in relative terms. In TAP, more than 50 percent of all documents that contain the term *režīm* also contain the collocate *šeheyūnistī*. Other collocates that are frequent in absolute terms occur in a smaller share of documents. Likewise, the overall most frequent collocate in absolute terms – *šeheyūnistī* in TalkBank – occurs in only just over 20 percent of all blogs containing the term *režīm*.

The overall weight count is not entirely accurate, though, for the field of politics. It is therefore necessary to make another calculation that will approximate the weight of each term regarding their political connotation. For this purpose, the term frequencies of the abovementioned non-political collocates will be subtracted from the frequencies of both *režīm* and *nezām*. In the case of *režīm*, this means eliminating *gāzāye* and *verzeš* from the count. In the case of *nezām*, it means excluding *ámūzešī*, *bānkī*, and *pardāht* from the calculation. The approximate adjusted weight of *nezām* is then 743,554 in TalkBank and 468,423 in TAP. For *režīm*, the value is 388,321 in TalkBank and 475,712 in TAP. This approximation does not

change the overall result. The approved Persian equivalent term *nezām* is still significantly heavier in blogs than the foreign term *reżīm* while the latter narrowly beats *nezām* in newspapers.

## 5.8 The Concept of *Bureaucracy* and Its Derivatives

When it comes to analyzing the term *bureaucracy*, it is important to distinguish the political science meaning of the concept from a connotation that *bureaucracy* is often associated with in everyday life and colloquial language. In the latter sense, the term and concept have acquired a negative connotation with *bureaucracy* standing for a manifestation of unfettered authority that aims to dominate the whole society through complicated rules, an excess of regulations, and red tape. As a theoretical concept, however, bureaucracy simply describes “a body of nonelected government officials” and “government characterized by specialization of functions, adherence to fixed rules, and a hierarchy of authority” (Webster, Bureaucracy, 2021). In the same vein, according to Mo‘īn, *būrūkrāsī* is “the rule of administrative organizations over society” (Mo‘īn M. , 1383 [2004]).

The foreign term *bureaucracy* entered Persian via French. It was borrowed from the French term *bureaucratie*, itself a combination of *bureau*, which means *office* or *desk*, and the French suffix *-cratie*, which goes back to the Greek term *kratia* meaning *power* or *strength*, and has come to be used to denote a type of a government (Webster, Bureaucracy, 2021).

The first use of the term *bureaucratie* in French was evidenced before 1759 when the French economist Vincent de Gournay coined the term at some point during his lifetime (1712-1759). However, while the emergence of the term is closely linked to the development of centralized states and state administrations in 18<sup>th</sup>-century Europe – the first being absolutist France – (CNRTL, Bureaucratie, subst. fém., 2012) and (Bürokratie, n.d.), historically, according to Abūlfazl Šādeqpur, the idea and the concept of bureaucracy predate the coinage of the term *bureaucracy* by millennia. A look at the great human civilizations such as those of ancient Iran, ancient China, and ancient Egypt reveals that the administration of such governments would not have been possible without the use of large administrative organizations – or bureaucracies (Šādeqpur, 1369 [1989], pp. 2-3) as cited in (Reżā-Zālī, 1374[1995], S. 1).

The APLL approved the term *dīvānsālārī* to replace the foreign term *būrūkrāsī* in the second volume of approved terms. It is defined by the Academy to denote a hierarchical administrative system used to administer a country’s affairs (Terminology group, 1384 [2005]).

In Persian, the term *dīvān* describes a higher administrative unit such as a government agency or the Court (with capitalized C) in a monarchy headed by the grand vizier or prime minister tasked with running a country's affairs (Sa'īdeyān, 1378 [1999]b., p. 2046). Another meaning of *dīvān* includes the place where the rulers gathered and sat down with their advisers in the sense of *maḡles-e šūrā* (Nafīsī, 1355[1977], p. 1579). Etymologically, it is widely accepted that the term *dīvān* is of Pahlavi origin and used to describe a collection of manuscripts, an archive, or a registry already in Iran's pre-Islamic era. The Old Persian prefix *dipi-* (inscription, document) can be traced via Akkadian *tuppu* to Sumerian *dub* (clay tablet). It may have evolved as *dipi-vān* (relating to documents) throughout the Middle Persian era and perhaps in exchange with Armenian before it was contracted to *dīvān* and (Tabrīzī M. Ḥ., 1342 [1963]b., p. 918) and (De Blois, 1995). As for the suffix *-sālārī*, the terminology group of the APLL follows the same pattern in coining similar terms by using the term *sālārī* to replace the suffix *-cracy* like with *mardomsālārī* instead of *demūkrāsī* and *yekesālārī* instead of *otūkrāsī*. *Dīvānsālārī* is therefore a loan translation according to Bateni's system of categorization.

The representative of *bureaucracy* is the *bureaucrat*, a term that entered Persian as *būrūkrāt*. The terminology group of the Academy accepted the Persian equivalent *dīvānī* to replace *būrūkrāt* in the second volume of approved terms. *Dīvānī* in the sense of *bureaucrat* is a combined loan translation and rendition where the morpheme *dīvān* refers to the prefix *bureau-* and where the suffix *-crat* is explained by the suffix *-ī*. The term *dīvānī* occurred in Persian dictionaries long before the Academy formally approved it. Dehḡodā refers to it as someone holding a position in the government apparatus, a government employee. In this sense, the APLL approval process added the meaning of *bureaucrat* to the dictionary entry of *dīvānī*. (Dehḡodā, 1377 [1998]b.)

The following table presents the term frequency and document frequency of the foreign terms *būrūkrāsī*, *būrūkrāt*, and *būrūkrātīk* and their approved Persian equivalents *dīvānsālārī* and *dīvānī* in the two corpora. The term pair *būrūkrāsī/dīvānsālārī* does not record a high frequency in either corpus. The highest use is counted in TalkBank with 1,181 hits. No other of the related terms in either corpus reaches a four-digit frequency.

In the direct comparison between the terms, it can be observed, on the one hand, that the foreign term *būrūkrāsī* is more frequent in absolute numbers than the approved equivalent *dīvānsālārī*. What stands out is the fact that in newspapers the foreign term is used only approximately five times more than the approved term whereas in blogs *būrūkrāsī* produces ten

times more results than *dīvānsālārī*. Even though the term is not in wide use to start with, this suggests that newspaper editors and journalists as opposed to bloggers and independent writers might feel more compelled or comfortable to use the “official” approved term.

*Dīvānī* on the other hand is almost nine times more frequent than *būrūkrāt* in TalkBank and almost twice as frequent in TAP. While again this shows a relative preference of using approved terms in newspapers, it also shows that the approved Persian equivalent is more frequent.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>būrūkrāsī</i>	بوروکراسی	1,181	73	865	62
<i>būrūkrāt</i>	بوروکرات	73	10	62	8
<i>būrūkrātīk</i>	بوروکراتیک	244	20	184	17
<i>dīvānsālārī</i>	دیوانسالاری	118	14	104	14
<i>dīvānī</i>	دیوانی	632	19	479	15

Table 35. Term frequency and document frequency of the terms in the semantic field of *bureaucracy*

As for the weight of all terms, the following values can be calculated. The adjusted TF-IDF value of *dīvānsālārī* is 991 in TalkBank and 580 in TAP. The foreign term *būrūkrāsī* weighs 7,715 in TalkBank and 12,817 in TAP thus clearly outweighing the approved Persian equivalent. For *būrūkrāt* and *būrūkrātīk* the results are less clear. They yield a weight of 653 and 1,939 in TalkBank respectively and 2,439 and 4,375 in TAP respectively. Here, the adjusted TF-IDF value of the Persian equivalent *dīvānī* is 4,470 in TalkBank and 4,236 in TAP. The weight of *dīvānī* in both newspapers and blogs is therefore heavier than that of *būrūkrāt(īk)*.

## 5.9 The Concept of *Ideology* and Its Derivatives

Another French loanword that entered Persian and is in wide use is *īde'ūlūzī* (ideology, sometimes spelt *īdeyūlūzī* in Persian). It is taken from the French word *idéologie*, which consists of the morpheme *idé(e)* meaning *thought* or *vision*, and the suffix *-logie* (from Greek *logos*), which refers to *knowledge*. According to Ášūrī, this term was introduced first during the French Revolution by the French thinker Destutt de Tracy, who used this term to denote the science of knowledge. More specifically, it referred to mission-driven knowledge with the goal of serving human beings and even saving them from ignorance by removing preconceptions and training them to accept and submit to reason. The spread of *idéologie* as a term and as a thought was not limited only to the French. Rather, it continued to spread well into the 19<sup>th</sup> century, which in itself became famous as an era of ideology (“‘ašr-e īde'ūlūzī”) when this term was reproduced



extensively in many countries and became a common word in other languages such as German, English, and Italian to name but a few (Ášūrī, 1376[1997], pp. 52-53).

It was only by the mid-19<sup>th</sup> century that the term acquired the meaning of an inspiring political doctrine (CNRTL, *Idéologie*: subst. fém., 2012), which is arguably the dominant connotation nowadays. Consequently, Mo‘īn defines it as a type of intellectual and philosophical systems including religious systems that effectively determine the policies, practice, or stance of individuals who believe in socio-political matters (Mo‘īn M. , 1386 [2007]a., p. 413). Such early politico-economic ideologies – or worldviews – would be Communism, Socialism, Capitalism, and Marxism, as well as Liberalism.

The Persian term *marām* is originally an Arabic noun. The Arabic term *al-marām* means *goal, target, objective, and purpose*, but not *ideology*, for which Arabic uses the loanword *aydiyūlūġiyā* itself. According to Mo‘īn and Dehḡodā, *marām* has the same meaning in Persian as in Arabic (Mo‘īn M. , 1386 [2007]d.) and (Dehḡodā, 1352[1973], p. 95). It is therefore not exclusively limited to the political field.

However, the Farhangestān appropriated the foreign term *marām* and added to it the meaning of *ideology* in Persian. This could be considered a loanshift in Bateni’s categorization. Its meaning in the sense of ideology illustrates the concept of the systematic set of beliefs and fixed ideas to explain or change the current situation and to prescribe specific social behavior as explained in the second volume of approved terms (Terminology group, 1384 [2005]), which corresponds to the following meanings: ideology, creed, and persuasion.

On a side note, in her master thesis, Mona Valinezhad introduces *ġahānbīnī* (“world vision”) as a third term in use as another proposed equivalent of *īde’ūlūżī* and *marām*. She conducted a survey among 22 translators, political scientists, and linguists, 21 of whom answered the question on their preference given the terms *īde’ūlūżī*, *marām*, and *ġahānbīnī*. Sixteen of the participants voted in favor of the loanword *īde’ūlūżī*, four opted for the neologism *ġahānbīnī*, and only one endorsed the approved term *marām* (Valinezhad, 2016).

The term pairs *īde’ūlūżī/marām* and *īde’ūlūżīkī/marāmī* differ from most other cases in this thesis since the terminology group of the APLL approved the two foreign terms side by side with their Persian equivalents. The terms *īde’ūlūżīkī* and *marāmī* serve as descriptions for anything related to *īde’ūlūżī* and *marām* respectively. Both term pairs were approved in the second volume of approved terms in the spelling with *hamza* as indicated by the transliteration

(*īde 'ulūžī/kī*). However, the approved spelling is only used in newspapers where more than 97 percent of all occurrences of *īde 'ulūžī* apply the spelling with *hamza*. The blogs in TalkBank, by contrast, use exclusively the spelling *īdeyūlūžī* with the letter *yāy* instead of *hamza*. This is another indication of a more formal and Farhangestān-friendly approach in newspapers whatever their political leaning.

However, the APLL has not yet approved a Persian term to denote the person who propagates a certain ideology. The term used in Persian for this concept is *īde 'ulūg* (ایدئولوگ) (Ášūrī, 1376[1997], p. 55), which is a loanword based on the French *idéologue*. The loanword *īde 'ulūg* has only 14 hits in seven documents in TAP. Searching for this term in TalkBank does not produce any result. It is therefore negligible.

The following table 36 presents various examples from both corpora with the examined terms.

<b>Example 1</b>	ایران در دهه 60 شمسی علاوه بر تهاجم خارجی و درگیر شدن با یک جنگ بزرگ و تحمیلی، محل تاخت و تاز گروهک ها و سازمانهای مسلح و تروریستی بود که قصد داشته و دارند که ایدئولوژی ادعایی خود را به زور اسلحه به مردم تحمیل کنند.
Transliteration	Īrān dar dah-e 60 'alāve bar tahāgom-e ḥāreḡī va dargīr šodan bā yek ḡang-e bozorg va taḥmīlī, maḥal-e tāht va tāz-e gorūhakhā va sazmānhā-ye mosalaḥ va terūrīstī būd ke qaṣd dāšte va dārand ke īde 'ulūžī-ye ede 'āyī-ye ḥūd rā be zūr-e asleḥe be mardom taḥmīl konand.
Translation	In the 1980s, in addition to foreign aggression and engaging in a major and imposed war, Iran was invaded by armed and terrorist groups and organizations that sought to impose their alleged ideology on the people by force.
<b>Example 2</b>	عجیب نیست که در زمان ریاست لاریجانی بر صدا و سیما بیشتر اندیشه‌های هایدگر ترویج می‌شد و لو گاه با لعاب‌های ایدئولوژیک.
Transliteration	'Aḡīb nīst ke dar zamān-e reyāsāt-e Lārīḡānī bar ṣedā va sīmā bīštar andīšehā-ye Heidegger tarviḡ mīšod va lū ḡāh bā la 'ābhā-ye īde 'ulūžīk.
Translation	It is not surprising that during Larijani's presidency, most of Heidegger's ideas were promoted on radio and television, sometimes with ideological glazes.
<b>Example 3</b>	بنابراین از اینجا به بعد دیگر باید مرام و مسلک رییس جمهور را از یک مدیر اجرایی عالی رتبه هم فراتر برد و او را در هیئت یک مکتب فکری یا سیاسی یا دست کم مدیریتی دید و بر اساس آن تحلیل و ارزیابی داشت.
Transliteration	Banābarīn az īnḡā be bā 'd-e dīḡar bāyad marām va maslak-e rayīs-e ḡomhūr rā az yek-e modīr-e eḡrāyī-ye 'ālī-rotbe ham farātar bord va ū rā dar haybat-e yek-e maktab-e fekrī yā seyāsī yā dast-e kam-e modīreyatī dīd va bar asās-e ān taḥlīl va arzeyābī dāšt.
Translation	Therefore, from now on, the ideology and profession of the president of the republic should be surpassed by a high-ranking executive director, and he should be seen in the awe of an intellectual or political school, or at least a managerial school, and analyzed and evaluated accordingly.

<b>Example 4</b>	ورود اصل شوراها به قانون اساسی جمهوری اسلامی ایران را میتوان تبلور مشی و مرام بنیانگذار فقید انقلاب اسلامی، حضرت امام خمینی(ره) و اعتقاد ایشان به میزان بودن رای ملت دانست.
Transliteration	Vorūd-e aṣl-e šūrāhā be qānūn-e āsāsī-ye ġomhūrī-ye eslāmī-ye Īrān rā mītāvān tabalvar-e mašī va marām-e bonyāngoẓār-e faqīd-e enqelāb-e eslāmī, ḥāzrat-e emām-e Khomeini (rh) va e' teqād-e īšān be mīzān būdan-e rā'ye mellat dānest.
Translation	The entry of the principle of councils into the constitution of the Islamic Republic of Iran can be considered as the crystallization of the policy and ideology of the late founder of the Islamic Revolution, Imam Khomeini (RA) and his belief in the extent of the nation's vote.
<b>Example 5</b>	استثنائگرایی در اندیشه و مرام تمدن جدید و بعضا موسسان آمریکای جدید دیده میشود که حاکی از روحیه نژادپرستی آمریکایی است.
Transliteration	Estetnāgarāyī dar andīše va marām-e tamadon-e ġadīd va ba'zā mo'asesān-e Āmerikā-ye ġadīd dīde mīšavad ke ḥākī az rūḥīh-e neẓādaparastī-ye Āmrīkā-ye ast.
Translation	Exceptionalism is seen in the thought and ideology of the new civilization and sometimes the founders of the new America, which reflects the American spirit of racism.
<b>Example 6</b>	اما برجستگی اصلاح طلبانه در انتخابات ماه مارس می‌تواند مرام و ایدئولوژی محافظه کارانه داخلی وی که سیاست‌های بین‌المللی را متأثر می‌سازد و بسیاری آن را با سال‌های اول انقلاب مقایسه می‌کنند به چالش بکشد.
Transliteration	Ammā bargastegī-ye eṣlāḥ-ṭalabāneh dar entehābāt-e māh-e māres mītāvānad marām va īdeyūlūzī-ye moḥāfezeh karāneh-ye dāḥelī-ye vey ke seyāsathā-ye baynolmelālī rā mota'āter mīsāzad va besyārī-ye ān rā bā sālhā-ye avval-e enqelāb moqāyaseh mīkonand be čāleš bekašad.
Translation	But the reformist prominence in the March election could undermine his conservative domestic ideology [marām va īdeyūlūzī] that influences international politics, which puts him in the challenging situation that people will compare him with the first years of the revolution.
<b>Example 7</b>	از سوی دیگر نگاهی به ایدئولوژیهای غیرمذهبی و طرز تلقی مردم ایران نسبت به آنان از عدم موفقیت این گونه مرام‌ها در این سرزمین حکایت می‌کند.
Transliteration	Az sū-ye dīgar, negāhī be īdeyūlūzīhā-ye ġayremāzhabī va tarz-e talaqī-ye mardom-e Īrān nesbat be ānān az 'adammovaffaqeyat-e īn ġūneh-e marāmhā dar īn sarzamīn ḥekāyat mīkonad.
Translation	From another perspective, a look at the non-religious ideologies and the attitude of Iranians towards them reflects the failure of such ideologies in this land.

Table 36. Examples of *īde'ulūzī* and *marām* in the corpora

A few observations could be made from looking at *īde'ulūzī* and *marām* in practical use. First, one of the collocations of *marām* is the term *maslak* (persuasion, مسلک). They occur together 1,416 times in 1,409 documents in TalkBank and only five times in TAP (see example 3 in table 36 above). This is noteworthy because *maslak* is the only high-frequency collocates of *marām* although it is used in a variety of contexts not limited to the political sphere. Second, it is worth mentioning that both *īde'ulūzī* and *marām* are mentioned in the same sentence 46 times in TalkBank (see examples 6 and 7). Even though in TAP the two terms do not meet at all in the same text, the combined use of both terms linked through “and” or “or” points at the possibility that they are used synonymously and interchangeably by some authors (for reasons of style) regardless of the fact whether one of them is approved or not. This is also a hint at the

lack of awareness on the part of those authors about the conflict that the Farhangestān as a governmental institution faces when it comes to the coining of neologisms.

A reading of the search results shows that the most common use of *marām* oscillates between two connotations. The first one relates to creed and (religious) persuasion. The second one has to do with political or social ideology.

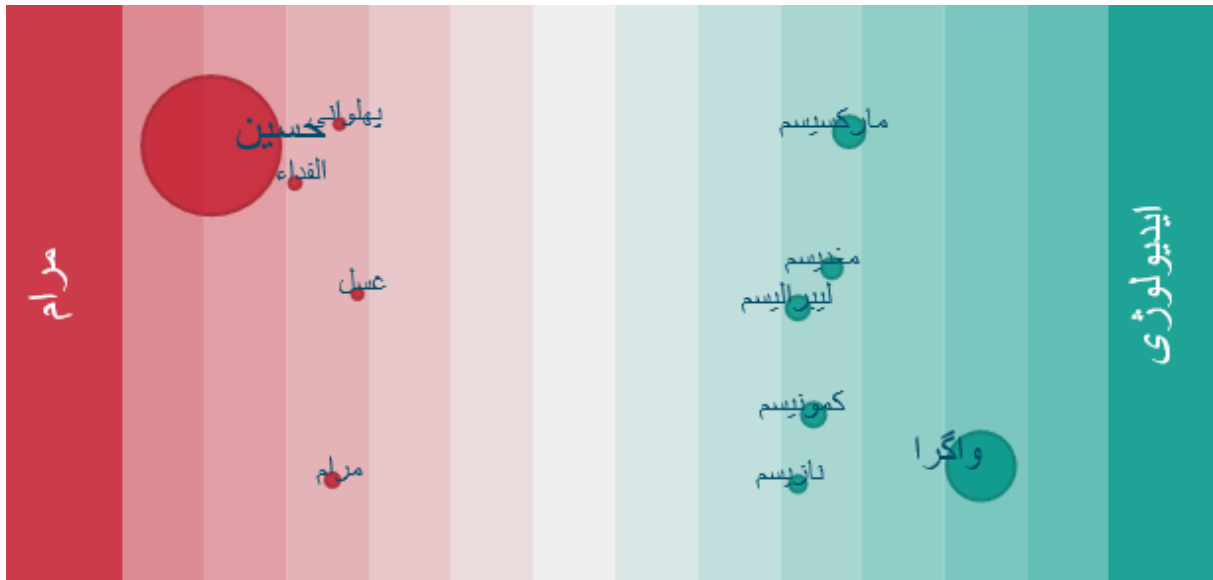


Figure 4. Visualization of the frequency of collocations compared between *ide'ulūzī* and *marām* in TalkBank

Figure 4 above visualizes the comparison of the use of both *ide'ulūzī* and *marām* regarding their collocations in TalkBank. Depending on the collocates' positions in the chart and the size of the circle, they are more closely associated with one node or the other. The further on the right side and the bigger the circle, the more a collocation comes with *ide'ulūzī*; the further on the left, the more a collocation occurs with *marām*. As can be seen from the chart, *ide'ulūzī* is frequently used to refer to ideologies that have existed throughout history such as Marxism, Liberalism, Communism, and Nazism. The term Maghdism, which occurs quite often with *ide'ulūzī*, seems to refer to a specialized discussion in a blog in TalkBank that distorts the results here. The most frequent collocate of *ide'ulūzī* is, however, *vāgerā* (divergent), which indicates that oftentimes bloggers denounce an ideology as divergent from one they presumably consider better.

*Marām*, on the other hand, frequently occurs in collocation with religious terms such as *ḡosl* (full ablution, ritual purification) or *al-fedā'* (redemption). The most frequent collocate is *Ḥoseyn* referring to the Shiite Imam venerated by many pious Iranians. These collocations point at a secondary meaning of *marām*. Not only does this approved term serve as an equivalent to

*īde`ūlūzī*, but it is also a term in its own right and in this sense means *ritual*. Another frequent collocate is *pahlavānī* (heroic) referring to heroic ideology often in sports.

The results for the foreign and approved terms can be found in table 37 below summarizing different spellings to the same effect in one row each.

Term forms		Term frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>īde`ūlūzī</i>	ایدنولوزی	7,143	494	3,834	172
<i>īde`ūlūzīk/ī</i>	ایدنولوزیک	4,231	460	2,613	214
<i>marām</i>	مرام	4,951	195	4,457	141
<i>marāmī</i>	مرامی	379	10	355	9

Table 37. Term frequency and document frequency of the terms in the semantic field of *ideology*

Both *īde`ūlūzī* and *marām* occur frequently in the corpora with a majority in favor of the foreign term in each corpus. While in TAP *īde`ūlūzī* makes up roughly 71.7 percent of the total number of hits for both terms, in TalkBank the majority is not so pronounced with a mere 59.1 percent for *īde`ūlūzī*. One interesting observation is that, on average, the two approved foreign terms occur between 1.6 and 2.9 times per document while the two approved Persian equivalents count only between one and 1.4 hits per document. In other words: When writers speak about *īde`ūlūzī* they do so a lot. This leads to the noteworthy reversal of term frequency and document frequency for *īde`ūlūzī* and *marām*. There are more documents in both corpora containing the term *marām* than there are containing the term *īde`ūlūzī* although the latter term is counted more often.

The adjusted weight of *īde`ūlūzī* and *īde`ūlūzīk/ī* is 36,935 and 23,361 respectively in TalkBank and 69,914 and 61,749 in TAP respectively. However, the approved terms *marām* and *marāmī* show low adjusted weight values compared to the foreign terms of 24,918 and 2,785 respectively in TalkBank and 28,891 and 2,400 in TAP respectively.

This does not yet account for a distinction in the use of *marām*, though. If one excludes the instances where *marām* refers to *ritual* instead of *ideology*, its weight is still less. The approved Persian equivalent *marām* has an approximate adjusted weight in a purely political sense – that is, discarding the collocates *Hoseyn, al-fedā`*, *pahlavānī*, and *gosl* – of 21,775 in TalkBank and 27,782 in TAP.

Clearly, in newspapers, the foreign terms *marām* and *marāmī* are between 2.5 and 25 times heavier than their Persian equivalents. The difference is not so pronounced in blogs where the foreign terms are only between 1.7 and 8.4 times heavier.

This suggests, on the one hand, that the discussion of religious rituals is more prevalent in blogs. On the other hand, bloggers generally seem to be more actively involved in discussing matters of ideology. But in a way, the low frequency and weight of *marām* in newspapers is counterintuitive compared to other findings where the officially licensed journalists would usually submit to the officially approved way of writing.

## 5.10 The Concept of *Terrorism* and Its Derivatives

Like so many other borrowed terms in Persian, *terūr* entered the Persian lexicon via French. Derivatives of *terūr* (terror) are *terūrīsm* (terrorism), *terūrīst* (terrorist), and the verb *terūr kardan* (to assassinate – political killing with equipment), all of which go back to the French *terreur* (Dehḥodā, 1343[1964], p. 636). Alternatively, Ášūrī uses the Persianized loan blending *terūrgarī* instead of *terūrīsm* for *terrorism* (Ášūrī, 1376[1997], p. 98) long before the Farhangestān approved an equivalent to the foreign term. The term *terūrgarī* does not produce any hits in the two corpora, though, and can therefore be considered anecdotal.

Etymologically, the terms *terror* and *terrorism* derive from the Latin term *terror* indicating a feeling of threat or intimidation (CNRTL, *Terrorisme*, 2012). The concept of *terror* gained wide recognition in Western Europe in the 1790s during the French Revolution (Ṭayyeb, 1380 [2001], p. 55), specifically during the so-called Reign of Terror (“ḥokūmat-e terūr” (Ášūrī, 1376[1997], p. 99)) in 1793/1794 – what Sa‘īdeyān calls the “era of terror” or “era of assassinations” (“dūre-ye terūr”, “dūre-ye vaḥšatī”) (Sa‘īdeyān, 1378 [1999]a., p. 1180) – when the French revolutionaries of the day persecuted their domestic enemies and had tens of thousands guillotined. The term and concept have since then expanded in meaning and scope where *terror* refers to either “governmental intimidation of citizens” or “clandestine attacks on governmental targets by domestic opponents” (Tilly, 2004, pp. 8-9). The transition from *terror* to *terrorism* is fluid. Where “[t]error has become a practice[, i]t has truly become terrorism, the systematic and primary use of randomly focused violence by organized groups against civilian targets to effectuate a political objective” (Forte, 1986, p. 42).

