



A Corpus-based Analysis of Orthographic Errors of Arabic Persian Learners at the Advanced Level in the Framework of Persian Script Instruction

Golnoosh Shafaei

PhD candidate of Linguistics, Islamic Azad University, South Tehran Branch, Tehran, Iran.
golnooshshafaei@gmail.com

Shahram Modarres Khiabani*

Corresponding author, PhD, Assistant Professor of English Language Teaching and Translation Department, Islamic Azad University, Karaj Branch, Karaj, Iran.
shmodarress@yahoo.com

Faeghe Shahhoseini

PhD Graduate in General Linguistics, Persian Language Instructor, International College of Tehran University of Medical Science, Tehran, Iran.
f.Shahhoseini@gmail.com

Abstract:

In recent decades, a growing number of non-Persian speakers have been visiting Iran, eager to gain a deeper understanding of Iranian culture and civilization. Simultaneously, there has been a notable influx of Arabic-speaking students enrolling in Iranian universities. Therefore, in this research, an attempt has been made to examine the Persian writing structure of advanced-level Arabic language learners from the corpus-based perspective. The study aims to discern spelling and technical errors in their written works, adhering to the writing guidelines established by the Persian Language and Literature Academy. To achieve this objective, a corpus consisting of 151 Persian essays from Arab learners from Iraq, who underwent the Persian proficiency test at Imam Reza International University, was compiled. The corpus consists of 3234 types and 32096 samples. Utilizing the Wordsmith software, word frequencies were extracted from the essays. Employing content analysis, the study scrutinized various error types, including spelling/stylistic errors, semi-space deletion errors, space deletion errors, space insertion errors, cross-typing, and separate writing. The findings highlighted that the most prevalent errors were in the category of space insertion, followed by semi-space deletion errors, spelling/stylistic errors, and spelling errors. Separate writing exhibited the lowest occurrence, with no errors identified in the category of space deletion. The analysis revealed that these errors predominantly originated from intra-language factors, specifically second-language learning and developmental errors.

*Cite this article: Shafaei, Golnoosh. Modarres Khiabani, Shahram. Shahhoseini, Faeghe. A Corpus-based Analysis of Orthographic Errors of Arabic Persian Learners at the Advanced Level in the Framework of Persian Script Instruction. *Vol. 12, No. 2 (Tome 26), October 2023, 207-240.*

DOI: 10.30479/JTPSOL.2024.19654.1658



Received on: 10/01/2024

Accepted on: 17/02/2024

© The Author(s).

Publisher: Imam Khomeini International University

Introduction

In recent decades, we have been encountering many non-Persian speakers travelling to Iran and gaining more knowledge about Iranian culture and civilization. Notably, in 2018, before the Coronavirus outbreak, Iraqis surpassed visitors from all other nations in their travels to Iran, a fact substantiated by statistics from the Ministry of Cultural Heritage, Tourism, and Handicrafts.

Beyond tourism, there has been a notable influx of Arabic-speaking students enrolling in Iranian universities. To accommodate this, various centers dedicated to teaching Persian to non-native speakers are actively engaged. These centers focus on developing proficiency in the four essential language skills-reading, listening, writing, and speaking as well as imparting knowledge of Persian script.

The teaching of Persian script is driven by the dual purpose of honoring the essence of the Persian language and safeguarding its official script status in Iran. Recognized as the official language and script of the country, the Persian script holds significance not only for Iranians but also plays a crucial role in presenting the history of Iranian culture and civilization to foreigners. According to the fifteenth article of the Constitution of the Islamic Republic of Iran, the Persian script is mandated for all official documents, correspondences, and textbooks. This necessitates the establishment of clear rules to ensure uniformity and preservation of the Persian script's integrity. The Academy of Persian Language and Literature, fulfilling its duty to protect the Persian language and script, has issued a comprehensive set of guidelines titled "Persian Script Instructions" to regulate and maintain the essence of the Persian script.

This research delves into an analysis of orthographic errors made by 151 non-native Persian language learners, specifically Iraqi participants in the Persian language proficiency test conducted by Imam Reza International University. This examination is conducted within the framework set by the Persian Language and Literature Academy.

Considering the contemporary emphasis on objective linguistic data in research and the pivotal role of corpora in linguistic studies, this research adopts the use of linguistic corpora for data analysis. Recognizing the effectiveness of corpora in second language teaching, this approach enables language teachers to observe and assess the learning process of language learners.

Methodology

This research focuses on 151 non-native Persian language learners, primarily from Iraq, who took part in the Persian language proficiency test at Imam Reza International University. The aim is to assess their proficiency in Persian for pursuing Master's degree studies.

In the next step, the compositions have been typed. Then, to enable computer processing, the formatted typesetting file was converted to a text file through the web-based software Espoz. After that, using WordSmith corpus analysis software, a list of words was prepared, including the frequency of each word along with its percentage and also the dispersion of each word in the entire text was determined. Then, these words and their frequencies were analyzed.

Categorized as corpus-based research, this study utilizes data extraction from a linguistic corpus, employing WordSmith corpus analysis software. The self-motivated method was employed for data collection. The corpus comprises 32,096 samples of Persian writings by Arabic language learners, encompassing 3,234 types. It's noteworthy that the calculation of types only pertains to vocabulary, excluding punctuation marks. The distinctive feature lies in the distance inserted on both sides, setting this research apart from other forms.

Results

In alignment with Modarres Khiabani's research (2016), this study examines errors in six categories: spelling/stylistic errors, insertion of space, deletion of space, deletion of half-space, cross-typing, and separate writing based on the Persian script and correct word forms.

Considering that by spelling error we mean that the intended word is spelled incorrectly, spelling errors of Persian learners were extracted from the corpus, which were examined in two areas phonetic (consonant and vowel) errors and calligraphic errors. In the phonetic error domain, we are faced with 421 samples, which is a relatively significant frequency. Incorrect usage of signs such as Hamza, and Tanwin and the use of Arabic-derived words by Arabic-speaking Persian learners resulted in 586 spelling errors that were examined in the calligraphic error group.

The stylistic error refers to cases where although the intended word is not misspelled in terms of writing, it is irrelevant to the formal style of the specific test we are considering, which here is the Persian language proficiency test. We encountered 221 instances in the corpus.

In the conducted examinations, two types of spacing errors in the learners' corpus were observed, due to the insertion of space before punctuation marks such as "dot," "comma" and "colon" as well as the insertion of space between compound words. The frequency of errors resulting from the insertion of space before these punctuation marks is 1459, 136, and 5 cases, respectively, in which the error related to the insertion of space before the dot is considerable.

Regarding the deletion of space error, no space has been inserted between two independent words. No cases of space deletion error were observed in the Persian learner corpus, and learners strongly tended to insert spaces between all words in their writing.

Since the written versions of Persian learners are handwritten, examining this type of error is difficult. However, among all the errors examined in the corpus, the error of deleting half-space is the most frequent. 1583 instances of errors indicate that Arabic-speaking Persian learners do not pay attention to the use of half-space in compound words and after some inflectional affixes such as "mi-" or "-ha" and "-haye".

Cross-typing error refers to cases that should be written separately but are written together, which we encountered with 138 errors in the corpus.

Separate writing error refers to cases that should be written together but are written separately, which we encountered with 53 errors in the corpus.

Analysis of each error type within the Persian language learner corpus reveals the highest frequency in the insertion of space category, followed by deletion of half-space, spelling/stylistic errors, and spelling errors. Interestingly, no errors were found in the field of space deletion.

Categorizing errors based on Zia Hosseini's classification, two types are considered: interlingual errors, stemming from the interference of the Persian learner's mother tongue, and intralingual errors, resulting from incorrect or incomplete learning of the target language. The study underscores that writing errors are significantly influenced by interlanguage-origin errors, highlighting the impact of learning and developing a second language on the prevalence and nature of these errors.

Conclusion

The findings reveal the highest frequency of errors in inserting a space and removing a semi-space, indicating a potential lack of training on issues related to proper spacing before punctuation marks and the use of semi-spaces in Persian. Notably, the advanced-level Farsi-Arabic learners exhibit a high frequency of spelling and stylistic errors, underscoring the need for more precise teaching points and additional training exercises. While half-space omission errors are more

frequent than spelling/stylistic errors, the latter holds greater importance in the realms of communication and comprehension. Despite the more frequent occurrence of half-space omission errors, the impact of spelling/stylistic errors on communication and understanding elevates their significance. Errors associated with transliteration and separate writing rank third, highlighting the complexity of these topics. This suggests a pressing need for more detailed training and increased practice for Farsi-Arabic language learners in these specific areas.

Considering the results of this research and since the Persian writing system is an alphabetic system, it seems that phonological awareness, as a component, has a significant impact on reading and writing in this language.

Based on the identified errors, it is expected that revisions will be made to Persian teaching resources, and exercises and classroom activities will be designed to reduce the errors of Persian learners.

In addition to using instructional resources, regular and periodic evaluations can be a fundamental step towards facilitating the learning process and reducing errors.

Conflict of Interest

There was no conflict of interest in this research.

Acknowledgment

Appreciation goes to all the people who helped us in doing this research, in particular Dr. Jalal Rahimian, who is a professor of linguistics in the Department of Foreign Languages and Linguistics at Shiraz University.

Keywords: Language proficiency test, language corpus, Persian script and language, learning Persian, linguistic error, advanced-level Arabic language learners.



بررسی پیکره‌بنیاد خطاهای نگارشی فارسی آموزان عرب‌زبان سطح پیشرفته در چارچوب دستور خط فارسی (پژوهشی)

گل‌نوش شفافی

دانشجوی دکتری زبان‌شناسی، دانشگاه آزاد اسلامی واحد تهران جنوب، تهران، ایران.

golnooshshafaei@gmail.com

شهرام مدرس خیابانی*

نویسنده مسئول، استادیار گروه آموزش و مترجمی زبان انگلیسی، دانشگاه آزاد اسلامی، کرج، ایران.

shmodarress@yahoo.com

فائقه شاه‌حسینی

دکتری زبان‌شناسی همگانی، مدرس زبان فارسی به غیر فارسی‌زبانان، کالج بین‌الملل دانشگاه علوم پزشکی تهران، تهران، ایران.

f.Shahhoseini@gmail.com

چکیده

در دهه‌های جدید، با خیل سفر غیرفارسی‌زبانان به ایران و آشنایی هرچه بیشتر آن‌ها با فرهنگ و تمدن ایرانی-اسلامی روبرو هستیم. در این میان، آمار دانشجویان عرب‌زبان مشغول به تحصیل در دانشگاه‌های ایران نیز قابل توجه است؛ لذا در این پژوهش سعی بر آن بوده است، ساختار نگارشی فارسی‌آموزان عرب‌زبان سطح پیشرفته، از منظری پیکره‌بنیاد بررسی شود؛ تا اشکالات املائی و فنی تولیدات نوشتاری آن‌ها، بر اساس دستورالعمل نگارش مصوب فرهنگستان زبان و ادب فارسی مشخص شود. برای دستیابی به این هدف، ابتدا پیکره‌ای از ۱۵۱ انشاء فارسی‌آموزان عرب‌زبان از کشور عراق که در آزمون بسندگی زبان فارسی دانشگاه بین‌المللی امام رضا علیه‌السلام شرکت کرده بودند، آماده شد. این پیکره، شامل ۳۲۰۹۶ نمونه است که از ۳۲۳۴ نوع تشکیل شده است؛ سپس به کمک نرم‌افزار ورداسمیت، فراوانی واژه‌های موجود در انشاهای استخراج شد. با روش تحلیل محتوا، اشکالات املائی، سبکی، خطای حذف نیم‌فاصله، خطاهای حذف فاصله، خطای درج فاصله، سرهم‌نویسی و جدانویسی بررسی شد. یافته‌ها نشان داد؛ بیشترین خطا در حوزه درج فاصله بود و حذف نیم‌فاصله، اشکالات املائی، سبکی و خطای سرهم‌نویسی در رده‌های بعدی و همچنین جدانویسی در مرتبه آخر قرار داشت و در نهایت هیچ مورد اشتباه در حوزه حذف فاصله به چشم نخورد. همچنین خطاهایی با منشأ درون‌زبانی که خطاهای یادگیری و رشد زبان دوم بودند، سهم بسزایی در تعیین منشأ بروز خطاها داشت.

کلیدواژه‌ها

آزمون بسندگی، پیکره زبانی، خط و زبان فارسی، خطای زبانی، عرب‌زبانان سطح پیشرفته.

*استاد: شفافی، گل‌نوش. مدرس خیابانی، شهرام، فائقه شاه‌حسینی. بررسی پیکره بنیاد خطاهای نگارشی فارسی‌آموزان عرب‌زبان سطح پیشرفته در چارچوب دستور خط فارسی، سال دوازدهم، شماره دوم (پیاپی ۲۶)، پاییز و زمستان ۱۴۰۲، ۲۰۷-۲۴۰.

