

Extended Abstract

Morphological Issues in Literary Translation: An Analytical Study of Persian-to-English Renderings

Mahsa Ardeshiri ¹ 

PhD Candidate in Linguistics, Department of Foreign Languages and Linguistics, Faculty of Letters and Humanities, Shiraz University, Shiraz, Iran.

mahsaardeshiri17@gmail.com

Jalal Rahimian ^{2*} 

Professor of Linguistics, Department of Foreign Languages and Linguistics, Faculty of Letters and Humanities, Shiraz University, Shiraz, Iran.

jrahimian@rose.shirazu.ac.ir

Rahman Sahragard ³ 

Professor of Applied Linguistics, Department of Foreign Languages and Linguistics, Faculty of Letters and Humanities, Shiraz University, Shiraz, Iran.

rsahragard@rose.shirazu.ac.ir

Review History: Received: 2026-04-20 Revise: 2026-05-16 Accepted: 2026-05-26	Abstract The present study examines morphological and lexical errors in the translation of literary texts from Persian into English by undergraduate translation students. Translation, as a process involving meaning transfer and structural reconstruction, requires a strong understanding of the lexical and morphological systems of both the source and target languages. Accordingly, this study seeks not only to identify the major patterns of errors but also to interpret the underlying causes of their occurrence. The sample consisted of 70 translation students selected through purposive sampling. The research data were collected from the translation of a Persian literary text, in which the translated sentences were compared with the original text, and all morphological and lexical errors were extracted and categorized. The results showed that 38% of the errors occurred in nouns and 61% in verbs, indicating significant challenges students face in transferring Persian verbal structures into English. In addition to the error analysis, semi-structured interviews were conducted with three expert instructors to gain insights into linguistic, pedagogical, and cognitive factors contributing to these errors. Findings from the interviews revealed that limited awareness of subtle semantic distinctions, differences in verbal inflection systems between the two languages, and overreliance on source-language structures are among the main causes of such errors. Overall, the findings of this study can inform the improvement of translation-training programs and the development of effective instructional strategies to enhance students' translation competence.	Keywords: Translation Error analysis in translation Linguistic errors Morphological errors Translation errors
<i>How to cite this article:</i> Ardeshiri, M., Rahimian, J. & Sahragard, R. (2026). Morphological Issues in Literary Translation: An Analytical Study of Persian-to-English Renderings. <i>Journal of Linguistics & Iranian Languages</i>, 11(2), 1-16. doi: https://doi.org/10.22099/jlil.2026.55951.1471		
 COPYRIGHTS ©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publisher.		

* Corresponding author

Introduction

Translation quality plays a crucial role in the localization of sciences and the advancement of knowledge (O'Brien, 2020; Kenny, 2022). Despite the remarkable growth in the number of translation graduates, many still face challenges in producing high-quality translations, highlighting the need for meaning-based and strategic pedagogical training (Pym, 2021; Li, 2020).

Among the existing challenges, morphological and lexical errors warrant particular attention, as they directly affect the accuracy of meaning transfer and the naturalness of the target text. Weaknesses in understanding the inflectional and morphological features of languages make it difficult to reproduce nominal and verbal structures accurately in English. A systematic analysis of these errors can reveal underlying linguistic weaknesses and the contributing factors behind them, thus paving the way for more effective instructional approaches.

Despite the importance of error analysis, much of the previous research has focused on syntactic errors or general theoretical perspectives (Arabi and Rostampour 2019; Yazdani, 2011), while more recent studies have mainly addressed machine translation (Sasani et al., 2023; Afrouz, 2022; Seifi, 2025). Consequently, morphological and lexical errors in students' translations remain underexplored.

Materials and Methods

Multiple studies indicate that students face persistent difficulties with morphological and inflectional features of language, particularly in selecting appropriate noun and verb forms. Inability to identify the semantic and grammatical functions of affixes, as well as cross-linguistic interference, is among the main sources of error. Research by Kabiri and Izadi (2025) and Ardeshiri and Zarafshan (2015) has shown that grammatical, lexical, and discourse-level errors reduce translation quality. Drawing on NaPhm's (2005) model, the present study examines nominal errors (plural marking, omission of plural marking, and possessive constructions) and verbal errors (verb formation and subject–verb agreement). By emphasizing context and communicative function, it also frames error analysis as a pedagogical tool in translator training.

The study employed a qualitative research design and was conducted in three stages: analysis of the translations produced by undergraduate students majoring in English translation, interviews with these students, and interviews with instructors. Morphological errors were categorized based on NaPhm's (2005) classification. Following Creswell (2008), the interviews were transcribed, coded, and organized into themes and subthemes. Two external raters reviewed and validated the coding and analysis to enhance the reliability and validity of the findings.

The aim of the present study was to identify and analyze morphological errors in translation students' English translations of Persian literary texts.

Results and Discussion

The primary objective of this study was to identify, describe, and analyze the morphological errors produced by senior translation students when translating literary texts from Persian into English. The analysis reveals that morphological errors remain a persistent linguistic challenge for students, suggesting that even at advanced academic levels, certain English morphological features are not fully internalized. Findings indicate that these errors cluster into two main categories—nominal and verbal—aligning with the results reported by Ansari and Seifi (2024) regarding interlanguage patterns in translation students.

Data analysis demonstrates that approximately 61% of morphological errors pertain to the verbal domain, while 38% involve nominal structures. Nominal errors primarily include the overgeneralization of plural markers (e.g., adding ‘-s’ to uncountable nouns), the omission of necessary plural markers, and the misuse of possessive forms. These errors often stem from overgeneralization and the transfer of Persian strategies to the target language. When attempting to represent concepts of “totality” or “plurality” found in Persian literary texts, students rely on familiar morphological tools, failing to account for the more restricted, rule-governed nature of English pluralization. This issue is particularly pronounced in literary translation, where abstract and metaphorical Persian concepts do not always correspond to English plural structures.

Furthermore, confusion regarding possessive markers, specifically the distinction between “its” and “it’s,” highlights a recurrent challenge. While students grasp the concept of possession, they struggle with the written form and orthographic realization. The absence of an equivalent apostrophe-based structure in Persian, combined with the cognitive load of the translation process, exacerbates these errors. Such mistakes illustrate that morphological challenges in advanced learners often arise from processing pressure and intralingual interference from a fundamental lack of competence.

Regarding verbal errors, two patterns dominated: incorrect verb formation and subject–verb agreement failure. Instances such as the use of ‘fined’ instead of ‘found’ illustrate the overgeneralization of regular verb formation rules. These errors are consistent with Selinker’s (1972) interlanguage theory, which posits that learners develop a unique linguistic system influenced by both the mother tongue and the target language. Difficulties with subject–verb agreement, particularly regarding the auxiliaries ‘be’ and ‘have’, remain persistent. This matches Lardiere’s (2017) findings, showing that even advanced learners may omit agreement markers during complex tasks with high cognitive demands, such as literary translation.

Finally, the study addresses the structural challenges within Iran’s translation education system. As noted by Afrouz (2022), a mismatch between the specialized training of instructors—often coming from TESOL or linguistics rather than translation studies as their primary specialization—and the practical needs of students contributes to these foundational gaps. This underscores a pressing need for professional development and a reassessment of pedagogical policies in translator training programs.

Conclusion

The primary aim of this study was to identify and analyze the morphological errors made by undergraduate translation students when translating from Persian into English. NaPham's (2005) framework was adopted as the main analytical model, allowing for a systematic categorization of errors and a more precise description of students' interlanguage patterns. The findings revealed that the highest concentration of errors occurred in the verbal domain, indicating persistent difficulties in selecting appropriate verb forms, mastering regular and irregular paradigms, and maintaining subject–verb agreement. In the nominal domain, recurrent problems included the overuse or omission of plural markers and the inappropriate use of possessive structures, suggesting that students struggle to transfer Persian nominal structures accurately into English, especially in contexts involving abstract or generic reference. Overall, the results show that morphological errors are not random or merely due to carelessness; rather, they emerge from a complex interaction of cross-linguistic transfer, structural properties of the target language, and cognitive factors such as processing load and reliance on simplification strategies. Interview analysis further indicated that weaknesses in source-text comprehension, insufficient awareness of structural differences between Persian and English, limited mastery of basic grammatical patterns, lack of specialized translation instructors, and curricular gaps all contribute to these errors, underscoring the need for pedagogical revision in translator training programs.

References

- Afrouz, H., & Ranjbar Mohammadi, R. (2022). Evaluating the quality of metaphor translation in Omar Khayyam's quatrains using House's model: A study on overt and covert translation. *Journal of Language and Translation*, 12(2), 59–71.
- Afrouz, M. (2022). Literary text-research from the perspective of literary translation studies: A critique of overt errors in literary text translation relying on linguistic text models. *Literary Text Research*, 26(94), 293–322.
- Ansari, N., & Seifi, T. (2024). Text and translation interaction in cultural reflection: A case study of *The Thorn and the Carnation* and Yahya Sinwar. *Translation Researches in Arabic Language and Literature*, (32), 289–322.
- Arabi, M., & Rostampour Molki, R. (2019). Translation errors of Arabic language and literature students and solutions for their removal relying on error analysis method. *Language Related Researches*, 10(5), 157–181.
- Ardehshiri, M., & Zarafshan, M. (2015). Students' causes of errors in translating pragmatic senses. *International Journal of English and Education*, 3(4), 238–253.
- Ardehshiri, M., Rahimi, M., & Foroutan, M. (2015). Students' errors in translating from Persian to English. *International Journal of Language Learning and Linguistics World*, 9(1), 233–244.
- Askari, M. I., Sattarian, A., & Ranjbar, M. (2019). Investigating covert and overt errors in machine translation of academic texts based on House's (2015) translation quality assessment model. *Journal of Language and Translation*, 11(3), 57–73.
- Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Sage.
- Freitag, M., Foster, G., Grangier, D., Ratnakar, V., & Tan, Q. (2021). Experts, errors, and context: A large-scale study of human evaluation for machine translation. *Transactions of*

- the Association for Computational Linguistics, 9, 1460–1474. https://doi.org/10.1162/tac1_a_00437
- Homayouni, S., & Fouladi, M. (2022). Analyzing the role of linguistic formulations and cultural domains in the translation of metaphorical collocations (from a cognitive linguistics perspective). *Translation Researches in Arabic Language and Literature* [in Persian], 12(27), 253–277.
- Jamalimanesh, A.R. (2009). Electronic sources: A glance at the ailing system of teaching translation in Iranian universities. <http://www.translationdirectory.com/articles/article2044.php> (30 December 2010).
- Kabiri, A., & Izadi, M. (2025). Error analysis models in translation and translator training in Iran. *Journal of Translation and Language Education*, 8(1), 45–68.
- Kenny, D. (2022). *Machine translation for everyone: Empowering users in the age of artificial intelligence*. Language Science Press.
- Lardiere, D. (2017). *Second language acquisition: A biolinguistic perspective*. Routledge.
- Li, S. (2020). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 70(3), 878–918.
- Lorscher, W. (1991). *Translation performance, translation process, and translation strategies: A psycholinguistic investigation*. Tübingen, Germany: Gunter Narr Verlag.
- Na, P. P. Q. (2005). Errors in the translation of topic-comment structures of Vietnamese into English. *Asian EFL Journal*, 7(3), 53–78.
- O'Brien, S. (2020). Translation, technology and the user. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 435–449). Routledge.
- O'Brien, S., & Ehrensberger-Dow, M. (2020). Machine translation and translator training: Evaluating the impact of automated tools on translation quality. *Journal of Specialised Translation*, 33, 1–20.
- Pournia, S., Chalak, A., & Heidari Tabrizi, H. (2025). Effective writing strategies among Iranian intermediate EFL learners: A Photovoice-based study. *Iranian Journal of English for Academic Purposes*, 14(4), 16–34.
- Pym, A. (2010). *Translation and text transfer: An essay on the principles of intercultural communication* (Rev. ed.). Intercultural Studies Group.
- Sasani, F., Farjah, M., Navabzadeh Shafiei, S., & Nafar, R. (2023). Updating the translation of cultural elements in literary translation relying on Newmark's theory. *Journal of Language Research*, 13(4), 579–594.
- Yazdani Farsad, M. (2011). *A study of cultural challenges in literary translation* [Unpublished master's thesis]. Tarbiat Modares University.

زبان‌شناسی و گویش‌های ایرانی

سال ۱۱، شماره ۲، پیاپی ۲۱ (تابستان ۱۴۰۵) شماره صفحات: ۱ - ۱۶

چالش‌های صرفی در ترجمه ادبی: مطالعه‌ای بر ترجمه فارسی به انگلیسی

مهسا اردشیری^۱، جلال رحیمیان^{۲*}، رحمان صحراگرد^۳

۱. دانشجوی دکتری زبان‌شناسی، بخش زبانهای خارجی و زبان‌شناسی، دانشکده ادبیات و علوم انسانی، دانشگاه شیراز، شیراز، ایران
۲. استاد زبان‌شناسی، بخش زبانهای خارجی و زبان‌شناسی، دانشکده ادبیات و علوم انسانی، دانشگاه شیراز، شیراز، ایران.
۳. استاد زبان‌شناسی کاربردی، بخش زبانهای خارجی و زبان‌شناسی، دانشکده ادبیات و علوم انسانی، دانشگاه شیراز، شیراز، ایران

چکیده

پژوهش حاضر به بررسی خطاهای صرفی در ترجمه متون ادبی از فارسی به انگلیسی توسط دانشجویان رشته مترجمی زبان انگلیسی می‌پردازد. ترجمه، به‌عنوان فرایندی که شامل انتقال معنا و بازآفرینی ساختار در زبان مقصد است، مستلزم درک نظام‌های ساختارهای هر دو زبان مبدأ و مقصد است. در این راستا، مطالعه حاضر می‌کوشد ضمن شناسایی الگوهای خطا، علل بروز آن‌ها را نیز تفسیر کند. نمونه تحقیق شامل ۷۰ دانشجوی رشته مترجمی زبان انگلیسی است که از طریق نمونه‌گیری هدفمند انتخاب شدند. داده‌های پژوهش از ترجمه یک متن ادبی فارسی گردآوری شدند؛ به این صورت که جملات ترجمه‌شده با متن اصلی مقایسه و خطاهای صرفی استخراج و طبقه‌بندی شدند. نتایج نشان داد که ۳۸ درصد از خطاهای صرفی در حوزه اسم و ۶۱ درصد در حوزه فعل رخ داده است که بیانگر چالش‌های دانشجویان در انتقال ساختارهای مربوط به فعل از زبان فارسی به انگلیسی است. علاوه بر تحلیل خطاها، برای تبیین علل آن‌ها، مصاحبه‌هایی نیمه‌ساختارمند با سه تن از استادان انجام شد تا دیدگاه‌های آنان درباره عوامل مؤثر بر خطاها گردآوری شود. یافته‌های این مصاحبه‌ها حاکی از آن است که تفاوت‌های نظام تصریف فعل در دو زبان، ناآگاهی از قواعد ساختارهای زبان مقصد، تعمیم‌بخشی افراطی و اتکای بیش از حد به ساختارهای زبان مبدأ از عوامل بروز خطا به شمار می‌روند. در مجموع، نتایج پژوهش حاضر می‌تواند در بهبود طراحی دوره‌های آموزش ترجمه و تدوین راهبردهای آموزشی مؤثر برای ارتقای آگاهی صرفی دانشجویان مورد استفاده قرار گیرد.

واژه‌های کلیدی:

ترجمه ادبی
تحلیل خطا در ترجمه
خطاهای زبانی
خطاهای صرفی
ترجمه فارسی به انگلیسی

تاریخچه مقاله: دریافت: ۳۱ فروردین ۱۴۰۵

بازنگری: ۲۶ اردیبهشت‌ماه ۱۴۰۵

پذیرش: ۵ خردادماه ۱۴۰۵

* آدرس ایمیل نویسنده مسئول: jarahimian@rose.shirazu.ac.ir

اردشیری، مهسا، رحیمیان، جلال و صحراگرد، رحمان (۱۴۰۵). چالش‌های صرفی در ترجمه ادبی: مطالعه‌ای بر ترجمه فارسی به انگلیسی. زبان‌شناسی و گویش‌های ایران، ۱۱(۲)، ۱-۱۶. <https://doi.org/10.22099/jlil.2026.55951.1471>

حق نشر ۲۰۲۶-نویسنده(گان). این یک مقاله با دسترسی آزاد است که تحت شرایط مجوز Creative Commons Attribution (CC BY-NC 4.0) منتشر شده است، استفاده، توزیع و تکثیر نامحدود را در هر رسانه‌ای مجاز می‌داند، مشروط بر اینکه به نویسندگان و منبع اصلی اشاره شود. هیچ اجازه‌ای از نویسندگان یا ناشر لازم نیست.

۱. مقدمه

گسترش تعاملات علمی، فرهنگی و تخصصی میان جوامع، جایگاه ترجمه را به‌عنوان یکی از ضرورت‌های بنیادین در تبادل دانش و دسترسی به منابع جهانی برجسته کرده است. ایران نیز برای همگامی با تحولات دانش‌محور معاصر، نیازمند تربیت مترجمان کارآمدی است که افزون بر تسلط زبانی، توان تحلیل متن و انتقال دقیق معنا را داشته باشند. کیفیت ترجمه در بومی‌سازی علوم و توسعه دانش نقش مهمی ایفا می‌کند (اوبراین^۱، ۲۰۲۰؛ کنی^۲، ۲۰۲۲). با وجود رشد چشمگیر فارغ‌التحصیلان مترجمی، بسیاری از آنان در ارائه ترجمه مطلوب با چالش‌هایی روبه‌رو هستند و این امر ضرورت آموزش‌های معنامحور و راهبردی را تقویت می‌کند (پیم^۳، ۲۰۲۱؛ لی^۴، ۲۰۲۰).

یافته‌های پژوهشی نیز بر ضعف‌های بنیادین در فرایند ترجمه دانشجویان تأکید کرده‌اند. اردشیری و زرافشان (۲۰۱۵) نشان دادند که بسیاری از دانشجویان در درک متن مبدأ، تشخیص عناصر فرهنگی و بازشناسی ویژگی‌های بافتی ضعف دارند و این امر به تولید ترجمه‌های لفظ‌گرا منجر می‌شود. همچنین، مشکلاتی همچون دایره واژگانی محدود و ضعف نگارشی از چالش‌های مهمی است که جمالی‌منش (۲۰۰۹) مطرح کرده است. این یافته‌ها نشان می‌دهد که بخشی از کاستی‌ها به ساختار سنتی آموزش ترجمه بازمی‌گردد.

در میان چالش‌های موجود، خطاهای صرفی/ساختواژی و واژگانی نقش ویژه‌ای دارند، زیرا بر دقت انتقال معنا و طبیعی‌بودن متن مقصد تأثیر مستقیم می‌گذارند. ضعف در فهم ویژگی‌های تصریفی و ساختواژی زبان‌ها موجب دشواری در بازآفرینی ساخت‌های اسمی و فعلی در زبان انگلیسی می‌شود. تحلیل نظام‌مند این خطاها می‌تواند الگوهای ضعف زبانی و عوامل مؤثر بر بروز خطا را آشکار کند و زمینه شکل‌گیری رویکردهای آموزشی کارآمدتر را فراهم آورد.

با وجود اهمیت تحلیل خطا، بخش قابل توجهی از پژوهش‌های داخلی بر خطاهای نحوی یا رویکردهای نظری عمومی تمرکز داشته‌اند (عسکری و همکاران، ۱۳۹۸؛ انصاری و سیفی، ۱۴۰۳) و مطالعات جدید نیز عمدتاً به ترجمه ماشینی پرداخته‌اند (یزدانی، ۱۳۹۰؛ ساسانی و همکاران، ۱۴۰۲؛ افروز، ۱۴۰۱). از این رو، خطاهای صرفی/ساختواژی و واژگانی در ترجمه دانشجویان کمتر واکاوی شده‌اند. پژوهش حاضر با هدف شناسایی و تحلیل این خطاها در ترجمه متون ادبی فارسی به انگلیسی توسط دانشجویان مترجمی انجام شده است.

۲. پیشینه پژوهش

پژوهش‌های متعددی از جمله برخی از منابعی که در پاراگراف قبل بیان شد، حاکی از آن هستند که دانشجویان در مواجهه با مسائل ساختواژی، تصریفی و انتخاب صورت‌های مناسب اسمی و فعلی با دشواری روبه‌رو هستند؛ به‌ویژه ناتوانی در تشخیص نقش‌های معنایی و دستوری وندها و تداخل زبانی از عوامل رایج بروز خطا به شمار می‌روند. لورش (۱۹۹۱) نیز سه دسته مشکل واژگانی، نحوی و میان‌رشته‌ای را در فرایند ترجمه شناسایی کرده است.

¹ S. O'Brien

² D. Kenny

³ A. Pym

⁴ S. Li

پژوهش‌هایی همچون انصاری و سیفی (۱۴۰۳) و رحیمی (۱۳۸۲) نشان داده‌اند که خطاهای دستوری، واژگانی و گفتمانی در افت کیفیت ترجمه نقش دارند و تفاوت‌های صرفی و ساختوازی می‌توانند فهم متون مبدأ و تولید متن مقصد را دشوار سازند. مطالعات جدیدتر نیز بر تحلیل هم‌زمان خطاهای صرفی و بررسی علل بین‌زبانی و درون‌زبانی آن‌ها تأکید کرده‌اند (اردشیری و همکاران، ۲۰۱۵؛ کبیری و ایزدی، ۱۴۰۴). مطابق با ریچاردز^۱ و همکاران (۱۹۷۴)، تحلیل خطا امکان سنجش میزان تسلط زبان‌آموز و کشف مشکلات رایج یادگیری را فراهم می‌کند.

در حوزه خطاهای واژگانی، افروز (۱۴۰۱) خطاهای ترجمه را ناشی از انتخاب ترکیبات نادرست و معادل‌های نامناسب می‌داند. پیم (۲۰۱۰) سه نوع لغزش، خطای نظام‌مند و خطای مهارتی را معرفی کرده است. سگینوت^۲ (۱۹۹۰) بر نقش خطاها در نمایاندن نحوه پردازش زبانی تأکید دارد. مطالعاتی مانند همایونی و فولادی (۱۴۰۱) نیز نشان داده‌اند که مواجهه با ساخت‌های دشوار می‌تواند موجب رشد مهارتی شود. پژوهش‌های اوبراین و ارنسبرگ‌ردو^۳ (۲۰۲۰) و گسپری و همکاران (۲۰۲۲) نیز اهمیت خطاهای مرتبط با انتقال معنا را برجسته کرده‌اند.

در مقایسه ترجمه انسانی و ماشینی، عسکری و همکاران (۱۳۹۸) نشان دادند که ترجمه انسانی در انتقال دقیق معنا برتری دارد. در تحلیل خطاهای بین‌زبانی، الجرف^۴ (۲۰۰۰) گزارش کرده است که ۶۲ درصد خطاها ناشی از تداخل زبانی‌اند. در حوزه متون ادبی نیز افروز و رنجبر محمدی (۱۴۰۱) و عربی و رستم پور (۱۳۹۸) به تفاوت‌های معنایی و سبک‌شناختی به‌عنوان منبع خطا اشاره کرده‌اند.

معیارهای کیفیت انجمن زبان‌شناسی رایانه‌ای شامل ابعاد معنایی، نحوی، واژگانی و انسجام است (فری‌تگ^۵ و همکاران، ۲۰۲۱). انجمن مترجمان آمریکا نیز فهرستی از خطاهای ترجمه ارائه می‌دهد و کاسماول^۶ (۱۹۹۵) نشان می‌دهد که ارزیابی خطا می‌تواند از دو منظر زبانی دستوری و ارتباطی کارکردی صورت گیرد. بر این اساس، مرور پیشینه نشان می‌دهد که هرچند خطاهای ترجمه از منظرهای گوناگون بررسی شده‌اند، تمرکز بر خطاهای صرفی و واژگانی در ترجمه ادبی فارسی به انگلیسی، به‌ویژه در ترجمه‌های دانشجویان مترجمی، همچنان نیازمند بررسی دقیق‌تر است.

۳. الگوی عملی کورد (۱۹۷۴) و چارچوب مفهومی نا (۲۰۰۵)

در پژوهش حاضر از دو الگوی نظری مکمل بهره گرفته شده است. الگوی عملی کورد^۷ (۱۹۷۴) چارچوبی شناخته‌شده برای تحلیل خطاست که پنج مرحله گردآوری داده، شناسایی، طبقه‌بندی، تبیین و ارزیابی خطا را در بر می‌گیرد. کورد (۱۹۷۴) تأکید می‌کند که خطاها ابزارهایی آموزشی و شناختی‌اند، زیرا راهبردهای ذهنی فراگیر را آشکار می‌کنند. این ویژگی‌ها موجب می‌شود الگوی او برای تحلیل خطاهای صرفی مترجمان مناسب باشد.

¹ J. C. Richards

² C. Seguinot

³ M. Ehrensberger

⁴ R. Al-Jarf

⁵ M. Freitag

⁶ P. Kussmaul

⁷ S. P. Corder

چارچوب نا^۱ (۲۰۰۵) نیز به تفکیک خطاهای ترجمه‌ای و خطاهای زبانی می‌پردازد و بر نقش بافت و کارکرد ارتباطی در تحلیل خطا تأکید دارد. این رویکرد، تحلیل خطا را از سطح صوری به سطح کاربردی گسترش می‌دهد و آن را به ابزاری مؤثر برای آموزش ترجمه تبدیل می‌کند.

۳-۱. خطاهای مربوط به مقوله اسم

خطاهای صرفی مربوط به اسم از خطاهای مهم در میان دانشجویان مترجمی زبان انگلیسی‌اند و عمدتاً به تفاوت‌های بنیادی نظام اسمی فارسی و انگلیسی بازمی‌گردند. در فارسی، نشانه جمع برای طیف گسترده‌ای از اسم‌ها قابل کاربرد است و نظام تمایز میان اسم‌های قابل‌شمارش و غیرقابل‌شمارش به صراحت زبان انگلیسی عمل نمی‌کند. همین تفاوت‌ها می‌تواند زمینه‌ساز بروز خطاهای بین‌زبانی شود.

۳-۱-۱. افزودن نادرست نشانه جمع

افزودن نشانه جمع S- به اسم‌های غیرقابل‌شمارش یا اسم‌هایی که در انگلیسی ذاتاً مفرد یا جمع محسوب می‌شوند، مانند *information* و *furniture* یکی از خطاهای رایج در میان دانشجویان است. ریشه این خطا اغلب به انتقال منفی از فارسی و نیز راهبرد ساده‌سازی ساختاری بازمی‌گردد. پژوهش‌ها نشان می‌دهند که تعمیم افراطی این قاعده در مراحل اولیه و میانی یادگیری انگلیسی بسیار شایع است.

۳-۱-۲. حذف نشانه جمع

در مقابل، حذف نشانه جمع در اسم‌های جمع نیز از خطاهای پرتکرار محسوب می‌شود. بر اساس «فرضیه نشان‌داری»، ساخت‌های بی‌نشان، مانند صورت مفرد اسم، آسان‌تر پردازش می‌شوند و زبان‌آموز در تولید زبان به این صورت‌ها گرایش دارد. به همین دلیل، حذف نشانه جمع S- یکی از خطاهای رایج نوشتاری زبان‌آموزان EFL است (دولای^۲، برت^۳ و کراشن^۴، ۱۹۸۲).

۳-۱-۳. استفاده نادرست از حالت ملکی

خطای دیگر، به‌کارگیری نادرست حالت ملکی در انگلیسی است. زبان فارسی مفهوم مالکیت را از طریق «اضافه» بیان می‌کند و فاقد نشانه صرفی معادل «'s» است؛ بنابراین، زبان‌آموزان در تعیین جایگاه و ساختار مالکیت در انگلیسی دچار سردرگمی می‌شوند. این خطا معمولاً نتیجه تداخل بین‌زبانی و همچنین پیچیدگی ساختارهای ملکی زبان انگلیسی است.

۳-۲. خطاهای مربوط به مقوله فعل

نظام فعلی زبان انگلیسی در مقایسه با فارسی، در حوزه‌های زمان، نمود و تطابق صرفی تفاوت‌های قابل توجهی دارد. بنابراین، خطاهای فعلی می‌توانند شاخصی مهم برای بررسی توانش دستوری زبان‌آموز به شمار آیند.

۳-۲-۱. ساختار نامناسب فعل

^۱ P. P. Q. Na

^۲ H. C. Dulay

^۳ M. K. Burt

^۴ S. D. Krashen

این خطا زمانی رخ می‌دهد که زبان‌آموز قادر به تولید صورت صحیح فعل مطابق با زمان، شخص یا نمود نباشد. استفاده نادرست از زمان دستوری، به‌کارگیری نادرست مصدر، یا شکل‌دهی نادرست افعال مرکب در این دسته قرار می‌گیرد. نظریه بین‌زبانی این خطاها را بخشی طبیعی از فرایند رشد زبان دوم می‌داند و نه الزاماً نشانه ضعف فردی.