In Persian, too, *terūr* means scaring or terrifying others. In addition, it refers to terrifying or illegal political actions by a government to scare its opposition. It seems that especially during the Constitutional Revolution of 1906 several groups in Iran that can be identified with left-wing or right-wing terrorism followed the examples set in France and used assassinations of political figures as their method to meet political ends. In Iranian history and throughout the

Middle East, many groups were in this sense considered terrorist, among them the Nizari Ismailis under Hassan Sabah, known as the Assassins (Ášūrī, 1376[1997], pp. 98-99).

‘Amīd draws from the Latin roots of *terūr*, which he defines as an act of terrifying and scaring innocents, and he builds on the historical development of the term when he continues his definition by indicating that *terūr* is also a premeditated action to kill opponents. Furthermore, he describes *terūr* as an act of dreading and causing panic as well as an attack and sudden kill of someone without giving them the chance to defend themselves (‘Amīd, 1365[1986]a., pp. 676-677).

*Terūrīsm* is outlined by Áqābaḥšī and Afšārīrād as the systematic use of terror (“kārbord-e sīstemātik-e *terūr*”), especially as a means of coercion. It is thus reminiscent of the Western definition mentioned above. In addition, the two Iranian authors see in *terūrīsm* the struggle of revolutionary intellectuals separate from the masses (Áqābaḥšī & Afšārīrād, 1383[2004], p. 677). ‘Amīd defines *terūrīsm* as a means of using terror to achieve a specific end for political or ideological reasons (‘Amīd, 1365[1986]a., p. 677).

Consequently, according to ‘Amīd, a *terūrīst* is a person who uses the method of killing and threatening people and creating fear and panic in any way to achieve his or her political goals, such as changing the government or taking control of affairs (‘Amīd, 1365[1986]a., p. 677). In the same vein, Áqābaḥšī and Afšārīrād identify as *terūrīst* an individual person or a group who uses violence and threats to advance their political goals. This category of people is usually planning to overthrow the current system. In their definition, they set two distinct criteria to differentiate between terrorists and freedom fighters (“mobārez-e rāh-e ázādī”). First, the ethics of war are different for adherents to each group. Second, revolutionaries and freedom fighters usually fight against an oppressive regime while, in this categorization, terrorists are allied with such oppressive regimes (Áqābaḥšī & Afšārīrād, 1383[2004], p. 678). In any case, it becomes clear from the academic discussions outlined above that any characterization of terrorism is almost certainly political, and it is hard to come by an objective definition beyond the purely etymological derivation.

Going more into detail, Sa‘īdeyān mentions two specific types of terrorism. The first one is *terūrīsm-e dūlatī* (state terrorism, تروریسم دولتی), which explains the interference of a government or governments in the internal or external affairs of another government with the aim of spreading terror through military operations in order to weaken and overthrow that government or its leadership (Sa‘īdeyān, 1378 [1999]a., p. 1181). The second expression is

*terūrīsm-e kūr* (blind terrorism, تروریسم کور), which is carried out through a set of operations to propagate and create a state of chaos among people in favor of the group or movement that acts on it (Sa'īdeyān, 1378 [1999]a., p. 1181).

How does this relate now to the approved terms *vahšatafkanī* and *vahšatafkan*? This term pair consists of two parts each: the noun *vahšat* and the present stem of the verb *afkandan*, *afkan*. The word *vahšat* does not necessarily have a political meaning. Like *terror*, it means scare, panic, and fear. It could also carry the meaning of loneliness or the fear of being lonely (Dehḥodā, 1350 [1971], p. 134) as well as grief, homesickness, pessimism, or suspicion depending on the context (Mo'īn M. , 1386 [2007]d., p. 3990) and ('Amīd, 1365[1986]c., p. 2451). The terminology group of the APLL approved the term *vahšat* to replace *terūr*. It is defined as great fear caused by politically motivated violence (Terminology group, 1384 [2005]).

The verb stem *afkan* also carries various meanings such as launch, throw, and toss as well as cancel, plunk, and project. It goes back to the Middle Persian verb *afgandan* (افگندن-) (اوگندن) (Mo'īn M. , 1386 [2007]a., pp. 318-319) or *apakandan* with the meaning of *andāḥtan* (to cast, to throw, انداختن) and *rahā kardan* (to abandon, to release, رها کردن) (Hassandoust, 1393[2014]a., p. 252). The terms *vahšatafkan* and *vahšatafkanī* are therefore best translated as “creating panic” or “projecting fear” and could be characterized as a combined loan translation and rendition in Bateni's categorization because they translate the meaning of *terror* as *vahšat* while adding to it the explanatory suffix *-afkan*.

*Vahšatafkan* – the spreader of fear and panic – was accepted by the APLL in the second volume of approved terms as an equivalent to *terūrīst* to indicate either a person or a group or organization that willfully causes panic in others to achieve a political goal. In the same group of derivatives, the terminology group approved *vahšatafkanī* to replace *terūrīsm* as an act of inducting extreme fear using violent means, including kidnapping, hostage taking, and murder with the aim of achieving political objectives. By contrast, the verb *terūr kardan* was replaced by the terminology group using instead *be-qatl rasāndan* (to assassinate) although the Academy in this case approved both the borrowed verb and the coined verb. *Terūr* was also approved alongside *qatl-e seyāsī* to mean *assassination* (Terminology group, 1384 [2005]).

The morpheme *afkan* is usually used in Persian to build compound words. Such is the case with *nefāq afkan* (hypocrite), *'aksafkan* (photographer), *bombafkan* (bomber), *selāḥafkan* (someone who gives up his or her weapon). An in-depth reading of the dictionaries of approved



terms reveals that the terminology group approved many terms that contain this morpheme in different fields. Aside from *vahšatafkan* there are, among others, *vahšatafkan-e rāyāneh* (cyber terrorism), *būmvahšatafkanī* (bioterrorism, ecoterrorism), and *aždarafkan* (torpedo launcher), all of which have implications in the political-military field, as well as, for example, *nūrafkan* (projector) and *garehafkanī* (complication, as in the plot of a theater play or literary story).

The borrowed terms *terūrīst* and *terūrīsm* are used quite frequently in the Persian language like in many of the world's languages that use phonetically very similar words based on the original Latin and French etymology. The following table illustrates the number of times both of these terms were mentioned in TalkBank and TAP in addition to the document frequency in which the terms were mentioned only once without repeat.

Term forms		Term frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>terūrīst</i>	تروریست	28,897	1,226	21,919	637
<i>terūrīsm</i>	تروریسم	15,069	1,090	8,837	562
<i>terūrīstī</i>	تروریستی	38,749	2,127	19,352	802
<i>vahšatafkan</i>	وحشت‌افکن	3	0	3	0
<i>vahšatafkanī</i>	وحشت‌افکنی	2	0	2	0

Table 38. Term frequency and document frequency of the terms in the semantic field of *terrorism*

Both *terūrīst* and *terūrīsm* as well as their derivative adjective *terūrīstī* have a high frequency each in the corpora considering the difference in size between both used corpora. Interestingly, while the terms *terūrīsm* and *terūrīstī* seem to be spread almost evenly across the two corpora (at ratios of roughly 1 in 31,000 words and 1 in 13,000 words respectively), the term *terūrīst* shows an overweight in blogs. Perhaps bloggers are more at liberty or simply more willing than journalists and newspaper editors to identify people or groups as terrorist. Or else, perhaps blogs are just a more open platform to exchange political arguments and publicize matters of political dissent. In any case, if bloggers or journalists choose to speak about terrorism and terrorists, they almost certainly do so by using the foreign term and not the approved Persian equivalent. No other term pair in this thesis even remotely shows this kind of dominance of the foreign term against the approved term.

Both *vahšatafkan* and *vahšatafkanī* do not occur even once in TAP and are virtually non-existent in TalkBank with only three and two hits respectively. By contrast, only the word *vahšat* occurs 14,191 times in TalkBank and 541 times in TAP indicating usages that do not necessarily carry the meaning of political terror or terrorism. Other terms that contain the suffix *-afkan* are mentioned in the corpus. Among them one finds *farafkanī/farafkāneh* (projecting emotions, فرافکنی/فرافکانه) with 2,614 in TalkBank and 213 in TAP, *sarafkan/sarafkandegī*

(shame, humiliation, contempt, سرافکن/سرافکنندگی) with 1,061 hits in TalkBank and 30 in TAP, *bombafkan* (bomber, بمب افکن) with 704 in TalkBank and 29 in TAP, *nūrafkan* (projector, نورافکن) with 366 in TalkBank and 43 in TAP, as well as *herāsafkan* (terrify, هراس افکن) with 67 in TalkBank and 25 in TAP.

In reviewing the literature, most secondary studies join in the conclusion that the Academy did not succeed in finding a popular equivalent for the borrowed term *terūrīsm*. The approved term *vaḥšatafkanī* did not show any success factors in most of the previous studies whether it was based on a quantitative corpus analysis or on a questionnaire among students of a specific category or professional translators. For instance, Mahdi Aghili found that *vaḥšatafkan* as an approved neologism was among the least acceptable coinages considering the effect of the level of education on the extent of transparency and acceptability of coinages among the participants in his survey. Participants were Iranian university students of English translation departments and with non-English majors both on the master and bachelor levels. In their replies, the approved term did not show any elements of success. That is, some of the interviewees did not even know the term whereas those who did said that it was not a term they used or would use (Aghili, 2010).

Akbar Hesabi et al. advocate the same results. The approved term *vaḥšatafkan* has a frequency of zero in the corpus that they used in their study in contrast to 10,626 hits for the foreign term *terūrīsm* (Hesabi, Amirian, & Nazari, 2017). Finally, another study by Azimeh Sattari matches the results observed in earlier studies. Her results on the approved term pair *vaḥšatafkan/vaḥšatafkanī* indicate the failure of the two terms in radio programs considering a variety of genres such as news and interviews (Sattari, 1388 [2009]).

The loanword *terūrīst*, as can be seen in table 33 above, appears 28,897 times in 21,919 documents in TalkBank which means that the adjusted TF-IDF weight of the term is 103,334 while it records an adjusted weight of 119,947 in TAP. The adjusted weight of the derivative adjective *terūrīstī* in TalkBank is 142,979 and in TAP is 191,747. The foreign term *terūrīsm* has an adjusted weight of 66,407 in TalkBank and 111,197 in TAP. To sum up, *terūrīst* is almost 2,952 times heavier than *vaḥšatafkan* with its adjusted weight of 35 and *terūrīsm* 2,767 times

heavier than *vahšatafkanī* (24), not to speak of the absence of any weight for the two approved Persian equivalents in TAP<sup>51</sup> – a clear success for the two loanwords.

### 5.11 The Concept of *Populism* and Its Derivatives

The borrowed term *pūpūlism* entered Persian also via the French language. It was first mentioned in French in 1912 referring to a political doctrine with origins in Russia, the United States, and Germany in the late 19<sup>th</sup> century and goes itself back to the Latin *populus*, which means *people* (CNRTL, *populisme*, 2012) (cf. (CNRTL, *populiste*, adj., 2012) . It garnered renewed attention in the 2000s and 2010s with developments all across Europe and elsewhere where often right-wing, anti-Islamic populist political parties or individuals managed to gain public support and win elections such as the FPÖ in Austria, the Front National in France, the PiS in Poland, Cinque Stelle in Italy, or Alternative for Deutschland in Germany, as well as Donald Trump in the United States to name but a few.

Published in the sixth volume of the dictionary of the approved terms, the term *pūpūlism* was replaced by the terminology group of the APLL using the coinage *‘avāmgarāyī*. In the same volume, *‘avāmgarā* was introduced as a replacement for the adherent of populism: *pūpūlīst*.

The approved coinage consists of two parts: the Arabic noun *‘avām* (common people, عوام), which is the plural form of the singular Arabic noun *al-‘amma* (the public, العامة) and carries the same meaning in Persian as well (Deḥodā, 1342(1963), volume 85, p. 415). The second part *garā* is the present stem of the verb *garāyīdan* (گراییدن), which means *to desire* or *to tend to/care for*. Consequently, the approved term could be translated as the care for the common people. In this sense, it is a loan translation in Bateni’s categorization if one accepts the lexical meaning of *populism* to be the political ideology (*ism*) that promotes the cause of the common people. If one compares the etymologies of both terms, though, the Persian equivalent could also be considered a loan rendition where *‘avām* is equivalent to the prefix *popul-* and *garāyīdan* explains the context.

In practice, the term *‘avāmgarāyī* denotes a movement or philosophy that seeks to attract the masses and shows traces of demagoguery. The term *‘avāmgarā* implies the one who believes in populism. Like this concept, the Academy has also approved *‘avāmfarībī* (demagoguery,

<sup>51</sup> It is not possible to calculate the weight in these two cases because the calculation according to the TF-IDF formula would require dividing by 0, which is not a permissible mathematical procedure.

(عوام‌فریبی) to infer the concept of gaining the political support of the masses by abusing their feelings and 'avāmgarīb (demagogue, عوام‌فریب) as the person who applies this concept.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>pūpūlīsm</i>	پوپولیسیم	456	48	225	34
<i>pūpūlīstī/pūpūlīst</i>	پوپولیسیتی/پوپولیسیت	918	126	659	90
'avāmgarāyī	عوام‌گرای	47	13	30	12
'avāmgarā	عوام‌گرا	18	0	17	0

Table 39. Term frequency and document frequency of the terms in the semantic field of *populism*

Table 39 above shows a clear majority for the foreign terms vis-à-vis the approved Persian equivalents. In TalkBank, the name for the doctrine, *pūpūlīsm*, is almost ten times more frequent than 'avāmgarāyī. The name for the adherent, *pūpūlīst*, occurs even more often with a ratio of approximately 50:1. In TAP, the difference is not so stark with a ratio of less than 4:1 in favor of *pūpūlīsm*. However, newspapers make exclusive use of the term *pūpūlīst/ī*. The Persian equivalent 'avāmgarā is not counted even once.

The weight calculation produces the following adjusted weight values for the terms in question. In TalkBank, *pūpūlīst/ī* has an adjusted weight of 6,225 and *pūpūlīsm* has an adjusted weight of 3,541. The Persian equivalents 'avāmgarāyī and 'avāmgarā have adjusted weights of only 452 and 182 respectively. In TAP, the adjusted weight of *pūpūlīst/ī* is 20,556 and that of *pūpūlīsm* is 9,390 whereas the adjusted weight of 'avāmgarāyī is 2,995. The weight calculation for 'avāmgarā is not possible because it has no hits in TAP. The relatively low weight values for these terms can be explained by the fact that the topic of populism is not discussed frequently in the newspapers or blogs saved to the corpora. Clearly, though, the foreign terms *pūpūlīst/ī* and *pūpūlīsm* are heavier in both TalkBank and TAP.

## 5.12 The Concept of *Capitalism* and Its Derivatives

*Kāpītāl* is the foreign term that entered Persian through French with the meaning of *capital*, *property*, or *fortune* (Mo'īn M., 1386[2007]c., p. 2780) from the French masculine version *le capital* as well as, in rare instances, *capital city* (*pāytaḥt*, *korsī-ye nešīn*, *ḥākem-e nešīn*) (Dehḥodā, 1336[1957], p. 34) from the French feminine version *la capitale*. It was first mentioned in French in the meaning of *debt* or *rent*, that is, in a monetary sense, in 1567 and, more precisely, as the totality of goods claimed by one owner in 1606. By 1832, it had acquired the meaning of the wealth that is understood as a means of production (CNRTL, *Capital*, 2012). From then on it was a short way to the idea of capitalism. Only ten years later, the French term *capitalisme* was first used to describe a system of capitalization (CNRTL, *Capitalisme*, subst.

masc, 2012). In English, it was attested in 1833 as “an economic system characterized by private or corporate ownership of capital goods, by investments that are determined by private decision, and by prices, production, and the distribution of goods that are determined mainly by competition in a free market” (Webster, Capitalism, 2021).

The ideology of capitalism appeared in America and Western Europe by the first third of the 19<sup>th</sup> century and then found its way to exist and develop in non-Western countries (Ášūrī, 1376[1997], p. 199). *Kāpītālīsm* is defined in ‘Amīd’s dictionary as a type of economic system based on the principles of individual freedom in economic activities and private property rights (‘Amīd, 1365[1986]c., p. 1915). Mo‘īn defines it as the supremacy of capitalists in industrial affairs (Mo‘īn M., 1386[2007]c., p. 2780). It is defined as well as an economic system in which real estate, factories, and the means of production belong to a small group of people (Áqābaḥšī & Afšārīrād, 1383[2004], p. 81).

Capitalism is closely linked to communism, an economic or political system that propagates the abolition of private property and the shared use or ownership of all means of production in any given nation. It is usually identified with the theories and ideas of the German 19<sup>th</sup> century philosopher Karl Marx, author of *The Communist Manifesto* and *Capital*, and his revolutionary thoughts on the initial communal ownership of capital by the working class and the eventual development of a classless society (Webster, Capitalism, 2021).

The terminology group of the APLL approved in the ninth volume of the approved terms *sarmāyedārī* (سرمايهداری) as an official replacement of *kāpītālīsm* to indicate an economic system based on wage labor, the production of goods, the provision of services for sale and exchange, and the earning of profits in dimensions that go beyond satisfying the immediate needs of the producer (Terminology group, 1391 [2019]). However, the approved term had already been in use with the same meaning alongside the foreign term *kāpītālīsm* before its formal approval by the Farhangestān. In the same volume, the term *sarmāyedār* (سرمايهدار) was also approved as an equivalent to *kāpītālīst* expressing the concept of the person who invests to make a profit. However, ideologically, it refers to anyone who believes in capitalism and not just those who are its beneficiaries.

The prefix *sarmāye*, which forms the first part of the approved term *sarmāyedārī*, consists of two morphemes. The first one is *sar* (سر) meaning *head*. The second one is *māye* (مايه) and has a lot of meanings both metaphorically and literally as it could mean money, capital, and fortune, as well as basis, foundation, and fundament. It could also carry the meaning

of power, authority, and ability, for which ‘Amīd references one of the verses of Ferdowsi that includes the word *māye* with the connotation of power. Influenced by one of the verses of Sa‘dī, ‘Amīd gives one more synonym to *māye* that means cause or evidence. In addition to the previous concepts of *māye*, it could also refer to vaccine in the medical field (‘Amīd, 1365[1986]b., p. 2152).

Hence, *sarmāye* is the value of goods that form the basis of a business. It could also mean profit, that is, the amount left after subtracting the production-related money (‘Amīd, 1365[1986]b., p. 1427). *Sarmāye* is thusly a direct loan translation from the Arabic term *ra’s al-māl* (literally: head money, i.e., capital).

The suffix *-dārī* stems from the verb *dāštan*, which means to own or to possess. By and large, therefore, *sarmāyedārī* (capital ownership) is a loan translation and rendition of *kāpītālīsm* in accordance with Bateni’s system of categorization.

The following table presents various contextualized examples of both the foreign term *kāpītālīsm* and its approved Persian equivalent *sarmāyedārī*.

<b>Example 1</b>	حتی اگر کاپیتالیسم را در نظر بگیریم، میتوان آن را برای بشریت حائز خدماتی دانست و این دولتمردان هستند که با اغراض و امراض شخصی خود به وجه ایدئولوژی ضربه میزنند.
Transliteration	Ḥatta agar kāpītālīsm rā dar nazar begīrīm, mītavān ān rā barāye bašareyyat-e ḥā’ez-e ḥadamātī dānest va īn dūlatmardān hastand ke bā agrāz va amrāz-e šaḥṣī-ye ḥūd be veġeh-e īde’ulūzī zarbe mīzanand.
Translation	Even capitalism can be viewed as a humanitarian service, and those statesmen are the ones that deform the picture of ideology with their personal agendas and ailments.
<b>Example 2</b>	نظام سرمایه داری یا کاپیتالیسم نظام اقتصادی است که با انقلاب صنعتی، نخست در اروپای غربی و آمریکا و سپس به کشورهای غیر غربی نیز راه یافت.
Transliteration	Nezām-e sarmāyedārī yā kāpītālīsm nezām-e eqtešādī ast ke bā enqelāb-e šan’atī, neḥost dar Orūpā-ye ġarbī va Āmrīkā va sepaš be kešvarhā-ye ġayr-e ġarbī nīz rāh yāft.
Translation	Capitalism [sarmāyedārī yā kāpītālīsm] is an economic system that, with the Industrial Revolution, spread first to Western Europe and the United States and then to non-Western countries.
<b>Example 3</b>	کشورهای سرمایه داری بزرگ جهان نیز تلاش می نمایند تا شتاب و نمود بیشتری به فعالیتهای انسان دوستانه خود بدهند.
Transliteration	Kešvarhā-ye sarmāyedārī-ye bozorg-e ġāhān nīz talāš mīnamāyand tā šetāb va namūd-e bīštārī be fa’āleyathā-ye ensān-e dūstāneh-e ḥūd bedahand.
Translation	The world’s leading capitalist countries are likewise attempting to bolster their humanitarian efforts.

Table 40. Examples of both *kāpītālīsm* and *sarmāyedārī* in the corpora

As can be seen in example 1 in table 40 above, the speaker views capitalism as one ideology among many; in the interview from which this excerpt was taken, the interviewee speaks about ideologies in general and how statesmen try to bring them down. Example 2 is

noteworthy because it enumerates both terms as synonyms where the author seems to feel that an explanation of the expression “*nezām-e sarmāyedārī*” is needed by calling it “*kāpītālīsm*”. It was published in *Hamshahri Online* on May 21, 2020, which is at least six years after formal approval of *sarmāyedārī* by the APLL. Even though the latter is already common in Persian, as will be seen below, the author here still uses the foreign term along with it, presumably for stylistic reasons.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>kāpītālīsm</i>	کاپیتاليسم	4,137	30	3,998	26
<i>kāpītālīst</i>	کاپیتاليست	229	31	196	18
<i>sarmāyedārī</i>	سرمايهدارى	5,591	765	2,348	213
<i>sarmāyedār</i>	سرمايهدار	1,981	969	1,528	308

Table 41. Term frequency and document frequency of the terms in the semantic field of *capitalism*

The quantitative analysis of table 41 above reveals very mixed results. While the comparison between *kāpītālīsm* and *sarmāyedārī* has a favorable outcome for the approved Persian equivalent of only 57.5 percent in TalkBank, it is dominant by 95.9 percent in TAP. At the same time, *kāpītālīsm* is used in more TalkBank documents than *sarmāyedārī* with a ratio of 63 percent. The foreign term *kāpītālīsm* is by and large used only once per document whereas *sarmāyedārī* occurs on average more than two times per document in TalkBank and more than three times in TAP.

The dominance of the Persian equivalent is even more pronounced when it comes to *kāpītālīst*, which is hardly in use in the corpora. Here, the approved Persian equivalent *sarmāyedār* is prevalent in both blogs and newspapers with 89.6 percent and 96.9 percent respectively. Overall, this indicates that authors feel more comfortable with using the Persian term and that it is easier to pronounce or easier to understand, and therefore more widespread in texts, be they blogs or newspaper articles.

In numbers, the adjusted weight of *kāpītālīsm* is 21,233 in TalkBank while its adjusted weight in TAP is 6,751. *Kāpītālīst* records an adjusted weight of 1,807 in TalkBank and 6,722 in TAP. These values are comparatively low when looking at the weight of the Persian equivalents. *Sarmāyedārī* and *sarmāyedār* show an adjusted weight of 31,417 and 11,910 respectively in TalkBank and 102,810 and 118,301 respectively in TAP.

A look at the adjusted weight values of all terms in question here reveals that the approved Persian equivalents are in each case heavier than the foreign terms. *Sarmāyedārī* is only 1.5 times heavier than *kāpītālīsm* in TalkBank, but 15 times heavier in TAP. Similarly, *sarmāyedār*

is 6.6 times heavier than *kāpītālīst* in TalkBank and 17.6 times heavier in TAP. Again, journalists are more likely to use the formalized Persian equivalent compared to bloggers that make a more liberal use of both terms.

### 5.13 The Concept of *Reformism* and Its Derivatives

Reformism can be defined as the belief that social, political, or economic change is achieved through reform (“eṣlāḥāt”) by using peaceful and gentle methods instead of revolution (“enqelāb”) through violent and revolutionary methods. In his definition, Ḥasan ‘Alīzāde focuses on the origins of the term in the late 19<sup>th</sup> / early 20<sup>th</sup> century global socialist movement, which abandoned the idea of revolutionary violence and called for the peaceful transformation of social institutions through democratic means. Social reformers hence deny the necessity of Marxist class struggle, the socialist revolution, and the dictatorship of the proletariat (“dīktātūrī-ye prūletāreyā”) and, instead, advocate for class cooperation to achieve a society based on public welfare. These include British “Fabianism”, French and German “Reformism”, as well as Russian “Economism” (‘Alīzāde, 1389[2010], pp. 252-253). In this sense, reformism is opposed to conservatism on the one hand and to revolution on the other.

In Iran, the term *reformism* (*eṣlāḥtalabī*) describes a political current that has developed from the late 1990s onwards. It entails the promotion of democratization, political participation, and ideological diversity as well as the political and social modernization of Iran and its integration into the international community. Reformism was particularly strong during the two terms of President Khatami (1997-2005). Similar to reformism in modern political history, the origins of Iran’s reformist movement can also be traced back to the so-called Islamic left and left-leaning parts of the elites. It was, in this sense, initiated by parts of the ruling elites, even if some of its adherents turned out to be quite independent-minded groups among the population, who used the reform narrative to liberalize political participation and coopt civil society actors (Rivetti, 2020, p. 3ff.).

*Reform*, *reformism*, and *reformist* as foreign terms entered the Persian dictionary following the French terms *réforme*, *réformisme*, and *réformiste* respectively. The Persian term that the Academy selected as a basis to coin an equivalent to this concept is *eṣlāḥ*. It is originally an Arabic word with multiple meanings and connotations revolving around reform, rehabilitation, fixing, renovation, correction, reconditioning, renovation, and many more that express different types of reform in various contexts. According to Dehḥodā, *eṣlāḥ* in Persian also has the same connotations as in Arabic. He bases his definition of this term on ancient



Persian literary resources. In the common daily vernacular, the term *eṣlāḥ* is used as a synonym for the concept of *reform*, which itself refers to the religious Reformation that developed in Germany after Martin Luther's 1517 publication of his 95 theses and quickly spread to other parts of Europe as well. Gradually, the term *eṣlāḥ/reform* was applied to any change be it religious, social, or political (Dehḥodā ' A., 1332[1953], p. 2825). The APLL coined the compound word *ḡanbeš-e eṣlāḥ-e dīn* (reformation, جنبش اصلاح دین) to refer to the religious movement initiated by Luther and his contemporaries that originated in 16<sup>th</sup>-century Western Europe with the goal of reforming Catholicism and restricting the power of the Church and that eventually led to the formation of Protestantism (Terminology group, 1384 [2005]).