شناسه دیجیتال (DOI): 10.30479/JTPSOL.2024.19654.1658

تاریخ پذیرش نهایی مقاله: ۱۴۰۲/۱۱/۲۸

تاریخ دریافت مقاله: ۱۴۰۲/۱۰/۲۰

ناشر: دانشگاه بین‌المللی امام خمینی (ع)

۱. مقدمه

در دهه‌های جدید، با خیل سفر غیرفارسی‌زبانان به ایران و آشنایی هرچه بیشتر آن‌ها با فرهنگ و تمدن ایرانی - اسلامی روبرو هستیم. طبق آمار وزارت میراث فرهنگی، گردشگری و صنایع دستی ایران در سال ۱۳۹۸ که مربوط به دوران پیش از همه‌گیری بیماری کرونا است، عراقی‌ها رکورددار تحصیل و سفر به ایران بودند. در این میان، آمار دانشجویان عرب‌زبان مشغول به تحصیل در دانشگاه‌های ایران نیز قابل توجه است. هم‌سو با نیازسنجی‌ها، مراکز آموزش زبان فارسی به غیرفارسی‌زبانان با هدف آموزش و ترویج زبان فارسی و فرهنگ ایرانی - اسلامی مشغول به کار هستند و به تدریس مهارت‌های چهارگانه و آموزش اصول نگارشی خط فارسی به فارسی‌آموزان می‌پردازند. از جمله دلایل اصلی آموزش اصول خط فارسی را می‌توان توجه به روح زبان فارسی و حفاظت و پاسداری از رسم خط و زبان رسمی کشور دانست. همچنین لازم است، خط فارسی به‌عنوان یکی از پایه‌های اصلی در نمود تاریخ فرهنگ و تمدن ایرانی برای خارجیان به عرصه نمایش گذاشته شود.

با توجه به این که خط فارسی، به‌موجب اصل پانزدهم قانون اساسی جمهوری اسلامی ایران، خط رسمی کشور است و کلیه اسناد رسمی، مکاتبات و کتاب‌های درسی باید به این خط نوشته شود، بنابراین چنین خطی باید ضوابطی داشته باشد؛ تا همگان با رعایت آن‌ها، چهره خط فارسی را حفظ کنند. در این خصوص، فرهنگستان زبان و ادب فارسی، به‌حکم وظیفه‌ای که برطبق اساسنامه خود در پاسداری از زبان و خط فارسی بر عهده دارد، مجموعه قواعد و ضوابطی را با عنوان «دستور خط فارسی» ارائه نموده است.

در این پژوهش، سعی بر آن است تا در چارچوب دستور خط فارسی فرهنگستان زبان و ادب فارسی، مهارت نگارش ۱۵۱ فارسی‌آموز عرب‌زبان از کشور عراق که در آزمون بسندگی زبان فارسی دانشگاه بین‌المللی امام رضا علیه‌السلام شرکت کرده بودند، مورد تجزیه و تحلیل قرار گیرد.

از سوی دیگر، با توجه به ضرورت تکیه بر داده‌های عینی زبانی در پژوهش‌های زبان‌شناختی کاربردی و نقش کارساز پیکره‌ها در این امر، پژوهش حاضر از طریق بسامدشماری خطاها به روش پیکره‌بنیاد انجام شده است. در ادامه، پرسش‌های زمینه‌ساز تحقیق حاضر آورده شده است.

۱- خطاهای نگارشی فارسی‌آموزان عرب‌زبان سطح پیشرفته، براساس الگوی مصوب فرهنگستان زبان و ادب فارسی کدام است؟

۲- پربسامدترین خطاهای نگارشی فارسی‌آموزان عرب‌زبان کدام است؟

۳- فراوانی منشأ خطاهای نگارشی موردنظر کدام است؟

۲. چارچوب نظری پژوهش

همان‌گونه که در بخش مقدمه اشاره شد، در این پژوهش به‌منظور دستیابی به خط‌های نگارشی فارسی‌آموزان عرب‌زبان، از دستاوردهای زبان‌شناسی پیکره‌ای استفاده می‌شود و ملاک ارزیابی ساختار نگارشی موردنظر، اصول مطرح در «ویراست جدید دستور خط فارسی» مصوب فرهنگستان زبان و ادب فارسی (چاپ اول ۱۴۰۱) است. بنابراین ابتدا نکاتی دربارهٔ پیکره و زبان‌شناسی پیکره‌ای مطرح می‌شود و سپس به مبحث تحلیل خطا اشاره‌ای خواهد شد و درنهایت به برخی اصول نگارشی مرتبط با این پژوهش در «دستور خط فارسی» اشاره می‌شود و براساس اصول مطرح شده، طبقه‌بندی از خط‌های نگارشی ارائه خواهد شد.

۱.۲. پیکره و زبان‌شناسی پیکره‌ای

در دهه‌های جدید، رویکردهای پیکره‌محور در زبان‌شناسی مورد استقبال قرار گرفته است؛ چرا که امکان تحلیلی بسیار دقیق را برای زبان فراهم می‌نماید. علاوه بر این، حوزهٔ آموزش زبان فارسی به غیرفارسی‌زبانان نیز در ابتدای مسیر پرفراز و نشیب خود است. در این مسیر، یکی از گام‌های اساسی و ضروری، جمع‌آوری و ثبت داده‌های خام زبانی و تهیه پیکرهٔ زبانی فارسی‌آموزان غیرایرانی و توصیف داده‌های آن با استفاده از نظریه‌های زبان‌شناسی است. کمترین فایدهٔ چنین اقدامی آن است که چنانچه نظریات زبان‌شناسی در گذر زمان، ارزش و اعتبار خود را از دست بدهند، پیکرهٔ زبانی در همه‌حال ارزش خود را حفظ کرده و ماندگار می‌ماند (Mirzaei Hesariyan & Pooladestoon, 2020). بنابراین در این بخش به معرفی پیکره و زبان‌شناسی پیکره‌ای می‌پردازیم.

درخصوص پیکره، تعاریف گوناگونی ارائه شده است. مکنری^۱ و همکاران (McEnery et al., 2006: 6) پیکره را بدنه‌ای از زبان طبیعی دانسته‌اند. به نقل از مدرس خیابانی (Modarress Khiyabani, 2007) پیکره، مجموعه‌ای از متون نوشتاری و گفتاری آوانویسی شده است که می‌توان در توصیف و تحلیل زبان از آن بهره گرفت. پروان و سرکارحسن‌خان (Parvan & Sarkar Hassankhan, 2018) پیکرهٔ متنی را مجموعه‌ای از متن‌ها یا تگه‌متن‌های گفتاری و شنیداری دیجیتالی شده می‌دانند که از نمونه‌های زبانی جمع‌آوری شده از حوزه‌های مختلف است و در پژوهش‌های تجربی مورد استفاده قرار می‌گیرد. این پژوهشگران مشروعیت پیکرهٔ متنی را مشروط به رعایت چهار اصل؛ قابلیت تعمیم، ثبات داده‌ها، حجم و محتوا می‌دانند. مکنری و ویلسون^۲ (McEnery & Wilson, 2001:29-32) علاوه بر این موارد، به برخورداری پیکره از قابلیت پردازش رایانه‌ای توجه ویژه دارند.

درخصوص انواع پیکره، بیکر^۳ و همکاران (Baker, 2006: 48-49) پیکره‌های عمومی، تخصصی، مرجع،

۱. T. McEnery

۲. A. Wilson

۳. P. Baker

چندزبانه، موازی، زبان‌آموز، ناظر و در زمانی را ذکر کرده‌اند. عاصی (Assi, 2000) انواع پیکره را از نظر گستره به سه گروه محدود، متوسط و عظیم تقسیم کرده است. وی مبحث نشانه‌گذاری پیکره‌ها^۱ را نیز مطرح می‌کند که به منظور گویاتر شدن پیکره و کاربردهای خاص آن انجام می‌شود. این کدگذاری پیکره‌ها به دو صورت زبانی و غیرزبانی انجام می‌شود.

به منظور استفاده از پیکره زبانی، شاخه‌ای در زبان‌شناسی با عنوان «زبان‌شناسی پیکره‌ای» پدید آمده است. احمدی و همکاران (Ahmadi et al., 2021) زبان‌شناسی پیکره‌ای را شاخه‌ای از زبان‌شناسی کاربردی می‌دانند که به بررسی و مقایسه جنبه‌های گوناگون داده‌های زبانی می‌پردازد. طاهری و گوهری (Taheri & Gowhari, 2021) زبان‌شناسی پیکره‌ای را یک روش علمی در تحلیل زبان می‌دانند که در این رویکرد، پژوهش‌گر در پی یافتن شواهد تجربی از زبان، جهت تأیید یا رد فرضیه‌های زبانی است. صفری (Safari, 2016) داده‌های زبانی اصیل، عینی و قابل‌پردازش را اساس پژوهش‌های زبان‌شناسی پیکره‌ای می‌داند که پژوهش‌گران را قادر می‌سازد، با استناد به چنین منابعی تحلیل دقیق‌تر از مسائل زبانی ارائه دهند. برزینا^۲ و گابلاسوا^۳ (Brezina & Gablasova, 2018) زبان‌شناسی پیکره‌ای را روشی چندمنظوره برای تحلیل محاسباتی زبان می‌داند که در حوزه‌های دستور زبان، گفتمان و یادگیری و آموزش زبان کاربرد دارد. مدرس خیابانی (Modarress Khiyabani, 2017) به نقل از لیچ^۴ ویژگی‌های مهم این رشته را چنین برمی‌شمارد:

۱- تمرکز بر کنش زبانی و نه توانش زبانی؛

۲- تمرکز بر توصیف زبانی و نه بر همگانی‌های زبان؛

۳- تمرکز بر الگوهای کمی زبانی همانند الگوهای کیفی آن؛

۴- تمرکز بر دیدگاه‌های تجربی (و نه عقلانی) در بررسی‌های علمی زبان؛

مدرس خیابانی (Modarress Khiyabani, 2007) به نقل از تاگنینی بونلی^۵ دو رهیافت متفاوت پیکره‌بنیاد و برگرفته از پیکره را در زبان‌شناسی پیکره‌ای معرفی می‌کند. در رهیافت پیکره‌بنیاد، از پیکره به عنوان وسیله‌ای برای تأیید نظریه‌ها استفاده می‌شود. در رهیافت برگرفته از پیکره نظریه بر مبنای اطلاعات استخراج شده و همچنین خود پیکره ارائه می‌شود. با رهیافت برگرفته از پیکره، توصیف زبان‌شناختی از مرحله مشاهده محض خارج می‌شود و رویدادهای مکرر در کنار بسامد توزیع آن‌ها، صورت‌های مختلف زبانی را شکل می‌دهند.

^۱. corpus annotation

^۲. V. Brezina

^۳. D. Gablasova

^۴. G. Leech

^۵. E. Tognini-Bonelli

۲.۲. تحلیل خطا

شاخهٔ تحلیل خطا در فراگیری زبان دوم، توسط کورد^۱ و همکارانش شکل گرفت. این شاخه از زبان‌شناسی کاربردی، به مطالعهٔ خطاها در فرایند یادگیری زبان دوم می‌پردازد. کورد (Corder, 1967: 127) معتقد است؛ تحلیل خطاها در امر آموزش زبان، بسیار مهم است؛ زیرا شواهدی از چگونگی یادگیری زبان را برای پژوهش‌گران فراهم می‌نماید و همچنین مشخص می‌کند که فراگیران چه روش‌هایی را در کشف یک زبان به کار می‌برند. جیمز^۲ (James, 1998: 62) تحلیل خطا را مطالعهٔ دانش ناآگاهی زبان‌آموز می‌داند. به عبارت دیگر، بررسی آنچه مردم نمی‌دانند و روشی که آن‌ها تلاش می‌کنند تا دانسته‌های محدود خود را به کار ببرند. تجزیه و تحلیل خطاهای زبانی، بر سه فرض کلی استوار است:

- بروز خطاهای زبانی اجتناب‌ناپذیر است. نمی‌توان یک زبان را بدون بروز هیچ نوع خطایی فراگرفت.
- خطاهای زبانی مهم هستند.
- همهٔ خطاها ریشه در زبان مادری زبان‌آموز ندارند (Keshavarz, 1994: 44).

لازم به ذکر است؛ طبق نظر کشاوریز (Keshavarz, 1999: 52) لازم است بین خطا و اشتباه تمایز قائل شد. معیار کلی برای تشخیص خطا^۳ از اشتباه^۴، بسامد رخداد آن است. مواردی که بسامد وقوع کمتری دارند، به عنوان اشتباه و مواردی که از بسامد بیشتری برخوردارند، به عنوان خطا در نظر گرفته می‌شوند. البته بسامد وقوع کم، حتماً به معنای اشتباه بودن نیست؛ زیرا گاهی بسامد کم برخی از خطاها، به دلیل رخداد کم الگوی دستوری و یا به دلیل راهبرد اجتناب است؛ درحالی‌که بسامد رخداد بالا، معیار خطاهای زبانی است.

۱.۲.۲. منابع خطاهای زبانی

کورد (Corder, 1967: 216) منشأ خطاهای زبانی را به سه دسته تقسیم می‌کند:

الف) خطاهای بین‌زبانی: خطاهایی که بر اثر تداخل زبان مادری به وجود می‌آیند و به آن تداخل منفی^۵ نیز می‌گویند.

ب) خطاهای درون‌زبانی: خطاهایی هستند که به دلیل تعمیم افراطی^۶ قواعد زبان مقصد توسط زبان‌آموز ایجاد می‌شوند.

پ) خطاهای مربوط به راهکارهای آموزشی: گاهی شیوهٔ تدریس یا مواد آموزشی نیز باعث بروز خطا می‌شود.