۳-۲-۲. خطاهای تطابق فاعل و فعل

نقض تطابق فاعل و فعل، به‌ویژه در افعال «have و be»، از خطاهای رایج صرفی است. از آنجا که در فارسی اطلاعات مربوط به فاعل در خود فعل بازنمایی می‌شود، زبان‌آموزان ممکن است اهمیت نشانگرهای تطابق در انگلیسی را نادیده بگیرند. پژوهش‌ها نشان داده‌اند که این خطا حتی در میان زبان‌آموزان پیشرفته نیز پایدار باقی می‌ماند (لاردیر، ۲۰۱۷).

ترکیب الگوی عملی کورد (۱۹۷۴) و چارچوب مفهومی نا بنیانی دووجهی و مکمل برای تحلیل دقیق خطاهای صرفی فراهم می‌کند؛ به‌گونه‌ای که الگوی کورد (۱۹۷۴) با ارائه فرایندی نظام‌مند برای شناسایی، طبقه‌بندی و تبیین خطاها، امکان توصیف علمی و منسجم آن‌ها را فراهم می‌آورد. در مقابل، چارچوب نا با تمایزگذاری میان خطاهای جمله‌محور و متن‌محور و تأکید بر نقش بافت ارتباطی، به پژوهشگر اجازه می‌دهد خطاها را در چارچوب کارکردی و معنایی آن‌ها تفسیر کند. این تلفیق، مبنایی نظری و کاربردی برای پژوهش حاضر ایجاد می‌کند و زمینه‌ای مناسب برای تحلیل جامع و دقیق خطاهای صرفی دانشجویان در ترجمه متون ادبی فراهم می‌سازد.

۵. تحلیل داده‌ها

فرایند تحلیل داده‌ها در این پژوهش در سه مرحله اصلی انجام شد: تحلیل متون ترجمه‌شده دانشجویان، تحلیل مصاحبه‌های عمیق با دانشجویان، و تحلیل مصاحبه با مدرسان. این سه منبع داده امکان دادند که خطاهای صرفی هم از منظر عملکرد واقعی دانشجویان در ترجمه و هم از دیدگاه دانشجویان و مدرسان درباره علل بروز این خطاها بررسی شوند. ترجمه‌های دانشجویان پس از گردآوری به‌دقت بررسی شدند و خطاهای شناسایی‌شده براساس طبقه‌بندی پیشنهادی نا (۲۰۰۵) سازمان‌دهی گردیدند. برای این منظور، جدولی طراحی شد که در آن ردیف‌ها انواع خطاهای ساختاری و ستون‌ها منابع و گونه‌های احتمالی خطا را نشان می‌دادند. هر نمونه خطا در خانه مربوط به خود علامت‌گذاری شد و سپس فراوانی و درصد وقوع آن‌ها به‌صورت دستی محاسبه گردید. ترجمه‌ها با نسخه معیار مقایسه شدند تا ماهیت دقیق خطاها و گستره آن‌ها روشن شود. خطاها همراه با توضیحات اولیه مستقیماً روی نسخه‌های ترجمه‌شده دانشجویان یادداشت شد تا در مرحله مصاحبه مورد مراجعه قرار گیرد و دانشجو بتواند درباره علت خطا توضیح ارائه دهد. برای افزایش روایی و اعتبار تحلیل، دو مدرس با سابقه تدریس تحلیل خطا در زبان‌شناسی و ترجمه، به‌عنوان ممیزان خارجی، دسته‌بندی‌ها، کدگذاری‌ها و تحلیل‌ها را بازبینی و تأیید کردند. مطابق با روش پیشنهادی کرسول (۲۰۰۸)، مصاحبه‌ها ابتدا ضبط و سپس به‌طور کامل پیاده‌سازی شدند. رونوشت‌های مصاحبه چندین بار خوانده شد و کدگذاری اولیه با تمرکز بر مضامین تکرارشونده، علل احتمالی خطاها و تجربه‌های ذهنی دانشجویان انجام گرفت. تمامی کدها در یک فهرست گردآوری شد و سپس کدهای مشابه به‌منظور کاهش افزونگی و تقویت انسجام تحلیلی، در گروه‌های مفهومی

یکپارچه شدند. از دل این کدها، موضوعات اصلی شناسایی شد؛ موضوعاتی که بیشترین بسامد را در اظهارات شرکت‌کنندگان داشتند. در ادامه، این موضوعات براساس منطق تحلیلی به صورت سلسله‌مراتبی از سطح اصلی به سطح فرعی سازمان‌دهی شدند. در نهایت، مطابق با دستورالعمل‌های کرسول^۱ (۲۰۰۸)، یک نمودار درختی سلسله‌مراتبی طراحی شد که ضمن نشان دادن روابط میان مضامین، اطلاعات کلیدی مربوط به هر شرکت‌کننده را نیز منعکس می‌کرد و ساختاری روشن برای گزارش‌دهی یافته‌ها فراهم می‌ساخت.

تحلیل مصاحبه‌های مدرسان نیز براساس روش مراحل‌مند کرسول (۲۰۰۸) انجام شد. ابتدا تمامی مصاحبه‌ها ضبط و پیاده‌سازی شد و سپس متن‌ها چندین بار مطالعه شدند تا مفاهیم کلیدی شناسایی گردد. کدگذاری اولیه با تمرکز بر الگوهای مشترک در تجربه مدرسان درباره عملکرد دانشجویان، منابع خطاهای صرفی و چالش‌های رایج در ترجمه انجام گرفت. پس از تهیه فهرستی جامع از کدها، کدهای مشابه ادغام و دسته‌بندی شدند تا ساختاری مفهومی و منسجم ایجاد شود. موضوعات اصلی براساس فراوانی اشاره و اهمیت آموزشی آن‌ها استخراج و سپس در قالب موضوعات اصلی و فرعی سازمان‌دهی شدند. در ادامه، همانند بخش تحلیل مصاحبه‌های دانشجویان، یک نمودار درختی سلسله‌مراتبی مطابق روش کرسول (۲۰۰۸) طراحی شد که روابط میان مضامین و اطلاعات بافتی هر شرکت‌کننده، یعنی مدرس، را نمایش می‌داد. برای افزایش دقت و روایی تحلیل، دو مدرس با تجربه در تدریس تحلیل خطا، داده‌های تحلیل‌شده را بررسی و تأیید نمودند.

۵-۱. خطاهای صرفی

آمار خطاها

بر اساس یافته‌های استخراج‌شده از متون ترجمه‌شده، ۲۷ خطای صرفی مربوط به اسم شناسایی شد که در سه زیرگروه جای می‌گیرند: (۱) اضافه کردن نابجای نشانه جمع، (۲) حذف نابجای نشانه جمع، (۳) استفاده نادرست از حالت ملکی. این دسته از خطاها ۳۸/۵۷ درصد کل خطاهای صرفی را تشکیل می‌دهند. همچنین، ۴۳ خطای صرفی مربوط به فعل استخراج شد که در دو زیرگروه اصلی قرار می‌گیرند: (۱) استفاده از ساختار نادرست فعل، (۲) رعایت نکردن تطابق فاعل و فعل. این خطاها ۶۱/۴۲ درصد کل خطاهای صرفی را تشکیل می‌دهند. نسبت توزیع خطاهای اسمی و فعلی در مقایسه با کل خطاها در جدول شماره ۳ ارائه شده است و نشان می‌دهد که خطاهای فعلی سهم بیشتری از چالش‌های ساختارهای دانشجویان را به خود اختصاص داده‌اند.

جدول (۱)

فراوانی و درصد خطاهای مربوط به اسم

مقوله	نوع خطا	فراوانی	درصد
اسم	اضافه کردن نشانه جمع	۷	۲۵٪/۹۲
اسم	حذف نشانه جمع	۴	۱۴٪/۸۱
اسم	استفاده نادرست از حالت ملکی	۱۶	۶۱٪/۵۹
جمع خطاها	—	۲۷	۱۰۰٪

^۱ J. W. Creswell

جدول (۲)

فراوانی و درصد خطاهای مربوط به فعل

فعل	نوع خطا	فراوانی	درصد
	ساختار نامناسب	۱۴	۳۲٪/۵۵
	عدم رعایت تطابق فاعل و فعل	۲۹	۶۷٪/۴۴
جمع خطاها		۴۳	۱۰۰٪

جدول (۳)

فراوانی و درصد کل خطاهای مربوط به اسم و فعل

کل خطاها	فراوانی	درصد
اسم	۲۷	۳۸٪/۵۷
فعل	۴۳	۶۱٪/۴۲
جمع	۷۰	۱۰۰٪

۵-۲. تحلیل خطاها

۵-۲-۱. افزودن نشانه جمع به اسم

اولین نوع خطای صرفی مشاهده شده در داده‌ها، افزودن نشانه جمع S- به اسم‌هایی است که در زبان انگلیسی ببقاعده هستند. این خطا عمدتاً در ترجمه از فارسی به انگلیسی رخ داده و نشان‌دهنده‌ی عدم آشنایی دانشجو به برخی از اسم‌های بی‌قاعده از جمله «tooth» است. در اینجا دانشجوی به‌جای استفاده از «tooth» به اشتباه از «*tooths» استفاده کرده است.

متن مبدأ: دندان‌هایش را داغ کرده‌اند، ...

ترجمه دانشجو:

Its tooths are seared

ترجمه پیشنهادی:

Its teeth had been burnt . . .

در این مثال، دانشجو با افزودن نشانه جمع به واژه «tooth»، مرتکب خطای صرفی شده است. این خطا را می‌توان از دو منظر تحلیل کرد. نخست، از منظر انتقال بین‌زبان^۱، بدین معنا که دانشجو یا این تصور که واژه مزبور به مانند واژه‌ی فارسی «دندان‌ها» یک اسم باقاعده است، نشانه جمع را به «tooth» افزوده و صورت ناموجود «*tooths» را ساخته است. یکی از دانشجویان علت خطای ترجمه خود را به شرح زیر بیان کرد: «نمی‌دانستم که معادل انگلیسی واژه دندان‌ها، یک واژه ببقاعده است، و آن را به شکل واژه‌های باقاعده جمع بستم.»

علت: دانش کم واژگانی

^۱ interlingual transfer

اضافه کردن نشانه ی جمع، تعداد ۷ خطا را به خود اختصاص داده است.

۵-۲-۲. حذف نشانه جمع از اسم

حذف نشانه جمع یکی از رایج‌ترین خطاهای صرفی در تولید نوشتاری زبان‌آموزان زبان انگلیسی است و زمانی رخ می‌دهد که دانشجویان نشانه جمع -s را در انتهای اسم‌های قابل شمارش جمع حذف می‌کنند، همان‌گونه که در مثال‌های (۲) و (۳) مشاهده می‌شود. این نوع خطا به‌طور عمده ناشی از تفاوت‌های بنیادین میان نظام جمع‌سازی در زبان فارسی و زبان انگلیسی است.

متن مبدأ: هر صفت بدی را که /اسب نباید داشته باشد ...

ترجمه دانشجو:

*It had many *defect.*

ترجمه پیشنهادی:

It had every bad trait a horse shouldn't have.

در این مثال، دانشجو اسم قابل شمارش «defect» را پس از کمیت‌نمای «many» بدون افزودن نشانه جمع به‌کار برده است. از آنجا که «many» در زبان انگلیسی صرفاً با اسم‌های جمع قابل شمارش می‌آید، حذف نشانه جمع در این بافت، یک خطای صرفی آشکار محسوب می‌شود. این خطا نشان می‌دهد که دانشجو به رابطه‌ی نحوی-صرفی میان کمیت‌نماها و اسم‌های جمع در زبان انگلیسی توجه کافی نداشته است. از دیدگاهی بین‌زبانی، در زبان فارسی کمیت‌نماهایی مانند «بسیاری از» یا «خیلی» الزاماً نیازمند نشان‌گذاری صریح جمع روی اسم نیستند و مفهوم جمع‌بودن اغلب از طریق بافت جمله درک می‌شود. در نتیجه، زبان‌آموز در فرایند ترجمه، این الگوی زبان مادری را به زبان مقصد منتقل کرده و از نشان‌گذاری صرفی جمع در زبان انگلیسی غفلت کرده است. این پدیده را می‌توان در چارچوب انتقال منفی بین‌زبانی^۱ تحلیل کرد.

متن مبدأ: چون او را با زین و یراق روی هم رفته همه را پنج تومن قیمت نهاد ...

ترجمه دانشجو:

*Since I had decorated him with saddle and *trapping, . . .*

ترجمه پیشنهادی:

But when he gave his price with the saddle and all at only five Tomans, . . .

در این مثال نیز حذف نشانه جمع در اسم «trapping» به اشتباه رخ داده است. این اسم در زبان انگلیسی معمولاً به‌صورت اسم جمع به کار می‌رود و اشاره به مجموعه‌ای از تجهیزات و لوازم جانبی دارد. استفاده از صورت مفرد این اسم نشان می‌دهد که دانشجو به ویژگی‌های صرفی «اسم‌های همواره جمع» در زبان انگلیسی تسلط کافی ندارد. این نوع خطا را می‌توان علاوه بر تداخل بین‌زبانی، در چارچوب ساده‌سازی ساختاری نیز تبیین کرد؛ به این معنا که زبان‌آموز برای کاهش پیچیدگی صرفی، از

¹ negative transfer

صورت بی‌نشان یعنی مفرد استفاده کرده است. این که بر اساس فرضیه نشان‌داری، پردازش صورت‌های بی‌نشان برای دانشجوی غیر بومی زبان انگلیسی ساده‌تر هستند، حذف نشانه‌های صرفی در یک متن ادبی بیشتر محتمل است (اکمن^۱، ۱۹۷۷). یک دانشجو علت خطای ترجمه خود را این‌گونه بیان کرد: «وقتی کلمه‌هایی مثل «trapping» یا «goods» را دیدم فکر کردم استفاده از صورت مفرد نیز اشتباه نیست».

علت: ناآگاهی از اسم‌های جمع ثابت (plural-only nouns)

در مجموع، مثال‌های (۲) و (۳) حاکی از آن است که حذف نشانه جمع نه تنها نتیجه‌ی عدم آگاهی از قواعد صرفی زبان انگلیسی است، بلکه بازتابی از ارتباط میان انتقال بین‌زبانی، تعمیم ناقص قواعد، و راهبردهای پردازشی دانشجویان در نظام زبانی در حال شکل‌گیری آن‌هاست. تکرار این نوع خطا در ترجمه دانشجویان سطح متوسط بیانگر آن است که نشان‌گذاری جمع همچنان یکی از چالش‌های پایدار در اکتساب زبان انگلیسی برای فارسی‌زبانان به شمار می‌رود. نتایج نشان داد که تعداد ۴ خطا متعلق به حذف نشانه‌ی جمع از اسم بوده اند.

۵-۲-۳. استفاده نادرست از حالت ملکی

سومین نوع خطای صرفی مشاهده‌شده در داده‌های این پژوهش، استفاده نادرست از حالت ملکی در اسم‌ها است. این نوع خطا زمانی به‌عنوان خطای صرفی تلقی می‌شود که زبان‌آموز تصمیم درستی برای بیان مفهوم مالکیت می‌گیرد، اما در به‌کارگیری قاعده‌ی صرفی مناسب در زبان مقصد، به‌ویژه در استفاده از «apostrophe» و تکواژ 's، دچار خطا می‌شود. مثال (۵) نمونه‌ای روشن از این نوع خطاست.

متن مبدأ: «دندان‌هایش را داغ کرده‌اند»

ترجمه دانشجوی:

It's teeth are seared.

ترجمه پیشنهادی:

Its teeth had been cauterized.

در این مثال، دانشجو تلاش کرده است رابطه مالکیت میان «سب» و «دندان‌ها» را با استفاده از ساختار ملکی در زبان انگلیسی نشان دهد؛ بنابراین تصمیم او برای استفاده از ساختار ملکی از نظر مفهومی درست است. با این حال، در سطح صرفی، دانشجو به اشتباه از صورت *It's* استفاده کرده است که شکل کوتاه‌شده *It is* است. صورت صحیح ملکی در این بافت *Its* بدون «apostrophe» است.

یکی از دانشجویان علت خطای خود را به این شکل بیان کرد: «بین *its* و *it's* همیشه گیج می‌شوم. تلفظشان یکی است و موقع نوشتن سریع متوجه تفاوتشان نمی‌شوم».

علت: ابهام میان اشکال هم‌آوا و ضعف در تمایز املایی-صرفی

این خطا نشان‌دهنده‌ی عدم تمایز میان اشکال هم‌آوا در زبان انگلیسی است؛ تمایزی که برای دانشجویان فارسی‌زبان انگلیسی گاهی دشوار است، زیرا زبان فارسی فاقد نشانه نوشتاری مشابه «apostrophe» برای بیان مالکیت است و رابطه‌ی مالکیت

¹ F. R. Eckman

معمولاً از طریق ساختار «اضافه» (ـِ) نشان داده می‌شود. در نتیجه، دانشجویان در انتقال این مفهوم به زبان انگلیسی، به قواعد املائی و صرفی نادرست متکی می‌شوند و دچار خطا می‌گردند. علاوه بر این، می‌توان این خطا را در چارچوب تداخل بین‌زبانی و نیز پیچیدگی درون‌زبانی قواعد صرفی زبان انگلیسی تحلیل کرد. در زبان انگلیسی، وجود اشکال مشابه نوشتاری و آوایی مانند *its* و *it's* فشار شناختی بیشتری را در فرایند تولید نوشتاری ایجاد می‌کند، به‌ویژه در ترجمه متون ادبی که زبان‌آموز هم‌زمان درگیر انتقال معنا، انتخاب واژگان و ساختارهای دستوری است. طبق نتایج بدست آمده، تعداد ۱۶ خطا در استفاده ی نادرست از حالت ملکی بوده است.

۵-۲-۴. ساختار نامناسب فعل

خطاهای «ساختار نامناسب فعل» زمانی رخ می‌دهند که دانشجویان قادر به تولید شکل صرفی صحیح فعل متناسب با زمان و ساختار نحوی جمله نباشند. این خطاها شامل استفاده نادرست از زمان‌ها، شکل گذشته ساده، اسم مفعول، یا ساخت‌های فعلی مرکب است. مثال (۶) نمونه‌ای روشن از این نوع خطا را نشان می‌دهد.

متن مبدأ: /از تندی و تیزی او در راه فهمیدم که مال بدی نیست

ترجمه دانشجوی:

I had finded that it is not a bad property because it was a nimble horse in the way.

ترجمه پیشنهادی:

From its speed and energy on the way, I could tell it was not a bad animal.

در این مثال، دانشجو برای بیان مفهوم گذشته‌ی کامل، به‌درستی از ساختار *had + verb* استفاده کرده است؛ با این حال، در تولید شکل صرفی فعل اصلی دچار خطا شده و به‌جای اسم مفعول صحیح «*found*»، از شکل ناموجود «**finded*» استفاده کرده است. این خطا نشان‌دهنده‌ی تعمیم‌بخشی افراطی در صرف فعل است؛ بدین معنا که دانشجو قاعده‌ی افزودن پسوند *-ed* را بدون توجه به بی‌قاعده بودن فعل «*find*» به کار برده است.

یکی از دانشجویان علت خطای ترجمه خود را به شرح زیر بیان کرد: «قاعده *-ed* را بلد بودم و فکر کردم می‌شود آن را به همه فعل‌ها تعمیم داد. یادم رفت بعضی فعل‌ها بی‌قاعده هستند».

علت: تعمیم‌بخشی افراطی

از منظر بین‌زبانی، این خطا می‌تواند تحت تأثیر نظام صرفی زبان فارسی نیز قرار گرفته باشد، زیرا در فارسی، تمایز صوری مشخصی میان افعال باقاعده و بی‌قاعده به‌شکلی که در انگلیسی وجود دارد، دیده نمی‌شود. در نتیجه، زبان‌آموزان در مراحل میانی یادگیری، تمایل دارند یک الگوی واحد صرفی را به همه‌ی افعال تعمیم دهند. این نوع خطا را می‌توان در چارچوب نظریه بین‌زبانی تحلیل کرد که خطا را به‌عنوان بخشی طبیعی از فرایند رشد زبان دوم تلقی می‌کند، نه نشانه‌ی ضعف صرفاً فردی. تعداد ۱۴ خطا متعلق به ساختار نامناسب فعل می‌باشد.

۵-۲-۶. تطابق فاعل/فعل

دسته‌ی دیگری از خطاهای صرفی مشاهده شده در داده‌ها مربوط به عدم رعایت مطابقه فاعل و فعل است. این نوع خطا زمانی رخ می‌دهد که زبان‌آموزان نتوانند افعال *be* یا *have* را متناسب با فاعل جمله به درستی صرف کنند، بدون آنکه حذف ضمیر سوم شخص مدنظر باشد. مثال (۶) نمونه‌ای از این نوع خطاست.

متن مبدأ: هر صفت بدی را که اسب نباید داشته باشد و برعکس از صفت و محسنات اسب خوب به کلی محروم بود.

ترجمه دانشجو:

There were not to be a breed worse than this.

ترجمه پیشنهادی:

It had every bad trait a horse shouldn't have and was completely lacking any of the good ones.

در این مثال، دانشجو از *there were* استفاده کرده، در حالی که گروه اسمی مرتبط با آن مفرد است. در ضمن ساختار انتخاب شده با معنای جمله مبدأ هم خوانی ندارد. از دیدگاهی صرفی، عدم تطابق میان صورت جمع فعل «were» و گروه اسمی مفرد جمله «a breed» یک خطای مطابقه محسوب می‌شود. این خطا نشان می‌دهد که دانشجو در تشخیص رابطه‌ی میان فاعل دستوری و صورت صرفی فعل دچار مشکل شده است.

یکی از دانشجویان در توجیه خطای خود چنین گفت: «چون متن ادبی بود، سعی کردم جمله‌ها را شبیه انگلیسی رسمی بنویسم، ولی گاهی ساختار درست فعل یا تطابقش را از دست دادم».

علت: عدم تسلط بر متون ادبی

این نوع خطا را می‌توان ناشی از انتقال ویژگی‌های زبان فارسی به انگلیسی به شکلی اشتباه دانست؛ زیرا در فارسی، صرف فعل اطلاعات مربوط به شخص و شمار فاعل را به‌طور کامل در خود دارد و وجود یا عدم وجود فاعل صریح در جمله تأثیر چندانی بر شکل فعل ندارد. در نتیجه، زبان‌آموزان ممکن است در زبان انگلیسی نیز به نشانه‌های صرفی تطابق توجه کافی نداشته باشند. علاوه بر این، از دیدگاه پردازشی، چنین خطاهایی می‌توانند نتیجه‌ی بار شناختی بالا در فرایند ترجمه باشند؛ به‌ویژه زمانی که زبان‌آموز هم‌زمان درگیر انتخاب واژگان، بازسازی ساختار جمله و انتقال معنای ادبی متن است. طبق نتایج به دست آمده، تعداد ۲۹ خطا متعلق به عدم رعایت تطابق فعل و فاعل بود.

به‌طور کلی، خطاهای صرفی مربوط به فعل در این پژوهش نشان می‌دهند که زبان‌آموزان در مراحل میانی یادگیری زبان انگلیسی، همچنان در به‌کارگیری اشکال صرفی صحیح افعال و رعایت مطابقه فاعل-فعل با چالش مواجه‌اند. این خطاها نه تنها بازتابی از انتقال بین‌زبانی هستند، بلکه نشان‌دهنده‌ی ویژگی‌های نظام میان‌زبانی زبان‌آموزان و راهبردهای پردازشی آن‌ها در ترجمه متون ادبی محسوب می‌شوند. اظهارات دانشجویان نشان می‌دهد که خطاهای صرفی مشاهده شده در ترجمه متون ادبی نه صرفاً ناشی از فقدان دانش دستوری، بلکه حاصل تعامل پیچیده‌ای از عوامل شناختی، بین‌زبانی و آموزشی هستند. انتقال ساختارهای زبان فارسی، تعمیم بیش‌ازحد قواعد، فشار زمانی، و تمرکز هم‌زمان بر معنا و صورت، همگی در شکل‌گیری این خطاها نقش داشته‌اند. این یافته‌ها از دیدگاه نظریه بین‌زبانی، خطاهای صرفی را به‌عنوان بخشی طبیعی از فرایند اکتساب زبان دوم تأیید می‌کنند.

۵-۳. خطاهای دانشجویان

ابتدا، خطاهای دانشجویان در ترجمه بر اساس چارچوب نا (۲۰۰۵) کشف و بررسی شدند. سپس توضیحاتی درباره جزییات و علل خطاهای دانشجویان در ترجمه ارائه شد. در مرحله بعد، مصاحبه‌ای با اساتید جهت مشخص کردن علل خطاها مزبور انجام شد. تحلیل داده‌ها توسط محقق، و همچنین نتایج به‌دست آمده از مصاحبه با اساتید حاکی از آن بود که رایج‌ترین علل خطاهای دانشجویان عبارتند از:

۱. ضعف برخی از دانشجویان در درک صحیح متن مبدأ
۲. نداشتن مهارت کافی برای انتقال جملات پیچیده
۳. سهل‌انگاری در توجه کافی به پیام اصلی متن مبدأ
۴. عدم توفیق در یافتن معادل‌های واژگانی مناسب در فرهنگ لغت
۵. عدم مطالعه دقیق متن مبدأ توسط دانشجو پیش از دست به قلم بردن جهت ترجمه
۶. نداشتن تجربه کافی یا حتی لازم در به‌کارگیری راهبردهای اساسی ترجمه
۷. نداشتن دانش دستوری کافی و اشراف لازم به ساختارهای صرفی-نحوی در هر دو زبان مبدأ و مقصد
۸. عدم بازبینی نهایی ترجمه برای یافتن خطاهای حاکی از بی‌دقتی پیش از تحویل دادن ترجمه به استاد.
۹. نداشتن دانش لازم نسبت به قواعد ساختوازی و تصریفی زبان مقصد
۱۰. انجام ترجمه کلمه به کلمه بدون توجه به ظرائف بافتی و سبکی متن مقصد
۱۱. ضعف‌های آموزشی به علت عدم بهره‌مندی دانشجو از روش‌های علمی به‌روز برای آموزش ترجمه
۱۲. مهارت عملی ناچیز دانشجو به جهت عدم شرکت در کارگاه‌های عملی ترجمه، و انجام ترجمه‌های حرفه‌ای
۱۳. عدم آموزش مفید و کارآمد به علت شمار بیش از حد معیار در کلاس درس
۱۵. عدم تعامل چهره به چهره میان دانشجو و استاد به دلیل حضور پرتعداد دانشجو در کلاس درس

۶. بحث و بررسی یافته‌ها

هدف اصلی این پژوهش شناسایی، توصیف و تحلیل خطاهای صرفی دانشجویان سال آخر رشته مترجمی زبان انگلیسی در ترجمه یک متن ادبی از فارسی به انگلیسی بود. تحلیل داده‌ها نشان داد که خطاهای صرفی همچنان بخش چشمگیری از خطاهای زبانی دانشجویان را تشکیل می‌دهند؛ امری که بیانگر آن است که حتی در سطوح پیشرفته‌تر تحصیلی نیز برخی مؤلفه‌های صرفی زبان انگلیسی در ذهن زبان‌آموزان به‌طور کامل تثبیت نشده است. یافته‌ها نشان داد که خطاهای صرفی در دو مقوله اصلی اسمی و فعلی طبقه‌بندی می‌شوند؛ الگویی که با نتایج پژوهش انصاری و سیفی (۱۴۰۳) نیز هم‌خوان است و در آن، رفتارهای میان‌زبانی مشابهی در میان دانشجویان مترجمی گزارش شده بود.

بر اساس داده‌ها، حدود ۶۱ درصد خطاهای صرفی مربوط به حوزه فعل و ۳۸ درصد مربوط به حوزه اسم بود. خطاهای اسمی عمدتاً شامل افزودن نادرست نشانه جمع -s به اسم‌های غیرقابل شمارش، حذف اشتباه نشانه جمع در موارد لازم، و استفاده نادرست از حالت ملکی بودند. این نوع خطاها غالباً حاصل نوعی تعمیم‌بخشی افراطی و انتقال راهبردهای آشنای زبان فارسی به زبان مقصد است. دانشجویان هنگام تلاش برای بازنمایی «کلیت»، «تعدد» یا «جمعیت معنایی» موجود در متن فارسی به

انگلیسی، به‌طور پیش‌فرض به ابزار صرفی آشنای خود یعنی نشانه جمع متوسل شده‌اند؛ در حالی که حوزه کاربردی آن در زبان انگلیسی محدودتر و قاعده‌مندتر است. این مسئله در متون ادبی برجسته‌تر می‌شود، زیرا زبان ادبی فارسی اغلب با مفاهیمی انتزاعی، کلی و گاه استعاری سروکار دارد که معادل آن‌ها در انگلیسی الزاماً ساختار جمع نمی‌گیرد.