The terminology group coined many terms to replace *reform* and all of its concepts, be they religious, social, political, or scientific. The term *eṣlāḥ*, in the second volume of the approved terms, is approved in two fields. First, in the political and public relations field it means any action taken to enhance the state of society by modifying certain aspects of economic, political, or social life without effecting fundamental or structural changes. The second implementation of the same term is used in the field of biology-genetics and biotechnology agriculture with the meaning of breeding.

The approved terms *eṣlāḥṭalabī* (reformism, اصلاح طلبی) and *eṣlāḥṭalab* (reformist, اصلاح طلب) are meant to signify the concept of an intellectual movement or movement that believes in reform of society where the reformist is an adherent of this school or this way of thoughts. Assuming once more that *eṣlāḥ* is considered a Persian term and not one of Arabic origin, in which case we would speak of a loan exchange, *eṣlāḥ* is a loan translation of the foreign term *reform* when looking at Bateni's categorization. *Eṣlāḥṭalabī* and *eṣlāḥṭalab* are loan translations and renditions of *reformism* and *reformist* respectively. Here, the suffix *-ism* is expressed through the Arabo-Persian morpheme *ṭalab(ī)*, which could be translated as *demand* or *request* thus effectively making *eṣlāḥṭalabī* a request or demand for reform.

Ášūrī uses the expression *behbūdḥāhī* (بهبودخواهی) alongside the term *eṣlāḥṭalabī* to denote reformism and *behbūdḥāh* (بهبودخواه) to refer to the individual reformer (Ášūrī, 1376[1997], p. 72). This alternative couple of terms does not, however, record a high frequency in the two corpora: *Behbūdḥāhī* produces only four hits in TalkBank and only one hit in TAP while *behbūdḥāh* does neither record a single hit in TalkBank nor in TAP. Thus, the term proposed by Ášūrī is not in frequent use with the same connotation neither in blogs nor in newspapers.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>reform</i>	ر فرم	347	8	243	8
<i>reformhā</i>	ر فرمها	10	0	10	0
<i>reformīst</i>	ر فرمیست	154	2	129	2
<i>reformīsm</i>	ر فرمیسم	39	0	27	0
<i>eṣlāḥ</i>	اصلاح	96,633	8,680	45,624	3,519
<i>eṣlāḥāt</i>	اصلاحات	40,192	4,191	20,204	1,545
<i>eṣlāḥṭalab</i>	اصلاح طلب	21,625	1,881	9,432	756
<i>eṣlāḥṭalabī</i>	اصلاح طلبی	1,853	568	985	236

Table 42. Frequency and document frequency of the foreign terms and their approved Persian equivalents in the semantic field of *reform*<sup>52</sup>

By comparing the foreign terms and the approved ones in the semantic field of *reform*, it is clear that the approved Persian equivalents are widely used compared to the loanwords. It can be seen in table 42 above that none of the foreign terms has a high frequency in either corpus to the extent where the term *reformīsm* shows no existence in TAP at all. By contrast, the approved Persian equivalents dominate discourse. The term *eṣlāḥ* has a share of 99.7 percent across the corpora compared to *reform*; the plural form *eṣlāḥāt* is even more dominant with 99.98 percent compared to *reformhā*. The numbers are hardly more convincing for the terms in question: The term *eṣlāḥṭalabī* is represented in 98.4 percent of all occurrences compared to *reformīsm*. Finally, *eṣlāḥṭalab* has a share of 99.3 percent compared to *reformīst*. While there is a minimum use of the foreign terms in blogs, newspapers show virtually no results for *reform*, *reformīst*, and *reformīsm*, the latter of which does not appear a single time in TAP.

Clearly, *eṣlāḥ*, though of Arabic origin, is a term well-rooted in Persian, which is also true for its derivatives. The term *eṣlāḥ* can be traced at least to one of the important Persian-Arabic dictionaries, *Montahā al-arab fī loḡat-e al-‘arab*, which was written in the 13<sup>th</sup> century after the Hijra (18<sup>th</sup>/19<sup>th</sup> century) by ‘Abdelreḥīm Ben ‘Abdelkerīm Ṣafī Pūrī Šīrāzī. This dictionary contains all the Arabic words that were used in Persian at that time. It also mentions the lemma *eṣlāḥ* defined as doing good (nīkūkardan, نیکو کردن), anti-corruption (zedd-e efsād), correction or reform of one another (hamdīgar nīkī kardan, هم دیگر نیکی کردن), reconciliation (āštī namūdan, آشتی نمودن), and providing a sense of national belonging for mutual reconciliation (farāham āmadan-e qūmī barāy-e taṣāloḥ bā ham va nīkū’ī namūdan, فراهم آمدن قومی برای مصالح با هم و نیکوئی نمودن) (Ṣafīpūr, 2001, p. 696).

These results are consistent with a previous study that dealt with the term *eṣlāḥ* and its derivative forms *eṣlāḥ-e ṭalab* and *eṣlāḥāt*. They are considered among the approved terms that

<sup>52</sup> The results for the derivatives of *reform* include the respective derivatives of *refūrm*.

are most successful in both written and spoken news. All three forms of the approved term are among the most accepted approved political terms by the APLL ranking fourth (*eṣlāḥ-e ṭalab*), seventh (*eṣlāḥāt*), and twelfth (*eṣlāḥ*) in a list of the most used approved terms in spoken news as well as fourth (*eṣlāḥ-e ṭalab*), sixth (*eṣlāḥāt*), and fifteenth (*eṣlāḥ*) in a list of the most used approved terms in written news. By contrast, the term *reform* or any of its derivatives do not feature in the study (Sattari, 1388 [2009]). The following table 43 shows some examples from both corpora.

<b>Example 1</b>	البته در این راه میتوان روشها و اهداف میانی را اصلاح کرد و طبق یکی از فرمایشات مقام معظم رهبری حتی این امکان وجود دارد که در صورت لزوم، برخی اصول غیر قابل تغییر قانون اساسی نیز مورد بررسی مجدد قرار داد، مانند ریاستی بودن ریاست جمهوری یا پارلمانی بودن آن
Transliteration	Albatte dar īn rāh mitavān ravešhā va aḥdāf-e meyanī rā eṣlāḥ kard va teḥq-e yekī az farmāyešāt-e maqam-e mo'azzam-e rahbarī ḥatta īn emkān voğūd dārad ke dar ṣūrat-e lozūm, barḥī-ye oṣūl-e ġayr-e qābel-e tağyīr-e qānūn-e asāsī nīz mūred-e barrasī-ye moğaddad qarār dād, mānand-e reyāsati-ye būdan-e reyāsāt-e ġomhūrī yā pārlemānī-ye būdan-e ān.
Translation	Of course, in this way, the average methods and goals can be reformed. And according to one of the orders of the supreme leader, it is possible to re-examine some of the established principles of the constitution, if necessary, such as the position of the presidency of the republic or the presidency of its parliament.
<b>Example 2</b>	اگرچه اخیرا محسن مهر علی زاده رئیس سازمان تربیت بدنی دولت اصلاحات گفته است که نامزد اصلاح طلبان در ۱۴۰۰ تأیید صلاحیت نخواهد شد.
Transliteration	Agarče āḥīran, Moḥsen Mehr'alī Zāde, ra'īs-e sāzmān-e tarbeyāt-e badanī-ye dūlat-e eṣlāḥāt gofte ast ke nāmzad-e eṣlāḥ-e ṭalabān dar 1400 ta'yīd-e ṣalāḥeyyat naḥāhad šod.
Translation	Recently, however, Moḥsen Mehr'alī Zāde, the head of the reformist government's physical education organization, said that the reformist candidate would not be approved in the year 1400.
<b>Example 3</b>	این جریان دارای تفکر اصولگرایی ناب است که با روش اصلاحطلبی به پیش میرود.
Transliteration	Īn garayān-e dārā-ye tafakor-e oṣūlgarā-ye nāb ast ke bā raveš-e eṣlāḥṭalabī be piš mīravād.
Translation	This current has a pure fundamentalist thinking that proceeds with the method of reformism.
<b>Example 4</b>	حقیقت را می‌گویند که نسبی است و نه « سیاه و سفید » ، انقلابی‌گری جای خود را به رفرمیسم داده است و نگاه اتوپیک با توتالیتراریسم نسبت برقرار کرده.
Transliteration	Ḥaḳīqat rā mīġūyand ke nesbī ast va nah "seyāh va safīd", enqelābīgarī ġāy-e ḥūd rā be reformīsm dādeh ast va negāh-e otūpīk bā tūtālītārīsm nesbat barqarār kardeh.
Translation	Truth is said to be relative, not "black and white", revolutionarism has given way to reformism, and the utopian view is associated with totalitarianism.

Table 43. Examples of *eṣlāḥ* and *reform* from the corpora

The term *eṣlāḥ* has an adjusted weight of 280,753 in TalkBank and 354,145 in TAP. Compared to other approved terms, this is an extraordinarily heavy weight. By contrast, *eṣlāḥ-e ṭalab* and *eṣlāḥ-e ṭalabī* have adjusted TF-IDF weights of 89,329 and 11,885 respectively. The foreign terms *reform*, *reformīst*, and *reformīsm* weigh only 2,670, 1,274, and 378 respectively and are therefore significantly less heavy than the approved Persian equivalents.

## 5.14 The Concept of *Doctrine* and Its Derivatives

The loanword *doktorīn* entered the Persian dictionary through the French lexicon. However, the term *doctrine* itself goes back to the Latin term *doctrina*, which stands for *theory*, *method*, or *teaching*. The French term was first mentioned in the twelfth century (CNRTL, *Doctrine*, subst. fém., 2012). A doctrine in the political sense is simply a statement of a government's essential policies. Mo'īn defined the term *doktorīn* with the synonyms *nazar*, *fekr*, and *andīše* meaning *thought*, *thinking*, *opinion*, *mind*, *theory*, and *vision* (Mo'īn M. , 1386 [2007]b., p. 1544).

In English, the term *doctrine* was first used in the 14<sup>th</sup> century and has since, among others, acquired the meaning of “a statement of fundamental government policy especially in international relations”. Two famous historical doctrines of that kind were the 1823 Monroe Doctrine stipulating that the United States would not allow European interference in the Western Hemisphere and the 1947 Truman Doctrine stating that the United States would contain Communism and support countries against the spread of Soviet influence, both of which were named for the ruling presidents at the time (Webster, *doctrine*, noun, n.d.).

With the term *rahnāme*, the terminology group of the APLL coined a Persian equivalent to the foreign term *doktorīn*. It is defined in the second volume of the approved terms as a written declaration or statement about the official policy of the government, especially in the field of political relations and foreign policy (Terminology group , 1382 [2004]). *Rahnāme* shares the same Middle Persian origins with *rāhnāme* (راهنامه). As a compound noun, it consists of the morpheme *rāh* (راه), which refers to any place on earth where people come and go, a way or road in addition to a law or approach (‘Amīd, 1365[1986]b., p. 1214). The second morpheme is *nāme* (نامه) originally stemming from the Middle Persian *nāmak* (نامک), which denotes content that is written to a specific person and that is usually delivered by mail and in an envelope and also carries the meaning of paper (‘Amīd, 1365[1986]c., p. 2276). The coined term would therefore roughly translate as *the path that specific content takes* or *the written approach*. It is an explanation of the meaning of the foreign term *doktorīn* and, in Bateni's categorization, would be characterized as a loan rendition.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>doktorīn</i>	دکترین	1,912	151	1,187	75
<i>rahnāme</i>	رهنامه	14	0	14	0

Table 44. Term frequency and document frequency of the terms in the semantic field of *doctrine*

Table 44 above represents the frequency and document frequency of the two terms *doktorīn* and *rahnāme*. It shows that the approved term *rahnāme* has a very low frequency in TalkBank with only fourteen mentions and no hits in TAP whereas the foreign term *doktorīn* has more than 2,000 hits in both corpora resulting in an overall majority of 99.3 percent.

This comes as no surprise because also previous findings show that *rahnāme* has no hits in a corpus used by Hesabi et al. while *doktorīn* produces 758 hits (Hesabi, Amirian, & Nazari, 2017). Another study of written and spoken news reflects the same findings. Here, *rahnāme* is listed among the most unsuccessful approved terms (Sattari, 1388 [2009]).

The adjusted weight of *doktorīn* and *rahnāme* is 11,937 and 144 respectively in TalkBank and 25,553 and 0 respectively in TAP. In both corpora, the foreign term *doktorīn* is therefore significantly heavier than the approved Persian equivalent *rahnāme*: more than 82 times heavier in TalkBank and infinitely heavier in TAP given the value zero. In other words, *rahnāme* is one of the most unsuccessful Persian equivalents approved by the Academy.

## 5.15 The Concept of *Strategy* and Its Derivatives

The Persian term *esterātežī* (strategy) has Greco-Latin roots and comes to Persian from the French dictionary (*stratégie*). It first appeared in French in the early 19<sup>th</sup> century during the Napoleonic Wars where it referred to the art of leading, commanding, or moving an army in the field. In Greek, it derived from *strategós* and *strategía* as command of an army, having the qualities of a commander-in-chief, and maneuver (CNRTL, *Stratégie*, subst. fém, 2012).

In the Persian dictionary, *esterātežī* is a military term that indicates the stage of assessing the military situation and developing the appropriate plan to ward off and defeat the opponent. However, it evolved in politics where it entails mobilizing all resources and shifting the situation in the appropriate direction in order to attain a fundamental goal. In short, a strategy is a broad approach to achieve a certain goal (Ášūrī, 1376[1997], p. 26). In Mo‘īn’s dictionary, *esterātežī* is any long-term plan for a specific purpose, for example, an economic or military strategy (Mo‘īn M. , 1386 [2007]a., pp. 246-247).

‘Amīd presents the term *seyāsatgozārī* (which the APLL incidentally approved to mean policymaking, that is, the process of finding a political solution to attain a specific goal) as a synonym to better grasp the political implications of *esterātežī*. This would explain the concept of abilities and facilities for planning to accomplish a specific (political) goal. However, in

military language, he gives preference to *esterātežī*, which refers to the science of managing the operations and movements of an army at war (‘Amīd, 1365[1986]a., p. 167).

The approved Persian equivalent *rāhbord* consists of two separate parts. The first part is *rāh*, which is taken from Middle Persian where it means *way*, *path*, or *passage* and derived from the Old Persian *rāt-* that means *to move* (Hassandoust, 1393[2014]b., p. 1423). The second part in the approved term is *bord*, which goes back to the verb *bordan*. This verb comes from the Middle Persian *burtan* and indicates the sense of carrying or transferring someone or something from one place to another (Hassandoust, 1393[2014]a., pp. 438-440). In this sense, *rāhbord* can be translated as a “movement or path for transferal (towards a specific goal)” thus constituting a loan rendition in Bateni’s categorization.

In the first volume of approved terms by the APLL, the terminology group established *rāhbord* as a replacement for the borrowed term *esterātežī* with two definitions. First, it is defined as the art of developing troops and military resources and using them in times of war to maximize national interests and increase the likelihood of victory. Second, it carries the meaning of planning and managing national resources – be they economic, political, or military – to achieve maximum national interests in peacetime and increase the likelihood of victory in time of war. Not only is *rāhbord* approved to express a strategy in the political or military fields. It is also approved as a morpheme in a lot of other compound words to describe various concepts in information technology, management, futurology, statistics, audiology, psychology, as well as maritime transportation (Terminology group , 1382 [2004]).

In the fields of political and military sciences, the terminology group approved

1. *rāhbord-e mellī* (national strategy, راهبرد ملی) and *kalān-e rāhbord* (grand strategy, کلان راهبرد) to indicate the science and art of developing and using a country’s diplomatic, economic, and intelligence facilities as well as its military capabilities in times of peace and war to achieve national goals,
2. *rāhbord-e nezāmī* (military strategy, راهبرد نظامی) that implies techniques and knowledge related to the use of a country’s armed forces to achieve national policy goals through the use of force or the threat of the use of force, as well as
3. *rāhbord-e ešterākāt-e eṭṭelā ‘āt* (information sharing strategy, راهبرد اشتراکات اطلاعات) and *rāhbord-e mošāarakat-e eṭṭelā ‘āt* (راهبرد مشارکت اطلاعات) to refer to the strategy by which information is shared between agents and local organizations.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>esterātežī</i>	استراتژی	20,535	1,171	11,204	688
<i>esterātežīk</i>	استراتژیک	27,966	1,392	18,502	884
<i>rāhbord</i>	راهبرد	11,728	1,476	6,840	815
<i>rāhbordī</i>	راهبردی	15,599	4,447	9,388	1,826

Table 45. Term frequency and document frequency of the terms in the semantic field of *strategy*

Overall, the foreign terms *esterātežī* and *esterātežīk* occur more frequently in the two corpora than their approved Persian equivalents. A look at the details reveals an interesting correlation, though, where *rāhbord* and *rāhbordī* are used more often in newspapers with a 69.9 percent share of the combined total use of all terms in TAP while *esterātežī* and *esterātežīk* are favored in blogs with a 64 percent share in TalkBank. Once more, newspapers prove to be a stronghold for a formal or official way of writing more so than blogs.

It is interesting to note that both the foreign and approved term are used in the same text 474 times in TalkBank and 35 times in TAP. Table 46 below shows examples of the use of both terms as synonyms in the corpora.

<b>Example 1</b>	در مورد تولید ، حمایت از تولید ملی و به تبع آن حمایت از کار و سرمایه ایرانی؛ مناسبترین راهبرد، استراتژی حفظ و بهبود است.
Transliteration	Dar mūred-e tūlīd, hemāyat az tūlīd-e mellī va tab‘ ān hemāyat az kār va sarmāyeh-e Īrānī, monāsebtarīn-e rāhbord, estrātežī-ye hefz bahbūd ast.
Translation	Regarding production, support for national production and thus support for Iranian labor and capital; The most appropriate strategy is the strategy of maintenance and improvement.
<b>Example 2</b>	اما در عرصه سیاست خارجی اینکه چگونه می توان با راهبرد و استراتژیهای مناسب زمینه های وحدت و همگرایی هر چه بیشتر ملت مسلمان را فراهم کرد خود بحث مهمی است که نیاز به تحقیق و تفصیل فراوان دارد.
Transliteration	Ammā dar ‘arṣeh-ye seyāsāt-e hāregī-ye īnke čegūne mītavān bā rāhbord va estrātežīhā-ye monāseb-e zamīnehā-ye veḥdat va hamgarāye harče bīstar-e mellat-e mosalmān rā farāham kard, hūd-e baḥt-e mohemmī ast ke neyāz be taḥqīq va tafzīl farāvān dārad.
Translation	But in the field of foreign policy, how to provide the grounds for the unity and rapprochement of the Islamic nation with appropriate strategies [rāhbord va estrātežīhā] is an important debate that needs a lot of research and elaboration.
<b>Example 3</b>	جست و جوی جایگزین برای تنگه استراتژیک هرمز، راهبرد چند سال اخیر امریکا، رژیم صهیونیستی و سعودیها در منطقه بوده است.
Transliteration	Gostūgū-ye gāygozīn barāye tangeh-e estrātežīk Hormoz, rāhbord-e čand-e sāl-e aḥīr-e Āmrīkā, režīm-e šehyūnistī va Se‘ūdīhā dar manteqe būde ast.
Translation	The search for an alternative to the strategic Strait of Hormuz has been the strategy of the United States, the Zionist regime, and the Saudis in the region in recent years.

<b>Example 4</b>	این اجلاس با طرح استراتژی امنیتی « ایران ، نقطه تفاهم امریکا و اسرائیل » کوشید تا راهبردهای دوگانه اسرائیل و ایالات متحده را برای فشار بر ایران ، به هم نزدیک نماید.
Transliteration	Īn eġlās bā tarḥ-e estrāteżī-ye āmneyatī “Īrān, noqteh-e tafāhom-e Āmrīkā va Esrāyīl” kūšīd tā rāhbordhā-ye dūgāneh-ye Esrāyīl va Eyālāt-e Motahēdeh rā barāye fešār bar Īrān be ham nazdīk namāyad.
Translation	The summit sought to bring closer the dual strategies of Israel and the United States to put pressure on Iran by proposing a security strategy called “Iran, the point of understanding between the United States and Israel.”
<b>Example 5</b>	بازیگران اروپایی نیز این استراتژی را با جدیت دنبال کرده و سیاست خارجی خود در قبال کشورمان را نیز بر مبنای همین چارچوب و راهبرد کلان استوار ساخته اند.
Transliteration	Bāzīgarān-e Orūpā-ye nīz īn estrāteżī rā bā ġeddeyat donbāl karde vā seyāsāt-e ḥāregī-ye ḥūd dar qebāl-e kešvaremān rā nīz bar mabnā-ye hamīn čārčūb va rāhbord-e kalān-e estevār sāḥtehand.
Translation	European performers have seriously pursued this strategy and have based their foreign policy towards our country on the same framework and macro strategy.

Table 46. Examples of the foreign term *esterāteżī* and the approved term *rāhbord* in context from TalkBank and TAP

Given the numbers in table 45 above, the adjusted weight of the foreign term *esterāteżī* is 86,038 in TalkBank and 182,612 in TAP. The adjusted weight of the adjective *esterāteżīk* in TalkBank is 104,34 while in TAP, the adjusted weight of *esterāteżīk* is 113,916. For *rāhbord*, the adjusted TF-IDF score is 54,432 in TalkBank and 278,607 in TAP. Calculating the adjusted weight of the adjective *rāhbordī* gives a figure of 67,402 in TalkBank and 277,420 in TAP.

In conclusion, the simple frequency count is corroborated showing that *rāhbord* and its derivative enjoy a heavier weight in TAP than *esterāteżī* making the approved Persian equivalent more successful in newspapers. Nevertheless, the picture remains mixed with a favorable outcome for the foreign term *esterāteżī*, which weighs heavier in the blogs compiled in TalkBank.

## 5.16 The Concept of *Protocol*

Also the foreign term *protokol* arrived in Persian via French although the origin of this term goes back to the Greek *prōtokollon*, which refers to the first (*proto-*) sheet of papyrus roll glued (*kolla*) to a document or register entry summarizing its date of creation and content (Webster, Potocol, 2021). The meaning of *protocol* evolved in French and English to have many connotations at different stages in history. In Persian, too, the two terms in question – *protokol* and *tašrīfāt* – have more than one connotation. They may have a political connotation, a general connotation, or also a specialized connotation as is the case in the field of electronics science.

After the 14<sup>th</sup> century, *protocol* came to denote not only a foreword or introduction, but also “a record of a document or transaction”. From the 17<sup>th</sup> century onwards, its meaning expanded to indicate the manners favored by the rulers of France in the sense of formalities in



writing and, increasingly, of etiquette in ceremonies and international relations (e.g., diplomatic protocol) (Webster, Potocol, 2021) (cf. (CNRTL, Protocole, subst. masc., 2012).

It is in another connotation, though, that *protokol* is applied in Persian. Since the early 18<sup>th</sup> century, but more so since the mid-19<sup>th</sup> century, *protocol* has been used to refer to the codified results of diplomatic deliberations, negotiations, and conferences that serve as a basis for the final acts or international treaty (CNRTL, Protocole, subst. masc., 2012). Famous protocols in this sense are, for example, the Kyoto Protocol (*peymān-e Kyūtū, protokol-e Kyūtū*) on climate protection, the Alexandria Protocol (*protokol-e Eskandereye*), which established the League of Arab States, and the Geneva Protocol (*protokol-e Ženev*) to ban the use of chemical weapons in warfare. One should note here the different translations of *Protocol* into Persian.<sup>53</sup> In Persian, *protokol* has mostly a political connotation that refers to the common procedures in international relations and the form of political negotiations as well as the draft contract or minutes of negotiations that will later become a treaty (Mo‘īn M. , 1386 [2007]a., p. 763).

In addition, as mentioned in *Merriam Webster*, a protocol generally is “a detailed plan of a scientific or medical experiment, treatment, or procedure”. It could also refer to the network protocols as a set of agreements governing the processing and the coordination of data in an electronic communication system (Webster, Potocol, 2021). The same is true in Persian where the term *protokol* could either mean the draft contract or minutes of negotiations as well as common procedures in international relations (‘ Amīd, 1365[1986]a., p. 544) or the transmission control protocol (TCP) and Internet protocol (IP), which refers to the set of procedures to transfer data between electronic devices. While the latter connotation is not (yet) part of any of the Persian dictionaries, it is substantiated through a reading of texts in the two corpora.

In the second volume of APLL approved terms, the terminology group introduced the term *tašrīfāt* to replace the term *protokol* in its political connotation. However, *tašrīfāt* is not limited to this political concept only, but it has itself other meanings and connotations. The approved term *tašrīfāt* is originally an Arabic word that carries the meaning of special mores, etiquettes, and customs in important and formal receptions like in Persian (Mo‘īn M. , 1386 [2007]a., p. 1086). In this sense, *tašrīfāt* is a loan rendition that explains the meaning of *protokol*

---

<sup>53</sup> The four compound words *peymān-e Kyūtū* and *prūtokol-e Kyūtū, protokol-e Eskandereye*, and *protokol-e Ženev* occur in the two corpora only a few dozen times. TalkBank records 34, 32, 16, and 55 hits respectively while in TAP they appear five, 16, zero, and two times respectively.

but finds a local Persian equivalent to replace it – provided that one accepts, like the Farhangestān, that its Arabic roots are considered Persian(ized).

The terminology group approved *tašrīfāt* in various contexts. In addition to its political meaning, the Academy defines *tašrīfāt* to represent the collection of diplomatic conduct, manners, treatment, and behavior in the relations between states and countries (Terminology group, 1384 [2005]). This term is also translated as ceremony in *tašrīfāt-e tanazzol-e pāyghāh* (status degradation ceremony, تشریفات تنزل پایگاه). Finally, it conveys the meaning of “fine” or “high class” in such compound words as *ḥadamāt-e ḥūdro-we tašrīfāt* (limousine service, خدمات خودرو تشریفات) as well as the synonyms *restūrān-e tašrīfātī* (fine dining restaurant, رستوران تشریفات) and *gāzāḥūre-ye tašrīfātī* (fine dining restaurant, غذاخوری تشریفات).

As both foreign and approved terms have various connotations, the following tables will present the frequency of each term in both corpora in addition to their collocations in order to narrow down the scope in which each term is used and compare both of them.