۱. S. P. Corder

۲. C. James

۳. error

۴. mistake

۵. negative transfer

۶. overgeneralization

در این پژوهش، درباره علت خطا بر مبنای تقسیم‌بندی ضیاء‌حسینی (Ziahosseiny, 1994: 116) دو نوع خطای میان‌زبانی و درون‌زبانی در نظر گرفته می‌شود. خطاهای میان‌زبانی، خطاهای ناشی از تداخل زبان مادری فارسی‌آموز است و خطاهای درون‌زبانی، ناشی از یادگیری نادرست یا ناقص زبان مقصد است.

۳.۲. اصول نگارش در دستور خط فارسی

همان‌گونه که در ابتدا اشاره شد، در این پژوهش ملاک ارزیابی ساختار نگارشی فارسی‌آموزان عرب‌زبان، کتاب «ویراست جدید دستور خط فارسی» مصوب فرهنگستان زبان و ادب فارسی (چاپ اول ۱۴۰۱) است. در ویراست جدید این کتاب، به حدود ۲۸ قاعده درباره پیوسته‌نویسی و یا جدانویسی ترکیبات در زبان فارسی و همچنین املاهای صحیح برخی کلمه‌ها اشاره شده است. برخی از این قواعد که به پژوهش حاضر ارتباط دارد، به شرح زیر است:

- «این» و «آن» با فاصله کامل از کلمه پس از خود نوشته می‌شود. استثنا: آنچه، آنکه، اینکه، اینجا، آنجا، وانگهی.
- «همین» و «همان» همواره جدا از کلمه پس از خود نوشته می‌شود؛ مگر در مواردی که با کلمه پس از خود واژگانی شده باشد که در این صورت، با نیم‌فاصله نوشته می‌شود: همان‌جا و همین‌که.
- «هیچ» همواره جدا از کلمه پس از خود نوشته می‌شود؛ مگر در مواردی که با کلمه پس از خود واژگانی شده باشد که در این صورت، بی‌فاصله نوشته می‌شود: هیچ‌کدام و هیچ‌یک.
- «چه» با فاصله کامل از کلمه پس از خود نوشته می‌شود؛ مگر در چرا، چطور، چقدر، آنچه، چنانچه و چگونه.
- «را» در همه‌جا جدا از کلمه پیش از خود نوشته می‌شود؛ به‌استثنای: چرا در معنی «برای چه؟» و در معنای «آری»، در پاسخ به پرسش منفی.
- «ابن»/«بن» و «بنت»، وقتی بین دو علم (اسم خاص اشخاص) واقع شود، با فاصله کامل از کلمه پیش و پس از خود نوشته می‌شود: حسین بن علی و حسین ابن علی.
- در فاصله‌گذاری «به» سه حالت وجود دارد: الف) هنگامی که بر سر فعل یا مصدر بیاید و همچنین در صورت‌های واژگانی شده و تخفیف‌یافته؛ مانند: «بخرد»، «بنام» پیوسته نوشته می‌شود. ب) «به» در صفت‌ها، قیدها و حروف اضافه مرکب با نیم‌فاصله نوشته می‌شود؛ به‌خصوص (صفت و قید)، به‌سختی. ج) «به» در نقش حرف اضافه و فعل‌های گروهی (عبارت‌های فعلی) که حرف‌افزافه پیش از فعل است با فاصله کامل از اسم پس از خود نوشته می‌شود؛ به برادرت گفتم، به کار گرفتم.
- «بی» در نقش پیشوند نفی، همیشه با نیم‌فاصله از کلمه پس از خود نوشته می‌شود؛ بی‌تردید و بی‌قرار. استثنا: بیچاره، بیزار، بیدل و بینوا. مگر «بی» در نقش حرف‌افزافه.

- «می-» و «همی-» نسبت به جزء اصلی فعل همواره با نیم‌فاصله نوشته می‌شود.
- «هم» در ترکیب همواره، با نیم‌فاصله از کلمهٔ پس از خود نوشته می‌شود. استثناء: همسایه و همکف.
- «تر» و «ترین» همواره جدا از کلمهٔ پیش از خود نوشته می‌شود؛ مگر در: بهتر، مهتر، کهنتر، بیشتر، کمتر.
- «ها» (نشانهٔ جمع) در ترکیب با کلمات، به هر دو صورت پیوسته و جدا صحیح است؛ اما در مواردی جدانویسی الزامی است: بعد از واژه‌های بیگانه، برجسته‌سازی اصل کلمه برای آموزش، مختوم به کلمات پردندانه، جمع اسامی خاص و واژه‌های مختوم به «های» غیرملفوظ.
- نشانه کسره اضافه در خط آورده نمی‌شود؛ مگر برای رفع ابهام در کلماتی با دشواری تولیدی: اسب سواری / اسب‌سواری
- ۱- کلماتی مانند رهرو، پرتو، جلو، در حالت مضاف، گاهی با صامت میانجی «ی» می‌آید. مانند: پرتوی آفتاب و گاهی بدون آن؛ مانند: پرتو آفتاب. آوردن یا نیاوردن صامت میانجی «ی» تابع تلفظ خواهد بود.
- ۲- برای کلمات مختوم به «های» غیرملفوظ، در حالت مضاف، از علامت «ء» استفاده می‌شود: خانهٔ من، نامهٔ او.
- در خصوص واژه‌ها و ترکیبات و عبارات مأخوذ از عربی:
 - ۱- چنانچه «ة» در آخر کلمه تلفظ شود، به صورت «ت» نوشته می‌شود: رحمت و جهت؛ به‌استثنای صلوة، مشکوة، زکوة.
 - ۲- چنانچه «ة» در آخر کلمه تلفظ نشود، به صورت «های» غیرملفوظ نوشته می‌شود: علاقه و نظاره.
- آوردن تنوین (در صورتی که تلفظ شود)، در همهٔ متن‌ها الزامی است.
- گذاشتن تشدید ضرورت ندارد؛ مگر در هم‌نگاشت‌هایی که نگذاشتن تشدید، موجب ابهام در خواندن آن‌ها شود: معین / معین؛ بنا / بنا.
- تبصره: در متن‌های آموزشی، برای نوآموزان و غیرفارسی‌زبانان و نیز در اسناد و متن‌های رسمی دولتی، گذاشتن تشدید در همهٔ موارد ضروری است.
- ترکیبات دارای پسوند همیشه پیوسته نوشته می‌شود، مگر: حرف پایانی جزء اول با حرف آغازی جزء دوم یکسان باشد؛ مانند: نظام‌مند؛ کلمه به «های» غیرملفوظ ختم شود. مانند علاقه‌مند؛ ترکیب بسیط-گونه باشد مانند آبشار.
- ترکیباتی که همیشه با نیم‌فاصله / بی‌فاصله نوشته می‌شود: ترکیب‌هایی که با پیشوند یا پیشوندواره‌ها ساخته می‌شود؛ مگر ترکیب‌هایی با پیشوندهای «به»، «بی» و «هم»؛ ترکیب‌های عطفی در معنای مجازی مانند آب‌وهوا؛ عبارت‌های عربی که تبدیل به اصطلاح شده‌اند؛ مانند ان‌شاءالله؛ اجزای اعداد

بدون واو مانند یازده هزار؛ صفت‌های مرکب؛ مانند خوش‌برخورد.

- ترکیب‌هایی که همیشه با فاصله کامل نوشته می‌شود: ترکیب‌های اضافی هنگامی که کسره یا نشانه اضافه تلفظ شود؛ مانند دوست صمیمی. اگر کسره اضافه تخفیف یافته یا محذوف باشد با نیم‌فاصله بی-فاصله نوشته می‌شود؛ مانند تخم‌مرغ.

۲.۴. طبقه‌بندی خطاهای نگارشی فارسی‌آموزان عرب‌زبان در چارچوب «دستور خط فارسی»

مطابق پژوهش مدرس خیابانی (Modarress Khiyabani, 2017) در این پژوهش نیز با الگوگیری از دستور خط فارسی و اساساً شکل صحیح نوشتاری واژه‌ها، خطاها در شش گروه اشکال املائی/سبکی، درج فاصله، حذف فاصله، حذف نیم‌فاصله، سرهم‌نویسی و جدانویسی بررسی می‌شود که در اینجا به توصیف هریک از انواع این خطاها می‌پردازیم:

الف) اشکال املائی/سبکی: منظور از اشکال املائی این است که واژه موردنظر به لحاظ املائی نادرست باشد؛ به‌عنوان مثال «هواء» به جای «هوای» اشکال املائی محسوب می‌شود. منظور از اشکال سبکی این است که گرچه واژه موردنظر به لحاظ نوشتاری غلط نیست، به سبک آزمون موردنظر که در اینجا آزمون بسندگی زبان فارسی است، نامرتب است. به‌عنوان مثال؛ استفاده از واژه «ارزون» به جای «ارزان» اشکال سبکی است (Modarress Khiyabani, 2017).

ب) درج فاصله: در مواردی که نیاز به درج فاصله نبوده، ولی فارسی‌آموز فاصله گذاشته است؛ برای نمونه پیش از علائم سجاوندی؛ مانند «:» یا «.» نباید فاصله درج شود؛ در غیراین صورت خطای درج فاصله رخ داده است (Modarress Khiyabani, 2017).

پ) حذف فاصله: در این مورد به جای درج فاصله، حذف فاصله صورت گرفته است؛ به‌عنوان مثال: نوشتن «باعرض» به جای «با عرض» (Modarress Khiyabani, 2017).

ت) حذف نیم‌فاصله: بر اساس دستورالعمل فرهنگستان، برخی پسوندها و پیشوندها به کمک نیم‌فاصله، بعد یا قبل از واژه پایه نوشته می‌شود. در مورد واژه‌های مرکبی که اجزای آن‌ها باید جدا نوشته شود نیز صدق می‌کند. به کمک نیم‌فاصله، جدا بودن واژه مشخص می‌شود و همچنین بخش موردنظر، به‌عنوان بخشی از خود واژه نشان داده می‌شود (Modarress Khiyabani, 2017).

ث) سرهم‌نویسی: این نوع خطا به مواردی مربوط می‌شود که باید جدا نوشته شوند، اما سرهم نوشته شده‌اند؛ به‌عنوان مثال: پیشوند «می» باید جدا نوشته شود؛ درحالی‌که به‌صورت پیوسته «میروم» نوشته شده است (Modarress Khiyabani, 2017).

ج) جدانویسی: این خطا به مواردی اطلاق شده است که باید سر هم نوشته می‌شدند؛ ولی جدا نوشته

شده‌اند، برای نمونه: واژه «پنج‌شنبه» به جای «پنجشنبه» (Modarress Khiyabani, 2017). نکته مهم این است که ممکن است در یک نمونه، شاهد وقوع بیش از یک خطا باشیم. به‌عنوان مثال؛ هنگام نوشتن واژه «میره»، به جای «می‌روم» دو خطای اشکال سبکی و خطای پیوسته‌نویسی رخ داده است (Modarress Khiyabani, 2017). لازم به توضیح است که این پژوهش، شامل انشاهای فارسی‌آموزان عرب-زبان است و دارای ۳۲۰۹۶ نمونه است که از ۳۲۳۴ نوع^۱ تشکیل شده است. منظور از نمونه؛ واژه یا عدد یا هر نوع علامت سجاوندی است و منظور از نماد، وقوع مکرر یک نمونه است. برای مثال؛ واژه «دانشگاه» نمونه‌ای است که ۵۸۸ بار در متن مشاهده شده است. در چنین شرایطی می‌توان گفت؛ این نمونه ۵۸۸ نوع است.

۳. پیشینه پژوهش

پژوهش پیش‌رو، پژوهشی پیکره‌بنیاد است که به بررسی ساختار نگارشی فارسی‌آموزان عرب‌زبان سطح پیشرفته در چارچوب دستور خط فارسی می‌پردازد. بر این اساس، ضمن اشاره به پژوهش‌های صورت‌گرفته در زمینه خط فارسی، به پژوهش‌های صورت‌گرفته درباره پیکره و زبان‌شناسی پیکره‌ای در بحث خط فارسی نیز می‌پردازیم و در پایان به معرفی پژوهش‌های پیکره‌بنیاد درباره ساختار نگارشی فارسی‌آموزان غیرایرانی خواهیم پرداخت.

۳.۱. پژوهش‌های صورت‌گرفته در زمینه خط فارسی

درباره خط فارسی و مشکلات آن، از دیرباز پژوهش‌های متعددی صورت گرفته است. در بیشتر این پژوهش‌ها به مشکلات خط فارسی پرداخته شده و ساز و کارهایی برای رفع مشکلات آن ارائه شده است که از جمله می‌توان به آثار بهمنیار (Bahmanyar, 1942)، ناتل خانلری (Natel Khanlari, 1963)، فرشیدور (Farshidvar, 1971)، کاظمی موسوی (Kazemi Moosavi, 1990)، صدر امیرجانلو (Sadr Amirjanloo, 2002)، اعطار شرقی (Atar Sharghi, 2014)، پورنوروز و حسینی (Pornorouz & Hosseiny, 2018) اشاره کرد. با توجه به این اشکالات عده‌ای بر تغییر خط فارسی به لاتین تأکید داشته‌اند اما پژوهشگرانی بر اصلاح خط فارسی تمرکز کرده‌اند که از این میان می‌توان نظرات بهمنیار (Bahmanyar, 1942)، مدنی (Madani, 1976)، آشوری (Ashoori, 1986)، برجیان (Borjiyan, 1993)، به‌تویی (Behtooei, 2001)، ضیا (Zia, 2001)، آذرنگ (Azarang, 2004)، شعاعی (Shoaei, 2012)، سجودی (Sojoodi, 2012) و ذاکری (Zakeri, 2013) را نام برد. در این میان فرهنگستان زبان و ادب فارسی، به حکم وظیفه‌ای که بر حسب اساس‌نامه خود در پاسداری از زبان و خط فارسی بر عهده دارد، مجموعه قواعد و ضوابطی را با عنوان «دستور خط فارسی» در سال ۱۳۸۰ تصویب و ارائه نمود.