یکی دیگر از خطاهای قابل توجه، استفاده نادرست از ساختار ملکی بود، به‌ویژه در تمایز میان شکل‌های *its* و *it's*. تحلیل این خطاها نشان داد که دانشجویان در سطح مفهومی مالکیت را درک می‌کنند، اما در سطح صوری و املائی دچار سردرگمی می‌شوند. نبود ساختاری معادل «apostrophe» در فارسی و ماهیت نوشتاری-شکلی این تمایز، از عوامل اصلی پیچیدگی آن برای زبان‌آموزان است. این خطاها نشان می‌دهد که برخی چالش‌های صرفی حتی در سطوح پیشرفته نه به دلیل فقدان دانش، بلکه به دلیل فشار پردازشی، تداخل درون‌زبانی و بار شناختی بالای فعالیت ترجمه رخ می‌دهند.

در حوزه افعال، دو نوع خطای غالب مشاهده شد: استفاده از ساختار نادرست فعل و رعایت نکردن تطابق فاعل-فعل. نمونه‌هایی مانند «*finded» به جای «found» نمایانگر تعمیم‌بخشی افراطی قاعده صرف افعال باقاعده است. این الگو با نظریه میان‌زبان الیس (۲۰۱۵) قابل تبیین است؛ بدین معنا که زبان‌آموز در حال ساخت نظامی مستقل است که ویژگی‌هایی برگرفته از زبان مادری و زبان مقصد را توأمان دارد. مشکل تطابق فاعل-فعل، به‌ویژه در افعال «be» و «have»، نیز از خطاهای پایدار و پرتکرار بود. یافته‌ها با پژوهش لاردیر (۲۰۱۷) هم‌راستا است که نشان می‌دهد حتی زبان‌آموزان پیشرفته در موقعیت‌های تولید پیچیده، به‌ویژه در فعالیت‌هایی با بار شناختی بالا مانند ترجمه ادبی، ممکن است نشانه‌های تطابق را حذف یا نادرست استفاده کنند.

بررسی جدول‌های (۱، ۲ و ۳) نشان داد که بسیاری از خطاهای دانشجویان از نوع خطاهای پایدار یا شبه‌پایدار هستند؛ بدین معنا که در طول زمان و علیرغم آموزش رسمی همچنان تکرار می‌شوند. تحلیل مصاحبه با اساتید نیز نشان داد که برخی از این خطاها ریشه در ضعف‌های ساختاری زبان‌آموزان در نحو و دستور زبان فارسی دارد. این نکته پیش‌تر در پژوهش افروز و رنجبرمحمدی (۱۴۰۱) نیز مورد تأکید قرار گرفته بود. اساتید همچنین اشاره کردند که اگرچه دانشجویان در طی دوره تحصیل با انواع متون تخصصی، علمی، سیاسی، مطبوعاتی، اسلامی و ادبی آشنا می‌شوند، اما شواهد نشان می‌دهد که بهره‌گیری آنان از اصول، روش‌ها و راهبردهای ترجمه نظام‌مند و عمیق نیست و بیشتر به ترجمه مبتنی بر شهود و تجربه شخصی تکیه می‌کنند. نکته قابل توجه دیگر به وضعیت نظام آموزشی مترجمی در ایران بازمی‌گردد. بخش قابل توجهی از مدرسان دوره‌های مترجمی دارای تحصیلات دانشگاهی در حوزه‌های آموزش زبان انگلیسی، زبان‌شناسی یا ادبیات انگلیسی هستند و تنها تعداد اندکی از آنان سه مقطع تحصیلی خود را در رشته مترجمی گذرانده‌اند. این موضوع، مطابق گزارش‌های جمالی‌منش (۲۰۰۹)، پورنیا و همکاران (۱۴۰۴) و افروز (۱۴۰۱)، می‌تواند یکی از عوامل ضعف‌های بنیادین در آموزش نظری و عملی ترجمه باشد؛ زیرا همسویی تخصصی میان رشته تحصیلی مدرس و نیازهای دانشی دانشجویان در حوزه ترجمه به‌طور کامل برقرار نیست. این مسئله ضرورت به‌روز نگه‌داشتن دانش نظری و حرفه‌ای مدرسان، و همچنین بازنگری سیاست‌گذاری آموزشی در انتخاب و تربیت مدرسین رشته مترجمی را برجسته می‌کند.

۷. نتیجه‌گیری

هدف اصلی پژوهش حاضر شناسایی، تبیین و تحلیل خطاهای صرفی دانشجویان رشته مترجمی زبان انگلیسی در ترجمه از فارسی به انگلیسی بود. برای این منظور، از چارچوب مفهومی نا (۲۰۰۵) به عنوان مبنای تحلیل استفاده شد که امکان طبقه‌بندی نظام‌مند خطاها و بررسی الگوهای میان‌زبانی دانشجویان را فراهم می‌کرد. یافته‌ها نشان دادند که بیشترین فراوانی خطاها در حوزه فعل رخ داده است؛ موضوعی که بیانگر چالش‌های مداوم دانشجویان در انتخاب شکل مناسب فعل، تسلط بر الگوهای صرفی افعال باقاعده و بی‌قاعده، و رعایت تطابق فاعل-فعل است. همچنین مشخص شد که دانشجویان در حوزه اسمی نیز با خطاهایی نظیر افزودن نادرست نشانه جمع، حذف اشتباه این نشانه، و کاربرد نامناسب حالت ملکی مواجه‌اند. تکرار این خطاها حاکی از آن است که دانشجویان در انتقال دقیق ساختارهای اسمی از فارسی به انگلیسی با دشواری‌هایی روبه‌رو هستند، به‌ویژه در بافت‌هایی که زبان فارسی از مفاهیم کلی، جمعیت معنایی یا ساخت‌های ملکی با شفافیت کمتر بهره می‌گیرد. نتایج کلی پژوهش نشان می‌دهد که خطاهای صرفی پدیده‌هایی تصادفی یا ناشی از کم‌دقتی ساده نیستند، بلکه حاصل تعامل پیچیده‌ای از انتقال بین‌زبانی، ویژگی‌های ساختاری و درون‌زبانی زبان مقصد، و عوامل شناختی همچون بار پردازشی و بهره‌گیری از راهبردهای ساده‌سازی هستند. بدین ترتیب، خطاهای صرفی را باید بخشی طبیعی و اجتناب‌ناپذیر از نظام میان‌زبان زبان‌آموزان دانست؛ نظامی پویا که در آن، زبان‌آموز به تدریج قواعد را بازسازی و اصلاح می‌کند.

در بخش دوم پژوهش، علل بروز خطاهای صرفی از دیدگاه دانشجویان و مدرسان مورد بررسی قرار گرفت. تحلیل مصاحبه‌ها نشان داد که عوامل مختلفی از جمله ضعف درک مطلب در زبان مبدأ به‌ویژه در مواجهه با جملات پیچیده، عدم توجه کافی به تفاوت‌های ساختاری زبان فارسی و انگلیسی، تسلط ناکافی بر ساختارهای پایه‌ای هر دو زبان، کمبود مدرسان متخصص در حوزه مطالعات ترجمه، و وجود خلأهایی در چارت درسی رشته مترجمی زبان انگلیسی در بروز خطاها نقش دارند. این عوامل، مجموعه‌ای از چالش‌های آموزشی و ساختاری را آشکار می‌سازند که مستقیماً بر عملکرد دانشجویان در ترجمه تأثیر می‌گذارند. یافته‌های پژوهش حاضر می‌تواند مبنایی برای بازنگری آموزشی در رشته مترجمی زبان انگلیسی قرار گیرد. نتایج پژوهش بر ضرورت توجه هدفمند به آموزش ساختار و پدیده‌های صرفی، به‌ویژه در بافت ترجمه متون ادبی که بار شناختی بالاتری دارد، تأکید می‌کند. بر اساس داده‌های حاضر، برنامه‌ریزی آموزشی باید شامل تقویت آموزش آگاهی‌بخش صرفی، تمرین‌های هدفمند ترجمه با تمرکز بر مقولات اسمی و فعلی، و بازطراحی بخشی از محتوای درسی بر اساس رویکردهای نوین تحلیل خطا و مطالعات میان‌زبانی باشد. برای پژوهش‌های آینده پیشنهاد می‌شود تأثیر عوامل زیر بر کاهش خطاهای صرفی بررسی شود: بازبینی‌های مرحله‌ای و چندلایه در فرایند ترجمه، آموزش آگاهی‌بخش درباره الگوهای صرفی زبان انگلیسی، مقایسه عملکرد دانشجویان در متون ادبی و غیرادبی، و به‌کارگیری روش‌های آموزشی نوین مبتنی بر تحلیل میان‌زبانی. چنین پژوهش‌هایی می‌تواند به درک عمیق‌تر از چگونگی شکل‌گیری خطاهای صرفی و راهکارهای مؤثر برای رفع آن‌ها منجر شود.

منابع فارسی

- افروز، حاجیه، و رنجبر محمدی، رویا (۱۴۰۱). «ارزیابی کیفیت ترجمه‌ی استعاره‌ها در رباعیات عمر خیام با استفاده از مدل هاوس: مطالعه‌ای درباره ترجمه‌ی آشکار و پنهان». فصلنامه‌ی زبان و ترجمه، ۲(۲)، ۵۹-۷۱.
- افروز، محمود (۱۴۰۱). «متن‌پژوهی ادبی از منظر ترجمه‌پژوهی ادبی: نقدی بر بروز خطاهای آشکار در ترجمه متون ادبی با تکیه بر مدل‌های زبان‌شناختی متن». متن‌پژوهی/ادبی، ۲۶(۹۴)، ۲۹۳-۳۲۲.

انصاری، نرگس، و سیفی، طیبه (۱۴۰۳). «تعامل متن و ترجمه در بازتاب فرهنگ: موردپژوهی خار و میخک و یحیی سنوار». *پژوهش‌های ترجمه در زبان و ادبیات عربی*، شماره ۳۲، ۲۸۹-۳۲۲.

پورنیا، سمیه، چالاک، عزیزه، و حیدری تبریزی، حسین (۱۴۰۴). «راهبردهای مؤثر نوشتاری در میان زبان‌آموزان ایرانی سطح متوسط زبان انگلیسی: مطالعه‌ای مبتنی بر فتوویس». *فصلنامه‌ی انگلیسی برای اهداف دانشگاهی ایران*، ۱۴(۴)، ۳۴-۱۶.

کبیری، احمد، و ایزدی، مرضیه (۱۴۰۴). «مدل‌های تحلیل خطا در ترجمه و آموزش مترجمی در ایران». *مجله‌ی علمی پژوهشی ترجمه و آموزش زبان*، ۸(۱)، ۴۵-۶۸.

رحیمی، رامین. (۱۳۸۲). «ترجمه متون خبری: مطالعه‌ای درباره مشکلات دانشجویان ایرانی ترجمه». *فصلنامه‌ی مطالعات ترجمه*، ۱(۲)، ۴۴-۲۹.

ساسانی، فرناز، فرجاه، مرجان، نوابزاده شفیعی، سپیده، و نفر، ریحانه (۱۴۰۲). «روزآمدی ترجمه‌ی عناصر فرهنگی در ترجمه‌ی ادبی با تکیه بر نظریه‌ی نیومارک». *پژوهش‌های زبان‌شناختی در زبان‌های خارجی*، ۱۳(۴)، ۵۷۹-۵۹۴.

عربی، مینا و رستم‌پور ملکی، رقیه (۱۳۹۸). «خطاهای ترجمه دانشجویان رشته زبان و ادبیات عربی و راهکارهای رفع آن‌ها با تکیه بر روش تجزیه و تحلیل خطا». *جستارهای زبانی*، ۵(۱۰)، ۱۵۷-۱۸۱.

عسکری، محمدایمان، ستاریان، عدنان، و رنجبر، مهسا (۱۳۹۸). «بررسی خطاهای نهان و آشکار در ترجمه ماشینی متون دانشگاهی بر اساس مدل ارزیابی کیفیت ترجمه‌ی هوس (۲۰۱۵)». *فصلنامه‌ی زبان و ترجمه*، ۱۱(۳)، ۷۳-۵۷.

همایونی، سعدالله، و فولادی، مریم (۱۴۰۱). «تحلیل نقش صورت‌بندی‌های زبانی و حوزه‌های فرهنگی در ترجمه‌ی ترکیب‌های استعاره‌ی (از منظر زبان‌شناسی شناختی)». *پژوهش‌های ترجمه در زبان و ادبیات عربی*، ۱۲(۲۷)، ۲۵۳-۲۷۷.

یزدانی فرساد، مهناز (۱۳۹۰). «مطالعه و بررسی چالش‌های فرهنگی در ترجمه متون ادبی». پایان‌نامه کارشناسی ارشد، دانشگاه تربیت مدرس.

- Afrouz, H., & Ranjbar Mohammadi, R. (2022). Evaluating the quality of metaphor translation in Omar Khayyam's quatrains using House's model: A study on overt and covert translation. *Journal of Language and Translation*, 12(2), 59-71.
- Afrouz, M. (2022). Literary text-research from the perspective of literary translation studies: A critique of overt errors in literary text translation relying on linguistic text models. *Literary Text Research*, 26(94), 293-322.
- Al-Jarf, R. (2000). Grammatical agreement errors in L1/L2 translations. *International review of applied linguistics in language teaching*. IRAL 38 (1). 1-16.
- Ansari, N., & Seifi, T. (2024). Text and translation interaction in cultural reflection: A case study of *The Thorn and the Carnation* and Yahya Sinwar. *Translation Researches in Arabic Language and Literature*, (32), 289-322.
- Arabi, M., & Rostampour Molki, R. (2019). Translation errors of Arabic language and literature students and solutions for their removal relying on error analysis method. *Language Related Researches*, 10(5), 157-181.
- Ardeshiri, M., & Zarafshan, M. (2015). Students' causes of errors in translating pragmatic senses. *International Journal of English and Education*, 3(4), 238-253.
- Ardeshiri, M., Rahimi, M., & Foroutan, M. (2015). Students' errors in translating from Persian to English. *International Journal of Language Learning and Linguistics World*, 9(1), 233-244.
- Askari, M. I., Sattarian, A., & Ranjbar, M. (2019). Investigating covert and overt errors in machine translation of academic texts based on House's (2015) translation quality assessment model. *Journal of Language and Translation*, 11(3), 57-73.
- Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Sage.

- Freitag, M., Foster, G., Grangier, D., Ratnakar, V., & Tan, Q. (2021). Experts, errors, and context: A large-scale study of human evaluation for machine translation. *Transactions of the Association for Computational Linguistics*, 9, 1460–1474. https://doi.org/10.1162/tacl_a_00437
- Homayouni, S., & Fouladi, M. (2022). Analyzing the role of linguistic formulations and cultural domains in the translation of metaphorical collocations (from a cognitive linguistics perspective). *Translation Researches in Arabic Language and Literature* [in Persian], 12(27), 253–277.
- Jamalimanesh, A.R. (2009). Electronic sources: A glance at the ailing system of teaching translation in Iranian universities. <http://www.translationdirectory.com/articles/article2044.php> (30 December 2010).
- Kabiri, A., & Izadi, M. (2025). Error analysis models in translation and translator training in Iran. *Journal of Translation and Language Education*, 8(1), 45–68.
- Kenny, D. (2022). *Machine translation for everyone: Empowering users in the age of artificial intelligence*. Language Science Press.
- Lardiere, D. (2017). *Second language acquisition: A biolinguistic perspective*. Routledge.
- Li, S. (2020). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 70(3), 878–918.
- Lorsch, W. (1991). *Translation performance, translation process, and translation strategies: A psycholinguistic investigation*. Tübingen, Germany: Gunter Narr Verlag.
- Na, P. P. Q. (2005). Errors in the translation of topic-comment structures of Vietnamese into English. *Asian EFL Journal*, 7(3), 53–78.
- O'Brien, S. (2020). Translation, technology and the user. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 435–449). Routledge.
- O'Brien, S., & Ehrensberger-Dow, M. (2020). Machine translation and translator training: Evaluating the impact of automated tools on translation quality. *Journal of Specialised Translation*, 33, 1–20.
- Pournia, S., Chalak, A., & Heidari Tabrizi, H. (2025). Effective writing strategies among Iranian intermediate EFL learners: A Photovoice-based study. *Iranian Journal of English for Academic Purposes*, 14(4), 16–34.
- Pym, A. (2010). *Translation and text transfer: An essay on the principles of intercultural communication* (Rev. ed.). Intercultural Studies Group.
- Sasani, F., Farjah, M., Navabzadeh Shafiei, S., & Nafar, R. (2023). Updating the translation of cultural elements in literary translation relying on Newmark's theory. *Journal of Language Research*, 13(4), 579–594.
- Yazdani Farsad, M. (2011). *A study of cultural challenges in literary translation* [Unpublished master's thesis]. Tarbiat Modares University.

Extended Abstract

Design-Test: A Semi-Automated LLM-Based Framework for Typological Analysis with Human-in-the-Loop – A Case Study of Agreement Systems in the Bakhtiari Dialect of Izah

Saeed Mazraeh ^{1*} 

Department of Computer Engineering, Dez.C.,
Islamic Azad University, Dezful, Iran.
saeed.mazraeh@iau.ac.ir

Naser AliFar ²

Department of Linguistics, Bu.C., Islamic Azad
University, Bushehr, Iran.
naser.alifar@iau.ac.ir

Review History: Received: 2026-05-14 Revise: 2026-06-13 Accepted: 2026-07-04	Abstract The rapid documentation of endangered Iranian dialects is hampered by a scalability crisis due to time-intensive manual methods. This study aimed to design and evaluate a semi-automated framework based on a Large Language Model (LLM) to accelerate typological analysis. Employing a human-in-the-loop workflow, the Llama 3.1 model and few-shot learning were used to extract agreement system patterns from a 5,400-sentence corpus of the Bakhtiari dialect of Izah. The model's performance was evaluated on a gold-standard annotated test set (450 sentences) using precision, recall, and F1-score metrics, followed by a qualitative error analysis. The model achieved an overall accuracy of 87.4% and an F1-score of 87.4. Qualitative analysis of the 12.6% errors revealed two main limitations: 1) the model's bias towards high-frequency structures, and 2) its difficulty in processing specific morphophonological phenomena (e.g., broken plurals). In terms of efficiency, the framework yielded a 78% reduction in analysis time and enabled the processing of 4.6 times more data per unit of time. The findings demonstrate that LLMs, when embedded within a human-centric collaborative system, can dramatically enhance the speed and scale of low-resource language documentation without compromising accuracy. This framework provides a practical model for accelerating fieldwork in Iranian dialectology.	Keywords: Human-In-The-Loop Iranian Dialects Large Language Models (LLMs) Natural Language Processing (NLP) Syntactic Typology
<i>How to cite this article:</i> Mazraeh, M. & AliFar, N. (2026). Design-Test: A Semi-Automated LLM-Based Framework for Typological Analysis with Human-in-the-Loop – A Case Study of Agreement Systems in the Bakhtiari Dialect of Izah. . <i>Journal of Linguistics & Iranian Languages</i>, 11(2), 17-32. doi: https://doi.org/10.22099/jlil.2026.55606.1473		
 COPYRIGHTS ©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publisher.		

* Corresponding author

Introduction

The documentation of endangered Iranian dialects faces a severe "scalability crisis" due to time-intensive manual methods (Dabir-Moghaddam, 2012). Traditional fieldwork can take years and is often limited to small speaker samples. In an era of accelerating language loss, we urgently need tools that enhance both speed and scale without sacrificing analytical depth (Crystal, 2000; Karimi, 2020). Large Language Models (LLMs) like Llama 3.1, with their few-shot learning capability, offer a historic opportunity for automated processing of low-resource languages (Brown et al., 2020; Wei et al., 2022). However, naive application risks imposing biases from dominant languages and data colonialism (Bender et al., 2021; Bird, 2020). This study aims to design and evaluate a semi-automated "human-in-the-loop" framework in which an LLM acts as an intelligent assistant while the researcher retains final interpretation and quality control. The case study focuses on the agreement system of the Bakhtiari dialect of Izah. This dialect exhibits four specific features (vocalic pluralization, ergative constructions, pro-drop, and local phonological processes affecting agreement markers) that make it a challenging test for general-purpose language models.

Materials, Methods, and Objectives

The literature shows detailed descriptive studies of Bakhtiari (Kalbasi, 2016; Anonby et al., 2021), but no systematic framework for leveraging LLMs in the typological analysis of Iranian dialects exists. Recent studies on low-resource languages indicate that LLMs still face challenges such as data scarcity, the curse of multilinguality, and dialectal variation (Unjum et al., 2026). This research aims to fill this dual gap (methodological and empirical).

The methodology was as follows:

- **Data collection:** 12 hours of natural speech from 15 elderly Bakhtiari speakers (age 60–75) in rural Izah were recorded and transcribed, yielding a 5,400-sentence corpus.
- **Gold standard:** 600 sentences were randomly selected and annotated by two independent linguists (Cohen's kappa = 0.94). Of these, 450 sentences formed the final test set.
- **Model and prompt engineering:** Llama 3.1 (8B) was employed using few-shot prompting (5 examples) and chain-of-thought reasoning. The optimal prompt was refined over 15 iterations on a validation set.
- **Evaluation:** Model performance was measured using precision, recall, and F1-score. Errors were qualitatively analyzed. Manual analysis time was compared to semi-automated workflow time.

Results and Discussion

Quantitative results showed that the model achieved 87.4% overall accuracy and an F1-score of 87.4, outperforming a simple rule-based baseline (62.3%) by 25%. This demonstrates the model's ability to infer abstract syntactic patterns from limited data, supporting the hypothesis of "latent universal linguistic knowledge" in LLMs (Wei et al., 2022).

Qualitative analysis of 57 errors (12.6% of the test set) revealed two systematic failure patterns:

- **Structural bias (42% of errors):** The model tended to "distort" low-resource dialect structures into the nearest high-frequency patterns of Persian or English (e.g., misinterpreting ergative passive constructions as simple active).
- **Morphophonological blindness (35% of errors):** The model failed to recognize dialect-specific phenomena such as broken plurals and the influence of local phonological processes on agreement markers.

In terms of efficiency, the proposed framework reduced analysis time from 210 seconds per sentence (manual) to 45 seconds per sentence, representing a 78% time saving and enabling the processing of 4.6 times more data per unit time. This shifts the researcher's role from a labor-intensive analyst to an intelligent manager and final interpreter.

Ethical considerations included informed consent, anonymization, collective intellectual property rights for the speaker community, and avoidance of data colonialism (Bird, 2020). Key limitations are dependence on prompt engineering, inherent model bias, computational cost, and the need for transcribed data.

Conclusion

This study demonstrates that LLMs, embedded within a human-in-the-loop framework, can serve as powerful collaborative partners for field linguists. The proposed framework increases analysis speed by 78% and scalable processing capacity by 4.6 times while maintaining 87.4% accuracy. However, the model alone cannot grasp dialect-specific morphophonological intricacies or rule-breaking exceptions. Thus, the optimal future lies in a "dialectal synergy" between the machine's inductive speed and the human's historical-cultural depth of interpretation.

Future research directions include: (1) developing a shared library of optimized prompts for various typological phenomena; (2) fine-tuning an open-source base model on a balanced corpus of Iranian dialects; (3) integrating speech-to-text models for direct analysis of raw audio; and (4) conducting cross-dialectal comparative studies.

References

- Anonby, E., & Asadi, A. (2018). Bakhtiari dialect. In *Encyclopædia Iranica*. Brill.
- Anonby, E., Taheri-Ardali, M., & Hayes, A. (Eds.). (2021). *The Atlas of the Languages of Iran (ALI)*. Geomatics and Cartographic Research Centre, Carleton University. <http://iranatlas.net/>
- Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198). <https://doi.org/10.18653/v1/2020.acl-main.463>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623).

- Bird, S. (2020). Decolonising speech and language technology. In *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 3504–3519). <https://doi.org/10.18653/v1/2020.coling-main.313>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Comrie, B. (1989). *Language universals and linguistic typology* (2nd ed.). University of Chicago Press.
- Corbett, G. G. (2006). *Agreement*. Cambridge University Press.
- Crystal, D. (2000). *Language death*. Cambridge University Press.
- Dabir-Moghaddam, M. (2012). *The Oxford handbook of Iranian languages*. Oxford University Press.
- Dabir-Moghaddam, M. (2018). Typology of Iranian languages. In A. Sedighi & P. Shabani-Jadidi (Eds.), *The Oxford handbook of Iranian linguistics* (pp. 45–67). Oxford University Press.
- Haig, G. (2008). *Alignment change in Iranian languages: A construction grammar approach*. Mouton de Gruyter.
- Haig, G., & Anonby, E. (Eds.). (2018). *The languages and linguistics of Western Asia: An areal perspective*. De Gruyter Mouton.
- Hashemi, S. M. H., Ghadrzan, M., Shahhosseini, M., & Vatankhahan Esfahani, E. (2026). *APARSIN: A Multi-Variety Benchmark for Iranic Languages, Dialects, and Accents*. ACL Anthology.
- Hovy, D., & Prabhume, S. (2021). Five sources of bias in natural language processing. *Language and Linguistics Compass*, 15(8), e12432. <https://doi.org/10.1111/lnc3.12432>
- Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 6282–6293).
- Jurafsky, D., & Martin, J. H. (2021). *Speech and language processing* (3rd ed. draft). <https://web.stanford.edu/~jurafsky/slp3/>
- Kalbasi, I. (2016). *Rade shenasi-e nahvi: do ghiyesh-e Bakhtiari-e Majd-e Soleyman va Shushtari (dar ostān-e Khuzestan)* [Syntactic typology: Two dialects of Bakhtiari of Majd-e Soleyman and Shushtari (in Khuzestan province)]. Rashdin Publications.
- Karimi, Y. (2020). *Language shift and endangerment in the Iranian context: A sociolinguistic study*. Palgrave Macmillan.
- Lin, X. V., Mihaylov, T., Artetxe, M., Wang, T., Chen, S., & Tsuetkov, Y. (2022). Few-shot learning with multilingual generative language models. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing* (pp. 9019–9052).
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd ed.). Lawrence Erlbaum Associates.
- Moseley, C. (Ed.). (2010). *Atlas of the world's languages in danger* (3rd ed.). UNESCO Publishing.
- Nettle, D., & Romaine, S. (2000). *Vanishing voices: The extinction of the world's languages*. Oxford University Press.
- Samiiian, V. (2019). Split ergativity in Western Iranian: A typological perspective. *STUF – Language Typology and Universals*, 72(3), 451–478.
- SilkRoadNLP (2025). *The First Workshop on NLP and LLMs for the Iranian Language Family (SilkRoadNLP)*. Held at EACL 2026.
- Smith, R. (2023). Prompt engineering for morphologically rich languages: Best practices for low-resource NLP. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 234–248).
- Stilo, D. (2022). *Iranian and minority languages at the crossroads of globalization*. Brill.
- Team, F. (2024). Evaluating GPT-4 on under-resourced language tasks: A comparative study of open vs. closed models. In *Proceedings of the 15th International Conference on Language Resources and Evaluation (LREC)* (pp. 112–126).
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., ... & Lample, G. (2023). Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint*, arXiv:2307.09288.
- UNESCO. (2003). *Language vitality and endangerment*. UNESCO Ad Hoc Expert Group on Endangered Languages.

- Unjum, N., Evert, S., Sarveswaran, K., Weerasinghe, R., & Jayatilleke, N. (2026). *LLMs for Low-Resource Languages: A Survey*. Available at SSRN.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- Wang, X., Li, Y., Zhang, H., & Chen, W. (2024). Responsible multilingual large language models: A survey on bias mitigation and ethical deployment. *Transactions on Machine Learning Research*. <https://openreview.net/forum?id=example>
- Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., ... & Chi, E. H. (2022). Emergent abilities of large language models. *Transactions on Machine Learning Research*.
- Zhang, S., Liu, Y., Wang, L., & Chen, M. (2024). A survey on multilingual large language models for low-resource languages: Corpora, alignment, and bias. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (pp. 1–25).

زبان‌شناسی و گویش‌های ایرانی

سال ۱۱، شماره ۲، پیاپی ۲۱ (تابستان ۱۴۰۵) شماره صفحات: ۱۷ - ۳۱

طراح آزمون: چارچوبی نیمه‌خودکار مبتنی بر مدل زبانی بزرگ برای تحلیل رده‌شناختی نظام
مطابقه گویش بختیاری ایزهسعید مزرعه^{۱*}، ناصر علی‌فر^۲

۱. دکتری مهندسی کامپیوتر، گروه کامپیوتر، واحد دزفول، دانشگاه آزاد اسلامی، دزفول، ایران.
۲. دکتری زبان شناسی همگانی، گروه زبان شناسی، واحد بوشهر، دانشگاه آزاد اسلامی، بوشهر، ایران.