Term forms		Frequency		Document frequency	
		TalkBank	TAP	TalkBank	TAP
<i>protokol</i>	پروتکل	3,882	5,487	2,110	2,327
<i>tašrīfāt</i>	تشریفات	4,147	462	3,031	269

Table 47. Term frequency and document frequency of the terms in the semantic field of *protokol*

A simple look at the pure term frequencies shows, for example, that *tašrīfāt* occurs a little more often in TalkBank than *protokol*; however, when looking at both corpora, *protokol* occurs more than twice as often as *tašrīfāt*. Interestingly, in contrast to what could be expected, *tašrīfāt* hardly exists in newspapers; less than four percent of all mentions can be found in TAP. This does not yet, however, account for different connotations of the two terms.

As various as the connotations for both terms are, as variously they are used in practice. For instance, in the blog-based corpus TalkBank, the term *protokol* is used in the context of constitutional law such as in the compound word *protokol-e elḥāqī* (additional protocol, پروتکل الحاقی), which produces 1,320 hits (see example 1 in table 48). It is also used 333 times in the field of Internet and technology as *protokol-e / protokolhā-ye īntarnet* (Internet protocol, پروتکل/پروتکل‌های اینترنت) and *prūtolkolhā-ye hasta'e-ye īntarnet* (Internet core protocols, پروتکل‌های هسته‌ای اینترنت) (see example 2). *Protokol* is also used with *šeheyūnistī* even if not very frequently with only 81 times (see example 3) probably referring to the antisemitic *Protocols*

of the Elders of Zion,<sup>54</sup> which can also be found several dozen times as *Protokolhā-ye hokamā-ye Şehyūn* (پروتکل های حکمای صهیون), *Protokolhā-ye bozorgān-e Şehyūn* (پروتکل های بزرگان صهیون), or *Protokolhā-ye dāneşvarān-e Şehyūn* (پروتکل های دانشوران صهیون) in TalkBank only.

In newspapers, the term *protokol* is used, among others, in the compound word *protokolhā-ye behdāštī* (health protocols, پروتکل های بهداشتی) 2,237 times (see example 4), which deals with the topic of the novel Corona virus. This makes up for 40 percent of all mentions in TAP. Here, the term *protokol* has spread widely during the Covid-19 pandemic in the sense of a medical protocol followed to treat patients infected with the Corona virus. The collocation *protokol-e elhāqī* (additional protocol, پروتکل الحاقی) is mentioned 1,125 times in TAP and thus roughly as often as in TalkBank.

<b>Example 1</b>	سند دوم پروتکل الحاقی است که شامل شرایط و ضوابط مورد قبول ایران و دیگر اعضای کار گروه جهت عضویت ایران در سازمان تجارت جهانی است که در مذاکرات چند جانبه کار گروه تدوین شده است.
Transliteration	Sanad-e dovvom-e prütokol-e elhāqī ast ke šāmel-e šarāyeṭ va žavābeṭ-e müred-e qobül-e Īrān va dīgar-e ā'zā-ye kārğorūh-e ġehat-e 'ożveyyat-e Īrān dar sāzmān-e teġārat-e ġahānī ast ke dar mozākārāt-e čand-e ġanebe-ye kārğorūh tadvīn šode ast.
Translation	The second document is the Additional Protocol, which includes the terms and conditions accepted by Iran and other members of the Working Group for Iran's membership in the World Trade Organization, which have been developed in the Group's multilateral negotiations.
<b>Example 2</b>	به گزارش « مهر IP-TV » یا همان تلویزیون اینترنتی به ارسال و پخش برنامه های تلویزیونی برای مشترکان از طریق زیرساخت باند وسیع و توسط پروتکل اینترنت گفته می شود.
Transliteration	Be gozāreš-e "Mehr IP-TV" yā hamān televīzyūn-e ĩntarneti be ersāl va paḡš-e barnāmeḡā-ye televīzyūnī barā-ye moštarekān az ṭarīq-e zīresāḡt-e bānd-e vasi' va tavasoṭ-e prütokol-e ĩntarnet gofteh mišavad.
Translation	According to Mehr, IP-TV or Internet TV is the sending and broadcasting of TV programs to subscribers through a broadband infrastructure and by the Internet Protocol.
<b>Example 3</b>	تلاش شبکه های رسانه ای نظام استبداد بین المللی را برای برجسته کردن الگوهای منحرف و سخیف، برگرفته از اهداف خطرناک پروتکل های صهیونیستی دانستند و افزودند: این الگوسازی شبکه ای، بر اساس سیاست های مشخص و مدون صورت می گیرد.
Transliteration	Talāš-e šabakeḡā-ye rasāne-ye nezām-e estebdād-e baynolmelalī rā barā-ye bargaste kardan-e algūḡā-ye monḡaref va saḡīf, bargerefteḡ az āḡdāf-e ḡaṭarnāk-e prütokolhā-ye šeheyūnistī dānestand va afzūdand: ĩn algūsāzī-ye šabakeyī, bar asās-e seyāsathā-ye mošaḡaḡ va modavvan sūrat mīġīrad.
Translation	They considered the efforts of the media networks of the international authoritarian regime to highlight the deviant and degraded patterns taken from the dangerous goals of the Zionist protocols and added: This network modeling is based on specific and planned policies.

<sup>54</sup> For more details on the background of the *Protocols of the Elders of Zion* see, for example, (Webman, 2011), and especially the chapter by (Rahimiyan, 2011).

<b>Example 4</b>	معاون درمان وزیر بهداشت با اشاره به شیوع بالای بیماریهای حاد تنفسی ناشی از کرونا در چند استان کشور، گفت: عدم رعایت پروتکل‌های بهداشتی و ناقلین بدون علامت، نقش زیادی در گسترش این ویروس دارد و یکی از مهمترین راههای پیشگیری استفاده از وسایل حفاظت فردی همچون ماسک است.
Transliteration	Mo'āven-e darmān-e vazīr-e behdāšt bā ešāreh be šeyū '-e balā-ye bīmārīhā-ye ḥād-e tanaffosī nāšī az korūnā dar čand-e estān-e kešvar, goft: 'adam re'āyat-e prūtokolhā-ye behdāštī va nāqelīn-e bedūn-e 'alāmat, naqš-e zeyādī dar gostareš-e īn vīrūs dārad va yekī az mohemtarīn-e rāhhā-ye pīšgīrī estefādeh az vasāyel-e hefāzat-e fardī hamčūn-e māsak ast.
Translation	Referring to the high prevalence of acute respiratory diseases caused by the Corona virus in several governorates of the country, the Deputy Minister of Health said: Failure to follow health protocols and asymptomatic carriers play a major role in the spread of the virus, and one of the most important ways to prevent it is to use personal protective equipment such as masks.

Table 48. Examples of the foreign term *protokol* in the corpora

The approved term *tašrīfāt* is also used with its several connotations. The most frequent collocates of the approved term *tašrīfāt* in TalkBank are as follows: *qānūnī* (law, قانونی, see example 3 in table 49 below), *gomrokī* (customs, گمرکی, see example 4), *edārī* (administrative, اداری), *dādrasī* (trial, jurisdiction, دادرسی), *qovve-ye qazāyeye* (judiciary, قوه قضاییه), and *rayīs-e ḡomhūr* (president, رییس جمهور, see examples 1 and 5). These could be summarized as belonging to the semantic field of politics or the political system. Generally, in both corpora, *tašrīfāt* indicates the same intended meaning, which is related to any protocol followed to achieve a specific target, whether it is a political protocol between countries, or a protocol followed in the judiciary or administrative authorities.

Another prominent collocation occurs in the context of marriage as a synonym of *marāsem* (ceremony, مراسم) with the meaning of formalities and ceremonies. Here, *ezdevāḡ* (marriage, ازدواج) stands out with 134 hits. However, in TAP, the collocate *ezdevāḡ* has only 13 hits (see examples 2 and 6). Other than this, *tašrīfāt* points to the concept of protocol.

<b>Example 1</b>	اینکه رئیس جمهور فعلی به دیدار رئیس جمهور منتخب برود، یک رسم تشریفاتی سیاسی است که در همه دوره ها برگزار شده است.
Transliteration	Īnke ra'īs-e ḡomhūr-e fe'lī be dīdār-e ra'īs-e ḡomhūr-e montaḥab beravad, yek rasm-e tašrīfātī-ye seyāsī ast ke dar hame dūrehā bargozār šode ast.
Translation	The current president's visit to the newly elected president is a political tradition [i.e., protocol] that has existed throughout history.
<b>Example 2</b>	خوشبختانه شیوع کرونا باعث شد تا همه تشریفات و برنامه‌های اضافی که در مراسم عروسی است حذف شود و من هم خیلی زود با همسرم عقد کردم و این کار بار سنگینی را از دوش خانواده ها برداشت.
Transliteration	Ḥūšbaḥtāneh-ye šeyū '-e korūnā bā'eṭ šod tā hameh-ye tašrīfāt va barnāmeḥā-ye ezāfī ke dar marāsem-e 'arūsī ast, ḥazf šūd va man ham ḥīlī zūd bā hamsaram 'aqd kardam va īn kār-e bār-e sangīnī rā az dūš ḥānvādehā bardāšt.
Translation	Luckily, the Corona outbreak eliminated all extra formalities [i.e., protocols] and additional programs at the wedding ceremony, and I got married to my wife very soon, this removed a very heavy burden from families as well.

<b>Example 3</b>	این توافقنامه پس از گذراندن مراحل <b>تشریفات قانونی</b> داخلی، پس از سی روز از تاریخ دریافت آخرین یادداشت رسمی از طریق مجاری دیپلماتیک به مدت نامحدودی لازم الاجرا است، مگر اینکه یکی از دو طرف به طور کتبی تصمیم خود را مبنی بر پایان دادن به اعتبار آن اعلام کند.
Transliteration	Īn tavāfoqnāme pas az gozrāndan-e marāḥel-e tašrīfāt-e qānūnī-ye dāḥelī, pas az sī rūz az tāriḥ-e daryāft-e āḥarīn-e yāddāšt-e rasmī az ṭarīq-e moḡārī-ye dīplomātīk be moddat-e nāmaḥdūdī-ye lāzem al-eḡrā ast, magar īnke yekī az dū ṭaraf be ṭur-e katbī tašmīm-e ḥūd rā mabnī bar pāyān dādan be e' tebār-e ān e' lām konad.
Translation	Following the conclusion of internal legal procedures [i.e., protocols], this Agreement will enter into force indefinitely thirty days after the date of receipt of the last official memorandum through diplomatic channels unless one of the parties notifies the other in writing that the contract is no longer valid.
<b>Example 4</b>	رای کابوتاژ <b>تشریفات گمرکی</b> لازم است و دو نوع اظهارنامه دارد که یکی هنگام خروج کالا از بندر یا مرز تنظیم می‌شود و یکی هم هنگام ورود به گمرک مرز یا بندر.
Transliteration	Rā-ye kābūtāž-e tašrīfāt-e gomrokī lāzem ast va dū nū 'e ezhār nameh dārad ke yekī hangām-e ḥorūḡ-e kālā az bandar yā marz tanzīm mīšavad va yekī ham hangām-e vorūd be gomrok-e marz yā bandar.
Translation	The coastal navigation vote is required for customs procedures [i.e., protocols] and has two types of declarations, one to be set when the goods leave the port or border and the other when the customs enter the border or port.
<b>Example 5</b>	مقام ریاست کشور مقامی <b>تشریفات</b> است و در عمل کارهای روزمره کشورهای قلمرو از طریق دولت انتخابی به ریاست نخست وزیر انجام می‌شود.
Transliteration	Maqām-e reyāsat-e kešvar-e maqāmī-ye tašrīfātī ast va dar 'amal-e kārhā-ye rūzmarreh-e kešvarhā-ye qalamrū az ṭarīq-e dūlat-e entehābī be reyāsat-e noḥost-e vazīr angām mīšavad.
Translation	The presidency is a ceremonial authority [i.e., based on formal protocol], and in practice the day-to-day affairs of the countries of the territory are carried out through an elected government headed by the prime minister.
<b>Example 6</b>	گزارش «ایران» از تأثیر کرونا بر تشویق خانواده ها به برپایی ازدواج آسان و به دور از تشریفات.
Transliteration	Gozāreš-e "Īrān" az ta'ṭīr-e korūnā bar tašvīq-e ḥānvādehā be bar pāyī-ye ezdevāḡ-e āsān va be dūr az tašrīfāt.
Translation	"Iran" report on the effect of Corona on encouraging families to have an easy marriage away from formalities [i.e., protocols].

Table 49. Examples of the approved term *tašrīfāt* in both corpora

If one looks at the overall adjusted weight only, then the results for the term pair *protokol/tašrīfāt* are mixed. While in TalkBank *protokol* is narrowly lighter than *tašrīfāt* with an adjusted weight of 22,194 compared to 22,335, *protokol* is significantly heavier in TAP with an adjusted weight of 299,601 compared to 58,491. These results are very similar to the ones found for the term pair *režīm/nezām* where the ratio of foreign term and Persian equivalent is close to 50:50.

It is therefore useful to approximate the weight of *tašrīfāt* in the field of politics. This means subtracting from the term frequency those instances where it occurs with the non-political collocates *ezdevāḡ* and *'arūsī*. This does not exclude many occurrences because if *tašrīfāt* is used, it is most often in a political sense. The approximated adjusted weight of *tašrīfāt* in TalkBank is then 22,160 and in TAP 57,902. This approximation changes nothing in case of

TAP. However, in TalkBank, the results are now reversed leaning slightly in favor of the foreign term although it is still scratching on the 50:50 ratio.

## Chapter 6

### Findings and Conclusion

#### 6.1 Findings: Describing the Impact of Coined Terms on Modern Persian

This study examined the term-formation process by one of the most prolific language-planning institutions in the world: the Academy of Persian Language and Literature (APLL). Throughout its history, it has been tasked with creating new Persian terms and replacing loanwords ridding Persian especially from English and French influence, often reflecting scientific and technological progress. The special focus of this thesis is the success – or lack thereof – of political terms coined by the APLL terminology group.

The APLL is the official institution supported by the government of the Islamic Republic of Iran that carries the responsibility for preserving the Persian language and Persian orthography. The main task of the terminology group of the Farhangestān is to coin new terms or even retrieve original older terms – in accordance with the regulations of the APLL – in order to replace terms that are foreign to the language as well as to form neologisms for all the new concepts that exist in line with the civilizational and technological development in all fields even before those foreign terms spread among Persian speakers. This study thusly contributes to the understanding of a hitherto understudied or even neglected institution in Western academia. In Iran, on the other hand, there have been several studies and papers on the APLL, for example, with regard to the new terms approved by the Academy and criticism thereof.

Previous studies that were conducted in order to evaluate the impact of the Academy in spreading the use of approved terms depended either on questionnaires among a specific group of Persian native speakers or on the *Hamshahri* corpus, which contains the archive of the Iranian newspaper *Hamshahri online* and does neither represent the actual accurate use of the approved terms nowadays nor even a sizeable number of words.

In order to answer the question whether the selected terms approved by the APLL have successfully found their way into Persian, this chapter summarizes the key findings of this thesis. It is methodologically based on the corpus analysis of two different corpora with more than 500 million words: one presenting the online editions of a broad range of Persian-language Iranian newspapers (TAP) and the second one presenting an even larger variety of Internet blogs written and published by Persian native speakers (TalkBank). While the newspapers are

subject to editorial oversight and, therefore, are prone to reflect formal language, blog posts can be less formal and reflect spoken language more than standardized written language, even including the occurrence of more spelling mistakes than in newspapers. Blogs are an outward reflection of the virtual private sphere that may not answer to many restrictions and users may enjoy a higher degree of freedom of expression.

What distinguishes the present study is that the corpora used contain an extensive large number of Iranian news items until the first half of 2021 and the personal writings and blogs of Iranians at least up until the year 2017. This diversity and difference in the types of corpora greatly serves this study to answer the research questions in a comprehensive and accurate manner. Each of the two corpora is a primary source for the actual use of the terms approved by the APLL and the foreign terms used in Persian, be it among bloggers as not necessarily specialized native speakers who may not have enjoyed formal training and education in a writing profession or be it among journalists as specialists who have presumably studied how to write and edit news in a professional manner.

The quantifying factor to measure the success of each term is its weight or TF-IDF value, which is calculated by using the term frequency (TF) and the inverse document frequency (IDF), which in turn rely on the total number of words and documents in the corpus. The weight of each term is calculated separately for each corpus so that regularly there are four weight values for each term pair that can be compared easily hinting not only at the success and the extent of the use of the approved terms, but also at the success of the work of the terminology group of the APLL. There are a few exceptions, though, where more than four weight values for a term pair are calculated, which will be reviewed below.

The findings of the comparative term analysis in chapter 5 will be discussed through a series of charts that visualize the respective weight values of the examined terms – both the approved Persian ones and their foreign equivalents. As mentioned in the beginning of chapter 5, not all terms carry the same linguistic and semantic features considering their usage. With some terms, both the approved terms and their foreign equivalents could refer to only one semantic interpretation, which makes it easy to be compared. With other terms, which will be discussed in separate charts below, either the approved term or its foreign equivalent or even both may indicate more than one meaning. This is the case where the collocations of the terms will be identified to help contextualize these terms.



The first chart below clarifies the first set of term pairs in order to answer the main question of this study about measuring the impact of the approved terms in contrast to their foreign equivalents in TalkBank. The chart gives an overview of those approved political terms and their corresponding loanwords that do not require focusing on collocations because each of those terms has only one connotation. They are in this sense unambiguous. In this as in all other cases, the horizontal axis indicates the term pair with green columns for foreign terms and red columns for their Persian equivalents. The vertical axis displays the adjusted weight value of each term. It is based on a logarithmic scale so that very low and very high weight values can be visualized in the same chart. This means that each horizontal gridline represents a ten-fold increase of the value of a column compared to the next lower horizontal gridline. Columns that appear to be of similar height might therefore in reality reflect starkly different weight values.

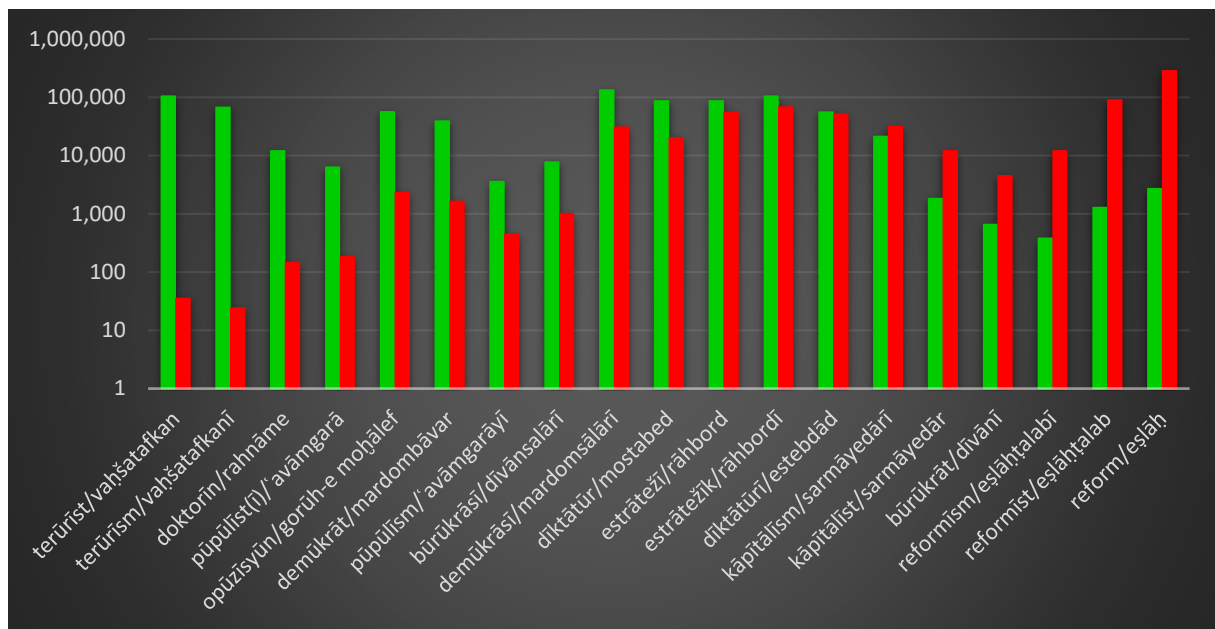


Chart 1. Comparison of adjusted weight values of foreign terms (green) and their Persian equivalents (red) in TalkBank

Chart 1 presents an overview of the comparison of the Persian terms approved by the APPL in contrast to the foreign terms that have been in use in Persian in TalkBank. The term pairs are sorted from the lowest weight ratio between foreign term and Persian equivalent on the left to the highest weight ratio on the right. That is, the most successful Persian equivalents can be found on the right side of the chart.

The data can be categorized into three groups. The first group includes the first ten term pairs in the chart which are *terūrīst/vaḥṣatafkan*, *terūrīsm/vaḥṣatafkanī*, *doktorīn/rahnāme*, *pūpūlīst(ī)/‘avāmgarā*, *opūzīsyūn/gorūh-e moḥālef*, *demūkrāt/mardombāvar*, *pūpūlīsm/‘avāmgarāyī*, *būrūkrāsī/dīvānsalārī*, *demūkrāsī/mardomsālārī*, and *dīktātūr/mostabed*. This category comprises the approved terms that were comparatively

unsuccessful. For example, the two loanwords *terūrīst* and *terūrīsm* record each an almost 3,000 times heavier weight compared to their Persian equivalents. By contrast, the approved terms *vahšatafkan* and *vahšatafkanī* show very low, two-digit only weight values – in fact the lowest among all terms in the chart. The remaining eight foreign terms in this group are between 4.2 and 83 times heavier than their respective Persian equivalents. With all of this, it is only of subordinated relevance whether any given term is widespread in the corpus or not because – thanks to the weight calculation – the absolute term and document frequencies are not by themselves necessary or sufficient to explain the impact of that term.

The second group comprises four term pairs and extends from *estrātežī/rāhbord* to *kāpītālīsm/sarmāyedārī*. The weight ratio for each term pair in this group lies between 1 and 2. A value of 1 means that the two terms enjoy the same weight whereas a value of 2 means that one term is twice as heavy as the other term. While this line is arguably drawn arbitrarily, it is still obvious that the term pairs in this group cannot clearly be categorized as favorable towards the success or failure of the Persian equivalents – even more so as all terms in this group occur quite frequently in the corpus.

This second group can be further divided into two subgroups. In case of the first three term pairs *estrātežī/rāhbord*, *estrātežīk/rāhbordī*, and *dīktātūrī/estebdād*, the foreign term is slightly heavier than its Persian equivalent. In case of the term pair *kāpītālīsm/sarmāyedārī*, the weight of the Persian equivalent is a little bit higher than that of the foreign term.

The third group, finally, represents five term pairs where the Persian equivalents prove to be noticeably more successful than the foreign terms. The weight ratios of *kāpītālīst/sarmāyedārī*, *būrūkrāt/dīvānī*, *reformīsm/eslāḥṭalabī*, *reformīst/eslāḥṭalab*, and *reform/eslāḥ* favor the Persian equivalents by a factor of between 6.5 and 106.

All in all, in the blogs collected in TalkBank, the Persian equivalents are not so strong vis-à-vis the foreign terms, neither in numbers – just five clearly successful terms compared to ten with little impact – nor in weight ratios. This was to be expected from a look at blogs that are a less formal way of disseminating news and information and, at the same time, a source of information that is not necessarily accountable to editorial oversight to the same extent as newspapers. It was therefore foreseeable that the examined Persian equivalents have a difficult standing in blogs. In the following, a closer look at TAP and the collected newspaper articles in that corpus will therefore put this set of findings in perspective.

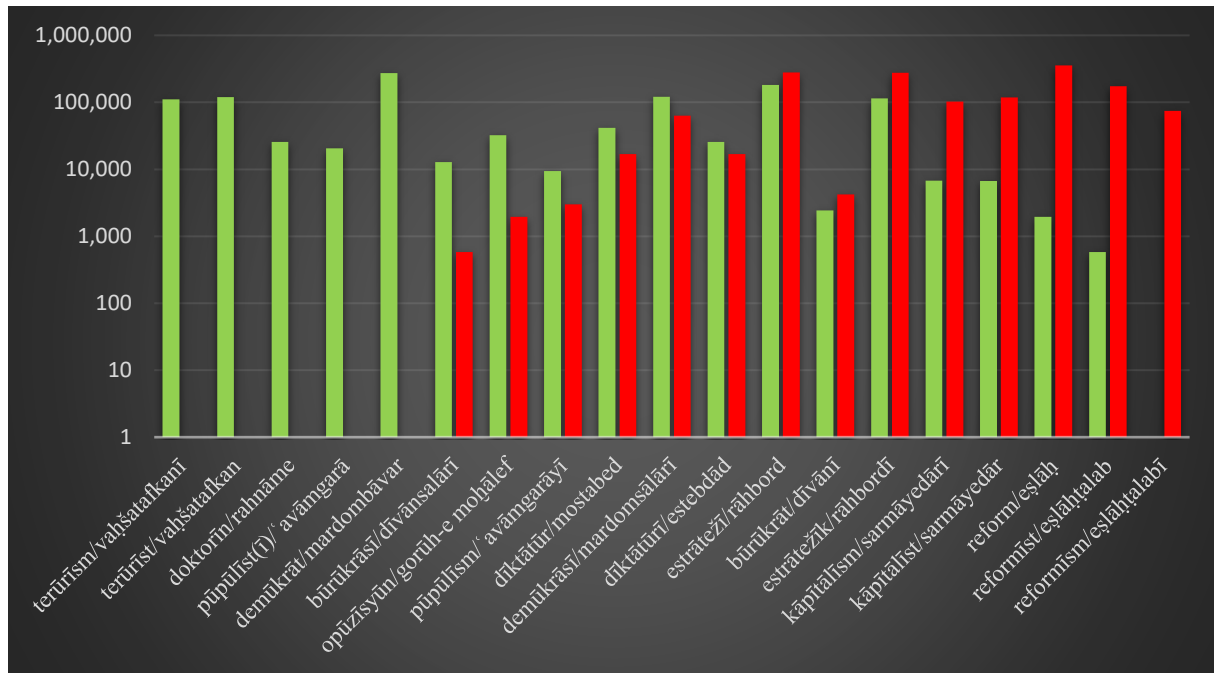


Chart 2. Comparison of adjusted weight values of foreign terms (green) and their Persian equivalents (red) in TAP

Chart 2 displays the same overview as chart 1 albeit based on TAP. Here, too, three groups of term pairs can be discerned that will briefly be described in the following. Like with TalkBank, the first group contains the term pairs where the Persian equivalents have a comparatively low weight. It contains nine term pairs, five of which are distinct in the way that there is no mention of the Persian equivalent in the corpus at all. The four remaining term pairs in this group have weight ratios between 2.4 and 23. These ratios are lower than those in TalkBank. This means either that the Persian equivalents tend to be heavier in newspapers than in blogs or that the foreign terms are lighter in newspapers than in blogs.