^۱. type

۲.۳. پژوهش‌های انجام شده درباره پیکره و زبان‌شناسی پیکره‌ای در بحث خط فارسی

در این بخش، هدف معرفی پژوهش‌های صورت‌گرفته درخصوص پیکره و زبان‌شناسی پیکره‌ای، در بحث خط فارسی است که علی‌رغم بررسی‌های انجام شده، به‌نظر می‌رسد؛ پژوهش‌های پیکره‌بنیاد معدودی در این زمینه انجام شده است. ستوده و هنرجویان (Sotude, & Honarjooyan, 2014) به بررسی تنوع و تأثیر آن بر جامعیت بازیابی اطلاعات در متون پیکره روزنامه همشهری پرداخته و میزان انطباق شیوه نگارش نگارندگان با دستور خط مصوب فرهنگستان و زبان و ادب فارسی مورد بررسی قرار گرفته است. نتایج پژوهش نشان می‌دهد؛ نگارندگان متون پیکره همشهری، تمایل به حذف یا جایگزینی نویسه‌های چالشی دارند. همچنین تحلیل‌های آماری حاکی از نداشتن انطباق گسترده الگوی نگارشی نگارندگان با دستور خط مصوب فرهنگستان زبان و ادب فارسی است. مدرس خیابانی (Modarress Khyabani, 2017) در مقاله‌ای با عنوان «آسیب‌شناسی نگارشی زیرنویس‌ها در شبکه‌های تلویزیونی خبر و آی‌فیلم» که بر اساس دستورالعمل نگارش مصوب فرهنگستان زبان و ادب فارسی انجام شده است، نشان می‌دهد که علاوه بر وجود انواع خطاهای نگارشی در هر دو شبکه، فراوانی خطاها در زیرنویس‌های شبکه آی‌فیلم بسیار زیاد است. بررسی میزان خطاهای نگارشی و مقایسه این نوع خطاها، در هر یک از دو شبکه نشان می‌دهد؛ میان انواع خطاها به لحاظ نوع و فراوانی، در بیشتر این جمله‌ها اختلاف معنی‌داری وجود دارد. خلیلی اردلی و همکاران (Khalili Ardali et al., 2017) وضعیت نگارش فارسی را در ۵۰ وبگاه پربیننده کشور، از منظر اصول ویرایش فنی به روش پیکره‌بنیاد بررسی کرده‌اند که اشتباهات شاخص در حوزه ویرایش فنی به ترتیب عبارتند از: دستور خط و هماهنگی در رعایت آن، نشانه‌گذاری‌ها، استقلال واژه، فاصله-گذاری‌ها، عددنویسی، غلط تایپی. همچنین به‌نظر می‌رسد که سه عامل آگاهی نداشتن، بی‌توجهی و استفاده نکردن از نرم افزارهای موجود ویراستاری، در تولید حجم بالای خطاهای نگارشی اثرگذار بوده‌اند. اسداللهی و آذرنیوار (Asadollahi & Azarnivar, 2022) به آسیب‌شناسی قانون مدنی از منظر دستور زبان فارسی و اصول نگارشی پرداخته‌اند. در این مسیر، از کتاب‌های فرهنگستان زبان و ادب فارسی، راهنمای نگارشی و ویرایشی و دیگر منابع مرتبط استفاده نموده‌اند. نتیجه پژوهش نشان داد؛ قانون مدنی به لحاظ زبانی، نگارشی و ویرایشی نیازمند بازنگری‌های دقیق زبانی است؛ زیرا علائم نگارشی در تمامی متون حقوقی و کتاب‌های قانون، به دلیل طولانی بودن بعضی از ماده‌ها که متشکل از چندین جمله است، نقش تعیین‌کننده‌ای در صدور حکم دارد و کوچک‌ترین اشتباه، موجب ابهام و عدم فهم صحیح قانون می‌شود.

۳.۳. پژوهش‌های پیکره‌بنیاد ساختار نگارشی فارسی‌آموزان غیرایرانی

اگرچه پیکره‌های زبانی می‌توانند سهم مهمی را در آموزش زبان دوم ایفا کنند؛ اما درخصوص ساختار نگارشی فارسی‌آموزان غیرایرانی، پژوهش‌های پیکره‌بنیاد محدودی انجام شده است. مطبوعی بنات (Matbooee

(Bonab, 2007) در پایان‌نامه خود به بررسی خطاهای نوشتاری زبان‌آموزان سطح مقدماتی انگلیسی‌زبان در یادگیری زبان فارسی پرداخته است. خطاهای نوشتاری این فارسی‌آموزان در چهار حوزه املایی، ساختواژی، نحوی و معنایی-واژگانی طبقه‌بندی شده است. نتایج تحقیق نشان می‌دهد که بیشترین تعداد خطا مربوط به املاء کلمات بوده است و حوزه‌های نحو، واژگانی-معنایی و ساختواژی به ترتیب در ردیف‌های بعد قرار گرفته است. مهم‌ترین عامل بروز خطا عامل انتقال درزبانی با ۵۴/۵ درصد بوده و انتقال میان‌زبانی فقط ۱۷/۵ درصد را به خود اختصاص داده است. نکته قابل توجه این است که تحقیق نامبرده، این فرضیه را که عمده‌ترین عامل بروز خطاهای نوشتاری در سطح مقدماتی تداخل زبان مادری است را رد می‌کند. احمدوند (Ahmadvand, 2010) تعداد ۱۴۸۱ مورد خطای نوشتاری فارسی‌آموزان آلمانی‌زبان سطح مقدماتی را تشخیص داده است و با بهره‌گیری از تقسیم‌بندی کشاورز (Keshavarz, 1999) این خطاها را در سه گروه املایی (واجی و خطی)، واژگانی - معنایی و دستوری طبقه‌بندی نموده است. در این پژوهش، بر مبنای تقسیم‌بندی ضیاء حسینی (Ziahosseiny, 1999) از دو نوع خطای درون‌زبانی و میان‌زبانی به‌عنوان منشأ تولید خطاها یاد شده است. نتایج این پژوهش حاکی از آن است که پربسامدترین خطاهای فارسی‌آموزان آلمانی‌زبان سطح مقدماتی، به ترتیب وقوع در حوزه‌های زیر تولید می‌شود: خطاهای واجی، خطاهای واژگانی-معنایی، خطاهای دستوری و خطاهای خطی. تداخل از زبان مادری، یکی از اصلی‌ترین عوامل در بروز خطا است. نکته قابل توجه آن است که این دستاورد برخلاف دستاورد به‌دست آمده از پژوهش مطبوعی بناب در حوزه فارسی‌آموزان انگلیسی‌زبان است که عمده‌ترین عامل بروز خطاهای نوشتاری در سطح مقدماتی را تداخل زبان مادری نمی‌داند.

طباطبایی (Tabatabaei, 2014) در پایان‌نامه خود با عنوان «تولید پیکره زبان‌آموز با برچسب‌گذاری خطاهای نوشتاری فارسی‌آموزان عرب‌زبان در سطح متوسط»، خطاهای انشا نویسی این گروه را تحلیل کرده است. نتایج تحقیق نشان داد؛ خطاهای نحوی، پربسامدترین و خطاهای واژگانی، کم‌بسامدترین خطاها هستند. در مقاله میردهقان و همکاران (Mirdehghan et al., 2014) نیز خطاهای املایی - واجی فارسی‌آموزان آلمانی‌زبان سطح مقدماتی مورد تحلیل قرار گرفته است. داده‌های تحقیق شامل ۸۵۰ خطای املایی - واجی مربوط به نویسه‌ها و حروف آن‌ها در واژگان است. این تحقیق نیز مشابه پژوهش احمدوند (Ahmadvand, 2010) نشان می‌دهد که زبان فارسی به‌عنوان زبان دوم، نسبت به تداخل زبان آلمانی به‌عنوان زبان مادری، تأثیر بیشتری در بروز خطاها دارد. متولیان و دهخدا (Motavallian Naeini & dekhoda, 2018) به بررسی انواع خطاهای املایی فارسی‌آموزان عرب‌زبان مدرسه جامع‌المصطفی‌العالمیه قم پرداخته‌اند. خطاهای

املائی سه سطح زبان آموزی؛ شامل خطاهای نظام واکه‌ای، خطاهای نظام همخوانی و خطاهای صوری شناسایی و دسته‌بندی شد. تحلیل خطاهای املائی این فراگیران نشان داد که از نظر مقوله، بیشترین میزان خطا در نظام واکه‌ای است. از لحاظ منشأ خطا، خطاهای میان‌زبانی بیشترین بسامد را دارند و خطاهای مبهم و درون‌زبانی در رتبه بعدی قرار گرفتند. ایجاد خطاهای بین‌زبانی بیشتر از جانب فارسی‌آموزان مبتدی صورت گرفته بود که با افزایش سطح آن‌ها، تعداد این خطاها کمتر شد. افشین‌پور و همکاران (Afshinpoor et al., 2018) در پژوهشی پیکره‌بنیاد به توصیف خطاهای هم‌نشینی واژگانی و بررسی عوامل دخیل در بروز این نوع خطاها در انشای فارسی‌آموزان عرب‌زبان دانشگاه مجازی جامعه‌المصطفی قم پرداخته‌اند. نتایج پژوهش نشان داد؛ عوامل مؤثر در بروز خطاهای هم‌نشینی واژگانی به ترتیب عبارتند از: به‌کارگیری راهبردهای یادگیری زبان دوم انتقال زبان مادری و تعمیم نابجا. سجادی و صحرايي (Sajjadi & Sahraei 2019) به بررسی توالی فراگیری بندهای موصولی از جنبه توصیفی و توضیحی بودن آن‌ها پرداخته‌اند. یافته‌های این پژوهش پیکره‌بنیاد، بیانگر آن است که فارسی‌آموزان حتی هم‌زمان با فراگیری بندهای هم‌پایه، به تولید بندهای ناهم‌پایه از نوع موصولی می‌پردازند و فارغ از جنسیت و ملیت، موصولی‌های توصیفی را پیش از موصولی‌های توضیحی فرامی‌گیرند و حتی پس از فراگیری، تمایل بیشتری در استفاده از بندهای موصولی توصیفی دارند. پژوهش میرزایی حصاریان و پولادستون (Mirzaei Hesariyan & Pooladestoon, 2020) با استفاده از پیکره‌های زبانی برگرفته از آزمون نگارش پایان دوره عمومی فارسی‌آموزان چینی انجام شده است. نتایج حاصل نشان می‌دهد که این دسته از فارسی‌آموزان، هنوز مهارت یا تمایل کافی برای تولید نوشتاری جمله‌های خوشه‌ای را آن‌گونه که برای جمله‌های هسته‌ای کسب کرده‌اند، ندارند. پژوهش محمدابراهیمی جهرمی و نیّری فلاح (Mohammed 2021) Ebrahimi Jahromi & Nayeri Fallah با هدف تعیین منبع خطاهای فعلی فارسی‌آموزان سطوح مبتدی، میانی و پیشرفته شکل گرفته است. نتایج این بررسی نشان می‌دهد؛ خطاهای فعلی فارسی‌آموزان در سطح مبتدی، از منبع برون‌زبانی نشأت می‌گیرد و در سطح میانی به راهبرد یادگیری و در سطح پیشرفته، به بافت یادگیری مربوط است. قابل‌ذکر است که این پژوهش، مشابه پژوهش هونگ (Hong, 2011) است.

۴. روش‌شناسی پژوهش

روش این پژوهش، رویکردی پیکره‌بنیاد است و مبنای تحلیل داده‌ها، بررسی ساختار نگارشی فارسی‌آموزان عرب‌زبان سطح پیشرفته در چارچوب ویراست جدید شیوه‌نامه «دستور خط فارسی» مصوب فرهنگستان زبان و ادب فارسی (۱۴۰۱) است. بر این اساس، ابتدا لازم است پیکره‌ای از انشاهای آماده کرده و سپس به کمک

ابزارهای رایانه‌ای، فراوانی واژه‌های موجود در تولیدات نوشتاری را استخراج کنیم و در نهایت، به بررسی مشکلات موجود در این تولیدات نوشتاری بپردازیم.

۱.۴. آزمودنی‌های پژوهش

شرکت‌کنندگان در این پژوهش، ۱۵۱ فارسی‌آموز خانم و آقای عرب‌زبان از کشور عراق هستند که در آزمون بسندگی زبان فارسی دانشگاه بین‌المللی امام رضا علیه‌السلام شرکت کرده بودند؛ تا میزان توانایی آن‌ها در کاربرد زبان فارسی برای ادامه تحصیل در مقطع کارشناسی ارشد مشخص شود. در این پژوهش، انشاهای این شرکت‌کنندگان از لحاظ مطابقت با اصول دستور خط فارسی، مورد تجزیه و تحلیل قرار گرفته است. بررسی‌های این پژوهش، فارغ از دو متغیر سن و جنسیت صورت گرفته است.