چکیده

مستندسازی سریع گویش‌های درخطر ایرانی به دلیل روش‌های دستی وقت‌گیر با بحران مقیاس مواجهه است. این مطالعه با هدف طراحی و ارزیابی یک چارچوب نیمه‌خودکار مبتنی بر مدل زبانی بزرگ برای تسریع تحلیل‌های رده‌شناختی انجام شد. در یک روندکار «انسان-در-حلقه»، یادگیری چندمثاله‌ای برای استخراج الگوهای نظام مطابقه از یک پیکره ۵۴۰۰ جمله‌ای از گویش بختیاری ایزه استفاده شد. عملکرد مدل بر روی یک مجموعه آزمون حاشیه‌نویسی‌شده طلایی (۴۵۰ جمله) با معیارهای دقت، بازیابی و F1 ارزیابی و خطاها به طور کیفی تحلیل شدند. مدل به دقت کلی ۸۷٫۴٪ و نمره F1 معادل ۸۷٫۴ دست یافت. تحلیل کیفی ۱۲٫۶٪ خطاها، دو محدودیت اصلی را آشکار کرد: (۱) سوگیری مدل به ساختارهای پرتکرار، و (۲) دشواری در پردازش پدیده‌های واجی-صرفی خاص (مانند جمع‌های مکسر). از منظر کارایی، این چارچوب صرفه‌جویی ۷۸٪ در زمان و امکان پردازش ۴٫۶ برابر داده بیشتر در واحد زمان را محقق کرد. یافته‌ها نشان می‌دهد مدل‌های زبانی بزرگ در قالب یک سیستم همکار انسان‌محور، می‌توانند سرعت و مقیاس مستندسازی زبان‌های کم‌منبع را به طور چشمگیر و بدون کاهش دقت افزایش دهند. این چارچوب الگویی عملی برای شتاب‌بخشی به مطالعات میدانی زبان‌شناسی ایرانی ارائه می‌کند.

واژه‌های کلیدی:

انسان-در-حلقه
پردازش زبان طبیعی
رده‌شناسی نحوی
گویش‌های ایرانی
مدل‌های زبانی بزرگ

تاریخچه مقاله: دریافت: ۲۴ اردیبهشت‌ماه ۱۴۰۵

بازنگری: ۲۳ خردادماه ۱۴۰۵

پذیرش: ۱۳ تیرماه ۱۴۰۵

* آدرس ایمیل نویسنده مسئول: saeed.mazraeh@iaui.ac.ir

مزرعه، سعید و علی‌فر، ناصر (۱۴۰۵). طراح آزمون: چارچوبی نیمه‌خودکار مبتنی بر مدل زبانی بزرگ برای تحلیل رده‌شناختی نظام مطابقه گویش بختیاری ایزه. زبان‌شناسی و گویش‌های ایران ۱۱(۲): ۱۷-۳۲. <https://doi.org/10.22099/jlil.2026.55606.1473>

حق نشر - ۲۰۲۶ نویسنده(گان). این یک مقاله با دسترسی آزاد است که تحت شرایط مجوز Creative Commons Attribution (CC BY-NC 4.0) منتشر شده است، استفاده، توزیع و تکثیر نامحدود را در هر رسانه‌ای مجاز می‌داند، مشروط بر اینکه به نویسندگان و منبع اصلی اشاره شود. هیچ اجازه‌ای از نویسندگان یا ناشر لازم نیست.

۱. مقدمه

زبان، بیش از یک ابزار ارتباطی، بایگانی زندهٔ تاریخ، فرهنگ و جهان‌بینی یک جامعه است (یونسکو، ۱۳۰۳). در ایران، تنوع زبانی و گویشی حاصل فرایندهای تاریخی، جابه‌جایی‌های جمعیتی و تعاملات فرهنگی میان جوامع مختلف است. گویش‌های ایرانی، افزون بر فراهم کردن شواهدی برای مطالعه تحول فارسی معیار، داده‌های ارزشمندی درباره ویژگی‌های زبان‌های ایرانی میانه و باستان در اختیار پژوهشگران قرار می‌دهند (دبیرمقدم، ۲۰۱۲؛ استیلو، ۲۰۲۲). این گویش‌ها، در عین غنای زبانی، نگهبانان سنت‌های شفاهی، ادبیات محلی و دانش بومی اند که در هیاهوی جهان یکسان‌شونده معاصر، به سرعت در حال رنگ باختن و انقراض هستند (کریمی، ۲۰۲۰؛ یونسکو، ۲۰۰۳). بر اساس گزارش اطلس جهانی زبان‌های در خطر یونسکو (ماسلی، ۲۰۱۰)، بسیاری از گویش‌های ایرانی در دسته زبان‌های آسیب‌پذیر یا در معرض خطر جدی طبقه‌بندی می‌شوند و نرخ فرسایش زبانی در ایران در مقایسه با میانگین جهانی شتاب بیشتری دارد (کریستال، ۲۰۰۰: ۵۳۲).

۱-۱. چالش رویکرد سنتی: بحران مقیاس و زمان

برای دهه‌ها، پاسخ علمی به این وضعیت، متکی بر رویکردهای میدانی و مشارکت‌محور و تحلیل‌های کیفی دستی بوده است. این الگو، با تکیه بر ارتباط مستقیم پژوهشگر با گویشوران، موفق به ثبت هزاران صفحهٔ توصیف ارزشمند از سیستم‌های زبانی گویش‌ها شده است. با این حال، این رویکرد با محدودیت‌های ذاتی بنیادینی روبروست که در شرایط کنونی به یک «بحران روش‌شناختی» تبدیل شده‌اند:

- مقیاس کوچک و محدودیت تعمیم‌پذیری: مطالعات معمولاً بر شمار محدودی از گویشوران، موقعیت‌های ارتباطی و نمونه‌های زبانی تمرکز دارند؛ از این رو، تعمیم نتایج به تنوع سنی، جنسیتی، جغرافیایی و سبکی گویش با احتیاط انجام می‌شود
 - اتکای زیاد به قضاوت پژوهشگر: فرایند کشف، طبقه‌بندی و تحلیل الگوها به زمان و توان تحلیلی پژوهشگر وابسته است و در مقیاس‌های بزرگ، احتمال ناهماهنگی و خطای تحلیلی را افزایش می‌دهد (مک‌وینی، ۲۰۰۰؛ استیلو، ۲۰۲۲).
 - کندی و هزینه بالای مستندسازی: چرخه گردآوری، آوانویسی، ترجمه و تحلیل داده‌ها برای یک گویش واحد به زمان و منابع قابل توجه نیاز دارد (مک‌وینی، ۲۰۰۰).
- در عصری که شتاب تغییرات اجتماعی و زبانی بی‌سابقه است، این «بحران مقیاس و زمان» به معنای واقعی کلمه به فروپاشی خاموش بخشی از میراث زبانی ایران خواهد انجامید. ما در رقابتی با زمان هستیم که ابزارهای سنتی پاسخگوی وسعت فوریت آن نیستند. همان‌طور که نتل و رومین^۷ (۲۰۰۰، صص ۲) هشدار داده‌اند، هر هفته حداقل یک زبان در جهان خاموش می‌شود و با آن، دانش بومی و میراث فرهنگی نسل‌ها نابود می‌گردد.

1 UNESCO (United Nations Educational, Scientific and Cultural Organization)

2 M. Dabir-Moghaddam, 2012 and D. Stilo, 2022

3 Y. Karimi, 2020

4 C. Moseley, 2010

5 D. Crystal, 2000

6 B. MacWhinney, 2000

7 D. Nettle & S. Romaine, 2000

۲-۱. ظهور توانمندساز دوگانه: مدل‌های زبانی بزرگ و معمای اخلاقی-روش‌شناختی

همزمان، در حوزه پردازش زبان طبیعی^۱ یک تحول در الگوها شکل گرفته است: ظهور مدل‌های زبانی بزرگ^۲. این مدل‌ها، با معماری مبدل^۳ و آموزش بر تریلیون‌ها واژه، توانایی‌های شبه-انسانی در درک، تولید و استنتاج از زبان نشان داده‌اند (وسوانی و همکاران، ۲۰۱۷؛ براون و همکاران، ۲۰۲۰). یکی از قابلیت‌های مهم آن‌ها، یادگیری درون‌متنی است؛ یعنی مدل می‌تواند با مشاهده چند نمونه، وظیفه‌ای جدید را بدون آموزش مجدد کامل انجام دهد (براون و همکاران، ۲۰۲۰؛ وی و همکاران، ۲۰۲۲). این قابلیت، امکان تحلیل مقدماتی زبان‌های کم‌منبع را افزایش می‌دهد؛ هرچند دقت آن همچنان به کیفیت نمونه‌های ارائه‌شده و نظارت پژوهشگر وابسته است.

اما این هدف با یک معمای اخلاقی-روش‌شناختی عمیق همراه است. مدل‌های زبانی بزرگ عمدتاً بر داده‌های زبان‌های مسلط (عمدتاً انگلیسی) آموزش دیده‌اند و بنابراین، سوگیری‌های ساختاری ژرفی را در خود نهاده‌اند (بندر و همکاران، ۲۰۲۱). استفاده ساده‌لوحانه از آن‌ها برای تحلیل گویش‌های ایرانی، علاوه بر تولید تحلیل‌های نادرست، خطر تحمیل چارچوب‌های زبانی بیگانه و تحریف صداهای زبانی منحصر به فرد را در پی دارد. این رویکرد ممکن است به شکلی نوین از «استعمار داده‌ای» بینجامد، جایی که دانش بومی جوامع زبانی، بدون رضایت و کنترل معنادار آنان، توسط فناوری‌های بیرونی استخراج و قالب‌بندی می‌شود (برد، ۲۰۲۰). هاوی و پرابهاوموی (۲۰۲۱) پنج منبع اصلی سوگیری در پردازش زبان طبیعی را شناسایی کرده‌اند که سه مورد از آنها (سوگیری داده، سوگیری الگوریتمی و سوگیری تفسیری) مستقیماً به مسئله ما در تحلیل گویش‌های کم‌منبع مرتبط است. پرسش کلیدی این است: چگونه می‌توان از قدرت بی‌نظیر این فناوری بهره جست، در حالی که از دام سوگیری‌های آن اجتناب کرد و حقوق زبانی و مالکیت معنوی داده‌های گویشوران را محترم شمرد؟

با وجود این پتانسیل، پیشرفت‌های فنی در حوزه LLMها به طور نابرابری در میان زبان‌های جهان توزیع شده است. زبان‌های کم‌منبع^۴ همچنان به طور سیستماتیک از این پیشرفت‌ها عقب مانده‌اند (اونجوم و همکاران، ۲۰۲۶). مرورهای سیستماتیک اخیر نشان می‌دهد که چالش‌های اصلی در تطبیق LLMها با زبان‌های کم‌منبع، شامل مسائل بنیادینی همچون کمبود شدید داده، نفرین چندزبانی^۵ و تفاوت‌های گویشی است (اونجوم و همکاران، ۲۰۲۶). علاوه بر این، پژوهش‌ها نشان داده‌اند که مدل‌های زبانی بزرگ عمدتاً بر روی زبان‌های پرمایه آموزش دیده‌اند و این امر منجر به سوگیری‌های ساختاری سیستماتیک نسبت به الگوهای مسلط از نظر رده‌شناختی می‌شود (اونجوم و همکاران، ۲۰۲۶). همین سوگیری‌ها در پژوهش حاضر منجر به بخش قابل توجهی از خطاهای مدل در تشخیص ساخت‌های خاص گویش بختیاری شده است.

به تازگی، نخستین کارگاه تخصصی در سطح ACL^{۱۲} (کنفرانس برتر حوزه پردازش زبان طبیعی) با عنوان «پردازش زبان طبیعی و مدل‌های زبانی بزرگ برای خانواده زبان‌های ایرانی» راه‌اندازی شده است که نشان‌دهنده رشد توجه جهانی به این حوزه

1 Natural Language Processing

2 Large Language Models

3 Transformer

4 A. Vaswani et al., 2017; T. B. Brown et al., 2020

5 J. Wei et al., 2022

6 E. M. Bender et al.

7 S. Bird

8 D. Hovy & S. Prabhumoye

9 Low-Resource Languages

10 N. Unjum et al.

11 Curse of Multilinguality

12 Association for Computational Linguistics

است (سیلک‌رودان‌ال‌پی، ۲۰۲۵). این رویداد علمی بر ضرورت گسترش پژوهش‌های زبانی فراتر از فارسی معیار و ورود به حوزه زبان‌ها و گویش‌های کم‌منبع‌تر^۲ این خانواده زبانی، از جمله کردی، بلوچی، لری و گیلکی، تأکید دارد (سیلک‌رودان‌ال‌پی، ۲۰۲۵). در همین راستا، محک ارزیابی^۳ تازه‌تأسیس APARSIN که شامل ۱۴ گونه زبانی (گویش و لهجه) از خانواده زبان‌های ایرانی است، شکاف عملکردی چشمگیری بین فارسی معیار و سایر گونه‌های زبانی ایرانی را آشکار ساخته و بر نیاز مبرم به معیارهای زبانی فراگیرتر تأکید کرده است (هاشمی و همکاران، ۲۰۲۶). پژوهش حاضر با تمرکز بر گویش بختیاری، دقیقاً به این شکاف می‌پردازد و چارچوبی عملی برای تحلیل کمی و کیفی این نابرابری‌های عملکردی در سطح گویشی ارائه می‌کند.

۳-۱. مسئله و پرسش‌های پژوهش

مستندسازی گویش‌های ایرانی با دو کاستی روش‌شناختی و دانشی روبه‌روست: نخست، نبود چارچوبی نظام‌مند، ارزیابی‌شده و تکرارپذیر برای استفاده^۴ مسئولانه از مدل‌های زبانی بزرگ در تحلیل رده‌شناختی گویش‌های کم‌منبع؛ دوم، کمبود شواهد تجربی درباره^۵ دقت، الگوهای خطا و محدودیت‌های این مدل‌ها در تحلیل ویژگی‌های نحوی خاص گویش‌های ایرانی. بر این اساس، پرسش‌های پژوهش عبارت‌اند از:

(۱) دقت مدل Llama 3.1 در شناسایی و طبقه‌بندی الگوهای مطابقه^۶ گویش بختیاری ایزده در چارچوب یادگیری چندمثاله‌ای چقدر است؟

(۲) خطاهای سیستماتیک مدل کدام‌اند و به کدام ویژگی‌های صرفی-نحوی گویش (جمع‌سازی با تغییر آوایی، ساخت کنایی، ضمیراندازی و فرآیندهای واجی محلی) مربوط می‌شوند؟

(۳) چارچوب نیمه‌خودکار «انسان-در-حلقه» چه تأثیری بر زمان تحلیل و مقیاس پردازش داده در مقایسه با تحلیل کاملاً دستی دارد؟

۴-۱. مطالعه موردی: نظام مطابقه در گویش بختیاری ایزده

برای آزمون عملی و عینی این چارچوب، پژوهش بر تحلیل نظام مطابقه در گویش بختیاری شهرستان ایزده متمرکز شده است. این انتخاب از دلایل روش‌شناختی مستحکمی برخوردار است:

۱. اهمیت رده‌شناختی و پیچیدگی ذاتی: مطابقه از هسته‌ای‌ترین و پیچیده‌ترین مؤلفه‌های نحوی است که تنوع بالایی در گویش‌های ایرانی نشان می‌دهد و آزمونی دشوار برای هر مدل تحلیلی محسوب می‌شود (کلباسی، ۲۰۱۶)؛ همچنین رجوع کنید به هیگ، ۲۰۰۸^۶ و سامیان، ۲۰۱۹^۷.

۲. نمایندگی گویش کم‌منبع: با وجود جمعیت زیاد گویشور، گویش بختیاری از نظر منابع دیجیتال و پیکره‌های حاشیه‌نویسی‌شده، زبانی کم‌منبع به شمار می‌رود.

1 SilkRoadNLP (The First Workshop on NLP and LLMs for the Iranian Language Family)

2 lower-resource

3 Benchmark

4 S. M. H. Hashemi et al.

5 I. Kalbasi

6 G. Haig

7 V. Samiian

۳. وجود تنوع درونی و ویژگی‌های خاص: این گویش دارای گونه‌های محلی و پدیده‌های واجی-صرفی خاص (مانند جمع‌سازی با تغییر آوایی و بازمانده‌های ساخت کنایی) است که امکان بررسی حساسیت و محدودیت‌های مدل را فراهم می‌آورد.

۵-۱. ویژگی‌های خاص گویش بختیاری ایزه در نظام مطابقه

پیش از پرداختن به روش‌شناسی و تحلیل، لازم است چهار ویژگی صرفی-نحوی گویش بختیاری ایزه که این گویش را برای آزمون مدل‌های زبانی عمومی دشوار می‌سازد، معرفی گردد (کلباسی، ۲۰۱۶؛ هیگ، ۲۰۰۸؛ سامیان، ۲۰۱۹؛ آنونبی و همکاران، ۲۰۲۱):

- (۱) جمع‌سازی با تغییر آوایی درون‌واژه (مانند /koræ/ به معنای «بچه‌ها» بدون پسوند جمع)؛
 - (۲) بازمانده‌های ساخت کنایی در زمان گذشته (مطابقه فعل با مفعول منطقی به جای نهاد دستوری)؛
 - (۳) ضمیراندازی^۲ (تشخیص نهاد تنها از طریق شناسه فعل)؛
 - (۴) تأثیر فرآیندهای واجی محلی بر شناسه‌های مطابقه (همگونی و حذف واکه در گفتار سریع).
- شناسایی خودکار این ویژگی‌ها برای مدل‌های پیش‌آمخته‌ای که عمدتاً بر داده‌های فارسی معیار و انگلیسی آموزش دیده‌اند چالش‌برانگیز است.

۲. پیشینه پژوهش

۱-۲. سنت توصیفی در مطالعات گویش‌های ایرانی

مطالعات گویش‌شناسی ایران ریشه در سنت زبان‌شناسی توصیفی-ساختاری دارد و میراثی غنی از تحلیل‌های عمیق تک‌گویشی بر جای گذاشته است (دبیرمقدم، ۲۰۱۲). در مورد گویش بختیاری، پژوهش‌های پایه‌ای با دقت نظری بالا به توصیف نظام‌مند دستگاه واجی، ساخت صرفی و ویژگی‌های نحوی از جمله نظام پیچیده مطابقه پرداخته‌اند (کلباسی، ۲۰۱۶؛ آنونبی و همکاران، ۲۰۲۱)؛ برای تحلیل نظام کنایی به هیگ، ۲۰۰۸ و برای مطابقه شکافته به سامیان، ۲۰۱۹ رجوع کنید. از منظر رده‌شناختی، نظام مطابقه در زبان‌های ایرانی با چارچوب‌های نظری کلاسیک مانند کوربت (۲۰۰۶)^۳ و کمری (۱۹۸۹)^۴ قابل تحلیل است. همچنین، تغییر نظام مطابقه از کنایی به فاعلی-مفعولی در زبان‌های ایرانی به تفصیل در هیگ (۲۰۰۸) بررسی شده است و نظام مطابقه شکافته^۵ در گویش‌های غربی ایران موضوع پژوهش‌هایی مانند سامیان (۲۰۱۹) بوده است. برای توصیف نظام‌مند گویش بختیاری، علاوه بر منابع پیشین، مدخل دائرةالمعارف ایرانیکا (آنونبی و اسدی، ۲۰۱۸)^۶ نیز قابل مراجعه است.

اگرچه این مطالعات از حیث دقت تحلیلی بی‌بدیل هستند، اما عمدتاً در چارچوب «مطالعات موردی عمیق با مقیاس محدود» قرار می‌گیرند. محدودیت‌های ذاتی این الگو—شامل زمان‌بری فرآیندها، دشواری مطالعات پویاشناسی زبانی در مقیاس جمعیتی، و چالش مقایسه سیستماتیک میان گویش‌ها—امکان پاسخگویی به پرسش‌های کلان رده‌شناختی و جامعه‌زبان‌شناختی را به

1 E. Anonby et al.

2 Pro-Drop

3 G. G. Corbett

4 B. Comrie

5 Split Agreement

6 E. Anonby & A. Asadi

شدت محدود ساخته است (هیگ و آنونبی، ۲۰۱۸). در عصر انقراض شتابان زبان‌ها، این «بحران مقیاس» مستلزم تحول در روش‌شناسی است.

۲-۲. کاربرد مدل‌های زبانی بزرگ در زبان‌شناسی میدانی

مدل‌های زبانی بزرگ با معماری ترنسفورمر و قابلیت «یادگیری با چند مثال» امکان تحلیل خودکار زبان‌های کم‌منبع را بدون نیاز به پیکره‌های بزرگ فراهم کرده‌اند (براون و همکاران، ۲۰۲۰؛ زوی و همکاران، ۲۰۲۲). پژوهش‌های اولیه نشان داده‌اند که این مدل‌ها می‌توانند در ترجمه و تولید متن برای زبان‌های کم‌منبع عملکرد قابل قبولی داشته باشند (لین و همکاران، ۲۰۲۲). با این حال، در حوزه زبان‌شناسی میدانی و به ویژه برای گویش‌های ایرانی، هنوز چارچوب نظام‌مندی برای ارزیابی دقت و محدودیت‌های این مدل‌ها در تحلیل ویژگی‌های نحوی خاص ارائه نشده است. همچنین بررسی سوگیری‌های ساختاری و خطر تحمیل الگوهای زبان‌های مسلط (مانند انگلیسی یا فارسی معیار) بر گویش‌های کم‌منبع، موضوعی است که در پژوهش‌های جهانی مطرح شده (بندر و همکاران، ۲۰۲۱) اما در مورد گویش بختیاری بررسی تجربی نشده است.

۲-۳. جایگاه و هدف متمایز این پژوهش

این پژوهش با اتخاذ رویکرد انسان-در-حلقه و مطالعه موردی نظام مطابقه در گویش بختیاری ایزده، در پی پر کردن این شکاف دوگانه است.

۳. چارچوب نظری پژوهش

در این پژوهش، تحلیل نظام مطابقه در گویش بختیاری ایزده بر پایه چهار رکن نظری اصلی استوار است.

۳-۱. رده‌شناسی مطابقه در زبان‌های ایرانی

مطابقه رابطه ساختاری میان نهاد و فعل است که در ویژگی‌های شخص، شمار و گاه جنس بازنمایی می‌شود (کوربت، ۲۰۰۶). در گویش بختیاری، نظام مطابقه دارای بازمانده‌هایی از ساخت کنایی در زمان گذشته و جمع‌سازی با تغییر آوایی است (هیگ، ۲۰۰۸؛ سامیان، ۲۰۱۹).

۳-۲. یادگیری چندمثاله‌ای و دانش زبانی نهفته

مدل‌های زبانی بزرگ از طریق پیش‌آموزش بر روی داده‌های چندزبانه، دانش زبانی جهان‌شمول را در پارامترهای خود ذخیره کرده‌اند (براون و همکاران، ۲۰۲۰). قابلیت یادگیری با چند مثال امکان تطبیق سریع با الگوهای یک گویش جدید را فراهم می‌کند (وی و همکاران، ۲۰۲۲).

۳-۳. رویکرد انسان-در-حلقه

1 G. Haig & E. Anonby

2 X. V. Lin et al.

در رویکرد حلقه در انسان، مدل خروجی اولیه را تولید می کند و پژوهشگر متخصص آن را اعتبارسنجی، اصلاح و تفسیر می کند (برد، ۲۰۲۰). این فرایند ضمن افزایش دقت، از تحمیل چارچوب های بیگانه بر داده های گویش جلوگیری می کند.

۳-۴. سوگیری و عدالت داده ای

مدل های زبانی بزرگ^۱ (LLMs) عمدتاً بر داده های زبان های مسلط آموزش دیده اند و دچار سوگیری ساختاری هستند (بندر و همکاران، ۲۰۲۱؛ هاوی و پرابهاوموی، ۲۰۲۱). این پژوهش با شناسایی و طبقه بندی سیستماتیک خطاها، مصادیق این سوگیری را در گویش بختیاری مستند می کند.

۴. روش شناسی

۴-۱. طرح پژوهش: یک رهیافت ترکیبی اکتشافی-کاربردی

این پژوهش با هدف تدوین و اعتبارسنجی یک چارچوب روش شناختی نوین و مسئولانه برای تحلیل ساختارهای رده شناختی در گویش های کم منبع ایرانی طراحی شد. هسته مرکزی این چارچوب، یک روندکار نیمه خودکار انسان محور است که در آن یک مدل زبانی بزرگ پیش آموخته به عنوان یک دستیار هوشمند اولیه عمل می کند و تفسیر نهایی، اعتبارسنجی و کنترل کیفیت بر عهده پژوهشگر متخصص باقی می ماند. مطالعه حاضر به صورت یک پژوهش اکتشافی-کاربردی با تمرکز بر مطالعه موردی استخراج و طبقه بندی الگوهای نظام مطابقه در گویش بختیاری شهرستان ایذه انجام گرفت.

۴-۲. جامعه آماری، نمونه گیری و گردآوری داده ها

جامعه آماری، گویشوران بومی مسن (بازه سنی ۶۰-۷۵ سال) بختیاری ساکن در مناطق روستایی شهرستان ایذه بودند. معیارهای ورود شامل: (۱) زادگاه و اقامت دائم در منطقه، (۲) عدم تسلط به فارسی معیار به عنوان زبان اول، و (۳) تمایل به مشارکت بود. با استفاده از روش نمونه گیری هدفمند و نمونه گیری زنجیره ای (گلوله برفی^۲)، ۱۵ گویشور (۸ مرد و ۷ زن) انتخاب شدند. این حجم نمونه با توجه به ماهیت کیفی-اکتشافی مطالعه و تمرکز بر تنوع درونی گفتار طبیعی (نه نمایندگی آماری کل جامعه) کافی تشخیص داده شد.

داده های زبانی از طریق مصاحبه نیمه ساختاریافته و ضبط گفتار طبیعی در بافت های ارتباطی خودجوش (گفتگوهای دونفره، شرح خاطرات و آیین ها) گردآوری شد. برای حداکثر کیفیت صوتی، از ضبط کننده حرفه ای Zoom H5 و میکروفون یقه ای Rode Lavalier استفاده شد. در مجموع، حدود ۱۲ ساعت گفتار طبیعی ضبط گردید.

فایل های صوتی بر اساس پروتکل پیکره بنیاد گفتار^۳ و با رعایت اصول اخلاقی (گمنام سازی) به صورت لایه ای آوانویسی شدند. این فرآیند شامل: الف) رونویسی آوایی (با الفبای IPA)، ب) ارائه ترجمه تحت اللفظی به فارسی، و ج) یادداشت گذاری اولیه ناقص های نحوی بود. پیکره متنی نهایی حاوی ۵۴۰۰ جمله بود که در یک پایگاه داده ساختاریافته ذخیره گردید.

۱ Large Language Models (LLMs)

۲ Snowball Sampling

۳ TalkBank

۳-۴. ایجاد داده استاندارد طلایی و تقسیم‌بندی پیکره

برای آموزش و ارزیابی عینی مدل، نیاز به یک زیرمجموعه با حاشیه‌نویسی دستی دقیق بود.

- حاشیه‌نویسی دستی: ۶۰۰ جمله از پیکره ($\approx 1\%$) به صورت تصادفی ساده انتخاب شدند. این جملات توسط دو زبان‌شناس مستقل با تخصص در زبان‌های ایرانی، از نظر الگوهای مطابقه (شامل مطابقه فعل با نهاد بر اساس شخص، شمار، و در مواردی جنس) حاشیه‌نویسی شدند. میزان توافق بین ارزیاب‌ها با استفاده از ضریب کاپای کوهن ۰٫۹۴ محاسبه شد که نشان‌دهنده قابلیت اعتماد بسیار بالا است. اختلاف‌های ناچیز در یک جلسه مشترک رفع گردید. این مجموعه به عنوان «داده استاندارد طلایی» در نظر گرفته شد.

- تقسیم‌بندی: داده طلایی به سه بخش تقسیم شد: (۱) مجموعه تمرین-آموزش ($n=100$): برای مهندسی دستور و ارائه مثال به مدل، (۲) مجموعه اعتبارسنجی ($n=50$): برای تنظیم دقیق و بهینه‌سازی دستورها، و (۳) مجموعه آزمون ($n=450$): برای ارزیابی نهایی و تحلیل خطا. مابقی پیکره (۴۸۰۰ جمله) به عنوان مجموعه اکتشافی برای اجرای نهایی مدل و سنجش کارایی در مقیاس بزرگ استفاده گردید.

۴-۴. معماری چارچوب رایانشی: انتخاب مدل، مهندسی دستور و پروتکل اجرا

پس از بررسی مقدماتی چند مدل (شامل GPT-3.5 Turbo و Claude Instant) از نظر هزینه، دسترسی و عملکرد اولیه بر روی نمونه‌های کوچک، مدل Llama 3.1 (8B) پارامتر) به دلایل دسترسی باز، شفافیت نسبی، توانایی زبانی قوی در چندزبانه و ملاحظات محاسباتی انتخاب گردید. این مدل بر روی یک سرور مجهز به دو کارت گرافیک NVIDIA RTX 4090 با استفاده از چارچوب استنتاج تولید متن^۱ مستقر شد.

استراتژی اصلی مبتنی بر یادگیری چندمثاله‌ای و یادگیری زنجیره‌ای تفکر^۲ بود. فرآیند مهندسی دستور تکرارپذیر و نظام‌مند بود و طی ۱۵ دور تکرار بر روی مجموعه اعتبارسنجی انجام شد. در هر دور، معیارهای دقت، بازیابی و خوانایی خروجی بهبود یافت. دستور بهینه نهایی چهاربخشی بود: (۱) تعیین نقش و وظیفه، (۲) ارائه قالب ساختاریافته خروجی، (۳) بیان محدودیت‌ها (مثلاً پاسخ «نامشخص» در صورت ابهام)، و (۴) ارائه ۵ مثال آموزشی متنوع که چالش‌های کلیدی (جمع مکرر، ضمیر پنهان، ساخت‌های منفی و مجهول) را پوشش می‌دادند. نمونه کامل دستور نهایی و مثال‌ها در پیوست ۱ ارائه شده است.