The second group covers the four term pairs *demūkrāsī/mardomsālārī*, *dīktātūrī/estebdād*, *estrātežī/rāhbord*, and *būrūkrāt/dīvānī*. Their weight ratios lie between 1.5 and 1.9. These four term pairs can therefore be considered to be evenly represented in the corpus. While for the first two term pairs the ratios lean in favor of the foreign terms, for the second two term pairs, they shift to the Persian equivalents.

Finally, the third group is made up of six term pairs, among which are all three term pairs relating to *reform*. In this group, too, there is one term pair where one term – the loanword *reformīsm* – does not occur in the corpus at all. The other three term pairs are *kāpītālīst/sarmāyedār*, *kāpītālīsm/sarmāyedārī*, and *estrātežīk/rāhbordī*. The weight ratios of the term pairs in group 3 – that is, weight of Persian equivalent divided by weight of foreign term – range between 2.4 and 299.

Comparing the two corpora with each other reveals a few noteworthy findings. What can be gathered from this comparative overview between the two corpora, between blogs and newspapers, is that the overall picture is the same no matter which primary source one looks at. The Persian equivalents that can be considered successful are almost the same in both corpora: *eslāḥ*, *eslāḥṭalab*, and *eslāḥṭalabī* are among the five (TalkBank) or six (TAP) terms with the highest weight ratio in favor of the Persian equivalents. The same is true for *sarmāyedār* as an equivalent to *kāpīṭalīst*. Similarly, the four least successful Persian equivalents are the same in both corpora: *vahšatafkanī*, *vahšatafkan*, *rahnāme*, and *‘avāmgarā*. They are not at all in use in the newspapers in TAP and in very low comparative use in the blogs in TalkBank.

The discussion has also shown that newspapers are generally more likely to use Persian equivalents. In fact, six of the approved Persian equivalents enjoy a more than two times higher use in newspapers than the respective foreign terms; in blogs, it is only five Persian equivalents with the same status. In the case of *eslāḥṭalabī*, the foreign term *reformism* does not even occur once in TAP. On the other hand, the representation of such an “extreme” value seems to be a feature of TAP because there are also five Persian equivalents that do not occur in the selected newspapers at all.

All of the above is true for terms which show no ambiguities in meaning. There is a second set of term pairs that merits its own discussion. It is concerned with terms that need collocations to identify their specific connotations among a variety of meanings. It is necessary to analyze their collocations in order to be able to extrapolate the semantic context of politics in which the term is mentioned. Only then will it be meaningful to compare the weight values of such foreign terms with their Persian equivalents. In this category six pairs of approved terms and loanwords can be identified:

1. the approved term *maḡles* and the loanword *pārlemān*,
2. the approved term *nāmzad* and the loanword *kāndīdā*,
3. the approved term *mellat* and the loanword *ommat*,
4. the approved term *nezām* and the loanword *režīm*,
5. the approved term *īde ‘ulūžī* and the loanword *marām* with the derivative term pair *īde ‘ulūžīkī/marāmī*,
6. and finally, the approved term *tašrīfāt* and the loanword *protūkūl*.

Each of the previous cases is explained in a single chart in the following. These charts are not logarithmic in nature so that the values displayed can be read in a linear manner where similar column heights reflect similar weight values.

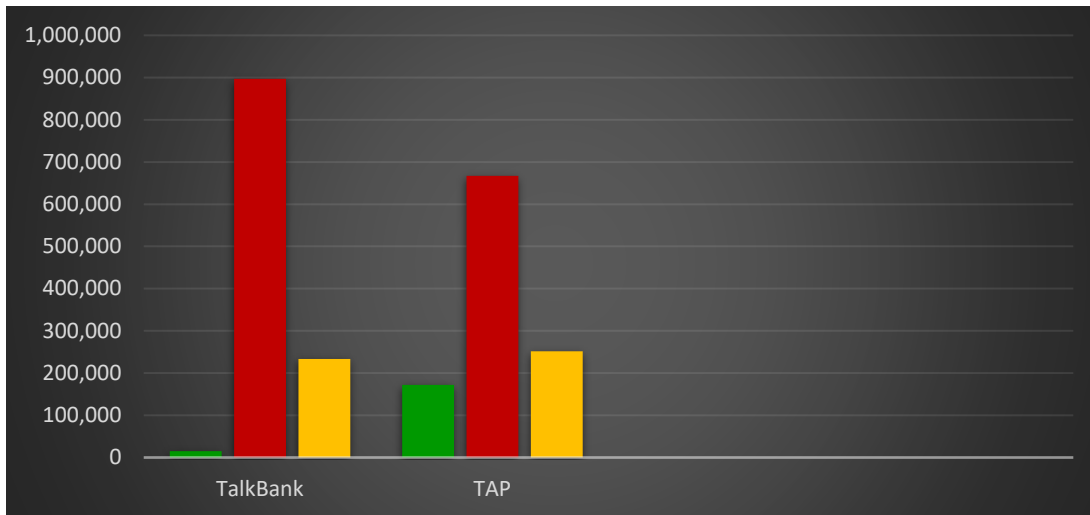


Chart 3. Comparison of adjusted weight values of the terms *p̄rlemān* (green) and *maġles* (red = absolute weight, yellow = approximate weight) in both corpora

The first term pair with ambiguities is *p̄rlemān/maġles*. The analysis of the blogs and newspapers collected in the two corpora reveals that each of the terms has a usage different from the other. This is due to several factors the most important of which is the name used by a sovereign state for its legislative body; some use a variation of the originally Arabic term *maġles* – such as Iran, Turkey, and several Arab countries – and others a different proper noun that is then oftentimes translated as *p̄rlemān*.

This is evident by reading the texts in the two corpora. Very often, in the case of written news about other countries and not about Iran, whether translated or not, authors retain the name used in the original text or the name used by the country itself. However, the results are slightly different when comparing between the newspaper-based corpus TAP and the blog posts-based corpus TalkBank.

Since the approved term *maġles*, which is meant to replace *p̄rlemān*, includes multiple meanings that are broader than the purely political meaning (as explained in chapter 5.1), it is necessary to approximate its weight in the political context and for the targeted meaning only. *Maġles* records a heavy weight in both corpora. The following chart presents a comparison between the weight of *maġles* and *parlemān* in general in the corpora and between *maġles* and *p̄rlemān* only when it refers to the political meaning.

The red columns in the chart present the weight of *maġles* in general in both corpora, which is the heaviest among all researched terms. The yellow columns visualize the weight of *maġles* only when used to refer to *parliament*. The green columns highlight the weight of *p̄rlemān*. From the chart it can be seen that *maġles* as a general term and as a political term has by far the heaviest weight regardless of the corpus.

Nevertheless, looking into the use of the term *mağles* in both corpora by checking its most common collocations, it is clear that *mağles* is used almost exclusively to denote the legislative bodies and political institutions in the Islamic Republic of Iran and its predecessor states such as *mağles-e šūrā-ye eslāmī*, *mağles-e šūrā-ye negahbān*, *mağles-e ħobregān*, and *mağles-e šūrā-ye mellī*. As for *pārlemān*, it is mentioned most often to refer to non-Iranian parliaments such as *pārlemān-e Orūpā*, *pārlemān-e Ītāleyā*, etc. This highlights that neither of the two terms *mağles* and *pārlemān* is explicitly a substitute for the respective other term. Despite the overwhelming numbers, *mağles* enjoys only limited success when it comes to its use for parliaments other than the Iranian one.

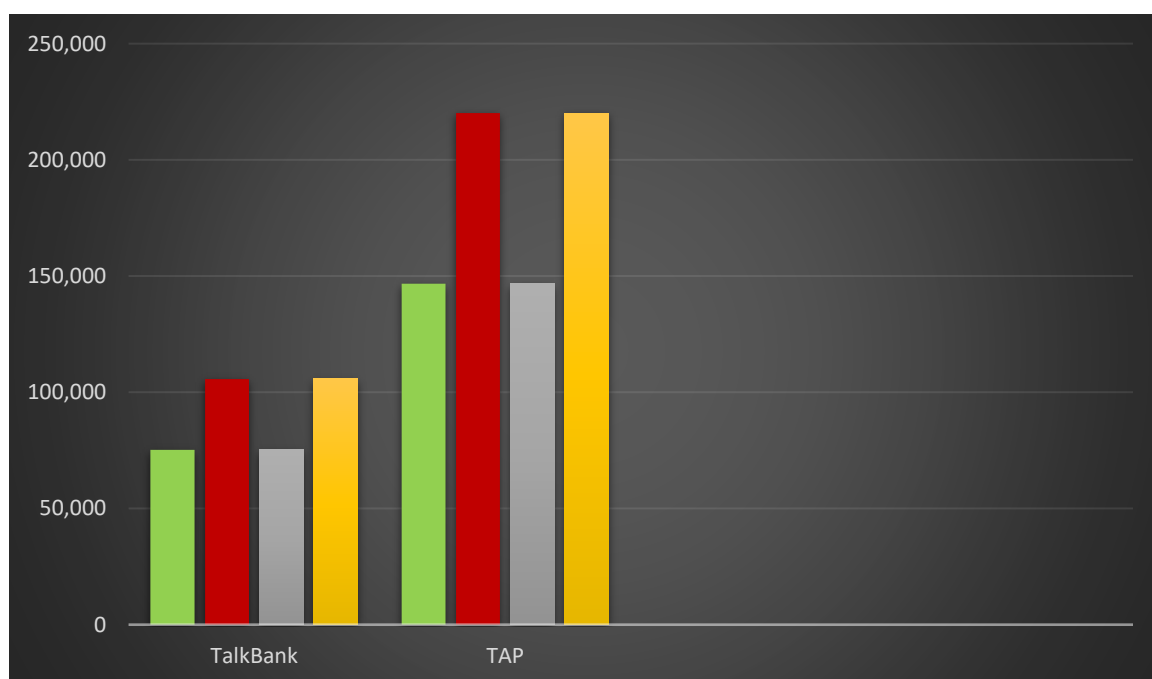


Chart 4. Comparison of adjusted weight values of the terms *kândīdā* (green) and *nāmzad* (red) and approximate weight values of *kândīdā* (gray) and *nāmzad* (yellow) in both corpora

The second ambiguous term pair that requires the use of collocations to evaluate its success in the political field is *nāmzad* and *kândīdā*. Especially the approved term *nāmzad* could refer to more than one meaning. To be able to get an approximate weight of both terms in a political context, their weight values are calculated through the most frequent collocations that point at a political context. Those terms are *enteḥābāt*, *reyāsat*, *ğomhūrī*, and *mağles*. It is possible in Sketch Engine to exclude duplicates, that is, double mentions of terms, so that for this approximation all texts that include *kândīdā* or *nāmzad* with one of the four collocates can be added. The resulting number of occurrences serves as the basis for the calculation of the weight.

It is clear from chart 4 that the approved term *nāmzad* is heavier than the loanword *kândīdā* with the indication of a political candidate in both corpora. In general, in the case of

these two terms, *nāmzad* is used more than the foreign term and weighs heavier than *kāndīdā*. Once more, the weight of the approved Persian equivalent *nāmzad* is heavier in newspapers than in blogs, both relative to the foreign term *kāndīdā* in newspapers as well as to the weight of *nāmzad* in blogs. Although the term *kāndīdā* enjoys a similarly heavy weight as *nāmzad* (with a weight ratio between 1 and 2), at first glance, *nāmzad* is a successful coining by the APLL. But since the ratio factor is less than 2 in both corpora, the term pair *nāmzad/kāndīdā* qualifies only as evenly represented in the two corpora.

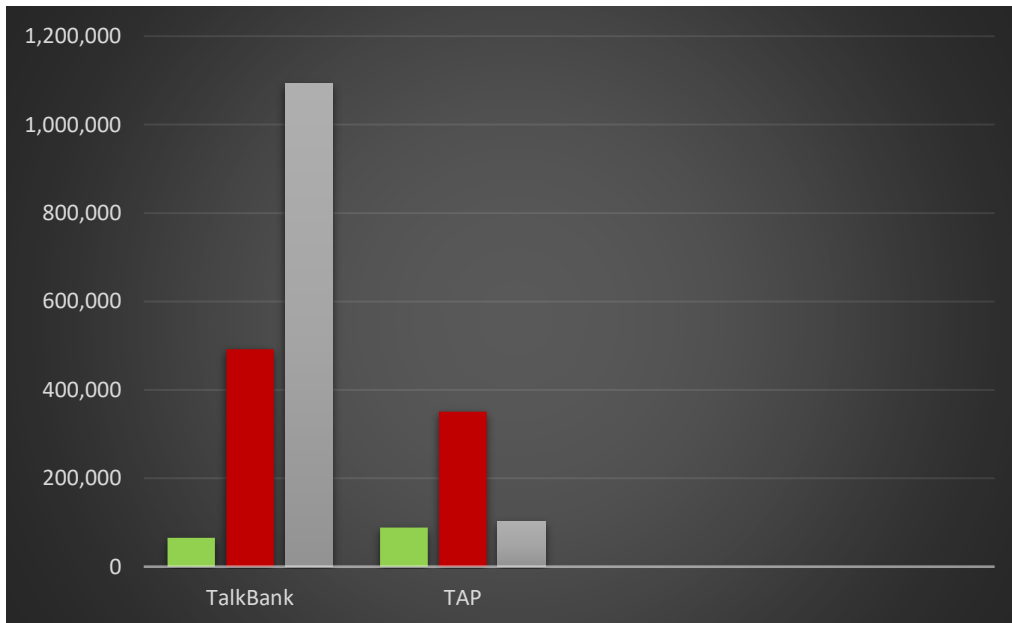


Chart 5. Comparison of adjusted weight values of the terms *ommat* (green), *mellat* (red), and *mardom* (gray) in both corpora

The concept of *nation* is not so much ambiguous as it is best described as a term triplet rather than a term pair. The comparison between *ommat* as a religiously connotated term on the one hand and *mellat* and *mardom* as two terms, on the other hand, that convey a worldly understanding of the *nation* and the *people* respectively shows that *ommat* plays a minor role in this triplet. *Ommat* occurs most frequently with collocations such as *eslām* and *mosalmān* while *mellat* and *mardom* are mostly represented with country names and other entities of territorial belonging. The weight of *mardom* in TalkBank exceeds by far the weight of both *mellat* and *ommat*. However, the approved term *mellat* is heaviest in TAP. This displays, once more, the Farhangestān-friendly approach of newspapers.

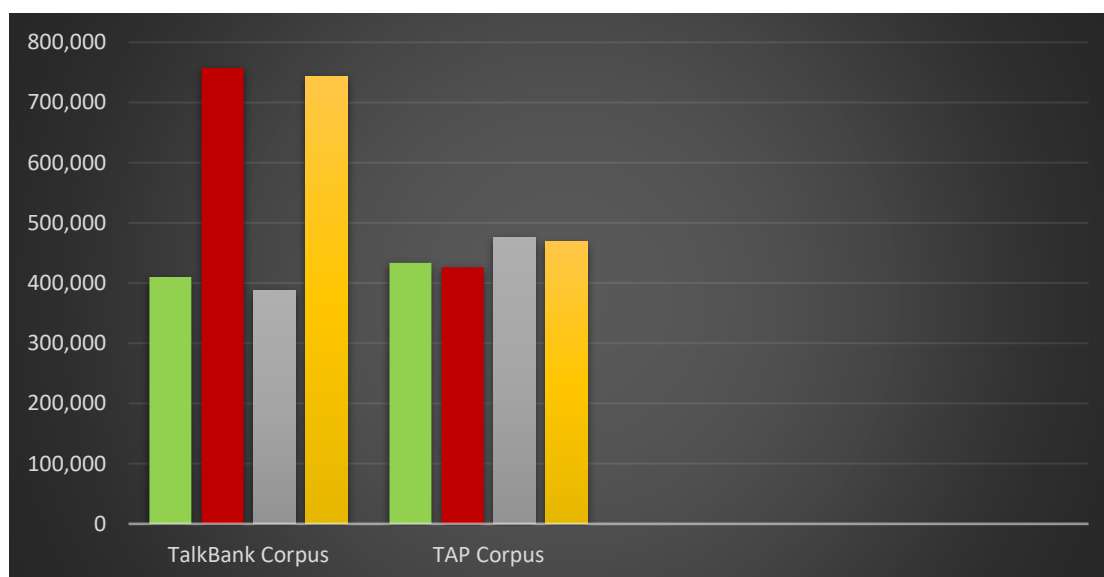


Chart 6. Comparison of weight values of the terms *reżīm* (green) and *neżām* (red) as well as of approximate weight values of the terms *reżīm* (gray) and *neżām* (yellow) in both corpora

In the case of the terms that refer to the concept of *regime* as a political system, both the foreign term and its Persian equivalent were approved by the APLL. However, the APLL approved the foreign term *reżīm* mostly in the field of food and nutrition with around 98 compound terms referring to the dietary regime that people follow for health reasons. In addition, the APLL approved *reżīm* four times in a political context. These compound terms that keep the foreign term *reżīm* are *reżīm-e baroštāde* (ancient regime), *reżīm-e neżāmī* (military regime), *reżīm-e solṭānī* (sultanic regime), and *reżīm-e dastnešānde* (puppet regime). The first three of these compound terms have a negative connotation as they point to the government of the Shah before the Iranian Revolution in 1979. The fourth compound term refers to a regime whose officials act in the interest of another government which carries a negative connotation as well. They can certainly be considered negative from the perspective of the Islamic Republic of Iran because they all refer to political systems unacceptable to the state. Moreover, *reżīm* shows a high frequency with the collocates *ṣehyūnīstī* (Zionist), *Esrāyīl* (Israel), *ašgālgar* (occupier), *nāmašrū* (illegal), as well as *Pahlavī* (Pahlavi) and *šāhenšāhī* (imperial) to name but a few (see chapter 5.7).

Unlike *reżīm*, the approved Persian equivalent *neżām* evokes more positive associations. Originally coming to Persian from Arabic, *neżām* could refer to more than just the meaning of political system; it is in this sense similar to *reżīm*. *Neżām* could refer to educational system, economic system, etc.

Despite the limited approval of *reżīm* in a political context and the fact that these approvals are negatively connotated, the foreign term *reżīm* is still widely used among native



speakers in blogs as well as in newspapers. That is in line with the observation that *reżīm* in the meaning of a political system shows a heavy weight in both corpora. It is also clear from chart 3 that the results are different between blogs and newspapers. In TalkBank, the weight of *neżām* is almost twice the weight of *reżīm*. Yet in TAP, the weight of both terms is very similar tipping slightly in favor of *reżīm*. Despite this, in analogy with the other terms discussed above, the term pair *reżīm/neżām* qualifies as evenly represented in the corpora because its weight ratio lies between 1 and 2.

The application of the two approved terms *reżīm* and *neżām* is interesting because the two terms – although describing the same concept, namely a political system – have acquired different connotations. These connotations are partly reflected and reinforced by the compound terms approved by the APLL. A case in point is that the term *reżīm* with its negative connotation is not mentioned even once to refer to the Islamic Republic of Iran in newspaper articles. However, it is mentioned around 8,000 times in blogs. This is an indication of blogs and bloggers being more outspoken on sensitive issues. It is also an indication of their being less susceptible to language planning.

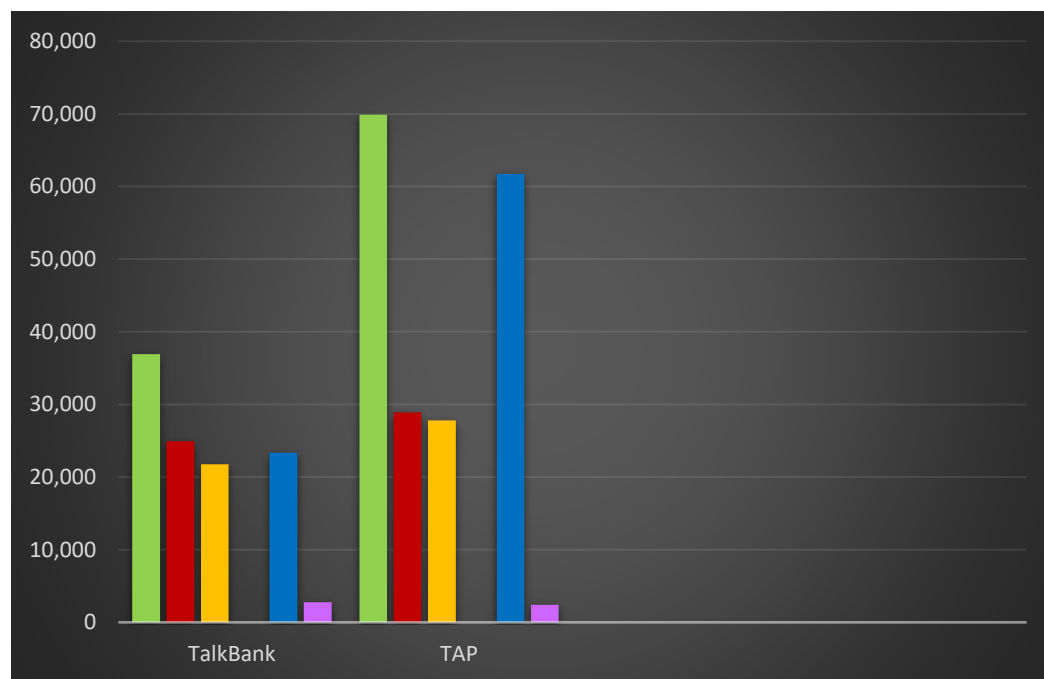


Chart 7. Comparison of adjusted weight values of the terms *ide'ulūzī* (green) and *marām* (red = absolute weight, yellow = approximate weight) as well as of the terms *ide'ulūzīkī* (blue) and *marāmī* (purple) in both corpora

The term pair *ide'ulūzī* and *marām* is special in the sense that both the foreign term and its Persian equivalent were approved by the APLL. While *ide'ulūzī* is only used as a political term referring to a range of different political and social ideologies, *marām*, aside from its political connotation, has a religious connotation, too. It can then be understood to mean *ritual*, a notion that should be excluded from the absolute weight count.

As can be seen in chart 6, the approximate weight does not change the final result that the foreign term is heavier than its Persian equivalent in both corpora. All it does is tilting the result slightly more in favor of *īde 'ūlūzī*. In TalkBank, *īde 'ūlūzī* is less than twice as heavy as *marām* and could therefore qualify as being evenly spread. However, in TAP, *īde 'ūlūzī* is more than twice as heavy as *marām*. As for the derivative term pair *īde 'ūlūzīkī/marāmī*, the case is clear. Here, the foreign term is by far heavier than the Persian equivalent.

The result is surprising insofar as it could be expected that the Persian equivalent enjoys a relative advantage in the newspapers corpus. But, to the contrary, by all measures, the foreign term is even heavier in newspapers than in blogs.

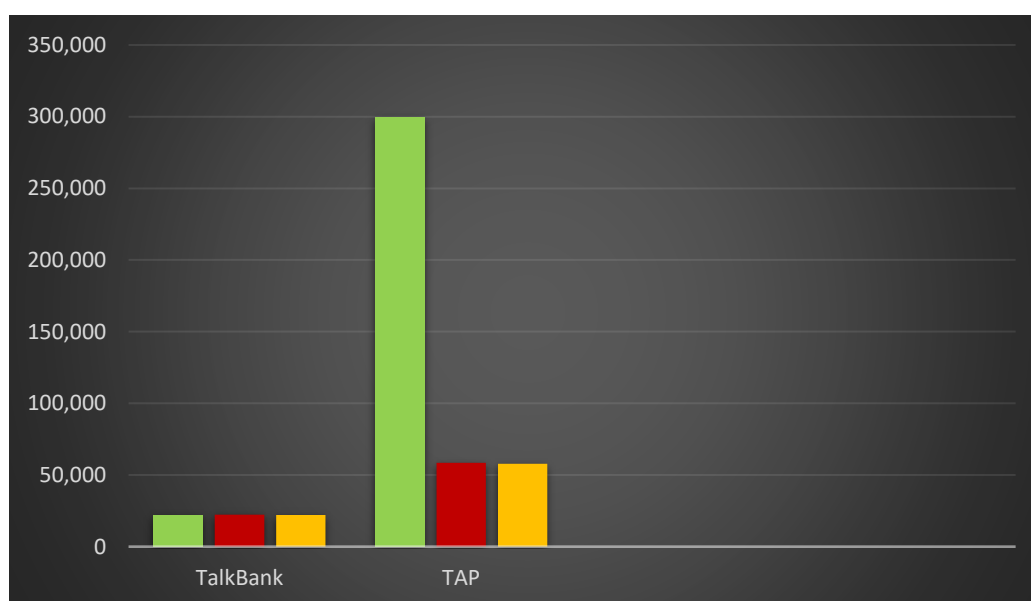


Chart 8. Comparison of adjusted weight values of the terms *prūtokol* (green) and *tašrīfāt* (red = absolute weight, yellow = approximate weight) in both corpora

Finally, the ambiguous term pair *prūtokol/tašrīfāt* produces two very different results for blogs and newspapers. In TalkBank, the weight values of both terms – whether absolute or approximate – are all within a narrow margin. *Prūtokol* has an adjusted weight of 22,194 in TalkBank whereas the adjusted weight of *tašrīfāt* in TalkBank is 22,335. However, the approximate adjusted weight of *tašrīfāt* in TalkBank is 22,160. Both terms can therefore be considered evenly spread in blogs. As can be seen in chart 8, the foreign term *prūtokol* is over five times heavier in TAP, though, regardless of the absolute or approximate weight of *tašrīfāt*. In this specific case, newspapers do not make greater use of the Persian equivalent than of the foreign term.

## 6.2 Conclusion: Explaining the Impact of Coined Terms on Modern Persian

In conclusion, the Persian equivalents approved by the APLL for political foreign terms are overall not very successful in the newspapers and blogs that were compiled in the two corpora TAP and TalkBank respectively. A more detailed look at the sources allows the conclusion that newspapers are generally more likely to accept approved terms than blogs. Maybe journalists and newspaper editors simply know more words because they enjoyed formal training as writers and authors and therefore can use them more liberally than bloggers. It is already clear, for example, that approved ways of spelling are comparatively more frequently applied correctly in newspapers than in blogs. This is an indication of a more formal and Farhangestān-friendly approach in newspapers. But in and by itself this is not a sufficient factor to predict the success of the approved terms. Instead, it might be helpful to turn to the type of borrowing as an auxiliary variable. It is, after all, conceivable that there is a correlation between type of source and type of borrowing that helps determine the success of approved Persian equivalents.