۲.۴. روش گردآوری داده‌های پژوهش

در گردآوری داده‌های پژوهش، از روش خودانگیزته^۱ استفاده شده است. احمدوند (Ahmadvand, 2010: 66) بیان می‌دارد که در این روش، از نوشتن آزاد و یا انشا استفاده می‌شود. محتوای آزمون نگارش فارسی‌آموزان این پژوهش به شرح ذیل بوده است: نگارش ۵ جمله در مورد یک تصویر، چهار خط در مورد فصل بهار، توصیف یکی از روزهای موردعلاقه در هفته، مقایسه آب و هوای شهر مشهد با شهر محل سکونت داوطلب، معرفی یک پارک یا دانشگاه بین‌المللی امام رضا علیه‌السلام.

با توجه به استخراج داده‌ها از پیکره زبانی در این پژوهش و استفاده از نرم‌افزار تحلیل پیکره‌ای ورداسمیت^۲، این پژوهش در زمره پژوهش‌های پیکره‌بنیاد قرار می‌گیرد. ابزار ورداسمیت یکی از بسته‌های نرم‌افزاری موجود برای محققان حوزه زبان‌شناسی پیکره‌ای است که در سال ۱۹۹۶ میلادی معرفی شد. این نرم‌افزار تحت ویندوز از مجموعه‌ای از ماژول‌های مختلف برای الگوهای جستجو در یک زبان استفاده می‌کند که آخرین نسخه آن (8.0.0) در این پژوهش مورد استفاده قرار گرفته است. سه بخش اصلی این نرم‌افزار؛ شامل فهرست واژه‌ها^۳، واژه‌های کلیدی^۴ و خطوط واژه‌نما^۵ است. در قسمت فهرست واژه‌ها که در این پژوهش مورد استفاده قرار گرفته است، فراوانی هر واژه به‌همراه درصد مربوط به آن و همچنین پراکندگی هر واژه در کل متن مشخص می‌شود. برای دستیابی به این هدف، در گام نخست انشاهای فارسی‌آموزان حروف‌چینی شد. در مسیر حروف‌چینی با مشکل بزرگ ناخوانا بودن خط فارسی‌آموزان روبرو بودیم. تلاش شد تا حد امکان، مطالب به همان‌گونه که

۱. self motivated

۲. Wordsmith

۳. wordlist

۴. keywords

۵. concordance

فارسی‌آموزان در برگه‌های خود نوشته‌اند، تایپ شود. در مرحله بعد، به‌منظور امکان اعمال پردازش رایانه‌ای، فرمت فایل حروف‌چینی شده از طریق نرم‌افزار اسپوز تحت وب^۱ به فایل متنی^۲ تبدیل شد. پس از آن، با استفاده از نرم‌افزار تحلیل پیکره‌ای ورد اسمیت، لیست واژه‌ها تهیه شد و فراوانی هر واژه به همراه درصد آن و همچنین پراکندگی هر واژه در کل متن مشخص شد. سپس این واژه‌ها و فراوانی آن‌ها مورد تحلیل قرار گرفت. در جدول زیر نمونه‌هایی به‌عنوان مثال ذکر شده است.

جدول ۱: بخشی از لیست واژگان پیکره در نرم‌افزار ورداسمیت

ردیف	واژه	فراوانی	درصد فراوانی	میزان پراکندگی هر واژه
۱	و	۱۹۲۲	۵٫۹۷٪	۰٫۹۶
۲	است	۱۴۱۹	۴٫۴۱٪	۰٫۹۵
۳	در	۱۳۳۲	۴٫۱۴٪	۰٫۹۵
۴	خیلی	۹۶۲	۲٫۹۹٪	۰٫۹۴
۵	دانشگاه	۵۸۸	۱٫۸۳٪	۰٫۹۴
۶	شهر	۵۷۳	۱٫۷۸٪	۰٫۹۵
۷	من	۵۳۵	۱٫۶۶٪	۰٫۹۴
۸	مشهد	۴۸۲	۱٫۵۰٪	۰٫۹۷
۹	فصل	۴۷۶	۱٫۴۸٪	۰٫۹۲
۱۰	امام	۴۴۴	۱٫۳۸٪	۰٫۹۴

پیکره استفاده شده در این پژوهش که شامل انشاهای فارسی‌آموزان عرب‌زبان می‌شود، دارای ۳۲۰۹۶ نمونه است که از ۳۲۳۴ نوع تشکیل شده است. نکته قابل توجه این است که هر نوع مورد بررسی، بر اساس فاصله درج شده در دو طرف آن، از دیگر صورت‌ها متمایز شده است. بنابراین موردی مثل نشانه جمع، چنانچه با فاصله نوشته شده باشد، دو نوع و اگر بدون فاصله باشد، یک نوع در نظر گرفته شده است: (بچه‌ها / بچه‌ها). نکته دیگر این است که نوع محاسبه شده، تنها مربوط به واژگان می‌شود و علائم سجاوندی را در بر نمی‌گیرد. به‌منظور تحلیل داده‌های مستخرج از این پیکره، از آنجا که احتمال می‌رود، واژه‌هایی با فراوانی یک یا دو بار به‌صورت تصادفی شکل گرفته باشند، این موارد از داده‌های موردبررسی حذف شد.

۵. ارائه و واکاوی داده‌ها

همان‌طور که پیش‌تر نیز گفته شد، در این پژوهش، مطابق پژوهش مدرس خیابانی (Modarress Khiyabani, 2017)، با الگویی از دستور خط فارسی و اساساً شکل صحیح نوشتاری واژه‌ها، خطاها در شش گروه خطای

^۱. www.aspose.app

^۲. txt

املایی/سبکی، درج فاصله، حذف فاصله، حذف نیم‌فاصله، سرهم‌نویسی و جدانویسی بررسی خواهد شد که در اینجا به توصیف و تجزیه و تحلیل انواع این خطاها که توسط فارسی‌آموز در مهارت نگارش آزمون بسندگی صورت گرفته است، خواهیم پرداخت. درخصوص علت خطا نیز بر مبنای تقسیم‌بندی ضیاء‌حسینی، دو نوع خطای میان‌زبانی و درون‌زبانی در نظر گرفته می‌شود.

۱.۵. خطای املایی در پیکره

باتوجه به این که منظور از خطای املایی این است که واژه موردنظر به لحاظ املایی نادرست باشد، خطاهای املایی فارسی‌آموزان از پیکره استخراج شد که در دو حوزه خطاهای واجی (همخوان و واکه‌ای) و خطاهای رسم‌الخطی قابل بررسی است.

۱.۱.۵. خطاهای ناشی از تفاوت‌های واجی دو زبان

گاهی زبان‌آموز به دلیل نبود واج‌هایی در زبان اول، دچار خطاهایی در هنگام تلفظ و نوشتن کلمات می‌شود. در این موارد، زبان‌آموز از یکی از واج‌های زبان مادری یا واجی شبیه آن استفاده می‌کند. خاتمی و همکاران (Khatami et al., 2019) با مقایسه نظام واجی زبان‌های عربی و فارسی دریافتند که چهار همخوان با صدای /tʃ/, /p/, /ʒ/, /g/ که معادل نویسه‌های «چ»، «پ»، «ژ» و «گ» است؛ در الفبا و زبان رسمی عربی وجود ندارد. بنابراین فارسی‌آموز عرب‌زبان ممکن است، در برخورد با چنین واج‌هایی که در زبان عربی وجود ندارد، از واج‌هایی استفاده کند که از نظر جایگاه و شیوه تولید نزدیک به واجی در زبان عربی باشد. جدول زیر شاهدهی بر این امر است که نشان می‌دهد، فارسی‌آموز عرب‌زبان از واج /b/ به جای /p/، از واج /k/ به جای /g/ و از واج /dʒ/ به جای /tʃ/ استفاده نموده است.

جدول ۲: ده نمونه پرسیامد اول از خطاهای املایی به علت تفاوت‌های واجی دو زبان

ردیف	خطای املایی	فراوانی	صورت صحیح
۱	بارک	۱۵۹	پارک
۲	دانشگاه	۶۴	دانشگاه
۳	کرم	۳۸	گرم
۴	بزرک	۳۸	بزرگ
۵	قشنگ	۲۰	قشنگ
۶	کریه	۱۷	گریه
۷	بارگ	۹	پارک
۸	دریاچه	۸	دریاچه
۹	توب	۷	توپ
۱۰	بیاده	۷	پیاده

نکته مهم دیگر این است که تمامی این خطاها در دسته خطاهای املائی قرار می‌گیرند. در اینجا با ۲۳ نوع خطا و ۴۲۱ نمونه با منشأ میان‌زبانی روبرو هستیم که این فراوانی تقریباً زیاد قابل توجه است.

از طرف دیگر، گاهی با وجود اینکه واج موردنظر فارسی‌آموز در زبان فارسی وجود دارد، او واجی را جایگزین واج صحیح در زبان دوم می‌کند. واژه «استاد» با ۶ نمونه خطای املائی میان‌زبانی، ناشی از جایگزینی نادرست همخوان است که صورت صحیح آن «استاد» است. به نظر می‌رسد، در اینجا فارسی‌آموز عرب‌زبان، معادلی را که در زبان عربی برای واژه «استاد» به کار می‌برد، از زبان مادری خود به زبان دوم وارد کرده است. بنابراین واژه «استاد» را به کار برده است و خطای قرض‌گیری صورت گرفته است.

گاهی فارسی‌آموز در هنگام نگارش واژه، همخوانی را به آن می‌افزاید و این عمل منجر به تولید خطای املائی می‌شود. استفاده از واژه «تمیز» به جای «تمیز» در مفهوم «نظافت و پاکی» با ۲۰ نمونه، خطای افزایش همخوان و از نوع میان‌زبانی محسوب می‌شود.

در هنگام نگارش واژه، گاهی فارسی‌آموز، همخوانی را از قلم می‌اندازد. در این حالت با خطای کاهش نادرست همخوان روبرو هستیم. دو واژه «رفیم» و «داشگاه»، با نمونه‌های ۴ و ۳ از این دست خطاهای املائی درون‌زبانی هستند که به عقیده زندی (Zandi, 1999) شنیداری است؛ اما به هر صورت خطای املائی محسوب می‌شود.

در ادامه به تجزیه و تحلیل خطاهای املائی فارسی‌آموزان عرب‌زبان در حوزه واژه‌ها می‌پردازیم. از آنجا که برای واژه‌های کوتاه در زبان فارسی نشانه‌ای نداریم، فارسی‌آموزان عرب‌زبان به اشتباه واژه‌های کوتاه و بلند را به جای هم استفاده می‌کنند. جایگزینی نادرست واژه بلند /a/ با واژه کوتاه /æ/ مانند استفاده از واژه «سختمان» به جای «ساختمان» با فراوانی ۱۳ مثالی از این نوع هست. همچنین جایگزینی نادرست واژه کوتاه /o/ با واژه بلند /u/ در واژه «مردوم» به جای صورت صحیح «مردم» با فراوانی ۱۶ مشاهده می‌شود. استفاده ناصحیح از واژه «خیل» به جای صورت صحیح «خیلی» مثالی از جایگزینی نادرست واژه بلند /i/ با واژه کوتاه /e/ با فراوانی ۶ است. علت بروز این خطاها تلفظ ناصحیح واژه قبل از نوشتن است. این خطاها که ریشه درون‌زبانی دارد، در جدول‌های زیر نشان داده شده است.

جدول ۳: خطاهای املائی براساس فراوانی به علت جایگزینی نادرست واژه بلند /a/ و واژه کوتاه /æ/

ردیف	خطای املائی	فراوانی	صورت صحیح
۱	سختمان	۱۳	ساختمان
۲	زیارات	۱۰	زیارت
۳	هو	۹	هوا
۴	مسافرات	۵	مسافرت
۵	هوی	۵	هوای
۶	نازدیک	۴	نزدیک
۷	بارف	۳	برف

جدول ۴: خط‌های املائی براساس فراوانی به علت جایگزینی نادرست واکه بلند /i/ و واکه کوتاه /e/

ردیف	خطای املائی	فراوانی	صورت صحیح
۱	خیل	۶	خیلی
۲	بی	۶	به
۳	مکنند	۵	می‌کنند
۴	تیرم	۴	ترم
۵	سیتا	۳	سه تا
۶	بنابرین	۳	بنابراین

۲.۱.۵. خط‌های رسم‌الخطی

کاربرد نادرست نشانه‌هایی مانند همزه، تنوین، مدّ و استفاده از کلمات مأخوذ از عربی توسط فارسی‌آموزان عرب‌زبان منجر به تولید خط‌های املائی شده است که در گروه خط‌های رسم‌الخط مورد بررسی قرار می‌گیرد. یکی از خط‌های املائی صورت‌گرفته، به‌دلیل درج همزه است. در ۴۱ مورد واژه «هوا» به‌صورت «هواء» و در ۲۱ مورد واژه «است» به‌صورت نادرست «أست» و با همزه نوشته شده است. دلیل این خطای میان‌زبانی آن است که فارسی‌آموز الگوی نوشتار «الف» در آغاز و پایان واژه در زبان عربی را، در نوشتار زبان فارسی نیز اعمال می‌کند.