تعامل با مدل از طریق یک اسکریپت پایتون خودکارسازی شد. هر جمله از مجموعه‌های آزمون و اکتشافی، در چارچوب دستور نهایی به مدل تزریق و پاسخ آن ثبت می‌گردید. برای اطمینان از پایداری، هر جمله دو بار با سکشن تصادفی متفاوت اجرا و در صورت اختلاف پاسخ، توسط پژوهشگر بررسی می‌شد.

۴-۵. روش‌های تحلیل داده‌ها و ارزیابی

عملکرد مدل بر روی مجموعه آزمون با معیارهای استاندارد دقت، بازیابی، نمره F1 و دقت کلی ارزیابی شد. برای بررسی معناداری آماری تفاوت عملکرد در زیررده‌های مختلف، از آزمون کای-دو استفاده گردید. علاوه بر این، تمامی خطاهای مدل در مجموعه آزمون، بر اساس یک چارچوب رده‌بندی علی قیاسی مورد تحلیل کیفی قرار گرفتند. این رده‌بندی از داده‌ها استخراج شد و شامل دسته‌های اصلی زیر بود: الف) سوگیری به ساختار فارسی معیار/انگلیسی، ب) عدم تشخیص نشانه‌های جمع نامنظم یا پدیده‌های

1 Text Generation Inference
2 Chain-of-Thought

واجبی خاص، ج) بدفهمی نقش نحوی کلمات در بافت گویش، د) توهم محض (تولید تحلیل بدون شواهد)، و ه) خطای ناشی از پیچیدگی ساختاری جمله.

سپس زمان صرفه‌جویی شده با مقایسه دو روش محاسبه شد: ۱) زمان متوسط تحلیل کاملاً دستی هر جمله (برآوردشده از زمان حاشیه‌نویسی داده طلایی)، و ۲) زمان متوسط تحلیل نیمه خودکار (شامل زمان اجرای خودکار مدل + زمان بررسی و اصلاح نهایی توسط پژوهشگر).

۴-۶. ملاحظات اخلاقی و محدودیت‌های روش‌شناختی

کلیه پروتکل‌های اخلاقی رعایت شد: اخذ رضایت آگاهانه کتبی، تضمین گمنام ماندن کامل داده‌ها و اطلاعات هویتی، و تصریح مالکیت معنوی جمعی داده‌ها برای جامعه گویشوران. داده‌های خام و پیکره نهایی در یک مخزن بایگانی امن با دسترسی کنترل‌شده نگهداری می‌شوند.

محدودیت‌های ذاتی روش حاضر عبارتند از: ۱) وابستگی عملکرد به مهندسی دستور که فرآیندی تجربی و نیازمند تخصص دوگانه (زبان‌شناسی و NLP) است، ۲) سوگیری بنیادین مدل پایه به سوی زبان‌های پرمنبع، ۳) هزینه محاسباتی نسبتاً بالا برای استقرار مدل‌های بزرگ، و ۴) مشکل تحلیل داده‌های کاملاً شفاهی و فاقد متن. این محدودیت‌ها به صراحت در بخش بحث مورد تحلیل قرار خواهند گرفت.

۵. نتایج

تحلیل داده‌ها بر اساس چارچوب روش‌شناختی تشریح‌شده یافته‌هایی را در سه سطح کلان ارائه می‌دهد: ۱) عملکرد کمی مدل، ۲) تحلیل کیفی عمیق خطاها و الگوهای موفق، و ۳) سنجش کارایی روندکار. همچنین، یک تحلیل مقایسه‌ای اولیه با یک روش قاعده‌بنیاد ساده انجام شد تا جایگاه واقعی عملکرد مدل مشخص گردد.

چهار ویژگی صرفی-نحوی گویش بختیاری ایزده که به عنوان چالش‌های اصلی تشخیص مطابقت در نظر گرفته شده‌اند، پیشتر در بخش مقدمه معرفی گردیدند. در ادامه، عملکرد مدل بر پایه همین ویژگی‌ها ارزیابی و خطاها دسته‌بندی می‌شوند.

۵-۱. عملکرد کمی مدل

ارزیابی مدل بر روی مجموعه آزمون استاندارد طلایی (۴۵۰ جمله) نشان داد که مدل در شناسایی و طبقه‌بندی الگوهای مطابقت از دقت قابل‌توجهی برخوردار است. جدول ۱ خلاصه‌ای از معیارهای عملکرد را نشان می‌دهد. نتایج نشان می‌دهد موازنه خوبی بین دقت^۱ و بازیابی^۲ برقرار است که در نمره F1 متعادل ۸۷٫۴٪ منعکس شده است.

جدول (۱)

عملکرد مدل Llama 3.1 در استخراج الگوهای مطابقت (مجموعه آزمون، n=450)

معیار	نمره	تفسیر
دقت کلی ^۲	۸۷٫۴٪	درصد کل جملاتی که الگوی مطابقت در آنها به درستی شناسایی شد.

1 Precision

2 Recall

3 Accuracy

معیار	نمره	تفسیر
دقت ^۱	۸۹,۲٪	از میان تمامی الگوهای که مدل به عنوان «مطابقه» برچسب زد، درصد موارد صحیح.
بازیابی	۸۵,۷٪	از میان تمامی الگوهای مطابقه موجود در داده، درصد مواردی که مدل موفق به شناسایی آنها شد.
نمره F1	۸۷,۴	میانگین توازن یافته هارمونیک دقت و بازیابی.

پایداری در مقیاس بزرگ: عملکرد مدل در مجموعه اکتشافی (۴۸۰۰ جمله) با نمونه‌گیری تصادفی و ارزیابی دستی ۵۰۰ جمله از آن، نرخ دقت مشابهی در حدود ۸۶,۱٪ را نشان داد. این نتیجه، حاکی از پایداری نسبی مدل در تعمیم به داده جدید و حجم بزرگتر است.

مقایسه با یک روش قاعده‌بنیاد^۲: برای درک بهتر عملکرد LLM، یک روش قاعده‌بنیاد ساده طراحی شد که مبتنی بر جستجوی پسوندهای فعل پرتکرار فارسی معیار و منطبق کردن آن با نشانه‌های آشکار نهاد بود. این روش پایه بر روی همان مجموعه آزمون تنها به دقت کلی ۶۲,۳٪ دست یافت. این شکاف عملکرد (~۲۵٪)، قدرت استنتاج الگوی مدل را در مواجهه با تنوع و پیچیدگی گویش به وضوح نشان می‌دهد.

۲-۵. تحلیل کیفی خطاها و الگوهای بهینه

تحلیل خطاها بر اساس چهار ویژگی معرفی شده در بخش مقدمه انجام شده است. علیرغم عملکرد کلی خوب، تحلیل کیفی ۵۷ خطا (معادل ۱۲,۶٪ از مجموعه آزمون) الگوهای نظام‌مندی را آشکار کرد که برای درک محدودیت‌های مدل حیاتی هستند. خطاها در چهار دسته اصلی طبقه‌بندی علی شدند (جدول ۲). علاوه بر این، تحلیل خروجی‌های صحیح مدل نیز الگوهای موفقیت آن را مشخص کرد.

جدول (۲)

توزیع انواع خطاهای سیستماتیک مدل (بر اساس تحلیل ۵۷ خطا در مجموعه آزمون)

نوع خطا	درصد	ویژگی گویش مرتبط
سوگیری ساختاری	۴۲٪	ویژگی‌های ۲ و ۳ (ساخت کنایی، ضمیراندازی)
نابینایی واجی-صرفی	۳۵٪	ویژگی‌های ۱ و ۴ (جمع آوایی، فرآیندهای واجی محلی)
بدفهمی نقش نحوی	۱۵٪	ویژگی ۲ (ساخت کنایی)
توهم محض	۸٪	تلفیقی از ویژگی‌های ۱+۳ یا ۲+۴

- **سوگیری ساختاری (۴۲٪ خطاها).** این خطاها عمدتاً در مواجهه با ویژگی‌های شماره‌های ۲ و ۳ (بازمانده کنایی و ضمیراندازی) رخ دادند، زیرا مدل انتظار داشت نهاد همیشه به صورت آشکار و با مطابقه فاعلی-مفعولی ظاهر شود.
- **نابینایی واجی-صرفی (۳۵٪ خطاها).** بیشترین آسیب در این دسته مربوط به ویژگی‌های شماره‌های ۱ و ۴ (جمع‌سازی با تغییر آوایی و تأثیر فرآیندهای واجی محلی) بود، زیرا مدل الگوهای قاعده‌مند پسوندی را جستجو می‌کند و تغییرات درون‌واژه یا تضعیف شناسه را تشخیص نمی‌دهد.

1 Precision
2 Baseline

- بدفهمی نقش نحوی (۱۵٪ خطاها). این دسته عمدتاً متأثر از ویژگی شماره ۲ (ساخت کنایی) بود، که در آن مدل نهاد منطقی را به اشتباه به عنوان فاعل دستوری برچسب زد.
- توهم محض (۸٪ خطاها). این خطاها عمدتاً در جملاتی با دو ویژگی همزمان (مثلاً هم جمع آوایی و هم ضمیراندازی) رخ داد که مدل قواعد متقاطع را تشخیص نداد.

۱-۲-۵. یافته‌های کلان از تحلیل خطا

- الگوی موفقیت: مدل در شناسایی الگوهای قاعده‌مند و پرتکرار (مانند مطابقه شخص و شمار در جملات ساده خبری) بسیار موفق است (دقت > ۹۵٪). این نشان می‌دهد مدل به سرعت می‌تواند نگاشت‌های پایه بین فرم و عملکرد در گویش هدف را یاد بگیرد.
- الگوی شکست سیستماتیک (الف): سوگیری به فارسی معیار. مدل تمایل داشت ساختارهای کم‌منبع یا نامتعارف گویش را به نزدیک‌ترین ساختار پرتکرار در فارسی معیار «تحریف» کند (مانند تفسیر ساخت مجهول ارگتیو به عنوان معلوم ساده). این خطاها به طور خاص متأثر از ویژگی شماره ۲ (ساخت کنایی) و ویژگی شماره ۳ (ضمیراندازی) بودند، زیرا مدل در داده‌های آموزش خود الگوی مطابقه فاعلی-مفعولی صریح دیده است و ساخت‌های کنایی یا نهادهای پنهان را به درستی پردازش نمی‌کند.
- الگوی شکست سیستماتیک (ب): نابینایی واجی-صرفی. مدل در تشخیص نشانه‌های جمع نامتعارف و تأثیر فرآیندهای واجی محلی (مانند ادغام یا حذف واکه) بر نشانه مطابقه، مستعد خطا بود. این دسته عمدتاً ناشی از ویژگی شماره ۱ (جمع‌سازی با تغییر آوایی) و ویژگی شماره ۴ (تأثیر فرآیندهای واجی محلی بر شناسه‌ها) بود. مدل به دنبال نشانه‌های جمع پسوندی (مانند «-ها» در فارسی) می‌گردد و تغییر واکه را نادیده می‌گیرد. همچنین تضعیف شناسه در گفتار سریع مانع بازشناسی مطابقه می‌شود.

مثال ۱ (خطای ناشی از جمع‌سازی با تغییر آوایی - ویژگی شماره ۱):

گویش (خط فارسی محلی): گُرا هُواغش شکست

/ˈkoræ hævɒ:ˈyɛʃ ʃikʰæst/

تجزیه واژه‌به‌واژه: /ˈkoræ/ (بچه) /hævɒ:ˈyɛʃ/ (اسباب‌بازی-ش) /ʃikʰæst/ (شکست. PST.3SG)

ترجمه آزاد: بچه‌ها اسباب‌بازی‌اش را شکستند.

تحلیل خطا: مدل نهاد /ˈkoræ/ را مفرد تشخیص داد و فعل /ʃikʰæst/ را سوم شخص مفرد گرفت در حالی که در این بافت /ˈkoræ/ معنای جمع دارد و فعل صورت سوم شخص جمع است. این خطا مصداق ویژگی شماره ۱ (جمع‌سازی با تغییر آوایی) است که مدل قادر به تشخیص آن نبود.

مثال ۲ (خطای ناشی از ساخت کنایی - ویژگی شماره ۲):

گویش (خط فارسی محلی): وِگِه مداره شو

/væˈgʲeʰ moˈdɒ:ræ ˈʃo/

تجزیه واژه‌به‌واژه: /væˈgʲeʰ/ (گویا) /moˈdɒ:ræ/ (مدیر) /ʃo/ (شد. PST.3SG)

ترجمه آزاد: گویا مدیر شد / منصوب شد.

تحلیل خطا: مدل نهاد /mo'dɒ:ræ/ را فاعل دستوری و فعل /ʃo'/ را مطابق با آن تشخیص داد. اما این ساخت مجهول کنایی با نشانه /væ'gje^h/ است و نهاد نحوی در نقش مفعول منطقی قرار دارد و فعل به صورت بی‌شخص به کار رفته است. این خطا مصداق ویژگی شماره ۲ (ساخت کنایی) است که مدل آن را به ساختار فاعلی-مفعولی فارسی معیار تحریف کرد.

مثال ۳ (خطای ناشی از ضمیراندازی - ویژگی شماره ۳):

گویش (خط فارسی محلی): دینام بازی خاھن

/dɪnɒ:m bɒ:zi: xɒ:hæn/

تجزیه واژه‌به‌واژه: /dɪnɒ:m/ (فردا) /bɒ:zi:/ (بازی) /xɒ:hæn/ (خواهیم کرد، FUT.1PL)

ترجمه آزاد: فردا بازی خواهیم کرد.

تحلیل خطا: در این جمله نهاد به صورت جداگانه ذکر نشده است و شخص و شمار تنها از طریق شناسه فعل (xɒ:hæn = اول شخص جمع) قابل تشخیص است. مدل نهاد را «نامشخص» برگرداند (در حالی که باید اول شخص جمع تشخیص می‌داد). این خطا مصداق ویژگی شماره ۳ (ضمیراندازی) است که مدل در شناسایی نهاد از طریق مطابقت فعل ناتوان است.

مثال ۴ (خطای ناشی از تلفیق دو ویژگی - هم جمع آوایی و هم ضمیراندازی):

گویش (خط فارسی محلی): رفته‌ن د شو

/ræftæn de ʃo/

تجزیه واژه‌به‌واژه: /ræftæn/ (رفته‌اند، PERF.3PL) de (در) ʃo (شب)

ترجمه آزاد: شب رفته‌اند.

تحلیل خطا: در این جمله نهاد آشکار وجود ندارد و فعل /ræftæn/ (رفته‌اند) خود حاوی شناسه سوم شخص جمع است. مدل به دلیل نبود نهاد ظاهری، «مطابق نامشخص» برگرداند. این خطا مصداق تلفیق ویژگی شماره ۳ (ضمیراندازی) و تأثیر فرآیند واجی (حذف شناسه در گفتار سریع) است.

جدول (۲) توزیع خطاها را بر اساس ویژگی‌های گویش نشان می‌دهد. همان‌طور که مشاهده می‌شود، ویژگی‌های ۱ و ۴ بیشترین سهم را در خطاهای نابینایی واجی-صرفی و ویژگی‌های ۲ و ۳ بیشترین سهم را در خطاهای سوگیری ساختاری دارند.

۳-۵. سنجش کارایی روندکار نیمه خودکار

مهم‌ترین یافته کاربردی، صرفه‌جویی چشمگیر در زمان تحلیل بود. برآورد زمان بر اساس زمان‌سنجی دقیق نشان داد:

- تحلیل کاملاً دستی: به طور متوسط ۲۱۰ ثانیه برای هر جمله (شامل گوش دادن دقیق، آوانگاری، تحلیل نحوی و ثبت).
- روندکار نیمه خودکار: به طور متوسط ۴۵ ثانیه برای هر جمله (شامل زمان پردازش خودکار مدل + زمان بررسی، اعتبارسنجی و تصحیح نهایی خروجی توسط پژوهشگر).
- این ارقام به معنای کاهش ۷۸٪ در زمان تحلیل هر جمله و امکان پردازش حجمی ۴,۶ برابری داده در واحد زمان مشابه است. کیفیت نهایی خروجی پس از مرحله نظارت پژوهشگر، معادل با تحلیل دستی کامل ارزیابی شد.

۶. بحث

در این بخش، یافته‌های پژوهش در پرتو چارچوب نظری تفسیر و پیامدهای روش‌شناختی و اخلاقی آن استخراج می‌شوند. یافته‌های این پژوهش نشان می‌دهد که مدل Llama 3.1 با استفاده از یادگیری چندمثاله‌ای می‌تواند الگوهای مطابقه را در گویش بختیاری با دقت ۸۷,۴٪ شناسایی کند که ۲۵٪ بهتر از یک روش قاعده‌بنیاد ساده است. با این حال، دو محدودیت عمده مشاهده شد: (۱) سوگیری ساختاری (۴۲٪ خطاها) ناشی از تحمیل الگوهای فارسی معیار بر ساخت‌های کنایی و ضمیرانداز، و (۲) نابینایی واجی-صرفی (۳۵٪ خطاها) ناشی از ناتوانی مدل در تشخیص جمع‌سازی با تغییر آوایی و تأثیر فرآیندهای واجی محلی. چارچوب «انسان-در-حلقه» پیشنهادی زمان تحلیل را ۷۸٪ کاهش می‌دهد و امکان پردازش ۴,۶ برابری داده را فراهم می‌کند، اما وابستگی به مهندسی دستور و سوگیری ذاتی مدل پایه از محدودیت‌های اصلی آن است.

۱-۶. تفسیر عملکرد مدل: دانش زبانی جهان‌شول، تعمیم‌پذیری محدود

عملکرد موفق مدل با نظریه یادگیری انتقالی^۱ قابل توضیح است: Llama 3.1 از طریق پیش‌آموزش بر داده‌های چندزبانه، مفاهیم نحوی جهان‌شول (مانند نهاد، فعل، شخص، شمار) را در پارامترهای خود ذخیره کرده است (براون و همکاران، ۲۰۲۰؛ وی و همکاران، ۲۰۲۲). یادگیری چندمثاله‌ای به مدل امکان می‌دهد این دانش را به گویش بختیاری نگاشت کند. اما تحلیل خطاها مرزهای این تعمیم‌پذیری را نشان می‌دهد. سوگیری ساختاری (۴۲٪ خطاها) ناشی از تحمیل الگوهای زبان‌های مسلط است (بندر و همکاران، ۲۰۲۱) و نابینایی واجی-صرفی (۳۵٪ خطاها) نشان می‌دهد که مدل در سطح نحو صوری متوقف می‌ماند و تعامل واج‌شناسی و صرف را درک نمی‌کند (بندر و کولر، ۲۰۲۰).

۲-۶. پیامدهای روش‌شناختی

صرفه‌جویی ۷۸٪ در زمان، سه پیامد اصلی دارد: (۱) تغییر نقش پژوهشگر از تحلیلگر دستی به مدیر هوشمند و مفسر نهایی. (۲) امکان مطالعات رده‌شناختی در مقیاس کلان (۴,۶ برابر داده بیشتر). (۳) دموکراتیک‌سازی فناوری با ارائه الگوهای دستوری از پیش بهینه‌شده (پیوست ۱) که بار فنی استفاده از LLM را برای زبان‌شناسان میدانی کاهش می‌دهد.

۳-۶. ملاحظات اخلاقی و مسئولیت اجتماعی

رعایت رضایت آگاهانه، گمنام‌سازی و مالکیت جمعی داده‌ها برای جامعه گویشوران از اصول این پژوهش بوده است. چارچوب «انسان-در-حلقه» با قراردادن پژوهشگر به عنوان دروازه‌بان نهایی، از «استعمار داده‌ای» جلوگیری می‌کند (برد، ۲۰۲۰). محصولات نهایی باید به جامعه بازگردند (کتاب‌های آموزشی، فرهنگ‌های دیجیتال).

۴-۶. محدودیت‌ها و نقشه راه پژوهشی آینده

مهم‌ترین محدودیت‌ها: وابستگی به مهندسی دستور (فرآیندی تجربی)، سوگیری ذاتی مدل پایه، هزینه محاسباتی و نیاز به داده آوانویسی‌شده. پژوهش‌های آینده باید به (۱) توسعه کتابخانه استاندارد از دستورهای بهینه‌شده برای پدیده‌های رده‌شناختی

1 Transfer Learning

2 E. M. Bender & A. Koller

مختلف، (۲) تنظیم دقیق^۱ یک مدل متن‌باز بر روی پیکره متوازن از گویش‌های ایرانی، و (۳) ادغام با مدل‌های صوتی-زبانی برای تحلیل مستقیم گفتار خام پردازند.

۷. نتیجه‌گیری

این پژوهش با هدف طراحی و اعتبارسنجی یک چارچوب نیمه خودکار «انسان-در-حلقه» مبتنی بر مدل زبانی بزرگ برای تحلیل نظام مطابقه در گویش بختیاری ایزده انجام شد. یافته‌های کمی نشان داد که مدل Llama 3.1 با استفاده از یادگیری چندمثاله‌ای به دقت کلی ۸۷٫۴٪ و نمره F1 معادل ۸۷٫۴ دست می‌یابد که ۲۵٪ برتر از یک روش قاعده‌بنیاد ساده است. تحلیل کیفی خطاها، دو محدودیت اصلی مدل را آشکار ساخت: (۱) سوگیری به ساختارهای پرتکرار فارسی معیار و انگلیسی (۴۲٪ خطاها) که منجر به تحریف ساخت‌های خاص گویش می‌شود، و (۲) نابینایی واجی-صرفی (۳۵٪ خطاها) که در مواجهه با پدیده‌هایی مانند جمع مکسر ظاهر می‌گردد. از منظر کاربردی، چارچوب پیشنهادی زمان تحلیل را ۷۸٪ کاهش داده و پردازش ۴٫۶ برابری داده را در واحد زمان ممکن ساخته است.

چارچوب ارائه‌شده سه پیامد مهم دارد: (۱) تغییر نقش پژوهشگر از تحلیلگر دستی به مدیر هوشمند و مفسر نهایی، (۲) امکان پذیر شدن مطالعات رده‌شناختی در مقیاس کلان، و (۳) دموکراتیک‌سازی فناوری از طریق ارائه الگوهای دستوری از پیش‌بهینه‌شده (پیوست ۱). با این حال، پژوهش با محدودیت‌هایی مانند وابستگی به مهندسی دستور، سوگیری ذاتی مدل پایه، هزینه محاسباتی و نیاز به داده‌های آوانویسی‌شده روبرو است. رعایت اصول اخلاقی (رضایت آگاهانه، گمنام‌سازی، مالکیت جمعی داده‌ها) و اجتناب از استعمار داده‌ای از ارکان اصلی این چارچوب بوده است.

در نهایت، این پژوهش اثبات می‌کند که مدل‌های زبانی بزرگ می‌توانند شریکی توانمند برای زبان‌شناسان میدانی باشند، مشروط بر اینکه در قالب یک همکاری دیالکتیکی میان سرعت استقرار ماشینی و عمق تفسیر تاریخی-فرهنگی انسان به کار گرفته شوند. پیشنهاد می‌شود پژوهش‌های آتی بر (۱) توسعه کتابخانه‌ای از دستورهای بهینه‌شده برای پدیده‌های رده‌شناختی مختلف، (۲) تنظیم دقیق یک مدل پایه بر روی پیکره متوازن از گویش‌های ایرانی، و (۳) ادغام با مدل‌های صوتی-زبانی برای تحلیل مستقیم گفتار خام متمرکز شوند.

منابع

- Anonby, E., & Asadi, A. (2018). Bakhtiari dialect. In *Encyclopædia Iranica*. Brill.
- Anonby, E., Taheri-Ardali, M., & Hayes, A. (Eds.). (2021). *The Atlas of the Languages of Iran (ALI)*. Geomatics and Cartographic Research Centre, Carleton University. <http://iranatlas.net/>
- Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185–5198). <https://doi.org/10.18653/v1/2020.acl-main.463>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623).
- Bird, S. (2020). Decolonising speech and language technology. In *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 3504–3519). <https://doi.org/10.18653/v1/2020.coling-main.313>

1 Fine-Tuning

- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Comrie, B. (1989). *Language universals and linguistic typology* (2nd ed.). University of Chicago Press.
- Corbett, G. G. (2006). *Agreement*. Cambridge University Press.
- Crystal, D. (2000). *Language death*. Cambridge University Press.
- Dabir-Moghaddam, M. (2012). *The Oxford handbook of Iranian languages*. Oxford University Press.
- Dabir-Moghaddam, M. (2018). Typology of Iranian languages. In A. Sedighi & P. Shabani-Jadidi (Eds.), *The Oxford handbook of Iranian linguistics* (pp. 45–67). Oxford University Press.
- Haig, G. (2008). *Alignment change in Iranian languages: A construction grammar approach*. Mouton de Gruyter.
- Haig, G., & Anonby, E. (Eds.). (2018). *The languages and linguistics of Western Asia: An areal perspective*. De Gruyter Mouton.
- Hashemi, S. M. H., Ghadrdan, M., Shahhosseini, M., & Vatankhahan Esfahani, E. (2026). *APARSIN: A Multi-Variety Benchmark for Iranic Languages, Dialects, and Accents*. ACL Anthology.
- Hovy, D., & Prabhumoye, S. (2021). Five sources of bias in natural language processing. *Language and Linguistics Compass*, 15(8), e12432. <https://doi.org/10.1111/lnc3.12432>
- Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 6282–6293).
- Jurafsky, D., & Martin, J. H. (2021). *Speech and language processing* (3rd ed. draft). <https://web.stanford.edu/~jurafsky/slp3/>
- Kalbasi, I. (2016). *Rade shenasi-e nahvi: do ghiyesh-e Bakhtiari-e Majd-e Soleyman va Shushtari (dar ostān-e Khuzestan)* [Syntactic typology: Two dialects of Bakhtiari of Majd-e Soleyman and Shushtari (in Khuzestan province)]. Rashdin Publications.
- Karimi, Y. (2020). *Language shift and endangerment in the Iranian context: A sociolinguistic study*. Palgrave Macmillan.
- Lin, X. V., Mihaylov, T., Artetxe, M., Wang, T., Chen, S., & Tsuetkov, Y. (2022). Few-shot learning with multilingual generative language models. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing* (pp. 9019–9052).
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd ed.). Lawrence Erlbaum Associates.
- Moseley, C. (Ed.). (2010). *Atlas of the world's languages in danger* (3rd ed.). UNESCO Publishing.
- Nettle, D., & Romaine, S. (2000). *Vanishing voices: The extinction of the world's languages*. Oxford University Press.
- Samiiian, V. (2019). Split ergativity in Western Iranian: A typological perspective. *STUF – Language Typology and Universals*, 72(3), 451–478.
- SilkRoadNLP (2025). *The First Workshop on NLP and LLMs for the Iranian Language Family (SilkRoadNLP)*. Held at EACL 2026.
- Smith, R. (2023). Prompt engineering for morphologically rich languages: Best practices for low-resource NLP. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 234–248).
- Stilo, D. (2022). *Iranian and minority languages at the crossroads of globalization*. Brill.
- Team, F. (2024). Evaluating GPT-4 on under-resourced language tasks: A comparative study of open vs. closed models. In *Proceedings of the 15th International Conference on Language Resources and Evaluation (LREC)* (pp. 112–126).
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., ... & Lample, G. (2023). Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint*, arXiv:2307.09288.
- UNESCO. (2003). *Language vitality and endangerment*. UNESCO Ad Hoc Expert Group on Endangered Languages.
- Unjum, N., Evert, S., Sarveswaran, K., Weerasinghe, R., & Jayatilleke, N. (2026). *LLMs for Low-Resource Languages: A Survey*. Available at SSRN.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.

- Wang, X., Li, Y., Zhang, H., & Chen, W. (2024). Responsible multilingual large language models: A survey on bias mitigation and ethical deployment. *Transactions on Machine Learning Research*. <https://openreview.net/forum?id=example>
- Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., ... & Chi, E. H. (2022). Emergent abilities of large language models. *Transactions on Machine Learning Research*.
- Zhang, S., Liu, Y., Wang, L., & Chen, M. (2024). A survey on multilingual large language models for low-resource languages: Corpora, alignment, and bias. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (pp. 1–25).