This thesis has looked at the types of borrowing according to Bateni's categorization, which includes eight different types of borrowing, only four of which are in use in the context of the examined terms. The following table highlights the researched terms according to the type of borrowing in Persian. It lists the term pairs in descending order of weight ratio in TalkBank – that is, weight of foreign term divided by weight of Persian equivalent. A weight ratio higher than 1 means that the foreign term is heavier than its Persian equivalent; a weight ratio lower than 1 means that the Persian equivalent is heavier in the corpus.

Term Pair (Foreign Term/Persian Equivalent)	Weight Ratio (Foreign Term/Persian Equivalent)		Type of Borrowing	Publishing Volume
	TalkBank	TAP		
terūrīst/vaḥṣatafkan	2,952.400	-/-	Loan translation and rendition	2
terūrīsm/vaḥṣatafkanī	2,766.958	-/-	Loan translation and rendition	2
doktorīn/rahnāme	82.896	-/-	Loan rendition	2
pūpūlist(ī)/'avāmgarā	34.203	-/-	Loan rendition	6
opūzīsyūn/gorūh-e moḥālef	24.434	16.527	Loan rendition	1
demūkrāt/mardombāvar	23.750	-/-	Loan translation and rendition	6
īde' ulūḡīk(ī)/marāmī	8.388	25.728	Loanshift	2
pūpūlīsm/'avāmgarāyī	7.834	3.135	Loan rendition	6
būrūkrāsī/dīvānsalārī	7.715	22.098	Loan translation	2

demūkrāsī/mardomsālārī	4.312	1.910	Loan translation	6
dīktātūr/mostabed	4.289	2.439	Loan translation	4
īde' ūlūžī/marām (approx.)	1.696	2.517	Loanshift	2
estrātežī/rāhbord	1.580	0.654	Loan rendition	1
estrātežīk/rāhbordī	1.537	0.410	Loan rendition	1
īde' ūlūžī/marām	1.482	2.419	Loanshift	2
dīktātūrī/estebdād	1.102	1.518	Loan translation	4
prūtokol/tašrīfāt (approx.)	1.002	5.174	Loan rendition	2
prūtokol/tašrīfāt	0.994	5.122	Loan rendition	2
kāndīdā/nāmzad (approx.)	0.712	0.667	Loan rendition	2
kāpītālīsm/sarmāyedārī	0.676	0.065	Loan translation and rendition	9
režīm/nezām	0.542	1.017	Loan translation	2/5
režīm/nezām (approx.)	0.522	1.016	Loan translation	2/5
kāndīdā/nāmzad	0.385	0.666	Loan rendition	2
kāpītālīst/sarmāyedār	0.152	0.056	Loan translation and rendition	9
būrūkrāt/dīvānī	0.146	0.575	Loan translation and rendition	2
ommat/mellat	0.132	0.253	Loanshift	5
reformīsm/ešlāḥṭalabī	0.032	-/-	Loan translation and rendition	2
reformīst/ešlāḥṭalab	0.014	0.003	Loan translation and rendition	2
reform/ešlāḥ	0.010	0.005	Loan translation	2
pārlemān/mağles	0.001	0.683	Loanshift	1
pārlemān/mağles (approx.)	0.001	0.683	Loanshift	1

Table 50. Weight ratios of the examined term pairs in TalkBank and TAP as well as the type of borrowing of the approved Persian equivalent term

Most of the researched political terms can be categorized as loan translations, loan renditions, a combination of loan translation and rendition, or loanshifts. The APLL tries to follow thusly its guidelines of keeping Persian equivalents simple and easy to understand. However, it is also difficult to discern a correlation between successful approved terms and their type of borrowing since borrowing follows a similar pattern for all terms in question.

This is consistent with the result of previous studies (Sattari (1388 [2009]) and Keshavarz (1388 [2009])) that dealt with the same perspective. There was no clear correlation between the type of borrowing and its effect on the success of the approved terms, whether positively or negatively. However, the dominance of the types is different from one another.

Sattari finds that the most common type used in the process of term formation is loan rendition while in the study of Keshavarz this dominance goes to loanshifts.

With a view to the list of approved terms, two observations can nonetheless be made: First, for an approved term to be successful it helps if that term is a combined loan translation and rendition. Second, if the translation part is rooted in a term that found its way to Persian from Arabic, it is likely more successful than if an (old) Persian term replaces the (originally French or English) loanword.

Three of the most successful approved political terms in Persian contain the morphemes *mağles*, *eşlāḥ*, or *nezām*, all of which are centuries-old Arabic terms in Persian. *Dīvān* is an interesting case because it is an originally Persian term. Its derivative *dīvānī* is the very successful approved Persian equivalent of *būrūkrāt* while *būrūkrāsī* prevails as a foreign term over the neologism *dīvānsālārī*, which contains the Persian suffix *-sālārī* to account for the *-cracy* part. In the same vein, neologisms such as *rahnāme* or *mardomsālārī*, which use Persian morphemes, have low weight values in the corpora.

Looking into the successful approved terms and the approved terms that are in use alongside the foreign ones, one finds that most of them were in use even before the APLL approved them for a certain meaning. This leads to the third factor that might influence the success of a newly coined term in Persian: time. The successful terms as displayed in table 50 were almost exclusively published in the first and second volumes of approved terms and have hence had almost two decades to spread among native speakers.

Moreover, the terms *dīvān*, *eşlāḥ*, *mağles*, and *nezām* can be found in old Persian dictionaries such as *Montahā al-arab fī loğat-e al-‘arab* that date back to at least the 19<sup>th</sup> century and hence long before the French political terms *bureaucratie*, *reformisme*, *parlement*, and *regime* came to replace them in the 20<sup>th</sup> and 21<sup>st</sup> centuries. What is more important, these terms even carry the same or very similar meanings to what they have come to describe in modern times.

According to the vice-chairwoman of the APLL terminology group Nesrīn Parvīzī, “one could measure the success of the terms with a change of generations.” The APLL thus acknowledges that language planning is a generational task. Looking back at the activities of the First Farhangestān proves that newly coined terms in the 1930s or 1940s may have caused conflicts with existing loanwords at the time, but nowadays most people would not even remember the loanwords that were replaced by the approved Persian equivalents (Parvīzī,

2021). In this sense, future research should compare its results with the findings of this study to illustrate the development and evolution of the relationship between foreign terms and their approved Persian equivalents.

What remains is that there are no monocausal explanations for the success of certain approved Persian terms. In a nutshell, if a neologism is coined as a loan translation and rendition with Arabic roots reaching back in history and presented to a newspaper it is more likely to be applied and thus to gain acceptance than if a neologism is coined as a loan-shift with Persian morphemes and presented to a blog. But there is, of course, room for other variables and factors that might influence the success of the terms approved by the APLL.

### **6.3 Recommendations: Ensuring the Positive Impact of Coined Terms on Modern Persian**

Languages are like living organisms. Language planning is a difficult task and its goal of purifying a language takes time. The Academy of Persian Language and Literature is an official entity that is concerned with language planning for the Persian language, which has an organized and specific work plan and operates according to a decreed framework. Despite this, it has not achieved the desired goal decisively in its many decades of existence. The Farhangestān should instead focus on redeveloping plans to facilitate the process of forming new terms and to reach people in a simple and uncomplicated manner.

A few recommendations can be made based on the findings of this study. These relate to the linguistic approach on the one hand and raising public awareness through teaching and public relations activities on the other.

The nature of the Persian language in terms of the ease of linguistic structures qualifies it for the process of forming new terms and re-forming unused terms thus increasing its vocabulary based on its compositional possibilities (cf. Ášūrī, 1392 [2014], pp. 158-159). But this is a step-by-step process that can be pursued over generations. Repeated experiences are likely necessary to reach the most accurate and simple term possible for a concept or to discover an inherent possibility of language to express a meaning or concept.

The issue of linguistic borrowing must be dealt with carefully. Borrowing is not a problem in and by itself as long as the borrowed terms are merged and integrated into the receiving language and absorbed in its body. This has been evidenced by the relationship between Persian and Arabic and by how Arabic words have become an integral part of the

Persian language. Abandoning those terms has become nearly impossible (Ášūrī, 1392 [2014], pp. 158-159).

In practice, the terminology group of the APLL should avoid using terms with multiple meanings in the process of coining neologisms as this might defer native speakers from using these terms in their daily conversations and it might negatively affect the extent to which those terms are used. Moreover, the Academy should consider re-coining terms that have failed to be used for many years. Many studies have proven the failure of such terms and year for year many studies were conducted that did not record any improvement in the use of some terms such as *vahšatafkan/ī*.

Another challenge for the Farhangestān lies in the issue that the approved terms are not compulsory in tertiary education and are rarely used by university teachers. One way to close this gap is to organize workshops, roundtables, expert meetings, and the like to convince professors and lecturers of the benefits of using approved language in their teaching activities. The Academy could also organize more workshops to target especially those active university teachers that produce a lot of books and papers regularly in a way to encourage them to use the coined and approved terms instead of foreign ones. It can be observed in newspapers and blogs that some authors use both the approved term and the foreign term in the same text or even the same sentence. This leaves the impression that authors use them as synonyms without considering the conflict of fighting loanwords in Persian. The loanwords in many cases are used to introduce the approved terms to the audience, which means they are mentioned in addition to the approved terms.

The terminology group and the Academy at large should also find ways to convince Iranians that even though the Farhangestān is an official institution approved by the government this does not mean that people should avoid using the approved terms as a way of political protest. The goal should be clear to people that preserving the Persian language is a lofty goal and has nothing to do with politics. This might help ensure the success of newly coined and approved terms in the daily use of modern Persian.





## Bibliography

- Abrahamian, E. (2008). *A History of Modern Iran*. New York: Cambridge University Press.
- Academy of the Arabic Language in Cairo (2004). *Al-mu'ğam Al-wasīf* (Vol. 4). Cairo: Al Shorouk Publications.
- Afshari, A. (1386 [2019]). *Sāht-e zabān-e Fārsī (The Structure Persian Language)*. Tehran: Samt Publishing Company.
- Aghili, M. (2010). *Transparency and Acceptability of Neologisms: A Case Study of Words Coined by The Academy of Persian Language and Literature*. Unpublished MA thesis. Isfahan: University of Isfahan.
- Aḥmadīpūr, T. (1389 [2010]). Barrasī-ye 'avāmel-e pīrūnzabānī-ye mo'atter bar pazīreš-e nū-vāzehā (The Effect of Non-linguistic Factors on the Acceptance of Neologisms). *Zabānšenāht-e pažūhešgāh-e 'olūm-e ensānī va moḡāla 'āt-e farhangī*, 1(1), 75-107.
- Aḥmadīpūr, T. (1390[2011]). Be sū-ye čārčūb-e seyāsāt-e zabān-e mellī (Towards a National Language Policy Framework). *Zabānšenāht*, 2(1), 1-16.
- Aizawa, A. (2003). An Information-Theoretic Perspective of TF-IDF Measures. *Information Processing and Management*. 39, pp. 45-65. Tokyo: National Institute of Informatics.
- Akbari, F. (2020). *Iran's Language Planning Confronting English Abbreviations "Persian Terminology Planning"*. Cham, Switzerland: Springer Publishing Company.
- Akbari, F. (2020). *Iran's Language Planning Confronting English Abbreviations*. *Persian Terminology Planning*. Cham, Switzerland: Springer.
- 'Alīzāde, H. (1389[2010]). *Farhang-e ḥāṣṣ-e 'olūm-e seyāsī: dar bargīrandeh-ye eṣṭelāḥātī ke be -isms ḥatm mīšavand (Specialized Dictionary for Political Science: A Case Study of the -isms terms)*. (Ş. Zībākalām, Ed.) Tehran: Nehādgarā.
- 'Amīd, H. (1365[1986]a.). *Farhang-e 'Amīd, ġald-e avval* (Vol. 1). Tehran: Amir Kabir Publishing Corporation.
- 'Amīd, H. (1365[1986]b.). *Farhang-e 'Amīd, ġald-e dovvom* (Vol. 2). Tehran: Amir Kabir Publishing Corporation.

- ‘Amīd, H. (1365[1986]c.). *Farhang-e ‘Amīd, ġald-e sevvom* (Vol. 3). Tehran: Amir Kabir Publishing Corporation.
- Áqābaḡšī, ‘A., & Afšārīrād, M. (1383[2004]). *Farhang-e ‘olūm-e seyāsī (A Dictionary of Political Science)*. Tehran: Čābār Publishers.
- Ariyae, K. (2019). Loanwords Adaption in Persian: A Core-Periphery Model Approach. *Toronto Working Papers in Linguistics (TWPL)*, 41(1).
- Ášūrī, D. (1364 [1985]). *Farhang-e seyāsī, šāmel: maktabhā, eṣṭelāḡhā, mohemtarīn-e sāzmānhāy-e baynolmellālī, ḡezbhā, mohemtarīn-e peymānhā (The Political Dictionary)* (Vol. 16). Tehran: Enteshārāt-e Morvārīd.
- Ášūrī, D. (1374[1995]). *Farhang-e ‘olūm-e ensānī (English-Persian Dictionary for Human Sciences)*. Tehran: Markaz Publishing Co.
- Ášūrī, D. (1376[1997]). *Dānešnāme-ye seyāsī (The Political Encyclopedia)*. Tehran: Enteshārāt-e Morvārīd.
- Ášūrī, D. (1386[2007/2008]). *Zabān-e Bāz: Pažūhešī-ye darbāreh-ye zabān va moderneyyat (Open Language: A Research on Language and Modernity)*. Tehran: Našr-e Markaz Publishing Co.
- Ášūrī, D. (1392 [2014]). *Bāzandīšī-ye zabān-e fārsī (Persian Language Rethought)*. Tehran: Našr-e Markaz.
- Ášūrī, D. (1394[2015]). *Farhang-e ‘olūm-e ensānī, goftegū bā Ášūrī darbāreh-ye raveš-e ū dar vāḡesāzī va nūsāze-ye zabānī (The Dictionary of Human Science, Interview with Ashouri about his Methods of Word-Formation and Language Modernization)*. (M. Jami, Interviewer)
- Badreh-ey, F. (1355[1976/1977]). *Gozārešī-e darbāreh-ye Farhangestān (A Report on The Iranian Academy)*. Tehran: Bonyād-e Shāhenšāhī-ye Farhangestānhā-ye Īrān, Farhangestān-e adab va honar-e Īrān.
- Bahadori, F., & Suzani, S. (2017). Investigating Dentistry Student’s perceived Applicability of Equivalents of Dentistry Terms proposed by the Academy of Persian Language and Literature: A Progressing Field of Interdisciplinary Research. *International Journal of English and Education*, 6(1), 128-137.

- Baradaran, A., & Akbarpour, M. (1395[2017]). Barrasī-ye mīzān-e qābeleyat-e qabūl-e mo'ādelhā-ye vāžehā-ye bānkī pešnahād šodeh va mošavvab-e Farhangestān-e zabān va adab-e fārsī dar meyān-e dānešgūyān-e kāršenāsī-ye aršad-e motarǧemī-ye zabān-e englīsī-ye dānešgah-e āzād-e eslāmī-ye Tehrān. *The first International and third National Conference on Management and Humanistic Research*, 3, pp. 1-13. Tehran. Retrieved August 8, 2019.
- Bāṭenī, M. R. (1385[2006/2007]). *Payrāmūn-e zabān va zabānšenāsī: maǧmū'-e maqālāt (About Language and linguistics: A Collection of Articles)*. Tehran: Agāh Publications.
- Bāṭenī, M. R. (1385[2007]a). *Čahārgoftār darbāreh-ye zabān (Four Speeches about Language)* (5 ed.). Tehran: Agāh Publishing House.
- Bāṭenī, M. R. (1385[2007]b). *Negāhī-ye tāze be dastūr-e zabān (A New Perspective of Grammar)* (12 ed.). Tehran: Agāh Publishing House.
- Bāṭenī, M. R. (1388 [2009/2010]). *Zabān va Tafakkor (Language and Thought)*. Tehran: Našr-e Agāh.
- Bāṭenī, M. R. (1388 [2009/2010]). *Zabān va Tafakkor (Language and Thought)*. Tehran: Našr Āgāh Publications.
- Beel, J., Langer, S., & Gipp, B. (2017). TF-IDF: A Novel Term-Weighting Scheme for User Modeling based on User's Personal Document Collections. *Proceedings of the iConference, 12*. Wuhan, China.
- Bešārat, F. (1397[2016]). Mīnvīsīm vāžeh, miḥānīm eštelāḥ: bāžḥānī-ye mafhūm-e vāžehgozīnī dar matn-e ošūl va žavābeṭ-e vāžehgozīnī, vīrāyeš-e sevvom (Written "Word", Read "term": A Study on the Concept of Term Formation in the Context of "Terminology Guideline", APLL 3rd Edition. 'Elm-e Zabān (Language Science), 5(7), pp. 198-221.
- Bloomfield, L. (1933). *Language*. London: Compton Printing LTD.
- Cabré, M. T. (1999). *Terminology. Theory, Methods, and Applications* (Vol. 1). (J. C. Sager, Ed., & J. A. DeCesaris, Trans.) Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Cabré, M. T. (2000). Elements for a Theory of Terminology: Towards an Alternative Paradigm. *Terminology*, 6(1), pp. 33-57.

- Cabré, M. T., Bagot, R. E., & Chelo, V. S. (2012). Neology in specialized communication. In: K. Kageura, & M.-C. L'Homme, *Terminology - International Journal of Theoretical and Applied Issues in Specialized Communication* (Vol. 18, pp. 1-8). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Calvet, L.-J. (1998). *Language Wars and Linguistic Borrowing*. (M. Petheram, Trans.) Oxford University Press.
- Campbell, L. (1999). *Historical Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- Chaika, E. (1982). *Language: The Social Mirror*. Rowley: Newbury House.
- Chatterjee, P. (1993). *The Nation and its Fragments: Colonial and Postcolonial Histories*. Princeton University Press.
- Cheng, L., Yang, Y., Zhao, G., & Gao, Z. (2018). Research and Improvement of TF-IDF Algorithm based on information theory. In Q. Liu, M. Misir, X. Wang, & W. Liu (Ed.), *The 8th International Conference on computer engineering and networks (CENet2018)* (pp. 608-616). Shanghai: Springer.
- Chisholm, H. (. (1911). Maupeou, René Nicolas Charles Augustin. In: *Encyclopædia Britannica* (11 ed., Vol. 17, pp. 906–907). Cambridge University Press. Retrieved May 2, 2021.
- Cooper, R. L. (1989). *Language Planning and Social Change*. New York: Cambridge University Press.
- Dastgard, V. (1327 [1948-1949]). Zabān-e fārsī rā daryābīd (Discovering Persian Language). *Ārmağān*, 2, pp. 41-45.
- De Bessé, B., Nkwenti-Azeh, B., & Sager, J. C. (1997). Glossary of Terms used in terminology. *Terminology*, 4(1), pp. 117-156.
- Dehḥodā, 'A. A. (1332[1953]). *Loğatnāme (Encyclopedic Dictionary)*. Tehran: Enteshārāt-e mağles-e šūrā-ye mellī.
- Dehḥodā, 'A. A. (1336[1957]). *Loğatnāme (Encyclopedic Dictionary)* (Vol. 34). Tehran: Enteshārāt-e mağles-e šūrā-ye Mellī.

- Dehḥodā, ‘A. A. (1337[1958]). *Loḡatnāme (Encyclopedic Dictionary) - Introduction* (Vol. 40). Tehran: Čaḡhāneh-e dūlatī-ye Īrān.
- Dehḥodā, ‘A. A. (1339[1960]). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 52). (M. Mo‘īn, Ed.) Tehran: Dūlatī Īrān.
- Dehḥodā, ‘A. A. (1341[1963]a.). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 80). (M. Mo‘īn, & S. Šahīdī, Eds.) Tehran: Entešārāt Publishings - University of Tehran Press.
- Dehḥodā, ‘A. A. (1342[1963]). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 86). (M. Mo‘īn, Ed.) Tehran: University of Tehran Press.
- Dehḥodā, ‘A. A. (1343[1964]). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 102). (M. Mo‘īn, Ed.) Tehran: Entešārāt Publishings - University of Tehran Press.
- Dehḥodā, ‘A. A. (1350 [1971]). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 175). (M. Mo‘īn, & S. Šahīdī, Eds.) Tehran: Entešārāt Publishings - University of Tehran Press.
- Dehḥodā, ‘A. A. (1351[1972]). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 181). (M. Mo‘īn, & S. Šahīdī, Eds.) Tehran: Entešārāt Publishings - University of Tehran Press.
- Dehḥodā, ‘A. A. (1352[1973]). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 200). (M. Mo‘īn, & S. Šahīdī, Eds.) Tehran: Entešārāt Publishings - University of Tehran Press.
- Dehḥodā, ‘A. A. (1352[1974]). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 206). (M. Mo‘īn, & Sayedḡa‘ far Šahīdī, Eds.) Tehran: Entešārāt Publishings - University of Tehran Press.
- Dehḥodā, ‘A. A. (1353[1974]). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 208). (M. Mo‘īn, & S. Šahīdī, Eds.) Tehran: Entešārāt Publishings - University of Tehran Press.
- Dehḥodā, ‘A. A. (n.d.). *Loḡatnāme (Encyclopedic Dictionary)* (Vol. 219). (M. Mo‘īn, & S. Šahīdī, Eds.) Tehrān: Čāb-e Sīrūs.
- Edward, J. (2009). *Language and Identity*. Cambridge: Cambridge University Press.
- Eṣfahānī, Š. F. (1337[1959]). *Vāžehnāmeḡ-ye fārsī: Baḡš-e čaharom-e me‘yār-e ḡamālī (Persian Dictionary)*. (Š. Keyā, Ed.) Tehran: Dānešḡāh-e Ṭehrān.
- Farhangestān-e Īrān (Ed.). (1319[1941]). *Vāžehā-ye nū ke tā pāyān-e sāl 1319 (Neologisms untill the End of 1319)*. Tehran: Dabīrḡāneh-ye Farhangestān.

- Fishman, J. A. (1970). *Sociolinguistics: A Brief Introduction*. Rowley: Newbury House Publishers .
- Fishman, J. A. (1973). Language Modernization and Planning in Comparison with Other Types of National Modernization and Planning. *Language in Society*, 2(1), pp. 23-43.
- Foroughi, M. A. (1354 [1975-1976]). *Payām-e man be Farhangestān (My Letter to the Academy)*. Tehran: Payam Publishing Company.
- Forte, D. F. (1986). Terror and Terrorism: There is a Difference. 39, pp. 39-44.
- Foundation for Iranian Studies. (n.d.). *Qānūn-e asāsī Īrān 1324 (Constitution of Iran 1324)*. Retrieved 2021, from Foundation for Iranian Studies: <https://fis-iran.org/fa/resources/legaldoc/iranconstitution>.
- Ĝārūtī, A.-A. (1356[1978]). *Farhang-e Mağmo‘at Alfors (The Dictionary of Persians)*. (‘.-A. ğovaynī, Ed.) Tehran: Bonyād-e Farhang-e Īrān.
- Gilani, A. S. (1374[1995]). Sābeqeh-ye Farhangestān dar Īrān (History of the Academy in Iran). *Nāme-ye Farhangestān*, 11, pp. 136-146.
- Gīvī, H. A., & Anvarī, H. (1390 [2011-2012]). *Dastūr-e zabān-e fārsī: 2 (Persian Language Grammar: 2)* (Vol. 4). Tehran: Fāṭemī Publications.
- Haddad-Adel, G. A. (1382 [2003-2004]). *Asāsnameh-e farhangestān-e zabān va adab-e fārsī (The Constitution of The Academy of Persian Language and Literature)*. Tehran: The Academy of Persian Language and Literature.
- Hassandoust, M. (1393[2014]a.). *Farhang-e rīše šīnāḩe-ye zabān-e fārsī (An etymological Dictionary of the Persian Language)*, (Vol. 1). Tehran: Našr-e Āṭār, The Academy of Persian Language and Literature.
- Hassandoust, M. (1393[2014]b.). *Farhang-e rīše šīnāḩe-ye zabān-e fārsī (An etymological Dictionary of the Persian Language)*, (Vol. 3). Tehran: Našr-e Āṭār, The Academy of Persian Language and Literature.
- Hassandoust, M. (1393[2014]c.). *Farhang-e rīše šīnāḩe-ye zabān-e fārsī (An etymological Dictionary of the Persian Language)* (Vol. 4). Tehran: Našr-e Āṭār, The Academy of Persian Language and Literature.

- Haugen, E. (1950). The Analysis of Linguistic Borrowing. *Language*, 26(2), pp. 210-231.
- Haugen, E. (1959). Planning for Standard Language in Modern Norway. *Anthropological Linguistics*, 35, pp. 109-123.
- Hazbavi, A. A. (2012). Investigating Iran's Success in Standardization of Terminologies of Computer Information and Technology. *Middle-East Journal of Scientific Research*, 11(8), pp. 1053- 1058.
- Hesabi, A., Amirian, Z., & Nazari, J. (2017). Political terms by APLL: Issues of Terminology implantation and acceptability. *Applied Research on English Language*, 6(3), pp. 339-362.
- Hidayat, R. Q. (1288 [1881]). *Farhang-e anğoman-e ārā-ye Nāşerī*. Tehran.
- Hochett, C. F. (1958). *A Course in Modern Linguistics*. New York: MacMillan Company.
- Ĥoseynī, B. (1380 [2001-2002]). *Hāmeşī bar dāneş-e zabānşenāşī (Every Thing about Linguistics)*. Tehran: Zhiar Publishing Company.
- Ibn Manżūr. (1922). *Līsān Al- 'Arab* (Vol. 12). Beirut: Dar Sader Publishers.
- Joan, R., & Jernudd, B. H. (1971). *Evaluation and Language Planning: Can Language be Planned?* Honolulu: University Hawai'i Press.
- Johnston, S. (2008). Persian Rap: The Voice of Modern Iran's Youth. *Journal of Persianate Studies*, 1(1), pp. 102-119.
- Kāfī, 'A. (1363 [1984-1985]). Gerāyeshā-ye gūnāgūn dar vāzehgozīnī (Different Orientations in Terminology). *Naşr-e Dāneş* .5(25), pp. 20-27.
- Kāfī, 'A. (1371[1992]). Barrasī-ye vāzehā-ye farhangestān-e avval (A Study of the Terms of the First Farhangestān). *Naşr-e Dāneş*, 69.
- Kāfī, 'A. (1386 [1989]). Barrasī-ye vāzehā-ye reyāzī-ye farhangestān-e avval (Studying the Mathematical Terms of The First Farhangestān ). *Naşr-e Dāneş*, 10(1), pp. 308-318.
- Kaplan, R. B., & Baldauf Jr., R. B. (1997). *Language Planning from Practice to Theory*. Bristol: Multilingual Matters.