مطابق دستور خط فارسی، آوردن تنوین (در صورتی که تلفظ شود)، در نوشته‌های رسمی و نیز در متون آموزشی الزامی است. با توجه به این که پیکره ما بر اساس مهارت نگارش آزمون بسندگی زبان فارسی که آزمونی رسمی است، تهیه شده است؛ آوردن تنوین در نوشتار الزامی است. در موارد زیر، فارسی‌آموزان عرب‌زبان، نشانه تنوین را درج نکرده‌اند که می‌توان نتیجه گرفت؛ چون تنوین خاص زبان عربی است و عرب‌زبانان بدون گذاشتن نشانه تنوین، آن را به‌آسانی تلفظ می‌کنند، فارسی‌آموزان ضرورت درج آن را احساس نکرده‌اند. این خطا در دسته خط‌های میان‌زبانی قرار می‌گیرد.

جدول ۵: خط‌های املائی براساس فراوانی به علت عدم درج نشانه تنوین

ردیف	خطای املائی	فراوانی	صورت صحیح
۱	مثلا	۱۹	مثلاً
۲	معمولا	۱۱	معمولاً
۳	تقریباً	۱۰	تقریباً
۴	کلا	۳	کلاً
۵	قبلا	۳	قبلاً
۶	اصلا	۳	اصلاً

تنها حروفی که در زبان عربی نشانه مدّ می‌گیرند، حروف «ا، و، ی» هستند که به آن‌ها «حروف مدّی» گفته می‌شود که وابسته به اعراب قبل از خود هستند. بنابراین، فارسی‌آموزان عرب‌زبان با الگوگیری از زبان عربی

علامت مدّ را درج نمی‌کنند. از آنجا که منشأ تولید این خطا زبان فارسی است، در گروه خطاهای میان‌زبانی قرار می‌گیرد. در اینجا با ۱۶ نوع خطا و ۲۲۲ نمونه روبرو هستیم که این فراوانی نسبتاً زیاد، قابل‌توجه است. در جدول شماره (۶) خطاهای املائی ناشی از عدم نگارش نشانه مدّ آورده شده است.

جدول ۶: ده نمونه پربسامد اول از خطاهای املائی به علت عدم نگارش نشانه مدّ

ردیف	خطای املائی	فراوانی	صورت صحیح
۱	اب	۱۲۱	آب
۲	ان	۲۵	آن
۳	اباد	۱۳	آباد
۴	الوده	۱۱	آلوده
۵	اموزش	۹	آموزش
۶	امدم	۸	آمدم
۷	انها	۵	آنها
۸	آبر	۴	ابر
۹	افتابی	۴	آفتابی
۱۰	استان	۴	آستان

مطابق با دستور خطّ فارسی «ة» در واژه‌ها و ترکیب‌ها و عبارت‌های برگرفته از زبان عربی، اگر در آخر کلمه تلفظ شود، به صورت «ت» نوشته می‌شود و اگر در آخر واژه تلفظ نشود، به صورت «های» غیرملفوظ نوشته می‌شود. براساس داده‌های جدول، فارسی‌آموز عرب‌زبان، همان‌طور که واژه را در زبان عربی تلفظ می‌کند، عیناً نوشته است. منشأ خطا نیز میان‌زبانی است. در جدول زیر خطاهای فارسی‌آموزان در این حوزه آورده شده است.

جدول ۷: خطاهای املائی براساس فراوانی به علت نگارش واژه‌های مأخوذ از زبان عربی

ردیف	خطای املائی	فراوانی	صورت صحیح
۱	زیاره	۱۰	زیارت
۲	ساعه	۵	ساعت
۳	جمعة	۸	جمعه
۴	تربيه	۴	تربیت
۵	استراحة	۴	استراحت

۲.۵. خطای سبکی در پیکره

منظور از خطای سبکی این است که گرچه واژه موردنظر به لحاظ نوشتاری غلط نیست، اما با سبک آزمون رسمی موردنظر ما که در اینجا آزمون بسندگی زبان فارسی است، نامرتب است. در پیکره با ۳۲ نوع خطا و ۲۲۱ نمونه روبرو هستیم که این فراوانی نسبتاً زیاد قابل‌توجه است. تعدادی از خطاهای سبکی به ترتیب فراوانی شامل موارد ذیل بود که منشأ آن‌ها برون‌زبانی است.

جدول ۸: ده نمونه پربسامد اول از خطاهای سبکی

ردیف	خطای املایی	فراوانی	صورت صحیح
۱	زیاده	۲۹	زیاد است
۲	گرمه	۱۷	گرم است
۳	مهربون	۱۴	مهربان
۴	داره	۱۲	دارد
۵	سرده	۱۱	سرد است
۶	قشنگه	۱۰	قشنگ است
۷	اونجا	۱۰	آنجا
۸	دوس	۸	دوست
۹	هستش	۷	است
۱۰	بعله	۷	بلی

۳.۵. خطای درج فاصله در پیکره

در بررسی‌های انجام شده، دو نوع خطای درج فاصله در پیکرهٔ زبان آموز مشاهده شد که به دلیل درج فاصله پیش از علامت‌های سجاوندی «نقطه»، «ویرگول» و «دو نقطه» و همچنین درج فاصله بین واژه‌های مرکب است. فراوانی خطای حاصل از درج فاصله پیش از این علائم سجاوندی «نقطه»، «ویرگول» و «دونقطه» به ترتیب ۱۴۵۹، ۱۳۶ و ۵ مورد است که خطای مربوط به درج فاصله پیش از نقطه، قابل تأمل است. به نظر می‌رسد، این مسئله از بی‌دقتی و همچنین تعمیم نادرست فارسی‌آموز در خصوص لزوم درج فاصله بین واژه‌ها، نشأت گرفته است. علاوه بر این که از لحاظ زبان‌شناختی، بین اجزا تشکیل‌دهندهٔ واژه مرکب، فاصله درج نمی‌شود. در پیکره فارسی‌آموز این مسئله رعایت نشده است که جدول زیر بیانگر این امر است.

جدول ۹: خطای درج فاصله در پیکره

ردیف	خطای املایی	فراوانی	صورت صحیح
۱	نقطه	۱۴۵۹	-
۲	ویرگول	۱۳۶	-
۳	دونقطه	۵	-
۴	پدر بزرگ	۱۰	پدربزرگ
۵	دو شنبه	۹	دوشنبه
۶	شهر بازی	۶	شهربازی
۷	باغ وحش	۵	باغ‌وحش
۸	چهار شنبه	۴	چهارشنبه
۹	تمیز تر	۴	تمیزتر
۱۰	رود خانه	۴	رودخانه

۴.۵. خطای حذف فاصله در پیکره

در این نوع خطا، میان دو واژه مستقل هیچ نوع فاصله‌ای درج نشده است. به عنوان نمونه؛ به جای «رود بزرگ»، «رود بزرگ» نوشته شده باشد. در پیکره فارسی‌آموز، هیچ موردی از خطای حذف فاصله دیده نشد و فارسی‌آموزان به شدت تمایل به درج فاصله بین تمامی کلمات نوشتار خود داشتند.

۵.۵. خطای حذف نیم‌فاصله در پیکره

از آنجایی که نسخه‌های نوشتاری فارسی‌آموزان دست‌نویس بوده است، بررسی این نوع خطا دشوار است؛ در عین حال، از میان همه خطاهای بررسی شده در پیکره، خطای حذف نیم‌فاصله، از بیشترین فراوانی برخوردار است. برای نمونه؛ در واژه‌های «دانش‌آموز»، «ترم‌افزار» و «بین‌المللی»، عدم درج نیم‌فاصله باعث شده است که «آموز»، «افزار» و «المللی» هرکدام یک واژه جدا تلقی شود. ۱۵۸۳ نمونه از ۱۶ نوع این خطا نشان می‌دهد که فارسی‌آموز عرب‌زبان به کاربرد نیم‌فاصله در واژه‌های مرکب و همچنین پس از برخی وندهای تصریفی مانند «می-» یا «-ها» و یا «-های» توجهی ندارد. به نظر می‌رسد، فارسی‌آموزان عرب‌زبان نسبت به جدانویسی برخی عناصر واژگانی؛ همچون وندهای «-ها» یا «می-» آگاهی دارد؛ ولی از نیم‌فاصله به دلیل عدم آموزش و ناآگاهی استفاده نمی‌کند. جدول زیر نشان‌دهنده موارد ذکر شده است.

جدول ۱۰: ده نمونه پربسامد اول از خطای حذف نیم‌فاصله در پیکره

ردیف	خطای املائی	فراوانی	صورت صحیح
۱	می-	۸۰۸	-
۲	ها-	۴۳۰	-
۳	های-	۱۱۹	-
۴	بین المللی	۱۰۸	بین‌المللی
۵	علیه السلام	۳۹	علیه‌السلام
۶	خوش اخلاق	۱۴	خوش‌اخلاق
۷	وکیل آباد	۱۳	وکیل‌آباد
۸	سه شنبه	۱۰	سه‌شنبه
۹	نمی-	۸	-
۱۰	تربیت بدنی	۷	تربیت‌بدنی

۵.۶. خطای سرهم‌نویسی در پیکره

این نوع خطا به مواردی مربوط می‌شود که باید جدا نوشته شوند؛ اما سرهم نوشته شده‌اند. در جدول زیر خطاهای سرهم‌نویسی آورده شده است.

جدول ۱۱: ده نمونه پربسامد اول از خطای سرهم‌نویسی

ردیف	خطای املائی	فراوانی	صورت صحیح
۱	بعلت	۳۳	به علت
۲	میکنم	۱۸	می‌کنم
۳	میکنند	۱۲	می‌کنند
۴	بخاطر	۱۱	به خاطر
۵	بنام	۱۰	به نام
۶	میکنند	۱۰	می‌کنند
۷	میروم	۸	می‌روم
۸	بزرگترین	۷	بزرگ‌ترین
۹	بنظر	۷	به نظر
۱۰	میروند	۴	می‌روند

همان‌طور که در جدول شماره (۱۱) مشاهده می‌شود، خطای پیوسته‌نویسی واژه «به‌علت» به‌صورت «بعلت» دارای بیشترین فراوانی است. طبق دستور خط فارسی «به» فقط در صورتی که «بای زینت» یا «صفت‌ساز» باشد، به‌صورت پیوسته نوشته می‌شود؛ در غیر این صورت باید جدا نوشته شود. همان‌طور که می‌دانیم، پیشوند فعلی «می-» همواره به‌صورت جدا از فعل پس از خود نوشته می‌شود. بنابراین فارسی‌آموز عرب‌زبان، درخصوص واژه‌ای مانند «میکنم» با ۱۸ نمونه دچار خطای پیوسته‌نویسی شده است.

۵.۷. خطای جدانویسی در پیکره

این خطا به مواردی اطلاق شده است که باید سر هم نوشته می‌شدند؛ ولی جدا نوشته شده‌اند. به‌طور مثال؛ صورت‌های نادرست «پنج شنبه» و «خوش حال» واژه‌هایی بسیط‌گونه هستند که مطابق دستور خط فارسی باید پیوسته نوشته شوند.

جدول ۱۲: خطای جدانویسی براساس فراوانی

ردیف	خطای املائی	فراوانی	صورت صحیح
۱	پنج شنبه	۲۷	پنجشنبه
۲	خوش مزه	۱۲	خوشمزه
۳	حساب داری	۷	حسابداری
۴	خوش حال	۴	خوشحال
۵	یک شنبه	۳	یکشنبه

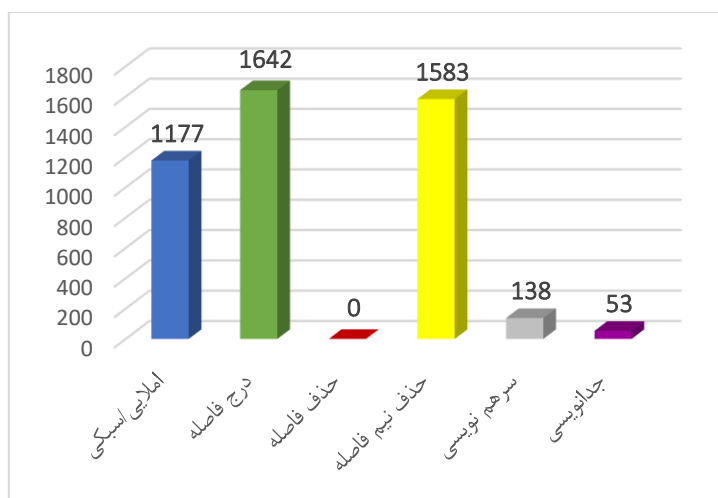
۵.۸. مقایسه فراوانی و منشأ خطاهای پیکره فارسی آموز

پس از بررسی هریک از خطاها در پیکره فارسی‌آموز، به مقایسه فراوانی نوع و نمونه خطاها در این پیکره می‌پردازیم که در جدول زیر نشان داده شده است.

جدول ۱۳: فراوانی نوع و نمونه انواع خطاهای نوشتاری فارسی‌آموزان عرب‌زبان

تعداد	نوع خطا	املائی / سبکی	درج فاصله	حذف فاصله	حذف نیم‌فاصله	سرهم- نویسی	جدانویسی	جمع کل اشکالات
نوع		۱۰۰	۱۰	۰	۱۶	۱۵	۵	۱۴۶
تعداد		۱۱۷۷	۱۶۴۲	۰	۱۵۸۳	۱۳۸	۵۳	۴۵۹۳

نتایج نشان می‌دهد که بیشترین خطا در حوزه درج فاصله است و حذف نیم‌فاصله، اشکالات املائی/ سبکی و خطای سرهم‌نویسی در رده‌های بعدی و همچنین جدانویسی در مرتبه آخر قرار دارد و در نهایت هیچ مورد اشتباه در حوزه حذف فاصله به چشم نمی‌خورد. در شکل شماره (۱) مقایسه بسامد خطاهای نوشتاری مشاهده می‌شود.