Extended Abstract

A Typological Examination of Moravcsik's Approach to Number in Persian

Zahra Teymoori

PhD student in Linguistics, Faculty of Literature, Alzahra University, Tehran, Iran,
zr.teymoori@gmail.com

Azita Abbasi* 

Assistant Professor, Department of Linguistics, Faculty of Literature, Alzahra University, Tehran, Iran
a.abbasi@alzahra.ac.ir

Abstract

"Number" has been extensively studied in linguistics from morphological, syntactic, semantic, pragmatic, and typological perspectives. This study examines the conceptualization of plurality in Persian based on Moravcsik's (2017) typological approach to Number. In this approach, the conceptualization of plurality is analyzed in terms of three parameters: individuation, numeration, and participation, with partonomy and taxonomy serving as important foundations for the analysis. The data consist of Persian examples collected primarily on the basis of the authors' linguistic judgments and are analyzed according to these three parameters. In the domain of individuation, both full and partial listing, as well as lumping, are found in Persian and can be expressed through syntactic, morphological, and lexical devices. The data also suggest that plural markers such as *-hā* and *-āt* may, in specific contexts, refer not only to sets of individuals but also to sets of kinds or types. In the domain of numeration, the expression of number interacts with the structure of the noun phrase, animacy, agreement, and definiteness. Moreover, colloquial uses of numerals indicate that numerical expressions may have approximate or exaggerative interpretations in discourse rather than always specifying an exact cardinality. In the domain of participation, collective and distributive interpretations are influenced by syntactic structure, verb semantics, adverbial modifiers, and discourse context. Overall, the findings suggest that Moravcsik's (2017) three-parameter framework provides a useful basis for analyzing the conceptualization of plurality in Persian, but a more comprehensive account requires consideration of its interaction with the morphological, syntactic, semantic, and pragmatic properties of Persian.

Keywords:

Typology
Singular
Plural
Number
Numeral
Quantifier
Participation

How to cite this article:

Teymoori, Z. & Abbasi, A. (2026). A Typological Examination of Moravcsik's Approach to Number in Persian. *Journal of Linguistics & Iranian Languages*, 11(2), 33-46. doi: <https://doi.org/10.22099/jlil.2026.56565.1492>



COPYRIGHTS ©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publisher.

* Corresponding author

Introduction

Number is an important topic that can be investigated in languages, including Persian, and numerous studies addressed it from different approaches in various areas of linguistics, including phonetics, morphology, syntax, semantics, and pragmatics. Number is not merely a formal category and needs to be considered on the basis of the partonomy, taxonomy, and the parameters of individuation, numeration, and participation. From this perspective, Moravcsik's (2017) approach provides a suitable framework. The present study, seeks to answer three questions: First, in what structures are the parameters of individuation, numeration, and participation realized in Persian? Second, in the realization of these parameters, what roles can be considered for factors such as word formation, definiteness, animacy, agreement, verb meaning, and discourse context? Third, to what extent are the Persian data consistent with this formulation, and in which cases are additional components needed for a more precise description of them? The main hypothesis is that this framework can account for the fundamental distinctions in the conceptualization of plurality in Persian; however, the linguistic realization of these distinctions depends on the specific characteristics of the Persian system.

Materials & Methods

Previous studies mainly focused on morphological, syntactic, and typological aspects and on the ways in which singular and plural are marked. Grammarians regarded *-hā* and *-ān* as the most important markers of plurality and examined the differences in their use (Gharib et al., 1971; Khazaeli, 1972; Khanlari, 1976; Shariat, 1993; Anvari & Ali, 2009). Alinezhad and Tayeb (2006) examined the Persian number system within Croft's (1990) typological framework and described morphological and syntactic strategies for forming the plural. Subsequent studies examined the functions of *-hā* (Darzi & Ghadiri, 2011; Bahrami, 2018), classifiers and the dependence of counting on noun type (Rezayi & Nazari, 2014), and the role of *-āt* in expressing concepts such as kind and group (Harandi & Safavi, 2020). Typological studies also demonstrated differences among number systems in different languages (Canta, 2016; Dbritz, 2021; Toshtemirovna, 2025). Studies of varieties of Persian, using Atlas of Iran database (Mostafavi, 2021) and Gilaki (Mahdavian & Abbasi, 2024), addressed formal variation and some cases of mismatch between formal and conceptual number. Overall, the conceptualization of sets, the relation among their members, and the role of individuation, numeration, and participation are less frequently examined within a unified framework; therefore, Moravcsik's (2017) approach can address this gap.

In this study, following a review of the previous literature and the theoretical framework, Persian data collected through the authors' linguistic intuition are examined within the framework of Moravcsik's (2017), followed by a summary and conclusion.

Results & Discussion

Following an examination of the conceptualization of plurality in Persian based on Moravcsik's (2017) typological approach in terms of the three parameters, it was found that these parameters

can account for a considerable part of the patterns of plurality conceptualization in Persian. Nevertheless, their realization in Persian does not always correspond to the mechanisms proposed in this framework and interacts with factors such as morphology, definiteness, animacy, agreement, verb type, and discourse context.

In individuation, she proposes two mechanisms, listing and lumping. Listing occurs in Persian in two forms: full listing and partial listing. Full listing is expressed through conjoining and comitative constructions, in which the order and ranking of members can be related to the syntactic organization of the sentence and verbal agreement. In a comitative construction, the individual or group preceding *bā* 'with' has a more prominent position, and the verb agrees with it. Partial listing is realized through syntactic, morphological, and lexical constructions. For example, constructions such as *lebās o chizhā-ye dige* 'clothes and other things' or *lebās mobās* 'clothes and suchlike' can refer to a set some members of which are not explicitly named. *Lebās mobās* shows a general word-formation process, namely reduplication/echo formation, can acquire a function in the conceptualization of plurality (Karimi Qahi, 2018).

In lumping, data such as *miveh-hā* 'fruits' show the plural in Persian can refer not only to a set of individuals or objects of the same kind, but also to a set of kinds or types belonging to a category. Therefore, *-hā* and *-āt* can, in certain contexts, play a role in expressing a set of kinds or types, although their lexical distribution is not identical (Harandi & Safavi, 2020).

Numeration in Persian is realized through lexical, morphological, and syntactic devices. In numeral constructions, the noun following the numeral remains in the singular form, and animacy can affect verbal agreement. *-hā* is associated with a definite reading in many contexts, while constructions such as *te 'dādi* 'a number of' are used to express an indefinite plural. Numerals do not always express an exact number in speech and are sometimes used for approximation or exaggeration; therefore, the concept of number also has a pragmatic dimension.

In participation, the distinction between collective and distributive interpretations in Persian is not determined by a single marker. *Bā* 'with' makes the participatory relation more explicit in comitative constructions, whereas in conjoining constructions, the type of participation also depends on the meaning of the verb and the context. Therefore, this framework makes it possible to classify various plural phenomena, but their precise analysis requires considering the interaction between the conceptualization of plurality and different levels of the Persian language.

Conclusion

In conclusion, the three parameters of individuation, numeration, and participation can classify and account for important distinctions related to the conceptualization of plurality in Persian; however, the realization of these parameters interacts with the structural and functional characteristics of the language. Individuation, full listing, partial listing, and lumping are realized in Persian through various syntactic, morphological, and lexical mechanisms.

Furthermore, the data concerning plural markers showed, in particular contexts, a distinction can be made between sets of individuals and sets of kinds or types.

In Numeration, the expression of number is related to noun-phrase structure, animacy, agreement, and definiteness, and the approximate or exaggerative interpretations of numerals also show the discourse uses of numerals cannot be reduced to the determination of an exact number. In participation, the distinction between collective and distributive interpretations, particularly in conjoining constructions, is influenced by verb meaning, context, and adverbial elements.

In addition to confirming the presence of Moravcsik's (2017) three parameters in Persian, this study also demonstrated how these parameters interact with mechanisms that are independent features of the Persian system. This finding can provide a basis for future comparative studies of the conceptualization of plurality in Persian and other Iranian languages.

References

- Alinezhad, B. & Tayeb, S. M. T. 2006. Number System in Contemporary Persian Nouns. *Journal of Social Sciences and Humanities, Shiraz University*. 25 (3). 157-170. [In Persian]
- Anvari, H. & Ali Abasabad, Y. 2009. *Persian Grammar 2*. Tehran: Payame Noor University. [In Persian]
- Bahrami, F. 2018. Morphosemantic Functions of “-hā” in Persian from a Typological Perspective. *Journal of Language Related Research*. 9 (3). 179-202. [In Persian]
- Canta, A. 2016. A contrastive analysis of the category of number in English and Albanian nominal system. *European Scientific Journal*. 12 (32). 33-53
- Croft, W. 1990. *Typology and Universals*. Cambridge: Cambridge University Press.
- Däbritz, C. L. 2021. Typology of number systems in languages of Western and Central Siberia. *Finnisch-Ugrische Forschungen*. 66. 85-138.
- Darzi, A. & Ghadiri, L. 2011. Quantification by the Plural Marker “-hā” in Persian. *Journal of Language Related Research*. 2 (1). 71-85. [In Persian]
- Fasihi Harandi, M. & Safavi, K. 2020. The Nature and Role of the Arabic Plural Morpheme “-āt” in Persian. *Journal of Comparative Literature*. 12 (22). 173-194. [In Persian]
- Ghadiri, L. 2014. Quantifiers in Persian. *Journal of Language Research*. 5 (1). 95-108. [In Persian]
- Ghomeshi, J. 2003. Plural marking, Indefiniteness, and the Noun phrase. *Studia Linguistica*. 57(2). 47-74.
- Karimi Qahi, M. 2018. Reduplication/Echo Formation in Persian Language and Literature. *Biannual Journal of Persian Language and Literature*. 26 (85). 115-134.
- Khazaeli, M. 1972. *Persian Grammar*. Tehran: Javidan Publishing and Printing Organization. [In Persian]
- Mahdavian, S. Z. & Abbasi, A. 2024. Number in Gilaki Language. *Persian Language and Iranian Dialects*. 9 (1). 7-30. [In Persian]
- Moravcsik, E. 2017. Number. *The Cambridge handbook of Linguistic Typology* (Alexandra Y., Aikhenvald and R. M. W. Dixon). Cambridge: Cambridge University Press. 440 -476.
- Mostafavi, P. 2020. Number Category in Languages Varieties of Some Cities in Kerman Province: A Typological study. *Journal of Language related Research*. 12 (6). 61-92. [In Persian]
- Natel Khanlari, P. 1976. *Historical Grammar of the Persian Language*. Tehran: Foundation for Iranian Culture Publications. [In Persian]
- Qarib, A., Bahar, M. S., Foruzanfar, B., Homaï, J. & Yasami, R. 1971. *Persian Grammar*. Tehran: Central Bookstore Publications. [In Persian]

- Safardoust, A., Ameri, H., Kord e Zafaranlu Kambuziya, A. & Golfam, A. 2023. The Category and Structure of Persian Numerals. *Journal of Comparative Linguistic Research*. 26. 197-222. [In Persian]
- Shariat, M. J. 1993. *Persian Grammar*. Tehran: Asatir Publishing. [In Persian]
- Shirzad, A. M. & Darzi, A. 2023. Type Shifting and the Number System in Persian. *Journal of Researches in Linguistics*. 15 (28). 137-158. [In Persian]
- Tafakkori Rezayi, S. & Nazari, K. 2014. The Syntactic Study of Numeral Classifier in Persian. *Journal of Language Research*. 5 (1). 1-20. [In Persian]
- Toshtemirovna, M. G. 2025. Comparative typology study of the category of number in Uzbek and English. *Western European Journal of Linguistics and Education*. 3 (3) .117-121.
- Wu, J. 2022. Plural words in Austronesian languages: Typology and history. *Oceanic linguistics*. 61(2). 721-756.

زبان‌شناسی و گویش‌های ایرانی

سال ۱۱، شماره ۲، پیاپی ۲۱ (تابستان ۱۴۰۵) شماره صفحات: ۳۳ - ۴۶

بررسی رویکرد رده‌شناختی موراوچیک به شمار در زبان فارسی

زهرا تیموری^۱، آریتا عباسی^{۲*} دانشجوی دکتری زبان‌شناسی، دانشکده ادبیات، دانشگاه الزهراء، تهران، ایران
استادیار گروه زبان‌شناسی، دانشکده ادبیات، دانشگاه الزهراء، تهران، ایران

چکیده

مقوله شمار یکی از حوزه‌های مهم پژوهش در زبان‌شناسی است که در زبان‌های مختلف از دیدگاه‌های صرفی، نحوی، معنایی، کاربردشناختی و رده‌شناختی بررسی شده است. پژوهش حاضر با هدف بررسی چگونگی مفهوم‌سازی جمع در زبان فارسی، رویکرد رده‌شناختی موراوچیک (۲۰۱۷) به شمار را به کار می‌گیرد. در این رویکرد، مفهوم‌سازی جمع بر اساس سه ویژگی «فردیت»، «شمارش» و «مشارکت» تحلیل می‌شود و رابطه جزء و کل و رده‌بندی از مبانی اصلی تحلیل به شمار می‌آیند. داده‌های پژوهش از نمونه‌های فارسی و با تکیه بر قضاوت زبانی نگارندگان گردآوری و بر اساس سه ویژگی مذکور تحلیل شده‌اند. در حوزه فردیت، فهرست‌بندی کامل و ناقص و نیز تجمیع در فارسی از طریق ساخت‌های نحوی، صرفی و واژگانی تحقق می‌یابند. همچنین، داده‌هایی مانند «میوه‌ها» نشان می‌دهند که نشانگرهای جمع در بافت‌های خاص می‌توانند علاوه بر مجموعه افراد، به مجموعه انواع نیز اشاره کنند. در حوزه شمارش، بیان تعداد با ویژگی‌هایی مانند ساخت گروه اسمی، جاندار، مطابقت و معرفتی تعامل دارد و کاربردهای تقریبی و اغراقی عدد نیز نشان می‌دهد که میان ارزش عددی و کاربرد گفتمانی اعداد رابطه یک‌به‌یک وجود ندارد. در حوزه مشارکت نیز تمایز جمعی و توزیعی تحت تأثیر ساخت نحوی، معنای فعل، عناصر قیدی و بافت قرار می‌گیرد. در مجموع، نتایج نشان می‌دهند که چارچوب موراوچیک (۲۰۱۷) برای طبقه‌بندی و تحلیل الگوهای مفهوم‌سازی جمع در فارسی کارآمد است، اما توصیف دقیق تحقق این ویژگی‌ها در فارسی مستلزم در نظر گرفتن تعامل آن‌ها با ویژگی‌های صرفی، نحوی، معنایی و کاربردشناختی زبان است.

واژه‌های کلیدی:

رده‌شناسی
مفرد
جمع
شمار
عدد
کمیت‌نما
مشارکت

تاریخچه مقاله: دریافت: ۱۴ تیرماه ۱۴۰۵

بازنگری: ۲۲ شهریورماه ۱۴۰۵

پذیرش: ۲۹ شهریورماه ۱۴۰۵

* آدرس ایمیل نویسنده مسئول: a.abbasi@alzahra.ac.ir

تیموری، زهرا و عباسی، آریتا (۱۴۰۵). بررسی رویکرد رده‌شناختی موراوچیک به شمار در زبان فارسی. زبان‌شناسی و گویش‌های ایران ۱۱(۲).

<https://doi.org/10.22099/jlil.2026.56565.1492> ۴۶-۳۳

حق نشر ۲۰۲۶ نویسنده(گان). این یک مقاله با دسترسی آزاد است که تحت شرایط مجوز Creative Commons Attribution (CC BY-NC 4.0) منتشر شده است، استفاده، توزیع و تکثیر نامحدود را در هر رسانه‌ای مجاز می‌داند، مشروط بر اینکه به نویسندگان و منبع اصلی اشاره شود. هیچ اجازه‌ای از نویسندگان یا ناشر لازم نیست.

۱. مقدمه

شمار از موضوعات مهم قابل بررسی در زبان‌های مختلف از جمله زبان فارسی است و مطالعات بسیاری با این موضوع با رویکردهای متفاوت در حوزه‌های مختلف زبان‌شناسی از جمله آواشناسی، صرف، نحو، معنی‌شناسی و کاربردشناسی صورت گرفته است. زبان فارسی به دلیل داشتن صورت مفرد و جمع در مقوله شمار، در رده دوم رده‌بندی کرافت^۱ (۱۹۹۰: ۹۷) قرار می‌گیرد. پربسامدترین نشانه جمع در فارسی «ها» است که می‌تواند منجر به معرفه شدن اسم نیز شود (درزی و قدیری، ۱۳۹۰). تکواژ عربی «-ات» نیز گاه در زبان فارسی برای جمع بستن واژه‌های عربی و فارسی استفاده می‌شود و معمولاً اشاره به دسته و گروه دارد (هرندی و صفوی، ۱۳۹۹). به علاوه، عددواژه‌های ساده و مرکب پیش از اسم، به طور عام، به مقوله معرفگر و به طور خاص، به زیرمقوله کمیت‌نما تعلق دارند (صفر دوست و همکاران، ۱۴۰۲).

زبان‌ها به جهت جنبه‌های مختلف دستوری شمار متفاوت هستند به طور مثال به گفته کانتا^۲ (۲۰۱۶)، اسم‌ها در دو زبان انگلیسی و آلبانیایی بر مبنای جنبه‌های مختلف مقوله دستوری شمار از جمله مفرد و جمع بودن، نحوه ساخت جمع و استفاده از حروف تعریف^۳ و اعداد تفاوت‌هایی را نشان می‌دهند و در ساخت جمع، اسم‌هایی که فقط شکل مفرد دارند^۴ و اسم‌هایی که فقط شکل جمع دارند^۵ به یکدیگر شباهت دارند.

در برخی از زبان‌ها علاوه بر آن که شمار به شکل صرفی یا نحوی نشان داده می‌شود، از طریق معنا یا کاربرد هم درک می‌شود. تاشتمیرونا^۶ (۲۰۲۵) معتقد است زبان ازبکی برای جمع بستن اسم‌ها از یک پسوند ثابت و یکسان استفاده می‌کند و هیچ صورت بی‌قاعده‌ای^۷ در این زبان مشاهده نمی‌شود. به طور کلی مقوله شمار در زبان ازبکی گاه از طریق معنا یا کاربرد درک می‌شود و نیازی به مشخصه دستوری نیست.

بیشتر مطالعات مربوط به شمار در فارسی مانند درزی و قدیری (۱۳۹۰) و رضائی و نظری (۱۳۹۳)، بر نمودهای صوری و دستوری این مقوله، از جمله نشانگرهای جمع، عددواژه‌ها، ممیزها و راهکارهای ساخت‌واژی و نحوی تمرکز داشته‌اند. برخی پژوهش‌های رده‌شناختی نیز تنوع نظام‌های شمار را در زبان‌های مختلف نشان داده‌اند مانند علی‌نژاد و طیب (۱۳۸۵) و بهرامی (۱۳۹۷)، اما در این مطالعات، مفهوم جمع و رابطه آن با سازمان‌دهی اعضای یک مجموعه و نحوه مشارکت آن‌ها کمتر در یک چارچوب واحد بررسی شده است. در نتیجه، اگرچه مطالعات پیشین تصویری نسبتاً روشن از چگونگی بازنمایی دستوری شمار در فارسی ارائه می‌کنند، هنوز این پرسش باقی است که این صورت‌های زبانی چگونه با نحوه مفهوم‌سازی جمع و روابط میان اعضای مجموعه ارتباط پیدا می‌کنند. رویکرد موراوچیک^۸ (۲۰۱۷) از این جهت چارچوب مناسبی فراهم می‌کند؛ زیرا شمار را صرفاً یک مقوله صوری نمی‌داند و با تکیه بر رابطه جزء و کل، رده‌بندی و ویژگی‌های فردیت، شمارش و مشارکت، امکان بررسی ابعاد مفهومی و کاربردشناختی آن را فراهم می‌سازد.

بر این اساس، پژوهش حاضر می‌کوشد به سه پرسش پاسخ دهد: نخست، ویژگی‌های فردیت، شمارش و مشارکت در چه ساخت‌ها و سازوکارهایی در زبان فارسی تحقق می‌یابند؟ دوم، در تحقق این ویژگی‌ها چه نقشی برای عواملی مانند ساخت واژه،

^۱ W. Croft

^۲ A. Canta

^۳ articles

^۴ singulariatantum

^۵ pluraliatantum

^۶ M. G. Toshtemirovna

^۷ irregular form

^۸ E. A. Moravcsik

معرفگی، جاننداری، مطابقت، معنای فعل و بافت گفتمانی می‌توان در نظر گرفت؟ سوم، داده‌های فارسی تا چه اندازه با صورت‌بندی موراوچیک (۲۰۱۷) همخوان‌اند و در چه مواردی برای توصیف دقیق‌تر آن‌ها به مؤلفه‌های تکمیلی نیاز است؟ فرض اصلی پژوهش آن است که چارچوب موراوچیک (۲۰۱۷) می‌تواند تمایزهای بنیادین مفهوم‌سازی جمع در فارسی را تبیین کند، اما تحقق زبانی این تمایزها به ویژگی‌های خاص نظام فارسی وابسته است.

در این پژوهش در ابتدا مطالعات پیشین صورت گرفته درباره مقوله شمار در فارسی و برخی دیگر از زبان‌های دنیا مطرح می‌شود، سپس در بخش چارچوب نظری نگاه رده‌شناختی موراوچیک (۲۰۱۷) که مبنای نظری این پژوهش است بیان می‌گردد و در ادامه داده‌های زبان فارسی که با شم زبانی نگارندگان گردآوری شده‌اند در چارچوب رویکرد موراوچیک (۲۰۱۷) مورد بررسی قرار گرفته و جمع‌بندی و نتیجه‌گیری صورت می‌پذیرد.

۲. پیشینه پژوهش

مقوله شمار در زبان فارسی از جنبه‌های مختلف ساخت‌واژی، نحوی، معنایی و رده‌شناختی بررسی شده است. بخش قابل توجهی از پژوهش‌های انجام‌شده بر توصیف صورت‌های دستوری شمار، شیوه‌های نشان‌گذاری جمع، نقش عددواژه‌ها و ممیزها و نیز تفاوت‌های شمار در گونه‌های مختلف فارسی تمرکز داشته‌اند. در این میان، برخی پژوهش‌ها علاوه بر نمود صوری شمار، به ارتباط آن با عواملی همچون جاننداری، معرفگی، نوع اسم و بافت نیز توجه کرده‌اند. با این حال، در بیشتر این مطالعات، «جمع» عمدتاً به‌عنوان یک مقوله دستوری یا یک ویژگی صوری بررسی شده و چگونگی مفهوم‌سازی مجموعه‌ها و رابطه میان اعضای آنها کمتر در یک چارچوب واحد مورد توجه قرار گرفته است.

دستورنویسان مختلفی به مقوله شمار در اسم پرداخته‌اند که اغلب تعاریف مشابهی ارائه می‌دهند. قریب و همکاران (۱۳۵۰: ۲۵)، خزائلی (۱۳۵۱: ۴۵)، خانلری (۱۳۵۵: ۵۱) شریعت (۱۳۷۲: ۱۹۸) معتقدند اسم مفرد بر شخص یا چیز واحد اشاره می‌کند و جمع بر بیش از یک فرد یا چیز دلالت می‌کند، «ها» و «ان» را نشانه جمع در فارسی می‌دانند که «ان» برای جمع اسم‌هایی که بر جاندار دلالت می‌کنند به کار می‌رود و به جمع‌های عربی و مکسر نیز می‌پردازند. انوری و عالی (۱۳۸۸: ۴۰) نیز مانند آرا ذکر شده «ها» و «ان» را نشانه جمع در فارسی می‌دانند و معتقدند «ها» از اواخر ادبیات پهلوی آمده است و در ابتدا برای جمع بستن اسما غیرذی‌روح به کار می‌رفت.

علی‌نژاد و طیب (۱۳۸۵) دستگاه شمار در مقوله اسم در زبان فارسی را از جنبه‌های مختلف زبان‌شناسی مورد تجزیه و تحلیل قرار داده‌اند. با توجه به این که زبان فارسی فقط دارای صورت مفرد و جمع برای مقوله اسم است، در رده دوم از ۴ رده مطرح شده از منظر کرافت (۱۹۹۰: ۹۷) قرار دارد. کرافت (۱۹۹۰: ۶۷-۸) زبان فارسی را در دسته زبان‌هایی می‌داند که تکواژ جمع در آن حضور دارد و فاقد تکواژ مفرد است. تثنیه نیز در فارسی وجود ندارد و در صورت لزوم عدد «دو» و یا کلمه «زوج» یا «جفت» را بر سر اسم آورده می‌شود. گاه نیز صیغه تثنیه عربی در فارسی به کار برده می‌شود. در زبان فارسی هر دو راهکار ساخت‌واژی و نحوی برای جمع بستن اسم‌ها وجود دارد ولی نمی‌توان هر دو را به صورت هم‌زمان به کار برد به همین دلیل «کتاب‌ها» و «دو کتاب» هر دو دستوری هستند در حالی که «دو کتاب‌ها» غیر دستوری است. این پژوهش توصیفی روشن از صورت‌های دستوری شمار و راهکارهای صرفی و نحوی بیان جمع در فارسی ارائه می‌کند و جایگاه فارسی را در یک رده‌بندی زبانی مشخص می‌سازد. با این حال، تمرکز اصلی آن بر نمود صوری و دستوری شمار است و چگونگی مفهوم‌سازی جمع در سطحی فراتر از تقابل مفرد و جمع در آن بررسی نمی‌شود. برای مثال، این که اعضای یک مجموعه چگونه از یکدیگر متمایز یا

به یک کل واحد مرتبط می‌شوند و یا در ساخت‌های جمعی چه نوع رابطه‌ای میان اعضای مجموعه برقرار است، در این توصیف مورد توجه قرار نمی‌گیرد.

درزی و قدیری (۱۳۹۰) بر این باورند که در زبان فارسی معاصر، چندین علامت جمع، فارغ از اصل و منشأ آن‌ها به کار می‌رود که پر بسامدترین و زیاترین آن‌ها، نشانه جمع «-ها» است. در واقع این تکواژ به عنوان نشانه جمع برای اسم‌های فارسی و عربی و هر نوع اسمی که در زبان فارسی به کار می‌رود استفاده می‌شود. آن‌ها به نقل از قمشی (۱۹۹۶) نشانه جمع «-ها» را در فارسی علاوه بر دلالت بر بیش از یک، باعث معرفه شدن اسم نیز می‌دانند. تمرکز این پژوهش بر تکواژ «-ها» سبب شده است که مفهوم جمع عمدتاً از طریق یک نشانگر صرفی بررسی شود و تفاوت میان «جمع به عنوان یک صورت دستوری» و «جمع به عنوان یک مفهوم» کمتر مورد توجه قرار گیرد. رویکرد موراوچیک (۲۰۱۷) با بررسی جمع در ارتباط با رابطه جزء و کل، رده‌بندی و نیز ویژگی‌هایی مانند شمارش و مشارکت، این امکان را فراهم می‌کند که فراتر از وجود یا عدم وجود نشانگر جمع، نحوه سازمان‌یافتن و تفسیر اعضای یک مجموعه در فارسی نیز بررسی شود.

رضائی و نظری (۱۳۹۳) ممیزها را از نظرگاه معنایی به دو نوع ممیز عددی و کمی تقسیم می‌کنند که ممیزهای عددی مانند «نفر، رأس، جلد، فقره» با اسم‌های قابل شمارش همراه می‌شوند و ممیزهای کمی مانند «کیلو، لیتر، گرم، تکه، قطعه» که با نام «نام‌واژه‌های اندازه‌گیری» نیز خوانده می‌شوند، با اسم‌های توده‌وار به کار می‌روند. این تقسیم‌بندی نشان می‌دهد که در فارسی، بیان کمیت و شمار به ماهیت اسمی که مورد شمارش قرار می‌گیرد وابسته است و از این رو، شمارش را نمی‌توان صرفاً پدیده‌ای صرفی دانست. با وجود این، تمرکز پژوهش بر ساخت نحوی ممیزهاست و چگونگی ارتباط شمارش با مفهوم جمع و سازمان‌دهی اعضای یک مجموعه در آن محور اصلی تحلیل نیست.

بهرامی (۱۳۹۷) کارکردهای معنایی تکواژ «ها» را در چارچوب رده‌شناختی آکواویوا^۱ (۲۰۰۸) بررسی کرده است و نتیجه‌گیری می‌کند که زبان فارسی از نظر رده‌شناختی از تنوعات کارکردی جمع‌سازی تبعیت می‌کند و تحت تاثیر ویژگی‌های معنایی اسم پایه، بیانگر یا تلفیقی از مفاهیم از جمله دلالت بر بیش از یک مصداق، بیان پیوستگی، بیان تعدد، القا خوانش ملموس از اسم‌های معنی و زمان/مکان، فراوانی کمی و کیفی، بیان پراکندگی و گستردگی موقعیت‌یافته برای اسم‌های توده‌ای، بیان مشابَهت و تاکید بر مصداق است. او از میان کارکردهای ذکر شده نقش‌های نمودی، شدت‌بخشی، تاکید و تقریب زمانی را ویژگی‌های معنایی مختص جمع‌سازی در زبان فارسی می‌داند که آکواویوا (۲۰۰۸) به آن‌ها اشاره‌ای نکرده است.