- Karam, F. X. (1974). Toward a Definition of Language Planning. In: J. A. Fishman (Ed.), *Advances in language planning* (pp. 103-124). Mouton: The Hauge.
- Kashefi, O., Mitra, N., & Kanani, K. (1389 [2010]). *Heṭāyābī-ye emlā'-e hūd-kār dar zabān-e fārsī (Towards Automatic Persian Spell Checking)*. Tehran: Supreme Council of Information and communication Technology (SCICT).
- Keshavarz, P. (1388[2009]). *Barrasī-ye mīzān-e kārbord-e vāzehā-ye seyāsī-ye moṣavvab-e farhangestān-e zabān dar baḥš-e seyāsī-ye maṭbū'āt. Unpublished MA thesis*. Tehran: Dānešgāh-e 'Alameh-ye Ṭabāṭabā'ī.
- Ketābī, S., 'Akīlī, M., & Abūlḥasanī, M. (1389 [2010]). Ğenseyat va rābeṭe-ye ān bā maqbūleyat va šafāfeyat-e vāzegān-e nū: barrasī-ye mūred-e vāzegān-e 'omūmī-ye moṣavvab-e farhangestān-e zabān va adab-e fārsī. *Zan dar farhang va honar (pažūheš-e zanān)*, 1(3), pp. 88-108.
- Keyā, Ṣ., & Gol-Golāb, Ḥ. (1355[1976-1977]). *Farhangestān-e Īrān va farhangestān-e zabān-e Īrān (the Iranian Academy and the Iranian language Academy)*. Tehran: Imperial Foundation of Iranian Academies.
- Khan, M. M. (15 Jumada Al Awwal 1311). *Andīsehā-ye Mīrzā Āqā Ḥān Kermānī*. Firaydun, Adamiyat.
- Khiabany, G. (2007). Is there an Islamic communication? The persistence of 'tradition' and the lure of modernity. *Critical Arts*, 21(1), pp. 106-124.
- Khiabany, G. (2010). *Iranian Media: the Paradox of Modernity*. New York: Routledge.
- Khiabany, G., & Sreberny, A. (2007)a. Becoming Intellectual: The Blogestan and Public Political Space in The Islamic Republic. *British Journal of Middle Eastern Studies*, 34(3), pp. 267-286.
- Khiabany, G., & Sreberny, A. (2007)b. The Politics of/in Blogging in Iran. *Comparative Studies of South Asia Africa and the Middle East*, 27(3), pp. 563-579.
- Kia, M. (1998). Persian Nationalism and The Campaign for Language Purification. *Middle Eastern Studies*, 34(2), pp. 9-36.



- Langrūdī, M. (1367 [1988]). *Termīnūlūžī-ye hoqūkī (Legal Terminology)*. Tehran: Ganġ-e Dāneš Publishers.
- Langrūdī, M. (1392[2013]). *Termīnūlūžī-ye hoqūkī (Legal Terminology)*. Tehran: Entesārāt-e Ganġ-e Dāneš.
- Lazard, G. (1384 [2006]). *Šaklegīrī-ye zabān-e fārsī (Formation of Persian Language)*. (M. Baħrīnī, Trans.) Tehran: Hermes Publication House.
- Maġd, O. (1388[2009]). Naqd-e čehel loġat-e tarġomeh šodeh-e farhangestān-e zabān va adab-e fārsī, hamrāh bā erā'eh-ye pīšnehādā-ye ġadīd va ravešhā-ye ešlāhī (Critique of Forty Terms Translated by the Academy of Persian Language and Literature). *Fašlenāmeħ-e 'elmī-pažūhešī "pažūheš-e zabān va adabeyāt-e fārsī"*, 13, pp. 87-106.
- Mansouri, R. (1376 [1997]). *Modīreyat-e važehgozīnī: taškīlāt, raveš-hā, avlaveyathā (Word Selection Management: Structure, Methods, and Piriorities)*. Tehran: The Academy of Persian Language and Literature.
- Mansouri, R. (1988). Standardization of Physics Terms in Persian. In: H. Czap, & C. Galinski, *TKE'87: Proceedings of International Congress on Terminology and Knowledge Engineering* (pp. 176-179). Frankfurt/M: INDEKS.
- Megerdoomian, K. (2010). Extending a Persian Morphological Analyzer to Blogs. In: H. Sameti, & M. Bijankhan, *Zabān-e fārsī va rāyāneh [Persian Language and Computer] (Volume II)* (pp. 20-35). Tehran: SAMT Publishers, Organization for Research and Editing of University Publications in the Humanities.
- Megerdoomian, K., & Hadjarian, A. (2010). Mining and Classification of Neologisms in Persian Blogs. *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 6-13). Los Angeles, California: Assosiation for Computational Linguistics.
- Meskoob, Š. (1385 [2006-2007]). *Hoveyat-e īranī va zabān-e fārsī (The Iranian Identity and Persian Language)* (3 ed.). Tehran: Farzan Publishers.
- Mirza, J. a.-D. (1871). *Nāmeħ-ye Khosravān*. 3 vols. Tehran.
- Modarresī, Y. (1368 [1989-1990]). *Darāmadī bar ġāme'ešenāsī (An Introduction to Sociology)*. Tehran: Institute for Humanities and Cultural Studies.

- Modarresī, Y. (1371[1992]). Masāyel-e zabānī va barnāmerīzī-ye zabān dar Īrān (Language Problems and Language Planning in Iran). *Farhang*, 13, pp. 177- 205.
- Mohāgerīzāde, Ğ. (1388 [2009]). Naqd-e vāzehā-ye mošavvab-e farhangestān, daftar-e paᅅom (Critique of The Approved Terms by The Academy Of Persian Language and Literature, volume 5). *Ketāb-e māh-e adabeyāt*, 143, pp. 65-74.
- Mo'īn, M. (1383 [2004]). *Farhang-e fārsī (Persian Dictionary)*. Tehran: Sarāyeš Publications.
- Mo'īn, M. (1386 [2007]a.). *Farhang-e fārsī, ġald-e avval (An Intermediate Persian Dictionary: First Volume)* (Vol. 1). Tehran: Amir Kabir Publishing Corporation.
- Mo'īn, M. (1386 [2007]b.). *Farhang-e fārsī, ġald-e dovvom (An Intermediate Persian Dictionary: Second volume)* (Vol. 2). Tehran: Amir Kabir Publishing Corporation.
- Mo'īn, M. (1386[2007]c.). *Farhang-e fārsī, ġald-e sevvom (An Intermediate Persian Dictionary: Third Volume)* (Vol. 3). Tehran: Amir Kabir Publishing Corporation.
- Mo'īn, M. (1386 [2007]d.). *Farhang-e fārsī, ġald-e čahārom (An Intermediate Persian Dictionary: Fourth Volume)* (Vol. 4). Tehran: Amir Kabir Publishing Corporation.
- Mo'meni, Š., & Faᅅᅅārzādeh, M. (1395 [2016]). Farāyandehā-ye mo'ādelyābī-ye farhangestān dar ᅅūze-ye rāyāneh va fannāvarī-ye eᅅtelā'āt va pazīreš-e mo'ādelhā dar kotob-e āmūzešī. *Našr-e adab va zabān*, 39(19), pp. 231-251.
- Monšī-zādeh, M., & Sharif, B. (1384 [2005]). Barrasī-ye vāzehā-ye 'omūmī-ye mošavvab-e farhangestān-e zabān va adab-e fārsī (Study the General Approved Terms by The Academy of Persian Language and Literature). *Maᅅaleh-e Zabān va Adab (The Journal of Language and Literature)*, 9(24), pp. 18-43.
- Moqaddam, A. Š. (1374 [1996]). Farhangestān-e dovvom (The Second Iranian Academy). *Nāmeᅅ-ye Farhangestān*, 2, pp. 132-142.
- Mustafa, I., & El Zayyat, A. (1960). *Al-mo'ᅅam Al-wasīfī*. Cairo: Maᅅma' Al-luᅅa Al-'arabīya.
- Nafīsī, ' . A. (1343 [1964]). *Farhang-e Nafīsī* (Vol. 1). Tehran: Ketābfrūšī ᅅayyām.
- Nafīsī, ' . A. (1355[1977]). *Farhang-e Nafīsī* (Vol. 2). Tehran: Ketābfrūšī ᅅayyām.
- Nafīsī, ' . A. (1355[1977]b.). *Farhang-e Nafīsī* (Vol. 5). Tehran: Ketābfrūšī ᅅayyām.

- Nağafī, A. (1361 [1983]). *Áyā zabān-e fārsī dar ḥaṭar ast?* (Is Persian Language in Danger?). 3(1), pp. 77-83.
- Naḥğavānī, M. b. (1355[1977]). *Şahāḥ Alfors: farhang-e loğāt-e fārsī az qarn-e haštom-e heğrī*. ('A. Ṭā'atī, Ed.) Tehran: Bangāh-e Tarğomeh va Naşr-e Ketāb.
- Najafī, A. (1390 [2011]). *Ġalaṭ nanevīsīm: farhang-e deşvārīhā-ye zabān-e fārsī (Let's avoid mistakes: A dictionary of Persian Language Difficulties)* (16 ed.). Tehran: Marakaz Naşr-e Daneşgāhī.
- Nāṭeqiyān Far, S., & Nāṭeqiyān Farī, F. (1389 [2010]). Mağāles-e 'aşr-e pahlavī. *Yād-e Ayyām*, 5(57).
- Nekvapil, J. (2011). *Handbook of Research in Second Language Teaching and Learning: The History and Theory of Language Planning* (Vol. 2). (E. Hinkel, Ed.) New York: Routledge.
- Nežād, S. 'A. (1388 [2010]). *Zabān-e fārsī dar čand-e goftogū (Persian Language in a Few Conversations)*. Tehran: Naşr-e Āgah.
- Pādeşāh, M. (1363 [1984]). *Farhang-e Ānnanderāğ: Farhang-e ġāme'-e fārsī*. (M. D. Sāqī, Ed.) Tehran: Ketābfrūşey-e ḥayyām.
- Paraskiewicz, K. (2015). English loanwords in New Persian. *FOLIA ORIENTALIA*, 52, pp. 281-292.
- Parvīzī, N. (1385 [2006-2007]). Bīst sāl vāzehgozīnī (Twenty Years of Terminology). *Nāme-e Farhangestān (moṭala'āt va zehgozīnī)*, 1, pp. 3-16.
- Parvīzī, N. (2021, 10 19). Letter from the Vice-Chairwoman of the Terminology Group, APLL, by e-mail (27/07/1400). (A. Shehata, Interviewer)
- Paul, L. (1999). "Iranian Nation" and Iranian-Islamic Revolutionary Ideology. *Die Welt des Islams*, 39(2), pp. 183-217.
- Paul, L. (2010). Iranian Language Reform in the Twentieth Century: Did the Farhangestān (1935-40) Succeed? *Journal of Persianate Studies*, 3, pp. 78-103.

- Perry, J. R. (1985). Language reform in Turkey and Iran. *International Journal of Middle East studies*, 17(3), pp. 295- 311.
- Pourdavoud, E. (1322 [1943-1944]). Kalāme-ye Farhangestān (The Words of the Academy of Persian Language and Literature). *Nāmeḥ-ye Farhangestān-e Qadīmī*, 5(8), pp. 690-693.
- Qāsemī, H., & Setarī, M. (1389[2010]). Mohemtarīn-e ‘avāmel va dūrehā-ye vorūd-e vāžehgān-e karzī dar zabān-e fārsī va ṭabaqebandī-ye ma‘nā-ye ānhā (The Important Factors and Periods of Entering Foreign Terms Into Persian and The Classification of Their Meanings). *Āndešehā-ye adabī (Literary Thoughts)*, 5, pp. 155-165.
- Rahimiyan, O. R. (2011). The Protocols of The Elders of Zion in Iranian Political and Cultural Discourse. In: E. Webman (Ed.), *The Global Impact of The Protocols of The Elders of Zion: A Century-Old Myth* (pp. 196-219). Abingdon: Routledge.
- Rajaraman, A., & Ullman, J. (2011). Data minning. In: *Mining on massive datasets* (pp. 1-17). Cambridge: Cambridge University Press.
- Ramos, J. (2003). *Using TF-IDF to Determine Work Relevance in Document Queries*. New Jersey: Rutgers University, Institute for computer sciences department.
- Ranjbar, A., Abbaspour, J., Sotudeh, H., & Moloodi, A. (1398 (2019)). A Study of The Conformity of Writing Behavior of The Authors and Users of The Persian Scientific Database to Guidelines of The Academy of Persian Language and Literature in Terms of Writing Words Continuously, Semi-Spacely and Separately. *Ketābdārī va eṭṭelā‘ rasānī (Library and Information Sciences)*, 22(3), pp. 163-187.
- Rezā-Zālī, M. (1374[1995]). Naqdī barāndīše-ye marg-e būrūkrāsī. *Taḥavvol-e Edārī*, 10 - 11, pp. 126-135.
- Richards, J. C., Schmidt, R., Kendricks, H., & Kim, Y. (2002). *Longman Dictionary of Language Teaching and Applied Linguistics*. London: Pearson Education.
- Rivetti, P. (2020). *Political Participation in Iran from Khatami to the Green Movement*. Switzerland: Springer.

- Rūstā'ī, M. (1377 [1998-1999]). Vāžegān-e 'elmī-ye moṣavvab-e farhangestān-e Īrān (The Approved Scientific Terms by The Iranian Academy). *Ganġīneh-ye Asnad (National Library and Archives of Iran)*, 8(31,32), pp. 76-88.
- Sa'adat, E. (1393 [2014-2015]). *Dānešnāmeḥ-ye zabān va adab-e fārsī (Encyclopedia of Persian Language and Literature)* (Vol. 5). Tehran: Farhangestān-e zabān va adab-e fārsī (Iranian Academy of Persian Language and Literature).
- Sadeghi, A. A. (1375 [1996]). Naqš-e farhangestān dar barnāmehrīzī barāye zabān-e fārsī (The Role of The Academy of Persian Language and Literature in Planning Persian language). *Nāmeḥ-ye Farhangestān*, 2(1), pp. 148-161.
- Sadeghi, A. A. (2001). Language Planning in Iran: A Historical Review. *International Journal of the Sociology of Language*, 148, pp. 19-30.
- Sadeghi, A. A. (2003). The Persian Language. In: E. İhsanoğlu (Ed.), *The Different Aspects of Islamic Culture: Culture and Learning of Islam* (Vol. 5, pp. 49-54). Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO).
- Sadeghi, R. (1370 [1991]). Taġrobehā-ye zabān-e fārsī dar 'elm (The Experiences of Persian Language in Science). *Maġmū'eh-ye maqālāt-e semīnār-e zabān-e fārsī va zabān-e 'elm (Proceedings of the Seminar on Persian Language and the Language of Science)*, pp. 72-129.
- Şādeqpūr, Ā. (1369 [1989]). *Naẓare-ye būrūkrāsī (The Theory of Bureaucracy)*. Tehran: Markaz-e āmūzešī-s modīreyyat-e dūlatī.
- Şafavī, k. (1374 [1995-1996]). Vāžehā-ye qarzī dar zabān-e fārsī (Borrowed Words in Persian Language). *Name-e Farhang*, 19, pp. 97-111.
- Şafavī, k. (1386 [2007]). *Āšenā-ye bā tāriḥ-e zabānhā-ye īrānī (A Short History of Iranian Languages)*. Tehran: Pezhvak Publication Company.
- Şafavī, K. (1388 [2009]). *Haft-e goftār darbāreh-ye tarġomeḥ*. Tehran: Ketāb-e Mād.
- Şafavī, K. (2005). Vāžehā-ye qarzī dar zabān-e fārsī (Borrowed Words in Persian). *Nāmeḥ-ye Farhang*, pp. 97-111.

- Şaffār Moqaddam, A. (1374 [1996]). Farhangestān-e dovvom : Pažūhešgāh-e važehgozīnī (The Second Academy: Terminology group). *Nāme-ye Farhangestān* (4), pp. 132-142.
- Şafīpūr, ‘A. (2001). *Montahā al-arab fī loğat-e al-‘arab* (Vol. 1 and 2). Tehran: Ketābhāne-ye sanā’ī.
- Sa‘īdeyān, ‘A. (1378 [1999]a.). *Dā‘erat al-ma‘āref-e nū (The New Encyclopedia)* (Vol. 2). (M. Sa‘īdeyān, B. Moğtabavī, & F. M. Ğalāl, Eds.) Tehran: Entesārāt-e ‘Elm va Zendeġī.
- Sa‘īdeyān, ‘A. (1378 [1999]b.). *Dā‘erat al-ma‘āref-e nū (The New Encyclopedia)* (Vol. 3). (M. Sa‘īdeyān, B. Moğtabavī, & F. M. Ğalāl, Eds.) Tehran: Entesārāt-e ‘Elm va Zendeġī.
- Samī‘ī Gīlānī, A. (1375 [1996]). Moqābaleh bā hoğūm-e važehā-ye bīganeh (Coping with the Influx of Foreign Words). *Nāme-ye Farhangestān* (5), 3.
- Sarlī, N. (1387 [2008-2009]). *Zabān-e fārsī me‘yār (Standard Persian Language)*. Tehran: Hermes Publishing Company.
- Sattari, A. (1388 [2009]). *Barrasī-ye mīzān-e kārbord-e vāžehā-ye seyāsī-ye moşavvab-e farhangestān-e zabān-e fārsī dar barnāmeᡁā-ye rādyū-ye ġomᡁūrī-ye eslāmī-ye Īrān. Unpublished MA thesis*. Tehran: Dāneşgāh-e Tehrān.
- Şeddīq, ‘Ī. (1322 [1943]). Tūziᡁ dar tāriᡁeᡁeh-ye farhangestān (Notes on the History of the Iranian Academy). *Nāme-ye Farhangestān*, 1(4), pp. 1-5.
- Shaghagi, V. (1391[2012-2013]). *Mabānī-ye şarf (An Introduction to Morphology)*. Tehran: SAMT.
- Sharifī, M. (1391[2013]). *Farhang adabeyyat-e fārsī (Dictionary of Persian literature)*. (M. R. Ğa‘farī, Ed.) Tehran: Naşr-e nū publication company.
- Shehata, A. (2021). Neologisms in Contemporary Persian Approved by the Academy of Persian Language and Literature: A Case Study of Epidemiology Terms. In: A. Aghdassi (Ed.), *Perspectives on Academic Persian* (pp. 81-103). Cham: Springer Nature.
- Shirazi, F. (2008). Social Network within Filtered ICT Networks: A Case Study of the Growth of Internet Usage within Iran. In: C. Avgerou, M. L. Smith, & P. van den Besselaar (Ed.), *Eighth International on Human Choice and Computers (HCC8),IFIPTC9*, (pp. 229-318). Pretoria, South Africa.

- Shohamy, E. (2006). *Language Policy: Hidden Agendas and New Approaches*. New York: Routledge.
- Spolsky, B., & Shohamy, E. (1999). *The Language of Israel: Policy, Ideology, and Practice*. Clevedon et al.: Multilingual Matters LTD.
- Strazny, P. (2011). *Encyclopedia of Linguistics*. New York: Fitzroy Dearborn.
- Ṭabāṭabā'ī, 'A. (1376 [1997/1998]). *Fe 'l-e basīṭ-e fārsī va vāžehsāzī (The Simple Persian Verb and Term-Formation)*. Tehran: Dānešgāhī Publications.
- Ṭabāṭabā'ī, 'A. (1386 [2007-2008]). Čeh vāžeh-ye fārsī ast? (Which terms are considered Persian?). *Vīžehnāmeḥ-ye nāmeḥ-ye Farhangestān*, 8, pp. 119-121.
- Tabrīzī, 'A. (1386 [2008]). *Masālek Al-Moḥsenīn (Paths of the Benefactors)*. Tehran: Omīd-e Fardā.
- Tabrīzī, M. Ḥ. (1342 [1963]a.). *Borhān-e Qāte'* (Vol. 1). Tehran: Ketābfrūšī-ye Ibn Sīnā.
- Tabrīzī, M. Ḥ. (1342 [1963]b.). *Borhān-e Qāte'* (Vol. 2). Tehran: Ketābfrūšī-ye Ibn Sīnā.
- Tabrīzī, M. Ḥ. (1342 [1963]c.). *Borhān-e Qāte'* (Vol. 3). Tehran: Ketābfrūšī-ye Ibn Sīnā.
- Tārīḥī az farhangestān-e Īrān (History of the Iranian Academy). (1322[1943]). *Nāme-ye Farhangestān (qadīmī)*, pp. 7-19.
- Tauli, V. (1968). *Introduction To A Theory Of Language Planning*. Uppsala: Almqvist and Wiksell.
- Tavakoli-Targhi, M. (1990). Refashioning Iran: Language and Culture during the Constitutional Revolution. *Iranian Studies*, 23(1/4), pp. 77-101.
- Ṭayyeb, 'A. R. (1380 [2001]). Terūrīsm dar ferāz va forūd-e tāriḥ (Terrorism in The Ups and Downs of History). *Rāhbord*, 21(4), pp. 55-67.
- Terminology group. (1382 [2004]). *Farhang-e vāžehā-ye moṣavvab-e farhangestān (A Collection of Terms, Approved by the Academy of Persian Language and Literature)* (Vol. volume 1). Tehran: The Academy of Persian Language and Literature.

- Terminology group. (1376 [1997]). *Vāžehā-ye mošavvab-e farhangestān va adab-e fārsī, Vāžehā-ye 'omūmī 1 (The approved terms by the Academy of Persian Language and Literature: General terms 1)*. Tehran: Čāb-e Rahgošā.
- Terminology group. (1376-1385 [1997-2007]). *Farhang-e vāžehā-ye mošavvab-e farhangestān (A Collection of Terms, Approved by the Academy of Persian Language and Literature)*. Tehran: The Academy of Persian Language and Literature.
- Terminology group. (1377[1998]). *Vāžehā-ye mošavvab-e farhangestān va adab-e fārsī, Vāžehā-ye 'omūmī 2 (The approved terms by the Academy of Persian Language and Literature: General terms 2)*. Tehran: Čāb-e Tandīs.
- Terminology group. (1382-1398 [2004-2021]). *Farhang-e vāžehā-ye mošavvab-e Farhangestān (A collection of Terms, Approved by the Academy of Persian Language and Literature) (Vols. 1-16)*. Tehran: The Academy of Persian Language and Literature.
- Terminology group. (1383 [2004]). *Farhang-e vāžehā-ye mošavvab-e farhangestān (A Collection of Approved Terms by The Academy of Persian Language and Literature) (Vol. 1)*. Tehran: Academy of Persian Language and Literature.
- Terminology group. (1384 [2005]). *Farhang-e vāžehā-ye mošavvab-e Farhangestān (A Collection of Terms, Approved by the Academy of Persian Language and Literature) (Vol. 2)*. Tehran: The Academy of Persian Language and Literature.
- Terminology group. (1387 [2008]). *Farhang-e vāžehā-ye mošavvab-e Farhangestān (A Collection of Terms, Approved by the Academy of Persian Language and Literature) (Vol. 5)*. Tehran: The Academy of Persian Language and Literature.
- Terminology group. (1388 [2009])a. *Farhang-e vāžehā-ye mošavvab-e Farhangestān (A Collection of Terms, Approved by the Academy of Persian Language and Literature) (Vol. 6)*. Tehran: The Academy of Persian Language and Literature.
- Terminology group. (1388 [2009])b. *Ošūl va žavābet-e vāžehgozīnī: hamrāh bā šarḥ va tūžīḥāt (The Principles and Regulations of Term Formation: With Explanations and Illustrative Examples) (third revised ed.)*. Tehran: The Academy of Persian Language and Literature.



- Terminology group. (1389 [2010]). *Farhang-e vāžehā-ye mošavvab-e Farhangestān (A Collection of Terms, Approved by the Academy of Persian Language and Literature)* (Vol. 7). Tehran: The Academy of Persian Language and Literature.
- Terminology group. (1391 [2019]). *Farhang-e vāžehā-ye mošavvab-e Farhangestān (A Collection of Terms, Approved by the Academy of Persian Language and Literature)* (Vol. 9). Tehran: The Academy of Persian Language and Literature.
- Terminology group. (1394 [2016]). *Farhang-e vāžehā-ye mošavvab-e Farhangestān (A Collection of Terms, Approved by the Academy of Persian Language and Literature)* (Vol. 12). Tehran: The Academy of Persian Language and Literature.
- The Academy of Persian Language and Literature. (1391 [2013/2014]). *Dastūr-e ḥaṭṭ-e fārsī (The Rules of the Persian Writing System)* (11 ed.). Tehran: Našr-e Āṭār.
- Tilly, C. (2004, March). Terror, Terrorism, Terrorists. *Sociological Theory, American Sociological Association*, 22(1), pp. 5-13.
- Ṭūsī, Ā. (1336[1958]). *Loḡat-e fors (Language of The Persians)*. (M. Dabīrseyāqī, Ed.) Tehran: Ketābhāneh-e Ṭahūrī.
- Vafaei, A.-A. (1390 (2011/2012)). *Dastor-e zabān-e Farsī (Grammar of Persian Language)*. Tehran: International Publishing CO.
- Valinezhad, M. (2016). *Investigating the rate of the acceptability of political equivalences suggested by the Academy of Persian Language and Literature. Unpublished MA thesis*. Tehran: Islamic Azad University.
- Webman, E. (Ed.). (2011). *The Global Impact of The Protocols of the Elders of Zion: A Century-Old Myth*. Abingdon: Routledge.
- Weinreich, U. (1968). *Languages in Contact: Findings and Problems*. The Hague: Mouton Publishers.
- Zarnikhi, A. (2010). Terminological Activities at the Academy of Persian Language and Literature. In: C. Heine, & J. Engberg, *Reconceptualizing LSP*. Aarhus: Online Proceedings of the XVII European LSP Symposium 2009.

Zarrinkoub, A. (1336[1957]). *Dū qarn-e sokūt (Two Centuries of Silence)*. Tehran: Amir Kabir Publishing Company.