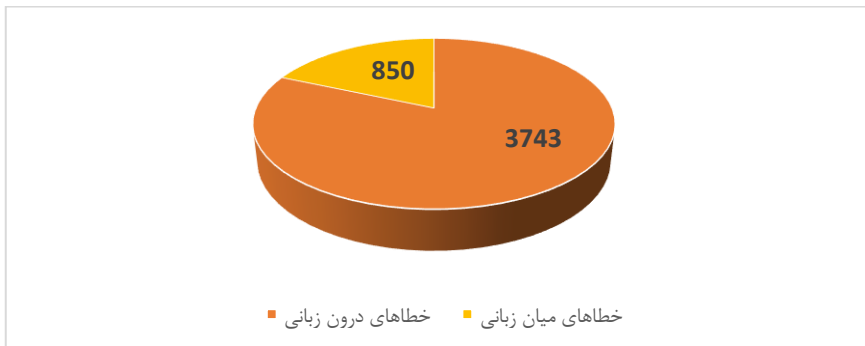
شکل ۱: مقایسه فراوانی خطاهای نوشتاری فارسی‌آموزان عرب‌زبان

در این پژوهش در کنار تقسیم‌بندی‌های صورت گرفته، منشأ خطاهای نوشتاری در دو گروه میان‌زبانی و درون-زبانی به شرح زیر بررسی شد.

جدول ۱۴: منشأ خطاهای نوشتاری (درون‌زبانی / میان‌زبانی)

منشأ خطا	نوع خطا	
میان‌زبانی	خطا در همخوان	خطای املايي ناشی از نبود یک واج در زبان مادری
میان‌زبانی		خطای املايي جایگزینی نادرست همخوانی دیگر
میان‌زبانی		خطای املايي افزایش نادرست همخوان
درون‌زبانی		خطای املايي کاهش نادرست همخوان
درون‌زبانی و میان‌زبانی	خطا در واکه	خطای املايي جایگزینی نادرست واکه بلند /a/ با واکه کوتاه /æ/
میان‌زبانی		خطای املايي جایگزینی نادرست واکه بلند /u/ و واکه کوتاه /o/
میان‌زبانی		خطای املايي جایگزینی نادرست واکه بلند /i/ با واکه کوتاه /e/
میان‌زبانی	خطاهای املايي رسم الخطی	خطای املايي ناشی از نگارش نادرست همزه
میان‌زبانی		خطای املايي ناشی از عدم نگارش تنوین
میان‌زبانی		خطای املايي ناشی از عدم نگارش نشانه مدّ (نویسه «آ»)
میان‌زبانی		خطای املايي ناشی از نگارش واژه‌های مأخوذ از زبان عربی
درون‌زبانی	-	خطای سبکی
درون‌زبانی	-	خطای درج فاصله
-	-	خطای حذف فاصله
درون‌زبانی	-	خطای حذف نیم‌فاصله
درون‌زبانی	-	خطای سرهم‌نویسی
درون‌زبانی	-	خطای جدانویسی

بر اساس اطلاعات جدول شماره (۱۴) و نمودار زیر، منشأ خطاها در دو گروه میان‌زبانی و درون‌زبانی مشخص شده‌اند. آمار نشان می‌دهد؛ خطاهایی با منشأ درون‌زبانی که خطاهای یادگیری و رشد زبان دوم بودند، سهم بسزایی در بروز خطاهای نگارشی داشته است.



شکل ۲: توزیع خطاهای نوشتاری فارسی‌آموزان برحسب منشأ خطاها

۶. نتیجه‌گیری و پیشنهادهای آموزشی و پژوهشی

در این بخش، به تبیین برخی مشکلات نوشتاری و ارائه پیشنهادهایی برای کاهش خطاهای نوشتاری خواهیم پرداخت.

خطاهای نگارشی فارسی‌آموزان عرب‌زبان سطح پیشرفته در چارچوب «دستور خط فارسی» و براساس پژوهش پیکره‌بنیاد مدرس خیابانی (Modarress Khiyabani, 2017) در شش گروه، اشکال‌های املائی/سبکی، درج فاصله، حذف نیم‌فاصله، سرهم‌نویسی و جدانویسی دسته‌بندی شدند که نتایج نشان می‌دهد، بیشترین خطا در حوزه درج فاصله است و خطای حذف نیم‌فاصله، اشکالات املائی/سبکی و خطای سرهم‌نویسی در رده‌های بعدی قرار می‌گیرد. همچنین خطای جدانویسی در مرتبه آخر قرار دارد و درنهایت هیچ مورد اشتباه در حوزه خطای حذف فاصله به چشم نمی‌خورد.

نتایج نشان می‌دهد که بیشترین فراوانی خطا در حوزه‌های درج فاصله و حذف نیم‌فاصله وجود دارد که به نظر می‌رسد، با توجه به عدم آموزش مسائل مربوط به عدم درج فاصله پیش از علائم سجاوندی و نیز درج نیم‌فاصله به فارسی‌آموزان عرب‌زبان، بروز این خطا امری بدیهی باشد. بسامد بالای خطاهای املائی/سبکی فارسی‌آموزان عرب‌زبان سطح پیشرفته نشان می‌دهد، علاوه بر دقت بیشتر در بحث آموزش نکات، لازم است، تمرین‌های آموزشی بیشتری مورد استفاده قرار گیرد. همان‌طور که جیمز (James, 1998) اهمیت نوع خطاها را براساس معیارهای مختلف؛ مانند بسامد، ارتباط، درک و غیره می‌سنجد. اگرچه خطاهای حذف نیم‌فاصله از فراوانی بیشتری نسبت به خطاهای املائی/سبکی برخوردار است، اما خطاهای املائی/سبکی در حوزه‌های ارتباط و درک از اهمیت بیشتری در مقایسه با خطاهای حذف نیم‌فاصله برخوردار است. خطاهای مربوط به دو حوزه سرهم‌نویسی و جدانویسی، در جایگاه سوم فراوانی قرار دارد که باتوجه به پیچیدگی مباحث مربوطه، به نظر می‌رسد؛ نیاز به آموزش دقیق‌تر و تمرین بیشتر فارسی‌آموزان عرب‌زبان در حوزه‌های سرهم‌نویسی و جدانویسی است.

باتوجه به یافته‌های پژوهش، خطاهایی با منشأ درون‌زبانی که خطاهای یادگیری و رشد زبان دوم بودند، سهم بسزایی در تعیین منشأ بروز خطاها دارند.

باتوجه به نتایج این پژوهش و از آنجا که نظام نوشتاری زبان فارسی یک نظام الفبایی همخوانی است، به نظر می‌رسد آگاهی واج‌شناختی^۱ به‌عنوان یک رکن، در خواندن و نوشتن این زبان بسیار اثرگذار است. آگاهی واج-شناختی درواقع توانایی تشخیص صداها در کلمات است و از این جهت که در املا و انشانویسی لازم است، فارسی‌آموز پیش از نوشتن بتواند واژه را به صداها سازنده آن بخش کند و از حروف مناسب برای نوشتاری کردن آن صداها استفاده کند، می‌توان گفت؛ هر نوع ضعف در مهارت شنیداری صداها و هجانبندی، تأثیر فراوانی در عدم درست‌نویسی فارسی‌آموز خواهد داشت.

۱. Phonological awareness

متناسب با خط‌هایی که شناسایی شد، انتظار می‌رود؛ در منابع آموزشی فارسی‌آموزان بازنگری شود و در راستای کاهش خطاهای فارسی‌آموزان، تمرین‌ها و فعالیت‌های کلاسی مناسب طراحی شود. علاوه بر استفاده از منابع کمک‌درسی، انجام ارزیابی‌های منظم و دوره‌ای می‌تواند گامی اساسی در جهت تسهیل فرایند یادگیری و کاهش خطاهای آنها باشد.

فهرست منابع:

- آذرنگ، عبدالحسین (۱۳۸۳). سیاست‌های فرهنگی، آموزشی، و ارتباطی، و اصلاح خط فارسی. نامه پارسی. ۳(۹)، ۱۴-۵.
- آشوری، داریوش (۱۳۶۵). چند پیشنهاد درباره روش نگارش و خط فارسی. نشر دانش. ۳۶، ۸-۲.
- احمدوند، احمد (۱۳۸۹). تجزیه و تحلیل خط‌های نوشتاری فارسی‌آموزان آلمانی زبان سطح مقدماتی. پایان‌نامه کارشناسی ارشد. رشته زبان‌شناسی گرایش آموزش زبان فارسی به غیرفارسی‌زبانان. دانشکده ادبیات و علوم انسانی دانشگاه شهید بهشتی.
- احمدی، طاهره و همکاران (۱۴۰۰). ارائه روشی مناسب برای برچسب‌گذاری واج‌گونه‌ای پیکره‌های گفتاری بر اساس سیستم IPA. فصلنامه زبان‌پژوهی. ۱۳(۳۸)، ۲۱۲-۱۸۵.
- اسداللهی، خدابخش و آذرنیوار، لیلا (۱۴۰۱). آسیب‌شناسی نگارشی و ویرایشی قانون مدنی با تکیه بر دستور زبان فارسی. پژوهش‌های دستوری و بلاغی. ۱۱(۲۰)، ۳۱۵-۲۸۷.
- اعطار شرقی، نوید (۱۳۹۵). بررسی تحلیل نگرش‌های داخلی و خارجی به مسئله تغییر نظام خطی زبان فارسی: موردهای اورور فارسی، یونی پرس و پارسیک. جستارهای زبانی. ۳۴(۷)، ۱۷۴-۱۴۳.
- افشین‌پور، مریم و همکاران (۱۳۹۷). بررسی خط‌های هم‌نشینی واژگانی فارسی‌آموزان عربی زبان. علم زبان. ۸(۵)، ۳۰-۷.
- برجیان، حبیب (۱۳۷۲). تجربه تاجیکستان در تغییر خط فارسی. ایران‌شناسی. ۱۷، ۱۸۰-۱۷۲.
- به‌تویی، محمد (۱۳۸۰). شیوه خط فارسی یک دیدگاه. چیستا. ۱۸۲ و ۱۸۳، ۱۸۴-۱۷۴.
- بهمنیار، احمد (۱۳۲۱). املاهای فارسی پیشنهاد به مقام فرهنگستان. نامه فرهنگستان. ۴، ۶۶-۴۲.
- پروان، حسن و سرکار حسن‌خان (۱۳۹۷). کاربرد پیکره‌های متنی در آموزش زبان آلمانی. پژوهش‌های زبان‌شناختی در زبان‌های خارجی، ۲(۸)، ۴۴۹-۴۷۴.
- پورنوروز، مسعود و حسینی، عارف (۱۳۹۷). وضعیت خط فارسی و تغییرات احتمالی آن؛ با توجه به فضاسازی ذهنی، رخسار زبان، ۶ و ۷، ۱۰۵-۱۲۱.
- خاتمی، فاطمه و همکاران (۱۳۹۸). مقایسه‌ی خطای درج واکه و همخوان در گفتار فارسی‌آموزان عرب‌زبان عراقی متوسطه و پیشرفته. نخستین همایش ملی بین‌رشته‌ای ایران‌شناسی، زبان‌شناسی و مطالعات ترجمه. دانشگاه ولی عصر رفسنجان. ۳۱۴-۳۳۴.
- خلیلی اردلی، وحید و همکاران (۱۳۹۶). وضعیت نگارش فارسی در وب‌سایت‌های پربیننده کشور از منظر اصول ویرایش فنی، فصلنامه پژوهش زبان و ادبیات فارسی، ۴۶، ۹۸-۶۹.
- ذاکری، مصطفی (۱۳۹۲). درباره تغییر خط فارسی. کتاب ماه ادبیات. ۱۹۳، ۶۱-۵۴.