هرندی و صفوی (۱۳۹۹) معتقدند تکواژ عربی «-ات» گاه در زبان فارسی برای جمع بستن واژه‌های عربی و فارسی استفاده می‌شود. هر چند بسیاری از صاحب‌نظران اعتقاد دارند در فارسی نباید از نشانه جمع عربی استفاده شود، برخی از صورت‌های جمع واژه‌های عربی مانند «اطلاعات» چنان رواجی یافته‌اند که اگر با نشانه‌های جمع فارسی به کار روند غریب، نامأنوس و دور از ذهن می‌نمایند. به علاوه برخی از صورت‌های مختوم به «-ات» مانند «ادبیات» صورت مفردی که در زبان فارسی کاربرد داشته باشد، ندارند. در حالت کلی، این تکواژ در فارسی برای جمع بستن اسامی غیرجاندار و انتزاعیات، بدون در نظر گرفتن جنسیت به کار می‌رود و دربرگیرنده معنای نوع، قسم و گروه می‌باشد به طوری که حتی اگر معنای جمع را در خود داشته باشد، این معنا و کاربرد با معنا و مفهوم دسته و گروه عجین است. به طور مثال «سبزیجات» به معنی «انواع سبزی‌ها» است و با «سبزی‌ها» مترادف نیست. بنابراین «-ات» را وام‌وندی در زبان فارسی می‌دانند که به صورت غیر مستقیم از زبان عربی وام‌گیری شده‌است و اینک عنصری در نظام ساخت‌واژی زبان فارسی است. یافته‌های این پژوهش از آن جهت اهمیت دارد که نشان می‌دهد مفهوم

¹ P. Acquaviva

جمع در فارسی همواره به معنای ساده «بیش از یک» نیست و گاه مفاهیمی مانند نوع، دسته و گروه نیز در آن دخالت دارند. با این حال، این تمایزها در پژوهش مذکور عمدتاً در ارتباط با یک تکواژ خاص بررسی شده‌اند و در چارچوبی کلی‌تر برای مفهوم‌سازی جمع قرار نگرفته‌اند.

مصطفوی (۱۴۰۰) به بررسی مقوله شمار با استفاده از داده‌های موجود در پایگاه داده‌های اطلس زبان‌شناسی ایران در پژوهشگاه میراث فرهنگی و گردشگری می‌پردازد. او معتقد است شمار دارای دو ارزش مفرد و جمع است که مفرد بدون نشانه و جمع نشان‌دار است. ابزار زبانی مورد استفاده برای نشانه‌گذاری جمع از نوع ساخت‌واژی است به طوری که پسوند جمع به اسم مفرد می‌پیوندد. هم‌چنین در کنار راهبرد ساخت‌واژی، راهبرد نحوی نیز در این گونه زبانی مشاهده می‌شود. این پژوهش نشان می‌دهد حتی در میان گونه‌های زبان فارسی نیز در نحوه بازنمایی شمار تنوع وجود دارد و راهکارهای ساخت‌واژی و نحوی می‌توانند در کنار یکدیگر به کار روند. با این حال، تمرکز اصلی آن بر شناسایی و توصیف راهکارهای صوری بیان مفرد و جمع است و ابعاد مفهومی و کاربردشناختی این تفاوت‌ها کمتر مورد بررسی قرار گرفته است.

صفر دوست و همکاران (۱۴۰۲) عددواژه‌های ساده و مرکب پیش از اسم را، به طور عام، به مقوله معرف‌گر و به طور خاص، به زیر مقوله کمیت‌نما متعلق می‌دانند. تحلیل عددواژه‌ها در این پژوهش عمدتاً از منظر مقوله دستوری و ساختار نحوی انجام شده است و ارتباط میان «شمارش» و «مفهوم جمع» در چارچوبی رده‌شناختی بررسی نشده است.

مهدویان و عباسی (۱۴۰۳) به چگونگی بازنمایی شمار در زبان گیلکی پرداخته‌اند و معتقدند در این زبان ضمائر به لحاظ شمار و در حالت‌های فاعلی، مفعولی و اضافی دارای صورت‌های متفاوتی هستند. حضور نشانگر جمع به همراه طبقه‌نما در گروه اسمی غیر معرفه موجب بدساختی می‌شود در حالی که در گروه اسمی معرفه حضور تکواژ جمع اجباری است. در ساخت هم‌پایه، با وجود اطلاق مفهوم جمع به هر دو عضو، اگر هر دو عضو به یک مجموعه تعلق داشته باشند، نشانگر جمع می‌تواند تنها بر روی سازه دوم ظاهر شود. دوگان‌سازی از فرآیندهای در تعامل با مفهوم شمار است. به اعتقاد نویسندگان، عدم تطابق صوری و مفهومی شمار در بافت‌هایی خاص می‌تواند به علت رعایت ادب و تکریم، تعمیم و جان‌داری باشد. این پژوهش از میان مطالعات بررسی‌شده، به رویکرد حاضر نزدیک‌تر است؛ زیرا علاوه بر نمود صوری شمار، به موارد عدم تطابق شمار صوری و مفهومی و نقش عواملی مانند جان‌داری، تعمیم و ادب و تکریم نیز توجه دارد. با این حال، این پدیده‌ها در چارچوب ویژگی‌های موراوچیک (۲۰۱۷) برای تحلیل مفهوم جمع، یعنی فردیت، شمارش و مشارکت، سازمان‌دهی و تبیین نشده‌اند.

کانتا (۲۰۱۶) به بررسی تفاوت‌ها و شباهت‌های مقوله شمار در این دو زبان که هر دو از اعضای شاخه زبان‌های هند و اروپایی هستند، می‌پردازد. داده‌های این پژوهش از کتاب‌های دستور این دو زبان جمع‌آوری شده است. طبق نتایج به دست آمده، اسم‌ها در این دو زبان بر مبنای جنبه‌های مختلف مقوله دستوری شمار از جمله مفرد و جمع بودن، نحوه ساخت جمع و استفاده از حروف تعریف و اعداد تفاوت‌هایی را نشان می‌دهند. اما بارزترین تفاوت، مربوط به نحوه نوشتن و عملکرد اسم‌های مجموعه^۱ و مرکب^۲ است. علی‌رغم تفاوت‌ها، شباهت‌هایی در این دو زبان وجود دارد که بیشتر به نحوه ساخت جمع، اسم‌هایی که فقط شکل مفرد دارند و اسم‌هایی که فقط شکل جمع دارند، مربوط می‌شود. اگرچه پژوهش مذکور امکان مقایسه نظام‌های شمار را فراهم می‌کند، چگونگی مفهوم‌سازی جمع و رابطه میان اعضای یک مجموعه در سطح مفهومی و کاربردشناختی در کانون تحلیل آن قرار ندارد.

^۱ collective nouns

^۲ compound nouns

دابریتز^۱ (۲۰۲۱) به بررسی مقوله شمار در هفت زبان خانتی شرقی^۲، نگاناسانی^۳، سلکویی^۴، دولگانی^۵، ترکی چولیم^۶، اونکی^۷ و کت^۸ می‌پردازد. طبق نتایج به دست آمده تمامی این زبان‌ها از نظر دستوری شمار را در دو حالت مفرد و جمع نشان می‌دهند و زبان‌های «زبان‌های اوگری^۹، خانتی شرقی، نگاناسانی، سلکویی» از حالت تشبیه به صورت منظم استفاده می‌کنند. در زبان‌های «دولگانی، ترکی چولیم، سلکویی، خانتی شرقی، نگاناسانی، اونکی» اسامی‌ای مشاهده می‌شود که از نظر صوری مفرد هستند اما معنای جمع می‌دهند و در حالت کلی، بیان عدد به مقوله جانداري مربوط است. این تحلیل عمدتاً در سطح توصیف و رده‌بندی نظام‌های شمار این زبان‌ها باقی می‌ماند و رابطه میان اعضای مجموعه، نحوه رده‌بندی آن‌ها و نوع مشارکت اعضا در یک ساخت جمعی را به صورت مستقل بررسی نمی‌کند.

تشتیمرونا (۲۰۲۵) به بررسی مقوله شمار در این زبان‌ها می‌پردازد. طبق نتایج به دست آمده، در زبان ازبکی فقط از پسوند “-lar” برای جمع بستن اسم‌ها استفاده می‌شود و بر خلاف انگلیسی هیچ صورت بی‌قاعده‌ای^{۱۰} در این زبان مشاهده نمی‌شود. همچنین اگر اسمی به گروهی از مردم اطلاق شود می‌تواند در معنی جمع بدون وجود پسوند جمع استفاده شود. در زبان ازبکی معمولاً مقوله شمار تاثیری بر روی فعل ندارد و فعل به صورت مفرد می‌آید. در این زبان مقوله شمار گاه از طریق معنا یا کاربرد درک می‌شود و نیازی به بازنمود دستوری نیست. این تحلیل عمدتاً بر مقایسه سازوکارهای دستوری و صرفی شمار در ازبکی و انگلیسی استوار است و رابطه میان این سازوکارها و نحوه مفهوم‌سازی یک مجموعه به صورت نظام‌مند بررسی نمی‌شود.

۳. چارچوب نظری

موراوچیک (۲۰۱۷) با نگاهی رده‌شناختی به بررسی مقوله شمار در زبان‌های مختلف می‌پردازد. او معتقد است که اگرچه شمارپذیری^{۱۱} در واقعیت آشکار است، اما چگونگی درک ما از جهان همیشه مطابق با تظاهر زبانی مفاهیم نیست. به طور مثال واژه “suggestion” به معنی «پیشنهاد» جمع بسته می‌شود در حالی که مترادف نزدیک آن “advice” به معنی «نصیحت» جمع بسته نمی‌شود.

موراوچیک (۲۰۱۷) تحلیل رابطه جزء و کل را مبنای مفهوم جمع، و رده‌بندی را دومین قدم تحلیلی می‌داند که به تشخیص اجزا از یکدیگر می‌پردازد. او رابطه جزء و کل را از نظر ادراکی^{۱۲}، بین‌زبانی^{۱۳}، تاریخی^{۱۴} و توسعه‌ای^{۱۵} بر رده‌بندی ارجح می‌داند. موراوچیک (۲۰۱۷) مفهوم‌سازی جمع را در قالب سه ویژگی فردیت^{۱۶}، شمارش^{۱۷} و مشارکت^{۱۸} مطرح می‌کند.

¹ C. L. Däbritz

² Eastern Khanty

³ Nganasan

⁴ Selkup

⁵ Dolgan

⁶ Chulym Turkic

⁷ Evenki

⁸ Ket

⁹ Uralic languages

¹⁰ irregular form

¹¹ countability

¹² conceptual

¹³ cross-linguistic

¹⁴ historical

¹⁵ developmental

¹⁶ individuation

¹⁷ numeration

¹⁸ participation

۳-۱. فردیت

موراوچیک (۲۰۱۷) ویژگی فردیت را به دو دسته فهرست‌بندی^۱ و تجمیع^۲ تقسیم می‌کند. فهرست‌بندی زمانی اتفاق می‌افتد که مجموعه جمع حداقل دارای دو زیرمجموعه باشد. در فهرست‌بندی ممکن است مانند مثال 1a با یک فهرست کامل^۳ روبه‌رو باشیم به طوری که تمام افراد نام برده شوند و یا مانند مثال 1b یک فهرست ناقص^۴ داشته باشیم به طوری که فرد یا افرادی نام برده می‌شوند مانند “John” و سایر افراد که به آن‌ها مرتبط هستند بدون این که نام تک تک آنها برده شود با یک مفهومی که به همه آن‌ها به صورت همزمان اشاره می‌کند مانند “others” ذکر می‌شوند.

1. a. John, Peter and Mary. (Moravcsik, 2017)

جان، پتر و مری

- b. John and the others. (Moravcsik, 2017)

جان و دیگران

در فهرست‌بندی، دو ویژگی مهم انسجام^۵ و مرتبه^۶ مطرح می‌شوند. انسجام می‌تواند به علت شباهت و یا عضویت در یک گروه یکسان باشد. مثال 1a یک فهرست کامل است که انسجام آن‌ها به واسطه شباهت است. مثال 1b یک فهرست ناقص است که مرتبه در آن لحاظ شده است زیرا تنها «جان» با نامش شناخته می‌شود. فهرست کامل را می‌توان به وسیله ساخت پیوندی^۷ یا همراهی^۸ به ترتیب مانند مثال‌های 2a و 2b بیان کرد:

2. a. John and his sister. (Moravcsik, 2017)

جان و خواهرش

- b. John with his sister. (Moravcsik, 2017)

جان با خواهرش

به اعتقاد موراوچیک (۲۰۱۷)، به نقل از استاسن^۹ (۲۰۰۰ و ۲۰۰۵) ساخت همراهی از ساخت پیوندی در زبان‌های جهان متداول‌تر است.

فهرست ناقص در زبان‌های مختلف به سه روش مختلف نحوی، صرفی و یا واژگانی ساخته می‌شود که به ترتیب در مثال‌های 3a-c نشان داده شده است.

3. a. John and the others. (Moravcsik, 2017)

جان و دیگران

- b. jama-jama ‘dresses and such’. BENGALI
dress-dress (Moravcsik, 2017)

لباس و این جور چیزها

- c. we → first non-singular person (Moravcsik, 2017)

ما (ضمیر اول شخص غیر مفرد)

¹ listing

² lumping

³ full listing

⁴ partial listing

⁵ cohesion

⁶ ranking

⁷ conjoining construction

⁸ comitative construction

⁹ L. Stassen

در تجمیع، به عنوان دسته دیگر فردیت، تمامی افراد در فهرست به دلیل شباهت و یا عضویت در یک گروه، در درون یک دسته قرار می‌گیرند. به این ترتیب افراد ذکر شده در مثال 1a به children در مثال 4 تغییر می‌کنند.

4. John, Peter, and Mary → children (Moravcsik, 2017)

جان، پیترو و مری

هر چند در تجمیع افراد مختلف با نوع یکسان به یک نام ارجاع داده می‌شوند، مثلاً به جای «سیب‌ها، گلابی‌ها و آلوها» از واژه «میوه‌ها» استفاده می‌شود، واژه انتخاب شده به تنهایی معنای تنوعی از یک نوع را می‌دهد، مثلاً «میوه‌ها» به معنی «انواعی از میوه‌ها» درک می‌شود.

۳-۲. شمارش

در فهرست کامل، تمامی افراد در مجموعه نام برده می‌شوند بنابراین تعداد آن‌ها مشخص است اما زمانی که افراد تجمیع می‌شوند، تعداد مشخص نیست و در صورت نیاز باید تعداد آن‌ها بیان شود. تعداد عناصر یک مجموعه^۱ به کمک ابزارهای واژگانی، صرفی و یا نحوی به ترتیب مانند مثال‌های 5a-c بیان می‌شوند:

5. a. pair, triumvirate, century (Moravcsik, 2017)

جفت، سه‌نفری، قرن

b. volk-ova 'two wolves' SLOVENE wolf-DU^۲ (Moravcsik, 2017)

دو گرگ

c. two friends (Moravcsik, 2017)

دو دوست

۳-۳. مشارکت

صرف نظر از این که در یک بافت زبانی، افراد به صورت فهرست و یا تجمیعی آمده باشند، میزان مشارکت آن‌ها می‌تواند جدا یا جمعی باشد. در مثال 6 می‌توان تصور کرد که «جان» و «مری» به صورت جداگانه یا به صورت هم‌زمان رسیده‌اند و این ابهام به کمک استفاده از قاموس‌واژه‌های^۳ جدا، مانند «each» و «together»^۴ و یا وندهای جمعی^۵ یا توزیعی^۵ قابل حل است. در زبان مجاری اعداد می‌توانند وندهای جمعی مانند مثال 6b و یا وندهای توزیعی مانند مثال 6c بگیرند.

6. a. John and Mary arrived early. (Moravcsik, 2017)

جان و مری زود رسیدند.

b. Ha'rm-anmentu'nk a moziba. HUNGARIAN three-lywe:went the to:cinema 'Three of us went to the cinema together.' (Moravcsik, 2017)

ما سه نفر با هم رفتیم سینما.

c. Ha'rm-an-ke'ntmentu'nk a moziba. HUNGARIAN three-ly-DISTRwe:went the cinema 'We went to the cinema in groups of three.' (Moravcsik, 2017)

^۱ cardinality of a plural set

^۲ dual-marker

^۳ lexemes

^۴ collective affixes

^۵ distributive affixes

ما به صورت گروه‌های سه نفره به سینما رفتیم.

در مثال 6a اگر تصور بر این باشد که دو فرد به صورت هم‌زمان رسیده‌اند، مشارکت جمعی و در غیر این صورت توزیعی است. موراوچیک (۲۰۱۷) هم‌چنین مشارکت افراد را وابسته به ماهیت فعل و بافت می‌داند. به عنوان نمونه در مثال 7a مشارکت به صورت جداگانه است اما در مثال 7b مشارکت جمعی است.

7. a. The apples rotted in the heat. (Moravcsik, 2017)

سیب‌ها در گرما گندیدند.

b. The apples formed a pile. (Moravcsik, 2017)

سیب‌ها یک توده تشکیل دادند.

۴. تجزیه و تحلیل داده‌ها

در این بخش تلاش بر این است ابعاد مختلف رویکرد رده‌شناختی موراوچیک (۲۰۱۷) در مورد شمار در زبان فارسی مورد مطالعه و بررسی قرار گیرد. با توجه به این که موراوچیک (۲۰۱۷) مفهوم‌سازی جمع را در قالب سه ویژگی «فردیت»، «شمارش» و «مشارکت» بررسی می‌کند، دسته‌بندی داده‌های زبان فارسی نیز بر این اساس خواهد بود.

بررسی داده‌های فارسی نشان می‌دهد که اگرچه سه ویژگی پیشنهادی موراوچیک (۲۰۱۷) می‌توانند بخش قابل توجهی از الگوهای مربوط به مفهوم‌سازی جمع در فارسی را تبیین کنند، تحقق این ویژگی‌ها در فارسی همواره با سازوکارهایی که در چارچوب وی توصیف شده‌اند یکسان نیست. به علاوه داده‌های فارسی نشان‌دهنده تعامل مفهوم جمع با ویژگی‌هایی همچون ساخت‌واژه، معرفگی، جانداري، مطابقت، نوع فعل، بافت گفتمانی است.

۴-۱. مفهوم‌سازی جمع

در این زیر بخش به بررسی سه ویژگی فردیت، شمارش و مشارکت مطرح شده توسط موراوچیک (۲۰۱۷) در زبان فارسی می‌پردازیم.

۴-۱-۱. فردیت

موراوچیک (۲۰۱۷) ویژگی فردیت را در دو زیربخش فهرست‌بندی و تجمیع مطرح و بررسی می‌کند. فهرست‌بندی نیز همان طور که ذکر شد، خود شامل دو زیر بخش فهرست کامل و ناقص است.

الف) فهرست‌بندی

دو نوع فهرست کامل و ناقص مطرح شده توسط موراوچیک (۲۰۱۷) به ترتیب مانند مثال‌های ۸ الف و ب در زبان فارسی نیز وجود دارند. طبق نظر موراوچیک (۲۰۱۷)، فهرست‌بندی بر مبنای انسجام و یا هم‌گروهی بودن شکل می‌گیرد. معمولاً اگر ذکر تمامی افراد اهمیت داشته باشد، تعداد آن‌ها محدود بوده و انسجام کلام را از بین نبرد، از فهرست کامل استفاده می‌شود اما در صورتی که مرتبه برخی افراد بالاتر بوده، دیگران به صورت منفرد اهمیتی نداشته باشند و یا تعداد افراد زیاد باشد به طوری که مطرح کردن آن‌ها انسجام کلام را از بین ببرد، از فهرست ناقص استفاده می‌شود.

۸. الف. مریم، علی و حسن مسابقه را به تیم روبرو باختند.

ب. رئیس جمهور و هیأت همراه از نمایشگاه دیدن کردند.

در زبان فارسی از هر دو ساخت پیوندی و همراهی برای بیان فهرست کامل استفاده می‌شود. در ساخت پیوندی زمانی که افراد مستقل از یکدیگر نام برده می‌شوند، مگر در شرایطی خاص و قراردادی بین گوینده و مخاطب، مرتبه خاصی بین افراد وجود ندارد، مانند مثال ۹ الف. حال اگر برخی از افراد وابسته به فرد یا افراد دیگری نام برده شوند به نظر می‌رسد که افراد مستقل می‌توانند مرتبه بالاتری داشته باشند مانند مثال ۹ ب که «من» و «علی» مرتبه بالاتری از سایر افراد دارند. هم‌چنین زمانی که از ساخت همراهی استفاده می‌شود، فرد یا افرادی که قبل از «با» قرار می‌گیرند مرتبه بالاتری دارند و مطابقت فعل با آن فرد یا افراد صورت می‌گیرد مانند مثال ۹ پ.

۹. الف. من، مریم و علی آمدیم.

ب. من و دوستم و علی و دوستش آمدیم.

پ. من با مریم و علی آمدم.

«مرتبه» در فهرست‌بندی صرفاً یک ویژگی معنایی یا مفهومی نیست، بلکه می‌تواند با سازمان نحوی جمله نیز ارتباط پیدا کند. به‌ویژه در ساخت همراهی، فرد واقع‌شده پیش از «با» از نظر نحوی جایگاه برجسته‌تری دارد و فعل نیز با همان فرد یا گروه اصلی مطابقت می‌کند. بنابراین، در فارسی، ساخت همراهی علاوه بر آن که وسیله‌ای برای تشکیل یک مجموعه است، می‌تواند در تعیین مرکزیت اعضای مجموعه و نحوه سازمان‌دهی نحوی آن‌ها نیز نقش داشته باشد. این ویژگی نشان می‌دهد که در فارسی، تمایزهای مربوط به فردیت می‌توانند با نظام نحوی و مطابقت نیز تعامل داشته باشند.

در فهرست ناقص مشخصاً مرتبه فردی که نام برده می‌شود بالاتر از سایر افرادی است که به عنوان یک دسته یا گروه تجمیع می‌شوند. در زبان فارسی فهرست ناقص به دو روش نحوی با ساخت پیوندی و همراهی، صرفی و واژگانی به ترتیب مانند مثال‌های ۱۰ الف - ت وجود دارد.

۱۰. الف. لباس و چیزهای دیگه خریدم.

ب. لباس با یه سری چیزهای دیگه خریدم.

پ. لباس لباس خریدم.

ت. ضمائر جمع: ما، شما (جمع)

در نمونه ۱۰ الف، «لباس» به‌صورت مشخص ذکر شده و «چیزهای دیگه» اعضای دیگری را که به‌صورت مشخص نام برده نشده‌اند دربرمی‌گیرد. در مثال ۱۰ ب نیز «لباس» در ساخت همراهی با گروهی از افراد یا اشیاء قرار گرفته است. در مثال ۱۰ پ، فهرست ناقص از طریق ساخت صرفی و در ۱۰ ت، از طریق ابزار واژگانی بیان شده است. بنابراین، در فارسی، مفهوم فهرست ناقص می‌تواند از طریق سازوکارهایی با سطوح دستوری متفاوت تحقق یابد.

روش صرفی در فهرست ناقص زیرمجموعه اتباع در واژه‌سازی است. به گفته کریمی قهی (۱۳۹۷) مرکب‌های اتباعی ترکیب‌هایی هستند که بر اثر تکرار قسمتی از پایه ساخته می‌شوند. این بخش مکرر معنی ندارد و هیچ‌گاه به تنهایی به کار نمی‌رود مانند «متاب» در «کتاب‌متاب». از این منظر، نمونه «کتاب‌متاب» یا «لباس لباس» نکته‌ای فراتر از رده‌بندی موراوچیک (۲۰۱۷) نشان می‌دهد. در چارچوب موراوچیک (۲۰۱۷)، چنین الگویی از نظر کارکردی می‌تواند در زمره سازوکارهای صرفی فهرست ناقص قرار گیرد، اما در فارسی منشأ آن در فرایند اتباع و واژه‌سازی است. بنابراین، در فارسی، یک فرایند واژه‌سازی می‌تواند در بافت مناسب، معنایی مشابه «و چیزهای دیگر از همین نوع» ایجاد کند. به عبارت دیگر، فهرست ناقص در فارسی لزوماً محصول یک نشانگر اختصاصی فهرست ناقص نیست، بلکه می‌تواند از فرایندهای عمومی‌تر واژه‌سازی نیز حاصل شود. اهمیت

نمونه‌هایی مانند «لباس مباس» در آن است که نشان می‌دهند رابطه میان یک فرایند صوری و یک کارکرد مفهومی الزاماً یک‌به‌یک نیست. فرایند اتباع در سطح واژه‌سازی تعریف می‌شود، اما می‌تواند در بافت مناسب کارکردی مشابه سازوکار فهرست‌بندی ناقص پیدا کند. از این رو، تحلیل داده‌های فارسی نشان می‌دهد که ویژگی‌های مفهوم‌سازی جمع ممکن است از طریق فرایندهایی تحقق یابند که در اصل برای بیان آن مفهوم ساخته نشده‌اند.

طبق نظر موراوچیک (۲۰۱۷) در رابطه با فهرست ناقص، در زبان فارسی نیز ضمیر اول شخص «ما» به معنی «من و بقیه» است و «بقیه» شامل همه افراد به‌جز اول شخص است. همچنین «شما» معنی «تو و بقیه» می‌دهد که «بقیه»، می‌تواند دوم شخص نباشند که در این صورت «ما» و «شما» از نوع واژگانی فهرست ناقص دارای مرتبه می‌باشند. حال اگر منظور از «شما» اشاره به افرادی باشد که همگی مخاطب گوینده هستند، در این صورت «شما» از نوع واژگانی فهرست همسان بدون مرتبه است و همچنین «آنها» نیز به دلیل اینکه مجموعه‌ای از سوم شخص‌های مفرد است در این گروه جای می‌گیرد.

ب) تجميع

به اعتقاد موراوچیک (۲۰۱۷)، تجميع زمانی اتفاق می‌افتد که افراد در زیر یک نوع قرار می‌گیرند و آن نوع اشاره به انواع آن دسته دارد. مثال ۱۱ موردی از تجميع در زبان فارسی است.

۱۱. الف. میوه‌ها را شستم.

این داده یکی از مواردی است که در آن زبان فارسی می‌تواند نکته‌ای فراتر از توصیف موراوچیک (۲۰۱۷) نشان دهد. در نمونه «میوه‌ها را شستم»، تفسیر جمع لزوماً به مجموعه‌ای از افراد یا اشیای هم‌نوع محدود نیست، بلکه در بافت مناسب می‌تواند مجموعه‌ای از انواع یا گونه‌های یک رده را دربرگیرد؛ برای مثال، «میوه‌ها» می‌تواند به سیب، موز، هلو و دیگر انواع مورد نظر در بافت اشاره داشته باشد. بنابراین، در فارسی باید میان «جمع افراد» و «جمع انواع» تمایز قائل شد. به این ترتیب، علاوه بر «-ات» (هرندی و صفوی، ۱۳۹۹)، «-ها» نیز می‌تواند به انواع اشاره کند که البته بافت کلامی و موقعیتی در این موضوع نقش مهمی دارد، ضمن آن که تمامی اسامی فارسی (مانند «میوه») قابل جمع بسته شدن با «-ات» نیستند. تفاوت میان «-ها» و «-ات» نشان می‌دهد که همه نشانگرهای جمع در فارسی رفتار یکسانی ندارند. در نتیجه، اگرچه هر دو می‌توانند در برخی بافت‌ها به مجموعه‌ای از انواع اشاره کنند، توزیع واژگانی آن‌ها یکسان نیست. از این رو، مفهوم تجميع در فارسی علاوه بر مؤلفه مفهومی مورد نظر موراوچیک (۲۰۱۷)، با محدودیت‌های صرفی و واژگانی نیز تعامل دارد.

۴-۱-۲. شمارش

به گفته مصطفوی (۱۴۰۰) در زبان فارسی جمع، نشان‌دار و مفرد بی‌نشان است. تحلیل داده‌های فارسی نشان می‌دهد تعداد افراد یک مجموعه به کمک ابزارهای واژگانی، صرفی (با پسوند «-ین» به معنی «دو») و یا نحوی به ترتیب مانند مثال‌های ۱۲ الف - پ بیان می‌شوند.

۱۲. الف. جفت، جین

ب. طرفین، جانبین

پ. دو دوست، چهار کتاب

در صورتی که از روش نحوی برای شمارش استفاده شود، اسم در گروه اسمی، مستقل از جاننداری، علی‌رغم دلالت بر معنای جمع، به صورت مفرد باقی می‌ماند. این در حالی است که جاننداری بر روی مطابقت تاثیر گذاشته و فعل با فاعل جاندار جمع مطابقت می‌کند اما با فاعل بی‌جان جمع، مگر در صورت استفاده از صنعت ادبی جان‌بخشی به اشیاء، مطابقت ندارد، مانند مثال‌های ۱۳ الف - پ.

۱۳. الف. دخترها / ۳ دختر در اتاق بازی می‌کنند.

ب. برگ‌های پاییزی با من سخن می‌گویند. (جان‌بخشی به اشیا)

پ. لیوان‌ها شکست.

تأثیر جاننداری بر مطابقت نشان می‌دهد که مفهوم شمار در فارسی با ویژگی‌های معنایی اسم نیز ارتباط دارد. در نتیجه، در فارسی باید میان نشان‌گذاری جمع در گروه اسمی و مطابقت دستوری تمایز قائل شد. از این رو، داده‌های فارسی نشان می‌دهند که مفهوم شمار در کاربرد زبانی خود با تعامل میان مؤلفه‌های معنایی، صرفی و نحوی همراه است و صرفاً به مشخص کردن تعداد اعضای مجموعه محدود نمی‌شود.