## Online Resources

- Académie Française.* (n.d.). Retrieved 12 July, 2020, from <http://www.academie-francaise.fr/linstitution/aperçu-historique>.
- Academy of The Arabic Language.* (2021, February 2). Retrieved from Historical Stub: <https://www.sis.gov.eg/newVR/acadmy/html/acadmay07.htm>
- Āḍarang, ‘A.-a.-Ḥ., & EIr. (2016). *Encyclopædia Iranica, online edition.* Retrieved from Parviz Khanlari: <https://www.iranicaonline.org/articles/khanlari-parviz>
- Afshar, I., & Fakhreddin, A. (1999, December 15). *Forūgī, Moḥammad- ‘Alī Dōkā Al-Molk.* Retrieved from Encyclopædia Iranica: <http://www.iranicaonline.org/articles/forugi-mohammad-ali#pt2>
- Ātār montašer šode dar gorūh-e vāzehgozīnī.* (1398 [2019]). Retrieved from Farhangestān-e zabān va adab-e fārsī: <https://apll.ir/1398/08/01/%d8%a2%d8%ab%d8%a7%d8%b1-%d9%85%d9%86%d8%aa%d8%b4%d8%b1%d8%b4%d8%af%d9%87-%d8%af%d8%b1-%da%af%d8%b1%d9%88%d9%87-%d9%88%d8%a7%da%98%d9%87%e2%80%8c%da%af%d8%b2%db%8c%d9%86%db%8c/>
- Bürokratie.* (n.d.). Retrieved September 2021, from Brockhaus: <https://brockhaus.de/ecs/enzy/article/burokratie#begriffsgeschichte>
- CNRTL. (2012). *Bureaucratie, subst. fém.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/bureaucratie>
- CNRTL. (2012). *Candidat.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/candidat>
- CNRTL. (2012). *Capital.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/capital>
- CNRTL. (2012). *Capitalisme, subst. masc.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/capitalisme>
- CNRTL. (2012). *Parlement, subst. masc.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie>

- CNRTL. (2012). *Populisme*. Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/populisme>
- CNRTL. (2012). *Populiste, adj.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/populiste>
- CNRTL. (2012). *Protocole, subst. masc.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/protocole>
- CNRTL. (2012). *Regime: subst. masc.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/regime>
- CNRTL. (2012). *Stratégie, subst. fém.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/strategie>
- CNRTL. (2012). *Terrorisme*. Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/terrorisme>
- CNRTL. (2012). *Idéologie: subst. fém.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/ideologie>
- CNRTL. (2012). *Aéroplane*. Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/A%C3%A9roplane>
- CNRTL. (2012). *Doctrine, subst. fém.* Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/doctrine>
- CNRTL. (2012). *Opposition*. Retrieved from Centre National de Ressources Textuelles et Lexicales: <https://www.cnrtl.fr/etymologie/opposition>
- De Blois, F. (1995). *Dīvān*. Retrieved from Encyclopædia Iranica, online edition: <https://iranicaonline.org/articles/divan>
- Darbāreh-e Magiran*. (2021). Retrieved from Magiran: <http://www.magiran.com/help#About>
- Dehkhodā, ‘A. A. (1377 [1998]a.). *Loghatnāme (Encyclopedic Dictionary) - Online Version* (Vol. 11). (M. Mo‘in, & J. Shahidi, Eds.) Tehran: Tehran University Publications. Retrieved from <http://dehkhoda.ut.ac.ir>

Dehḡodā, ‘A. A. (1377 [1998]b.). *Dīvānī*. Retrieved from Logatnāme (Encyclopedic Dictionary) - Online Version: <https://dekhoda.ut.ac.ir/fa/dictionary/156052/%D8%AF%DB%8C%D9%88%D8%A7%D9%86%DB%8C>

Dehḡodā, ‘A. A. (1377 [1998]). *Logatnāme (Encyclopedic Dictionary) - Online Version* (Vol. 11). (M. Mo‘in, & J. Shahidi, Eds.) Tehran: Tehran University Publications. Retrieved from <http://dekhoda.ut.ac.ir>

*Dostūr (Law, part 4)*. (1295 [1878]). Istanbul: Maṭba‘eh-e ‘āmereh. Retrieved from <https://web.archive.org/web/20170719020352/https://anayasa.tbmm.gov.tr/docs/1876/1876ilkmecinler/1876-ilkhal-osmanlica.pdf>

*Editorial Board*. (n.d.). (Institute for Humanities and Cultural Studies) Retrieved July 28, 2020, from Journal of the Linguistic Society of Iran: <http://lsi-linguistics.ihs.ac.ir/>

*Iran’s 1906 Constitution*. (1906, December 30). Retrieved from Foundation for Iranian Studies: <https://fis-iran.org/>

Jazayeri, M. A. (1999). *Farhangestān (The Academy)*. Retrieved from Encyclopedia Iranica: <https://www.iranicaonline.org/articles/farhangestan>

Mahtafar, T., Khiabany, M., & Niknejad, G. (2009, October 10). Tehran Bureau. Retrieved from <https://www.pbs.org/wgbh/pages/frontline/tehranbureau/2009/12/media-guide.html>

Neḡād, S. ‘A., & Rošan, S. (2008, June 2). *B.B.C Persian*. Retrieved from Gozāreš-e yek zendegī: Doktor Mohammad-Reza Bateni (The Biography of Mohammad Reza Bateni): [https://www.bbc.com/persian/arts/story/2008/06/080602\\_an-bateni-preface](https://www.bbc.com/persian/arts/story/2008/06/080602_an-bateni-preface)

*Payvand*. (2018, August 22). Retrieved from Radio Farda: <http://www.payvand.com/news/18/aug/1103.html>

*Pažūheškadeh-ye moṭāla ‘āt-e vāžehgozīnī (Research Institute for Terminology Studies)*. (n.d.). Retrieved from Farhangestān-e zabān va adab-e fārsī (The Academy of Persian and Literature): <https://terminology.apll.ir/fa/%d8%a2%d9%85%d9%88%d8%b2%d8%b4/#i>

The Academy of Persian Language and Literature. (2020, May 20). *Tūzīḥ-e farhangestān darbāreh-ye vāzehā-ye mortabeḥ bā bīmārī-ye kūvīd-19 (Explanation of the Academy about the terms related to the Covid-19 pandemic)*. Retrieved from Farhangestān-e zabān va adab-e fārsī (The Academy of Persian Language and Literature): <https://apll.ir/1399/02/31/%d8%aa%d9%88%d8%b6%db%8c%d8%ad-%d9%81%d8%b1%d9%87%d9%86%da%af%d8%b3%d8%aa%d8%a7%d9%86-%d8%af%d8%b1%d8%a8%d8%a7%d8%b1%db%80-%d9%88%d8%a7%da%98%d9%87%e2%80%8c%d9%87%d8%a7%db%8c-%d9%85%d8%b1%d8%aa%d8%a8/>

The Academy of Persian Language and Literature. (2019). *Farhangestān-e zabān va adab-e fārsī (The Academy of Persian Language and Literature)*. Retrieved October 8, 2019, from darbāreh-ye farhangestān-e zabān va adab-e fārsī (About The Academy of Persian Language and Literature): <https://apll.ir/>

The Academy of the Hebrew Language. (2021). *History*. Retrieved from The Academy of the Hebrew Language: <https://en.hebrew-academy.org.il/about-the-academy/history-of-the-academy/>

The Constitution of the Islamic Republic of Iran. (1989). *Constitution*. Retrieved April 2020, from Islamic Parliament of Iran: <https://en.parliran.ir/eng/en/Constitution>

*The Ottoman Constitution (23 December 1876)*. (n.d.). (A. I. Boğaziçi University, Producer) Retrieved 2021, June 29, from Türk Anayasa Hukuku Sitesi: <https://www.anayasa.gen.tr/1876constitution.htm>

TRT. (2021, 10 3). *Ĝoz'eyāt-e dīdārhā-ye ra'īs-e maġles-e mellī-ye kabīr-e Torkīye dar kazāḥestān (Details of the Meetings of the Speaker of the Turkish Grand National Assembly in Kazakhstan)*. Retrieved from TRT: <https://www.trt.net.tr/persian/trkhyh-1/2021/09/27/jzy-yt-dydrhy-ry-ys-mjls-mly-khbyr-trkhyh-dr-qzqstn-1711080>

*Supreme Council for Islamic Revolution in Iraq (SCIRI)*. (2015, June). Retrieved from National Consortium for the Study of Terrorism and Responses to Terrorism: <https://www.start.umd.edu/baad/narratives/supreme-council-islamic-revolution-iraq-sciri>

- Vajehyab*. (n.d.). Retrieved from peymūdan:  
<https://www.vajehyab.com/?q=%D9%BE%DB%8C%D9%85%D9%88%D8%AF%D9%86&d=en>
- Webster, M. (2021, August 26). Retrieved from Protocol: <https://www.merriam-webster.com/dictionary/protocol>
- Webster, M. (2021, May 2). *Aeroplane*. Retrieved from Merriam-Webster: <https://www.merriam-webster.com/dictionary/aeroplane#learn-more>
- Webster, M. (2021, September 21). *Bureaucracy*. Retrieved from Merriam-Webster: <https://www.merriam-webster.com/dictionary/bureaucracy>
- Webster, M. (2021). *Bureaucracy*. Retrieved from Merriam Webster: <https://www.merriam-webster.com/dictionary/bureaucracy>
- Webster, M. (2021, October 23). *Capitalism*. Retrieved from Merriam-Webster: <https://www.merriam-webster.com/dictionary/capitalism>
- Webster, M. (2021, September 25). *Dictatorship*. Retrieved from Merriam-Webster: <https://www.merriam-webster.com/dictionary/dictatorship#h1>
- Webster, M. (2021, June 26). *Floor*. Retrieved from Merriam-Webster: <https://www.merriam-webster.com/dictionary/floor#etymology>
- Webster, M. (2021, September 31). *Opposition*. Retrieved from Merriam-Webster: <https://www.merriam-webster.com/dictionary/opposition#h1>
- Webster, M. (n.d.). *Doctrine, noun*. Retrieved 10 25, 2021, from Merriam-Webster: <https://www.merriam-webster.com/dictionary/doctrine>
- Yarshater, E. (1989). *Persia or Iran, Persian or Farsi*. Retrieved 2 16, 2019, from Iran Heritage: [http://www.iran-heritage.org/?page\\_id=345](http://www.iran-heritage.org/?page_id=345)
- Yazıcı, T. (2002). *Ḥabīb Esfahāni*. Retrieved November 2020, from Encyclopædia Iranica, online edition: <https://www.iranicaonline.org/articles/habib-esfahani>
- Yūsufī, Ğ. Ḥ. (1988, December 15). *Bayhaqī, Abu'l-Faẓl*. Retrieved from Encyclopaedia Iranica: <https://iranicaonline.org/articles/bayhaqi-abul-fazl-mohammad-b>

(n.d.). Retrieved July 7, 2021, from National Assembly of IR of Afghanistan:  
<http://www.parliament.af/>

(n.d.). Retrieved from Türkiye Büyük Millet Meclisi (The Grand National Assembly of Turkey): <https://global.tbmm.gov.tr/>

Council of Ministers. (1999). *Islamic Parliament Research Center Of The Islamic Republic Of IRAN*. (H. Habibi, Editor) Retrieved February 9, 2021, from *Áyinnāmeḥ-e eḡrā'ī-ye qānūn-e mamnū'eyat-e bīkārgīrī-ye 'asāmī, 'anāvīn va eṣṭelāḥāt* (Executive Regulations to Ban the Use of Foreign Names, Titles, and Terms): <https://rc.majlis.ir/fa/law/show/119218>

(2019, 9 3). Retrieved from Sāzmān-e mellī-ye estāndārd-e Īrān (The Institute of Standards and Industrial Research of Iran): <http://isiri.gov.ir/portal/home/>

*Zنده bād Moḥammad Reżā Bāṭenī*. (n.d.). Retrieved from The Academy of Persian Language and Literature: <https://apll.ir/1400/02/22/%d9%be%db%8c%d8%a7%d9%85-%d8%aa%d8%b3%d9%84%db%8c%d8%aa-%d9%81%d8%b1%d9%87%d9%86%da%af%d8%b3%d8%aa%d8%a7%d9%86-%d8%b2%d8%a8%d8%a7%d9%86-%d9%88-%d8%a7%d8%af%d8%a8-%d9%81%d8%a7%d8%b1%d8%b3%db%8c-2/>



## Appendix

Terms	TF-IDF in TAP Corpus	TF-IDF in TalkBank Corpus
pārlemān	171,874	250
maḡles	666,804	897,288
maḡles (customized)	251,503	233,333
kāndīdā	217,416	150,205
kāndīdā (customized)	146,735	75,260
nāmzad	363,152	205,958
nāmzad (customized)	220,081	105,732
opūzīsyūn	32,245	56,272
gorūh-e moḡālef	1,951	2,303
demūkrāsī	121,253	131,672
mardomsālārī	63,457	30,534
demūkrāt	274,513	38,594
mardombāvar	0	1,625
dīktātur	41,373	86,120
mostabed	16,959	20,078
dīktātūrī	25,757	55,481
estebdād	16,959	50,358
režīm	433,642	410,420
režīm (customized)	475,712	388,321
nežām	426,197	757,585
nežām (customized)	468,423	743,554
būrūkrāsī	12,817	7,715
dīvānsālārī	580	1,000
būrūkrāt	2,439	653
dīvānī	4,236	4,470
īde'ulūžī	69,914	36,935
marām	28,891	24,918
marām (customized)	27,782	21,775
īde'ulūžīk/ī	61,749	23,361
marāmī	2,400	2,785
terūrīsm	111,197	66,407
vaḡšatafkanī	0	24
terūrīst	119,947	103,334
vaḡšatafkan	0	35
pūpūlīst	20,556	6,225
'avāmgarā	0	182
pūpūlīsm	9,390	3,541
'avāmgarāyī	2,995	452
kāpītālīsm	6,751	21,233
sarmāyedārī	102,810	31,417
kāpītālīst	6,722	1,807
sarmāyedār	118,300	11,910
reform	1,915	2,670
eṡlāḡ	354,145	280,753
reformīst	580	1,274

eṣlāḥ-e ṭalab	173,297	89,329
reformīsm	0	378
eṣlāḥ-e ṭalabī	74,391	11,885
doktorīn	25,553	11,937
rahnāme	0	144
esterāteżī	182,612	87,037
rāhbord	278,807	54,432
esterāteżīk	113,916	104,340
rāhbordī	277,420	67,880
protokol	299,601	22,194
tašrīfāt	58,491	22,335
tašrīfāt (customized)	57,902	22,160
ommat	88,606	65,070
mellat	350,734	491,876

## **Abstract**

Language is an organism. It is in constant change from the outside and inside as a matter of international relations, lifestyle, technology, culture, entertainment, science, etc. As language is influenced by the civilization of the people speaking it, it also intertwines and overlaps with other languages as they interact directly or indirectly with each other. Amid the openness we witness in the world, it is difficult to control this mixing of foreign terms and loanwords that enter the vocabulary of any language especially in the presence of the Internet and with the given ease of travel. However, some countries are still setting out to codify the use of foreign terms and maintain their language and national identity by way of creating language academies.

Iran established one of the most active and resourceful language academies of the world in terms of structure, organization, level of expertise, and current affairs: the Academy of Persian Language and Literature (APLL) or Farhangestān. This thesis examines the impact of the terms in the fields of Political Science and International Relations coined by the Farhangestān on the Persian language used in newspapers and blogs. Based on a set of quantitative, qualitative, and contextual questions, the study looks at the sources or type of publication, the type of linguistic borrowing, and the time that has passed since the formal introduction of the neologisms in question to ask whether they have been successful or not.

The thesis is divided into six chapters. After the Introduction in chapter 1, chapter 2 moves along three main axes. Starting, first, with a discussion of the historical background of language academies in general, the chapter then goes on in its second section to focus on the issue of linguistic borrowing. In doing so, it answers the question of how linguistic borrowing affects the nature of languages in general and sheds light on the eventful history of linguistic borrowing in Persian in particular. Moreover, it looks at how this led to the importance of the existence of language planning and the centralization of language policy, which formed a strong basis for the formation of language academies seeking to preserve the originality of their respective language.

The third axis of chapter 2 scrutinizes the Farhangestān as well as its goals in all its stages. Moreover, it takes a closer look at the APLL terminology group and the framework within which the terminology group acts to form neologisms. As the unit concerned with studying the foreign terms that entered Persian, it works on finding authentic linguistic alternatives in the language to replace those terms and on keeping the same meaning with the aim of gaining their acceptance by native speakers. The chapter concludes with a glance at the

guidelines of Persian term formation created and followed by the APLL. Chapter 3 then covers selected previous studies and efforts that deal with evaluating the impact of the terms coined and approved by the Farhangestān on modern Persian.

The fundamental focus of this study is covered in chapters 4 and 5. Chapter 4 concisely presents the methodology used in this study. The analysis is methodologically based (1) on a selection of foreign and approved terms in the field of politics that will be examined in the study, (2) on an extensive corpus analysis, and (3) on the weight calculation for all examined terms. After presenting the selected terms, this chapter introduces and contextualizes the two corpora TAP and TalkBank with more than 500 million words through which the terms are evaluated. It also presents the so-called TF-IDF (term frequency/inverse document frequency) equation used to calculate the weight of the terms in both corpora, which will help later in comparing the approved equivalent terms to the foreign terms they were meant to replace. This is crucial to answering the leading question of this study about the impact of the approved terms on modern Persian. In other words: Thanks to the weight calculation it is possible to measure the extent to which the Academy succeeded in making the neologisms it coined acceptable in newspapers and blogs, that is, among journalists and other proficient language users.

Chapter 5 deals with the analysis of the selected terms. Each approved Persian term is presented alongside the foreign term it was meant to replace. In the beginning of each section, the foreign terms and their Persian equivalents are presented separately. The terms are explained etymologically and, where possible, the term history is summarized and the date of their first use or entry in the Persian language is recorded. To be able to trace the terms, their history, and their evolution, some of the oldest Persian dictionaries and Iranian encyclopedias are consulted. Depending on their density, the terms are further contextualized, for example, by presenting derivatives (e.g., *demūkrāt* with *demūkrāsī*), related terms approved by the APLL, and compound terms formed with the terms in question.

The second part of each section in chapter 5 highlights the quantitative and statistical properties of each term in both corpora, that is essentially the term frequency and the document frequency. This is necessary to calculate their weight and put them in relation with each other. Some terms such as *mağles* (parliament) or *nāmzad* (candidate) have more than one meaning. In these cases, the historical-etymological discussion and the quantitative analysis are supported by a collocation analysis (context analysis) that allows for a more accurate weight calculation thanks to the use of approximated frequencies. Furthermore, where ambiguous, the terms in

question will be contextualized by providing example texts from the corpora to make a better understanding of their everyday use possible.

Each section of chapter 5 concludes with a discussion of the results and implications of the weight calculation. This ascertains the actual use of each Persianized term in newspapers and blogs and allows for conclusions regarding its success or lack thereof vis-à-vis the foreign term it came to replace.

Chapter 6, finally, presents the findings, conclusions, and recommendations of the research. Taken together, the analysis of each term according to its type of source or publication, the determination of its type of borrowing, as well as the consideration of the time that has passed since the (formal) introduction of that term reveal and put in perspective the impact of the Farhangestān on modern Persian.

## **Abstract (German)**

Sprache ist ein Organismus. Sie unterliegt einem ständigen Wandel von außen und innen in Bezug auf internationale Beziehungen, Lebensstil, Technologie, Kultur, Unterhaltung, Wissenschaft usw. Da Sprache von der Zivilisation der Menschen beeinflusst wird, die sie spricht, verschränkt und überschneidet sie sich auch mit anderen Sprachen, die unmittelbar oder mittelbar miteinander interagieren. Inmitten der Offenheit, die wir in der Welt erleben, ist es schwierig, diese Mischung aus Fremdwörtern und Lehnwörtern zu kontrollieren, die in den Wortschatz jeder Sprache eingehen, insbesondere angesichts des Internets und der gegebenen Leichtigkeit des Reisens. Einige Länder versuchen jedoch noch immer, die Verwendung von Fremdwörtern zu beschränken und ihre Sprache und nationale Identität durch die Einrichtung von Sprachakademien zu bewahren.

Iran hat eine der aktivsten und am besten aufgestellten Sprachakademien der Welt in Bezug auf Struktur, Organisation, Fachwissen und aktuelle Ereignisse gegründet: die Academy of Persian Language and Literature (APLL) oder Farhangestān. Diese Arbeit untersucht den Einfluss der von der Farhangestān geschöpften Begriffe aus den Bereichen Politikwissenschaft und Internationale Beziehungen auf die persische Sprache in Zeitungen und Blogs. Anhand einer Reihe quantitativer, qualitativer und kontextbezogener Fragen untersucht diese Studie die Quellen bzw. die Art der Veröffentlichung, die Art der sprachlichen Entlehnung und die Zeit, die seit der formalen Einführung der fraglichen Neologismen vergangen ist, um zu fragen, ob die Begriffe erfolgreich waren oder nicht.

Diese Dissertation gliedert sich in sechs Kapitel. Nach der Einführung in Kapitel 1 bewegt sich Kapitel 2 entlang dreier Hauptachsen. Ausgehend von einer Diskussion des historischen Hintergrunds von Sprachakademien geht das Kapitel dann im zweiten Abschnitt auf die Frage der sprachlichen Entlehnung ein. Es beantwortet damit die Frage, wie sich die Sprachentlehnung auf das Wesen von Sprachen im Allgemeinen auswirkt und beleuchtet die bewegte Geschichte der Sprachentlehnung im Persischen im Besonderen. Darüber hinaus wird untersucht, wie dies zur Bedeutung der Sprachplanung und der Zentralisierung der Sprachpolitik führte, die eine starke Grundlage für die Bildung von Sprachakademien bildeten, die die Originalität ihrer jeweiligen Sprache bewahren wollten.

Die dritte Achse von Kapitel 2 hinterfragt die Farhangestān sowie ihre Ziele in all ihren Phasen. Darüber hinaus werden die APLL-Terminologiegruppe und der Rahmen, in dem die Terminologiegruppe zur Bildung von Neologismen agiert, genauer betrachtet. Als Einheit, die

sich mit der Erforschung der persischen Fremdwörter beschäftigt, arbeitet sie daran, authentische linguistische Alternativen in der Sprache zu finden, um diese Begriffe zu ersetzen und ihre Bedeutung beizubehalten, mit dem Ziel, ihre Akzeptanz bei den MuttersprachlerInnen zu gewinnen. Das Kapitel schließt mit einem Blick auf die Richtlinien der persischen Begriffsbildung, die von der APLL erstellt und befolgt wurden. Kapitel 3 behandelt dann ausgewählte frühere Studien und Arbeiten, die sich mit der Bewertung der Auswirkungen der von der Farhangestān geprägten und anerkannten Begriffe auf das moderne Persisch befassen.

Die grundlegenden Schwerpunkte dieser Studie werden in den Kapiteln 4 und 5 behandelt. Kapitel 4 stellt die in dieser Studie verwendete Methodik prägnant vor. Methodisch basiert die Analyse (1) auf einer Auswahl ausländischer und anerkannter Begriffe aus dem Politikbereich, die in der Studie untersucht werden, (2) auf einer umfangreichen Korpusanalyse und (3) auf der Gewichtungsberechnung für alle untersuchten Begriffe. Nach der Vorstellung der ausgewählten Begriffe werden in diesem Kapitel die beiden Korpora TAP und TalkBank mit insgesamt mehr als 500 Millionen Wörtern, anhand derer die Begriffe bewertet werden, vorgestellt und kontextualisiert. Es stellt auch die sogenannte TF-IDF-Gleichung (Term Frequency/Inverse Document Frequency) vor, die verwendet wird, um das Gewicht der Begriffe in beiden Korpora zu berechnen, was später beim Vergleich der anerkannten äquivalenten Begriffe mit den ausländischen Begriffen hilft, die sie ersetzen sollten. Dies ist entscheidend für die Beantwortung der Leitfrage dieser Studie nach den Auswirkungen der anerkannten Begriffe auf das moderne Persisch. Mit anderen Worten: Dank der Gewichtungsberechnung lässt sich messen, inwieweit es der Akademie gelungen ist, die von ihr geprägten Wortschöpfungen in Zeitungen und Blogs, also bei JournalistInnen und anderen fähigen SprachnutzerInnen, salonfähig zu machen.

Kapitel 5 befasst sich mit der Analyse der ausgewählten Begriffe. Jeder zugelassene persische Begriff wird neben dem ausländischen Begriff angezeigt, der ersetzt werden sollte. Zu Beginn jedes Abschnitts werden die ausländischen Begriffe und ihre persischen Äquivalente separat vorgestellt. Die Begriffe werden etymologisch erklärt sowie, soweit möglich, die Begriffsgeschichte zusammengefasst und das Datum ihrer erstmaligen Verwendung bzw. Eintragung in die persische Sprache festgehalten. Um die Begriffe, ihre Geschichte und ihre Entwicklung nachvollziehen zu können, werden einige der ältesten persischen Wörterbücher und iranischen Enzyklopädien herangezogen. Je nach Dichte werden die Begriffe weiter kontextualisiert, beispielsweise durch die Darstellung von Derivaten.

Der zweite Teil jedes Abschnitts in Kapitel 5 beleuchtet die quantitativen und statistischen Eigenschaften jedes Begriffs in beiden Korpora, das heißt, im Wesentlichen die Begriffshäufigkeit und die Dokumentenhäufigkeit. Dies ist notwendig, um ihr Gewicht zu berechnen und sie in Beziehung zu setzen. Einige Begriffe wie *mağles* (Parlament) oder *nāmzad* (Kandidat) haben mehr als eine Bedeutung. In diesen Fällen werden die historisch-etymologische Diskussion und die quantitative Analyse durch eine Kollokationsanalyse (Kontextanalyse) unterstützt, die durch die Verwendung angenäherter Häufigkeiten eine genauere Gewichtsrechnung ermöglicht. Darüber hinaus werden die betreffenden Begriffe bei Uneindeutigkeiten durch die Bereitstellung von Beispieltextrn aus den Korpora kontextualisiert, um ein besseres Verständnis ihres alltäglichen Gebrauchs zu ermöglichen.

Jeder Abschnitt von Kapitel 5 schließt mit einer Diskussion der Ergebnisse und Auswirkungen der Gewichtsrechnung ab. Dadurch wird die tatsächliche Verwendung jedes persianisierten Begriffs in Zeitungen und Blogs ermittelt und lässt Rückschlüsse auf seinen Erfolg oder Misserfolg gegenüber dem durch ihn ersetzten Fremdbegriff zu.

Kapitel 6 schließlich präsentiert die Ergebnisse, Schlussfolgerungen und Empfehlungen der Forschung. Die Analyse jedes Begriffs nach seiner Art der Quelle oder Veröffentlichung, die Bestimmung der Art der Entlehnung sowie die Betrachtung der Zeit, die seit der (förmlichen) Einführung des Begriffs verstrichen ist, lassen den Einfluss der Farhangestān auf das moderne Persisch erkennen und setzen ihn in Beziehung zu seinem Entstehungskontext.