- زندى، بهمن (۱۳۷۸). آموزش مکالمه فارسی. تهران: دفتر پژوهش و برنامه‌ریزی.
- ستوده، هاجر و هنرجویان، زهره (۱۳۹۳). بررسی تنوع الگوهای نگارش فارسی و تأثیر آن بر جامعیت بازیابی اطلاعات (مطالعه موردی: پیکره همشهری). نشریه کتابداری و اطلاع‌رسانی، ۶۶، ۴۹-۳۱.
- سجادی، شهره سادات و صحرائی، رضامراد (۱۳۹۸). بندهای موصولی توصیفی و توضیحی: نگاهی تحلیلی بر توالی فراگیری بندهای موصولی فارسی‌آموزان غیر ایرانی. پژوهش‌نامه آموزش زبان فارسی به غیرفارسی‌زبانان. ۸(۱۸)، ۱۳۶-۱۱۹.
- سجودی، فرزانه (۱۳۹۱). چرا نباید خط فارسی را تغییر دهیم. دریافت از وبگاه: ۹۱۱۱۱۴۰۷۷۸۶۷ <https://www.isna.ir/news/>
- شعاعی، مالک (۱۳۹۱). برون‌رفت‌هایی برای چالش‌های فراروی خط فارسی. فصلنامه زبان و ادب فارسی. ۱۱(۴)، ۱۰۳-۸۷.
- صدر امیرجانلو، اصغر (۱۳۸۱). کاستی‌های خط فارسی و پیامدهای آن در آموزش زبان فارسی. نشریه دانشکده ادبیات و علوم انسانی دانشگاه تهران. ۲(۴۹)، ۴۰۸-۳۹۳.
- صفری، سعید (۱۳۹۵). پیکره زبان‌آموز: مبانی، روش‌شناسی، الگوی طراحی و تولید، مجموعه مقالات دومین همایش ملی زبان‌شناسی پیکره‌ای. تهران: نشر نویسه‌پارسی، ۱۲۳-۹۳.
- ضیاء، اسماعیل (۱۳۸۱). شیوه خط فارسی. دیدگاهی دیگر. نشریه چیستا. ۱۹۰، ۷۹۷-۷۹۴.
- طاهری، خدیجه و گوهری، حبیب (۱۴۰۰). ساخت اطلاع و ساخت آغازگری در زبان فارسی، یک مطالعه پیکره-محور. فصلنامه زبان‌پژوهی. ۱۳(۳۸)، ۲۴۲-۲۱۳.
- طباطبایی، عزت (۱۳۹۲). تولید پیکره زبان‌آموز با برچسب‌گذاری خطاهای نوشتاری فارسی‌آموزان عرب‌زبان در سطح متوسط. پایان‌نامه کارشناسی ارشد. رشته آموزش زبان فارسی به غیرفارسی‌زبانان. دانشکده ادبیات و علوم انسانی دانشگاه فردوسی مشهد.
- عاصی، مصطفی (۱۳۷۹). از پیکره زبانی تا زبان‌شناسی پیکره‌ای. همایش زبان‌شناسی ایران.
- فرشیدورد، خسرو (۱۳۵۰). بحثی درباره خط فارسی و پیشنهادهایی برای یکسان‌کردن آن. وحید. ۹(۹)، ۱۳۳۰-۱۳۱۸.
- فرهنگستان زبان و ادب فارسی و ویراست جدید (۱۴۰۱). دستور خط فارسی. چاپ اول. تهران: نشر آثار.
- کاظمی موسوی، احمد (۱۳۶۹). گذری و نظری: پیشنهادی برای به‌کرد خط فارسی. ایران‌نامه. ۳۳، ۱۵۸-۱۵۱.
- محمدابراهیمی‌جهرمی، زینب و نیری فلاح، سیده نرجس (۱۴۰۰). منبع خطاهای فعلی فارسی‌آموزان سطوح مبتدی، میانی و پیشرفته. فصلنامه پژوهش‌های ادبی. ۱۸(۷۱)، ۱۵۰-۱۲۳.
- مدرس خیابانی، شهرام (۱۳۸۶). بررسی باهم‌آیی واژگانی در زبان فارسی. رساله دکتری رشته زبان‌شناسی. دانشکده ادبیات فارسی و زبان‌های خارجی دانشگاه علامه طباطبایی.
- مدرس خیابانی، شهرام (۱۳۹۶). آسیب‌شناسی نگارشی زیرنویس‌ها در شبکه‌های تلویزیونی خبر و آی‌فیلم: پژوهشی پیکره‌بنیاد. نشریه رسانه‌های دیداری و شنیداری. ۲۷، ۶۱-۳۲.
- متولیان، رضوان و دهخدايي، زهرا (۱۳۹۶). تحلیل خطاهای املائی عرب‌زبانان فارسی‌آموز، دومانه‌نامه جستارهای زبانی، ۴۲(۸)، ۲۳۳-۲۶۴.
- مدنی، میرحمید (۱۳۵۵). مسائل خط فارسی. نشریه کاوه. ۶۲، ۲۰-۵.

- مطبوعی بناب، مجید (۱۳۸۵). بررسی خط‌های نوشتاری سطح مقدماتی انگلیسی‌زبان در یادگیری زبان فارسی. پایان‌نامه کارشناسی ارشد. رشته آموزش زبان فارسی به غیر فارسی‌زبانان. دانشکده ادبیات و علوم انسانی دانشگاه شهید بهشتی.
- میردهقان، مهین‌ناز و همکاران (۱۳۹۳). خط‌های نوشتاری فارسی آموزان آلمانی‌زبان در سطح مقدماتی: خط‌های املایی - واجی، پژوهش‌نامه آموزش زبان فارسی به غیر فارسی‌زبانان، ۶(۳)، ۹۱-۱۱۶.
- میرزایی حصاریان، محمدباقر و پولادستون، حمیده (۱۳۹۹). واکاووی دستوری نوشتار فارسی آموزان چینی سطح فرومیانی (A2). تحلیلی بر پایه دستور مقوله و میزان. پژوهش‌نامه آموزش زبان فارسی به غیرفارسی‌زبانان، ۹(۲۰)، ۱۱۵-۱۳۶.
- ناتل خانلری، پرویز (۱۳۴۲). زبان‌شناسی و زبان فارسی. تهران: بنیاد فرهنگ ایران.
- هونگ، هان ساتی (۱۳۸۹). بررسی خط‌های نحوی ویتنامی‌زبانان در یادگیری زبان فارسی. پایان‌نامه کارشناسی ارشد. رشته آموزش زبان فارسی به غیرفارسی‌زبانان. دانشکده ادبیات فارسی و زبان‌های خارجی دانشگاه علامه طباطبایی.

References:

- Afshinpoor, M., et al., (2018). The Lexical syntagmatic errors among Arabic speaking Farsi learners. *Language Science*, 8 (5), 7-30. [In Persian]
- Ahmadi, T., et al., (2021). Providing a suitable method for allophonic labeling of speech corpuses of according to the IPA system. *Scientific Journal of Language Research*, 13 (38), 185-212. [In Persian]
- Ahmadvand, A., (2010). *An orthographic error analysis of German learners of Persian at the elementary level* (Master Dissertation). Shahid Beheshti university, Tehran, Iran. [In Persian]
- Asadollahi, Kh., & Azarnivar, L. (2022). Pathology of Writing and Editing of Civil Law texts based on Persian Grammar. *Rhetoric and grammar studies*, 11(20), 287-315. [In Persian]
- Ashoori, D., (1986). Some suggestions about writing method and Persian script. *Publication of knowledge*, 36, 2-8. [In Persian]
- Assi, M., (2000). *From language corpora to corpus linguistics*. Iranian Linguistics Conference. [In Persian]
- Atar Sharghi, N., (2014). Analytical review of internal and external approaches to the linear regime change in Persian: Avrvfarsy case, ionic inquiry Parseek. *Language Related Research*, 34 (7), 143-174. [In Persian]
- Azarang, A., (2004). Cultural, educational, and communication policies, and reforming the Persian script. *Persian Letter*, 3 (9), 5-14. [In Persian]
- Bahmanyar, A., (1942). Persian spelling, recommended to the Academy of Persian Language and Literature. *Nameh Farhangestan*, 4, 42-66. [In Persian]
- Baker, P. Hardie. A. & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Scotland: Edinburgh University Press.
- Behtooei, M., (2001). The style of Persian calligraphy. *Cheesta Magazine*, 182-183, 174-184. [In Persian]
- Borjiyan, H., (1993). Tajikistan's experience in changing the Persian script. *Iranology*, 17, 172-180. [In Persian]
- Brezina, V. & Gablasova, D. (2018). *The corpus method*. available at <https://eprints.lancs.ac.uk/>

- Corder, S. P.** (1967). *The Significance of Learners' Errors*. International Review of Applied Linguistics in Language Teaching, 5, 161-170.
- Farshidvar, Kh.,** (1971). A discussion about the Persian script and suggestions to standardize it, *Vahid*, 9 (9), 1318-1330. [In Persian]
- Hong, H.,** (2011). *An analysis of the syntactic errors of Vietnamese Speakers in learning Persian* (Master Dissertation). Allameh Tabatabaei university, Tehran, Iran. [In Persian]
- James, C.** (1998). *Errors in language learning and use: Exploring error analysis*. England: Pearson Education.
- Kazemi Moosavi, A.,** (۱۹۹۰). A suggestion to improve Persian script. *Iran nameh*, 33, 151-158. [In Persian]
- Keshavarz, M.H.** (1994). *Contrastive analysis and error analysis*. Tehran: Rahnama Publications.
- Khalili Ardali, V. et al.,** (2017). Persian Writing in the most viewed websites in Iran from a technical editing point of view. *Quarterly Journal of Reserch in Persian Language and Litreture*, 46, 69-98. [In Persian]
- Khatami, F. et al.,** (۲۰۱۹). Comparison of vowel and consonant insertion errors in speech of intermediate and advanced Iraqi students. *The first national interdisciplinary conference on Iran studies, linguistics and translation studies*, Valiasr University. 314-334.
- Madani, M.,** (1976). Persian script problems. *Kaveh Journal*, 62, 5-20. [In Persian]
- Matbooe Bonab, M.** (2007). *The analysis of written errors of English-speakers who lear Persian at the elementary level* (Master Dissertation). Shahid Beheshti university, Tehran, Iran. [In Persian]
- McEnery, T. & Wilson, A.** (2001). *Corpus Linguistics: An Introduction*, United Kingdom: Edinburgh University Press.
- McEnery, T. Xiao, R. & Tono. Y.** (2006). *Corpus-Based Language Studies*. London: Routledge.
- Mirdehghan, M. et al.,** (2014). Written Errors of German-speaking Learners of Persian at Elementary Level: An Orthophonemic Analysis. *Journal of Teaching Persian to Speakers of Other Languages*, 6(3), 91-116. [In Persian]
- Mirzaei Hesariyan, Mb., Pooladestoon, H.,** (2020). Grammatical Analysis of chinese Persian learner's Writings (A2): a Study Based on Category and Scale Grammar. *Journal of Teaching Persian to Speakers of Other Languages*, 9(20), 115-136. [In Persian]
- Modarress Khiyabani, Sh.,** (2007). *The investigation of lexical collocation in Persian* (PhD. Dissertation). Allameh Tabatabaei university, Tehran, Iran. [In Persian]
- Modarress Khiyabani, Sh.,** (2017). A pathology of orthography of subtitles at IRINN news network, and IFilm TV network: a Corpus-Based study. *Quarterly Scientific Journal of Audio-Visual Media*, 27, 32-61. [In Persian]
- Mohammed Ebrahimi Jahromi, Z. & Nayeri Fallah, N.,** (2021). The current errors source of beginner, intermediate and advanced Persian learners. *Literary Research*, 18 (71). 123-150. [In Persian]
- Motavallian Naeini R, dekhoda Z.** (2018). Spelling Error Analysis of Arab Learners of Persian Language. *Language Related Research*, 42 (8), 233-264. [In Persian]
- Natel Khanlari, P.,** (۱۹۶۳). *Linguistics & Persian language*. Tehran: Iranian Culture Foundation. [In Persian]
- Parvan, H. & Sarkar Hassankhan, H.,** (2018). Using Text Corpora in Teaching of

- German Language. *Journal of Foreign Language Research*, 2 (8), 449-474. [In Persian]
- Pornorouz, M., & Hosseiny, A.,** (2018). The status of the Persian script and its possible changes, According to mental spaces. *Scientific Journal for linguistic and literary criticism studies*, 6-7, 105-121. [In Persian]
- Sadr Amirjanloo, A.,** (2002). Shortcomings of Persian script and its consequences in teaching Persian language. *The Journal of the Faculty of Literatures and Humanities of Tehran University*, 2-3 (49), 393-408. [In Persian]
- Safari, S.,** (2016). Learner corpora: basics, methodology, design and production pattern. *The articles collections of the 2nd International Corpus Linguistics Conference*. Tehran: Nevis Parsi Publishing, 93-123. [In Persian]
- Sajjadi, SH., & Sahraei, R.,** (2019). Restrictive and Non-Restrictive Relative Clauses in Persian Language: Sequence of Acquisition by Non-Iranian Persian Learners. *Journal of Teaching Persian to Speakers of Other Languages*, 8 (18), 119-136. [In Persian]
- Shoaei, M.,** (2012). Some solutions for the challenges facing Persian language script. *Quarterly Journals of Persian Language and Literature*, 11 (4), 87-103. [In Persian]
- Sojoodi, F.,** (2012). *Why should we not change the Persian script*. Retrieved from <https://www.isna.ir/news/۹۱۱۱۱۴۰۷۷۸۶۷>. [In Persian]
- Sotude, H. & Honarjooyan, Z.,** (2014). The Study of Persian Writing Style Variations and their Impacts on Information Retrieval: The case of Hamshahri Corpus. *Library and Information Science Journal*, 66, 31-49. [In Persian]
- Tabatabaei, E.,** (2014). *An Analysis of Writing Errors of Arab Learners of Persian: Compiling a Persian Learner Corpus for Intermediate Arab Learners of Persian* (Master Dissertation). Ferdowsi university, Mashhad, Iran. [In Persian]
- Taheri, Kh., & Gowhari, H.,** (2021). Information structure and thematic structure of Persian language, a corpus-based analysis. *Scientific Journal of Research Language*, 13 (38), 213-242. [In Persian]
- The Academy of Persian Language and Literature,** (2022). *Persian Script Instructions*. Tehran: Asar publication.
- Zakeri, M.,** (2013). About changing the Persian script. *Mah-e-Adabiyat*. 193, 54-61. [In Persian]
- Zandi, B.,** (1999). *Persian conversation teaching*. Tehran: Research and planning office. [In Persian]
- Zia, E.,** (2001). The style of Persian calligraphy. *Cheesta Magazine*, 190, 794-797. [In Persian]