با توجه به پژوهش‌های پیشین درباره رابطه نشانگر جمع «-ها» و معرفگی، حضور این نشانگر در بسیاری از بافت‌ها با خوانش معرفه گروه اسمی همراه است. از این رو، در برخی بافت‌ها که گوینده قصد اشاره به یک مجموعه نکره دارد، می‌توان از سازوکارهای دیگری مانند «تعدادی» استفاده کرد؛ مانند مثال‌های ۱۴ الف و ب. بنابراین، در فارسی، بیان جمع و تعداد با وضعیت معرفگی گروه اسمی تعامل دارد و انتخاب یک ساخت شمارشی می‌تواند تحت تأثیر این ویژگی قرار گیرد.

۱۴. الف. بسته‌ها رسیدند. (معرفه)

ب. تعدادی بسته رسید. (نکره)

در سطح نظام عددی، عددواژه‌های فارسی دارای ارزش عددی مشخصی هستند؛ باین‌حال، کاربرد یک عبارت عددی در گفتار الزاماً همیشه به معنای ارائهٔ تعداد دقیق نیست. در برخی بافت‌های محاوره‌ای، گاه افراد برای اشاره به کم بودن تعداد از «دو، دو سه، چهار، سه چهار» استفاده می‌کنند و گاه برای بیان احساسات نسبت به زیاد بودن تعداد چیزی یا اغراق در آن، بدون دانستن دقیق تعداد، بسته به ویژگی‌های آن اسم از یک عدد بزرگ استفاده می‌کنند، به ترتیب مانند مثال‌های ۱۵ الف - پ.

۱۵. الف. چهار تا سوال پرسیدم انقدر عصبانیت نداره!

ب. دو سه تا دونه پفک بهم میدی.

پ. چهارصد تا ادم امروز به من زنگ زد!

در زبان فارسی تمامی اعداد ترتیبی با اضافه شدن پسوند «م» به عدد اصلی ساخته می‌شوند مانند «یکم، دوم، سوم» و صرفاً «یکم» دارای مترادف پر بسامدتر «اول» است. بنابراین، داده‌های فارسی نشان می‌دهند که «شمارش» علاوه بر تعیین تعداد دقیق، می‌تواند دارای کارکردهای تقریبی و اغراقی نیز باشد. از این منظر، داده‌های محاوره‌ای فارسی جنبه کاربردشناختی مفهوم شمار را برجسته می‌کنند که در صورت‌بندی اصلی ویژگی شمارش در مدل موراوچیک (۲۰۱۷) محور اصلی تحلیل نیست.

۳-۱-۴. مشارکت

در زبان فارسی در بسیاری از بافت‌ها، ساخت همراهی با «با» به مشارکت جمعی دلالت می‌کند. مانند مثال ۱۶ الف اما در یک فهرست با ساخت پیوندی، مشارکت توزیعی و جمعی افراد به بافت و فعل بستگی دارد و گاه بدون حضور قید ابهام ایجاد می‌شود مانند مثال ۱۶ ب که دارای ابهام است و به صورت مثال‌های ۱۶ پ و ت می‌توان ابهام را برطرف کرد. اما مثال‌های ۱۷ الف و ب به دلیل ماهیت معنایی فعل ابهام ندارند و به ترتیب نشان‌دهنده مشارکت توزیعی و جمعی هستند.

۱۶. الف. مریم با علی رسید.

ب. مریم و علی رسیدند.

پ. مریم و علی باهم رسیدند.

ت. مریم و علی جداگانه رسیدند.

۱۷. الف. سیب‌ها گندیدند.

ب. سیب‌ها به شکل تپه درآمدند.

داده‌های فارسی نشان می‌دهند که تمایز میان مشارکت جمعی و توزیعی در همه ساخت‌ها از طریق یک نشانگر صرفی یا نحوی واحد بیان نمی‌شود. در ساخت همراهی، حرف اضافه «با» معمولاً رابطه همراهی میان اعضا را به صورت آشکارتری نشان می‌دهد، در حالی که در ساخت پیوندی، «و» به تنهایی نوع مشارکت را تعیین نمی‌کند و تفسیر جمله می‌تواند به عوامل دیگری از جمله معنای فعل وابسته باشد. برای مثال، در «سیب‌ها گندیدند» خوانش توزیعی، یعنی گندیدن هر یک از سیب‌ها، طبیعی‌تر است؛ در حالی که در «سیب‌ها به شکل تپه درآمدند» خوانش جمعی برجسته‌تر است، زیرا وضعیت توصیف‌شده مستلزم شکل‌گیری یک آرایش مشترک میان سیب‌هاست. یافته‌های بخش مشارکت نشان می‌دهند که مدل سه‌گانه موراوچیک (۲۰۱۷) برای توصیف تمایزهای اصلی در زبان فارسی کارآمد است، اما توصیف دقیق تحقق این تمایزها در فارسی به تعامل میان ساخت نحوی، عناصر واژگانی و قیدی، معنای فعل و بافت گفتمانی وابسته است.

تحلیل داده‌های فارسی نشان داد که سه ویژگی «فردیت»، «شمارش» و «مشارکت» در چارچوب موراوچیک (۲۰۱۷) می‌توانند برای طبقه‌بندی بخش مهمی از الگوهای مفهوم‌سازی جمع در فارسی به کار روند. با این حال، تحقق زبانی این ویژگی‌ها در بسیاری از موارد با ویژگی‌های مستقل نظام زبان فارسی تعامل دارد. از این منظر، رابطه میان چارچوب نظری و داده‌های فارسی را نمی‌توان صرفاً رابطه‌ای مبتنی بر انطباق کامل دانست.

بر اساس یافته‌ها، می‌توان گفت چارچوب موراوچیک (۲۰۱۷) برای تحلیل فارسی از دو جهت سودمند است: نخست، امکان می‌دهد پدیده‌های ظاهراً متفاوت مربوط به جمع در قالب چند تمایز مفهومی مشترک طبقه‌بندی شوند؛ دوم، امکان آشکار شدن ویژگی‌هایی را فراهم می‌کند که در تحلیل صرفاً صرفی یا نحوی شمار کمتر مورد توجه قرار می‌گیرند. در مقابل، داده‌های فارسی نشان می‌دهند که این چارچوب برای توصیف یک زبان خاص باید با ویژگی‌های ساختاری و کاربردی همان زبان تکمیل شود. سه ویژگی موراوچیک (۲۰۱۷) را می‌توان در فارسی شناسایی کرد، اما نحوه تحقق آن‌ها مستقل از سایر سطوح نظام زبانی نیست. به بیان دیگر، داده‌های فارسی بیش از آن که چارچوب موراوچیک (۲۰۱۷) را رد کنند، نشان می‌دهند که برای کاربرد آن در تحلیل یک زبان خاص باید تعامل میان مفهوم‌سازی جمع و ساختارهای صرفی، نحوی، معنایی و کاربردشناختی آن زبان نیز در نظر گرفته شود.

۵. جمع‌بندی و نتیجه‌گیری

پژوهش حاضر با هدف بررسی مفهوم‌سازی جمع در زبان فارسی بر اساس رویکرد رده‌شناختی موراوچیک (۲۰۱۷) انجام شد. نتایج نشان داد که سه ویژگی «فردیت»، «شمارش» و «مشارکت» می‌توانند تمایزهای مهم مربوط به مفهوم‌سازی جمع در فارسی را طبقه‌بندی و تبیین کنند، اما تحقق این ویژگی‌ها در فارسی با ویژگی‌های ساختاری و کاربردی زبان تعامل دارد. در حوزه فردیت، فهرست‌بندی کامل و ناقص و تجمیع در فارسی از طریق سازوکارهای مختلف نحوی، صرفی و واژگانی تحقق می‌یابند. همچنین، داده‌های مربوط به نشانگرهای جمع نشان دادند که در بافت‌های خاص می‌توان میان جمع افراد و جمع انواع تمایز گذاشت. در حوزه شمارش، بیان تعداد با ساخت گروه اسمی، جانداري، مطابقت و معرفتی ارتباط دارد و کاربردهای تقریبی و اغراقی عدد نیز نشان می‌دهد که کاربرد گفتمانی اعداد را نمی‌توان به تعیین تعداد دقیق فروکاست. در حوزه مشارکت نیز تمایز جمعی و توزیعی به‌ویژه در ساخت‌های پیوندی تحت تأثیر معنای فعل، بافت و عناصر قیدی قرار می‌گیرد.

بر این اساس، چارچوب موراوچیک (۲۰۱۷) را می‌توان برای تحلیل رده‌شناختی مفهوم‌سازی جمع در فارسی چارچوبی کارآمد دانست، اما داده‌های فارسی نشان می‌دهند که این چارچوب در کاربرد برای یک زبان خاص باید در تعامل با ویژگی‌های صرفی، نحوی، معنایی و کاربردشناختی آن زبان بررسی شود. بنابراین، یافته اصلی پژوهش نه صرفاً تأیید حضور سه ویژگی موراوچیک (۲۰۱۷) در فارسی، بلکه نشان دادن نحوه تعامل این ویژگیها با سازوکارهای مستقل نظام فارسی است. این نتیجه می‌تواند زمینه‌ای برای پژوهش‌های تطبیقی آینده درباره مفهوم‌سازی جمع در فارسی و دیگر زبان‌های ایرانی فراهم کند.

منابع و مآخذ

- انوری، حسن و عالی عباس آباد، یوسف (۱۳۸۸) دستور زبان فارسی ۲. تهران: دانشگاه پیام‌نور
- بهرامی، فاطمه (۱۳۹۷). کارکردهای صرفی – معنایی «ها» در زبان فارسی از منظر رده‌شناختی. *جستارهای زبانی*. دوره ۹. شماره ۳. پیاپی ۴۵. صص ۱۷۹-۲۰۲.
- تفکری رضائی، شجاع و نظری، کبری (۱۳۹۳). بررسی نحوی ممیز عدد در زبان فارسی. *پژوهش‌های زبانی*. سال ۵. شماره ۱. صص ۲۰-۱.
- خزائی، محمد (۱۳۵۱). دستور زبان فارسی. تهران: سازمان چاپ و انتشارات جاویدان.
- درزی، علی و قدیری، لیلا (۱۳۹۰). کمیت‌نمایی نشانه جمع «ها» در فارسی. *جستارهای زبانی*. دوره ۲. شماره ۱. صص ۷۱-۸۵.
- شریعت، م. ج (۱۳۷۲). دستور زبان فارسی. تهران: نشر اساطیر.
- شیرزاد، امیرمحمد و درزی، علی (۱۴۰۲). مقوله‌گردانی و نظام شمار در زبان فارسی. *نشریه پژوهش‌های زبان‌شناسی*. سال ۱۵، شماره ۲۸. ۱۳۷-۱۵۸.
- صفردوست، عباس؛ عامری، حیات؛ کرد زعفرانلو کلمبوزیا، عالی و گلفام، ارسلان (۱۴۰۲). مقوله و ساختار عددواژه‌های فارسی. *نشریه پژوهش‌های زبان‌شناسی تطبیقی*. شماره ۲۶. صص ۱۹۷-۲۲۲.
- علی‌نژاد، بتول و طیب، سید محمد تقی (۱۳۸۵). نگاهی به دستگاه شمار اسم در فارسی معاصر. *مجله علوم اجتماعی و انسانی دانشگاه شیراز*. دوره ۲۵. شماره ۳. پیاپی ۴۸. صص ۱۵۷-۱۷۰.
- فصیحی هرندی، مجید و صفوی، کوروش (۱۳۹۹). نقد و تحلیل ماهیت و نقش تکواژ جمع‌ساز «-ات» عربی در فارسی. *نشریه ادبیات تطبیقی*. سال ۱۲. شماره ۲۲. صص ۱۷۳-۱۹۴.
- قدیری، لیلا (۱۳۹۳). کمیت‌نماها در زبان فارسی. *مجله پژوهش‌های زبانی*. سال ۵. شماره ۱. صص ۹۵-۱۰۸.
- کریمی قهی، منصوره (۱۳۹۷). اتباع در زبان و ادب فارسی. *دوفصلنامه زبان و ادبیات فارسی*. سال ۲۶. شماره ۸۵. صص ۱۱۵-۱۳۴.
- قریب، عبدالعظیم؛ بهار، مالک الشعرا؛ فروزانفر، بدیع‌الزمان؛ همائی، جلال و یاسمی، رشید (۱۳۵۰). دستور زبان فارسی. تهران: انتشارات کتاب‌فروشی مرکزی.
- مصطفوی، پونه (۱۴۰۰). مقوله شمار در گونه‌های زبانی برخی از شهرستان‌های استان کرمان: بررسی‌ای رده‌شناختی. *دوماهنامه بین‌المللی جستارهای زبانی*. دوره ۱۲. شماره ۶. پیاپی ۶۶. صص ۶۱-۹۲.
- مهدویان، سیده زینب و عباسی، آریتا (۱۴۰۳). شمار در زبان گیلکی. *زبان فارسی و گویش‌های ایران*. دوره ۹. شماره ۱. پیاپی ۱۷. صص ۷-۳۰.
- ناتل خانلری، پرویز (۱۳۵۵). دستور تاریخی زبان فارسی. تهران: انتشارات بنیاد فرهنگ ایران.
- Canta, A. 2016. A contrastive analysis of the category of number in English and Albanian nominal system. *European Scientific Journal*. 12 (32). 33-53
- Croft, W. 1990. *Typology and Universals*. Cambridge: Cambridge University Press.
- Däbritz, C. L. 2021. Typology of number systems in languages of Western and Central Siberia. *Finnisch-Ugrische Forschungen*. 66. 85-138.
- Ghomeshi, J. 2003. Plural marking, Indefiniteness, and the Noun phrase. *Studia Linguistica*. 57(2). 47-74.
- Moravcsik, E. 2017. Number. *The Cambridge handbook of Linguistic Typology* (Alexandra Y., Aikhenvald and R. M. W. Dixon). Cambridge: Cambridge University Press. 440-476.
- Toshtemirovna, M. G. 2025. Comparative typology study of the category of number in Uzbek and English. *Western European Journal of Linguistics and Education*. 3 (3). 117-121.
- Wu, J. 2022. Plural words in Austronesian languages: Typology and history. *Oceanic linguistics*. 61(2). 721-756.

Extended Abstract

Kurdish, Turkish, and Other Languages in *Sharafname-ye Bedlisi*

Reza Amini* 

Associate Professor in Linguistics, Research Institute for Cultural Heritage and Tourism, Tehran, Iran.

r.amini@richt.ir

Review History: Received: 2026-08-23 Revise: 2026-09-07 Accepted: 2026-09-20	Abstract A scientific and multifaceted examination of important historical Persian texts is essential for their correct interpretation. Many such texts contain words from languages other than Persian, and without understanding the meanings of these non-Persian elements, a precise and complete grasp of the text remains unattainable. One of the significant Persian texts that features a notable use of non-Persian utterances is <i>Sharafname-ye Bedlisi</i> (11th century AH / 17th century CE). The author, a Kurd by origin, was raised in the Safavid court and later attained the ranks of Amir and Amir al-Umara (“Mir Mirani”) in both Iran and the Ottoman Empire. This unique background has resulted in the influence of multiple languages on his book: Kurdish, Turkish, and the languages of the Christian peoples with whom the Ottomans had centuries of contact in Eastern and Central Europe and the Caucasus. This article examines <i>Sharafname-ye Bedlisi</i> precisely from this multilingual perspective. The findings attest to the presence of utterances from Kurdish, Turkish, and other languages within the book, as well as the broader influence of these languages on its text. On another level, these linguistic traces also reflect the rich cultural exchanges among the speakers of these languages and the multilingual environment in which the author of <i>Sharafname</i> lived. The effects of these linguistic exchanges are still more or less present in the Persian language, although some of them have disappeared from Persian due to changing conditions in Iran over the past few centuries.	Keywords: <i>Sharafname-ye Bedlisi</i> Kurdish Turkish Mongolian Georgian Eastern European languages
<i>How to cite this article:</i> Amini, R. (2026). Kurdish, Turkish, and Other Languages in <i>Sharafname-ye Bedlisi</i>. <i>Journal of Linguistics & Iranian Languages</i>, 11(2), 47-65. doi: https://doi.org/10.22099/jlil.2026.56920.1498		
 COPYRIGHTS ©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publisher.		

* Corresponding author

Introduction

This article examines the *Sharafname-ye Bedlisi*, composed by Sharaf Khan Bedlisi in 1005 AH/ 1596–97 CE, as one of the significant Persian-language historical works produced during the Safavid era and pertaining to the Kurds, one of Iran's largest ethnic groups. On the one hand, drawing upon the author's Kurdish identity, the study seeks to identify traces of Kurdish language influence within the text. On the other hand, given Sharaf Khan's engagement with both the Iranian and Ottoman polities - in which the Turkish language held formal status - the article explores the extent and nature of Turkish elements present in the *Sharafname-ye Bedlisi*. Furthermore, considering the Ottoman state's contacts and interactions with Christian communities in Eastern and Central Europe and the Caucasus, the research also attempts to detect lexical borrowings from the languages of these regions within the work.

Material & Method

For the purposes of this research, the complete text of the *Sharafname-ye Bedlisi* was examined in full, and extracts containing Kurdish, Turkish, and other non-Persian lexical items were systematically collected and analyzed. In order to address the research questions with greater precision, several contemporaneous texts from the same compositional period were also consulted. Methodologically, the study is grounded in textual analysis, description, and interpretive explanation. The edition utilized throughout is that of Vladimir Vélyaminov Zernov, republished in Iran by Asatir Publications.

Results & Discussion

The data extracted from the *Sharafname-ye Bedlisi* reveal clear influences from both Kurdish and Turkish, manifesting in the form of Kurdish lexical items as well as Turkish words and phrases embedded within the Persian text. The author's treatment of "Kurds" and the "Kurdish language" reflects, on the one hand, an objective and impartial perspective toward this Iranian ethnic group and its language - understandable given his own Kurdish origin and native command of the language - and, on the other hand, attests to his considerable linguistic awareness. As discussed in the article, Sharaf Khan's explanation of the variant orthographic forms *Ruzaki*, *Ruzhaki*, and *Rujaki* may be regarded - to the best of this writer's knowledge - as the earliest systematic articulation of the distinction between the Northern and Southern branches of the Western Iranian languages. Furthermore, the examination of the data indicates that a portion of the vocabulary in the *Sharafname-ye Bedlisi* derives from the languages of various Christian peoples with whom the Ottoman state maintained contact.

Conclusion

The findings of this study demonstrate that Kurdish, Turkish, and several languages of Christian people with whom the Ottomans had contacts of interactions all left discernible imprints on the Persian of the *Sharafname-ye Bedlisi*. In addition to individual lexical borrowings, Turkish poems and sentences appear, and some Persian constructions in the text exhibit Turkish

syntactic patterns - a phenomenon attributable to the longstanding historical interaction between the Turkish and Persian languages. Upon closer examination, a significant proportion of the Turkish-Mongolian vocabulary in the *Sharafname-ye Bedlisi* pertains to military and administrative affairs, which may be understood as a consequence of the Turco-Mongol character of the dynasties that, for centuries prior to the composition of the *Sharafname-ye Bedlisi*, ruled over “Turan” and “Iran”, and beyond. Lexical items originating from the languages of Christian peoples consist primarily of proper names, toponyms, official titles, and occasionally other terms that entered Ottoman Turkish through contact, and were subsequently recorded in the *Sharafname-ye Bedlisi*. Given that the Ottomans’ interactions with Christian people - at least during the period in which the *Sharafname-ye Bedlisi* was written - occurred predominantly in the Balkans, Eastern and Central Europe, and the Caucasus, the relevant loanwords in the text derive largely from Slavic languages (Russian, Ukrainian, Czech, Serbian, etc.), Armenian, Georgian, and German/ Austrian.

References

- Aghamajidi, Leila, Soltani Kouhbanani, Hassan, Shohani, Alireza, Shiri, Qahraman (2022). “Structural and Content Stylistics of Badlisi Sharafnameh by Sharafkhan Badlisi”, *Journal of the stylistic of Persian poem and prose*, 15(number 75), 249-267. DOI :10.22034/bahareadab.2022 .15 .6421 [in Persian]
- Amini, R. (2022). “An Analysis of the Entrance of Turkish Words into Persian: The Case Study of “Dastur-e Shahriaran”. *Journal of Linguistics & Iranian Languages*, 7(1), 21-39. <https://doi.org/10.22099/jill.2022.44446.1324> [in Persian]
- Amini, R. (unpublished). “The Concept of Iran in Sharafname-ye Bedlisi: A Comprehensive History of Kurdistan”. [in Persian]
- Amini, R. (unpublished). “A Textual Review of Sharafname-ye Bedlisi”. [in Persian]
- Bedlisi, Sharaf Khan ibn Shams-al-din (2022). *Sharafname The Comprehensive History of Kurdistan*, Vladimir Vélyaminov Zernov (ed.), Tehran: Asatir Publications. [in Persian]
- Doerfer, Gerhard. (1975). *Türkische und Mongolische Elemente im Neupersischen*, Franz Steiner Verlag GMBH Wiesbaden.
- Farahi, Abunaser (1982). *Nesab-al-Sebyan*, Mohamad Javad Mashkur (ed.), Tehran: Ashrafi. [in Persian]
- Hatam, Gholam Ali, Qadernezhad, Mehdi, Aghaie, Abdollah (2016). “The study of paintings of Amir Sharaf Khan Bidlisi’s Sharafnameh”, *Negareh Journal*, 11(38), 48-60. <https://doi.org/10.22070/negareh.2016.359> [in Persian]
- Jamshidian, Sajad (2010). “Bedlisi’s Sharafname and the Emergence of Kurdish Historiography”, *KETAB-E MAH-E TĀRIKH VA JOGHRAFIĀ*, 158, 102-105. [in Persian] <https://B2n.ir/qk9942>
- Janhunen, Juha. (2003). *The Mongolic Languages*, London: Routledge.
- Johanson, Lars, Csató, Éva Á. (2022). *The Turkic Languages*, second edition, London: Routledge.
- Mackenzie, D. N. (1961). “The Origins of Kurdish”, *Transactions of the Philological Society*, 86, 68-86.
- Menges, Karl H. (1994). *The Turkic Languages and People, An Introduction to Turkic Studies*, 2nd revised edition, Göttingen: Harrassowitz Verlag.
- Mohammadi, Z. and Feizi, K. (2020). “An Analysis of Sharafkhan Badlisi's Approach to Historical Perspective and Historiography in Sharafnameh”. *Historical Perspective & Historiography*, 30(25), 269-295. <https://doi.org/10.22051/hph.2021.28380.1393> [in Persian]
- Mohamadi, Z., Surni, B., Azizian, Z. (2013). “The Refraction of thKurds in “The History of Ālamāray-e Abasi””, *PAYAM-E BAHARESTAN*, 24, 311-326.[in Persian] <https://B2n.ir/qu5237>
- Mohamadi, Y. and Parsa, S. A. (2022). “Introducing the Sources of Persian Poems in Bidlisi’s Sharafnameh”. *Journal of Kurdish Literature*, 7(2), 109-133. <https://dor.isc.ac/dor/20.1001.1.26453657.1400.7.2.8.1> [in Persian]

- Oranski, I. M. (2000). *Introduction to Iranian Philology*, translated into Persian by Karim Keshavarz, Tehran: Payam Publications. [in Persian]
- Saidi, Modares, Zargarinezhad, Gholam Hosein (2017). "The Evolution and Development of the Meaning of the Term "Kurd" in the Sources of the Islamic Period (1th to 9th Century AH). *Tārīkh-i Īrān*", 10(1), 91-116. <https://dor.isc.ac/dor/20.1001.1.20087357.1396.10.1.5.9> [in Persian]
- Soltani, Anwar. (1998). "The Sharafnama of Bitlisi: manuscript copies, translations and appendixes", presented at the *International Conference on the Sharafnama in Berlin*, 1-3 May, 1998: <https://B2n.ir/xg5041>.
- Torkaman, Eskandarbeyg (2022). *TARIKH-E ĀLAMĀRĀY-E ABĀSI*, Tehran: Amirkabir. [in Persian]
- Windfuhr, Gernot. (2009). *The Iranian Languages*, London: Routledge.
- Ziahosseiny, S. M. (2024). *Contrastive Analysis & English Error Analysis*, 3rd edition, Tehran: Rahnama Press.

زبان‌شناسی و گویش‌های ایرانی

سال ۱۱، شماره ۲، پیاپی ۲۱ (تابستان ۱۴۰۵) شماره صفحات: ۴۷ - ۶۵

کردی، ترکی و زبان‌های دیگر در شرف‌نامه بدلیسی

رضا امینی 

دانشیار زبان‌شناسی، پژوهشگاه میراث فرهنگی و گردشگری، تهران، ایران.

چکیده

بررسی علمی چندسویه متن‌های مهم گذشته زبان فارسی به خوانش درست آنها یاری می‌رساند. بسیاری از این متن‌ها دربردارنده واژه‌هایی از زبان‌های دیگر هستند که بدون دانستن معنای آنها، دریافت معنای دقیق و کامل متن شدنی نیست. یکی از متن‌های مهم زبان فارسی که در آن پاره‌گفتارهای نافارسی چشم‌گیری به کار رفته شرف‌نامه بدلیسی (سده یازدهم هجری قمری) است. نویسنده این کتاب کُرد است، دربار صفویه پرورش یافته، و در ایران و عثمانی، به مقام امیری و امیرالأمرای «میرمیرانی» رسیده است. چنین پیشینه‌ای موجب شده که در کتاب او هم اثر زبان کُردی دیده شود، هم اثر زبان ترکی، هم اثر زبان‌های مردمان مسیحی که عثمانی‌ها در اروپای خاوری و مرکزی و قفقاز با آنها برخوردی چندسده‌ای داشته‌اند. در این مقاله، شرف‌نامه بدلیسی از این دید بررسی شده است. یافته‌های پژوهش، گواه وجود پاره‌گفتارهای کُردی، ترکی و زبان‌های دیگر در این کتاب و اثر این زبان‌ها بر متن آن است. این امر، در سطحی دیگر، نشان‌دهنده دادوستدهای فرهنگی گویشوران این زبان‌ها و فضای چندزبانه‌ای است که صاحب شرف‌نامه بدلیسی در آن زندگی می‌کرده است. اثرهای این دادوستدهای زبانی هنوز در زبان فارسی کم و بیش پا برجاست، هرچند که به خاطر دگرگون شدن شرایط ایران در چند سده گذشته، برخی از این اثرها از زبان فارسی ناپدید شده است.

واژه‌های کلیدی:

شرف‌نامه بدلیسی
کُردی
ترکی
مغولی
گرجی
اروپای خاوری

تاریخچه مقاله: دریافت: ۱ شهریورماه ۱۴۰۵

بازنگری: ۱۶ شهریورماه ۱۴۰۵

پذیرش: ۲۹ شهریورماه ۱۴۰۵

* نشانی رایانامه نویسنده مسئول: r.amini@richt.ir

امینی، رضا (۱۴۰۵). کُردی، ترکی و زبان‌های دیگر در شرف‌نامه بدلیسی. زبان‌شناسی و گویش‌های ایرانی ۱۱(۲)، ۴۷-۶۵.

<https://doi.org/10.22099/jlil.2026.56920.1498>

حق نشر ۲۰۲۶-نویسنده(گان). این یک مقاله با دسترسی آزاد است که تحت شرایط مجوز Creative Commons Attribution (CC BY-NC 4.0) منتشر شده است، استفاده، توزیع و تکثیر نامحدود را در هر رسانه‌ای مجاز می‌داند، مشروط بر اینکه به نویسندگان و منبع اصلی اشاره شود. هیچ اجازه‌ای از نویسندگان یا ناشر لازم نیست.

۱. مقدمه

«وام‌گیری»^۱ زبانی پدیده‌ای جهانی است، یعنی در همهٔ زبان‌ها، کم‌وبیش، اثر آن دیده می‌شود. این پدیده، در وهلهٔ نخست، در سطح واژگان روی می‌دهد، اما در برخوردهای درازمدت زبان‌ها، به سطح‌های آوایی-واجی^۲، ساخت‌واژی^۳ و حتی نحوی^۴ زبان هم گسترش می‌یابد (مدرسی، ۱۳۹۳). بنابراین، وام‌گیری که امری زبانی است، اما پیامد برخوردهای اجتماعی و اقتصادی و سیاسی گویشوران دو یا چند زبان است، می‌تواند سنجه‌ای باشد برای فهم چند و چون این برخوردها و آگاهی از نوع رابطه‌ای که میان گویشوران زبان‌های موقعیت برخورد وجود داشته است.

در جهان امروز که آمدوشدهای میان جامعه‌ها و کشورها بیش‌تر و آسان‌تر شده و گویشوران زبان‌های بیش‌تری در برخورد و تعامل با هم قرار گرفته‌اند، برخوردهای اجتماعی، اقتصادی و سیاسی بیش‌تر به چشم می‌آید و شاید دادوستدهای یک‌سویه یا دوسویهٔ زبانی گویشوران زبان‌ها هم بیش‌تر شده باشد. با این همه، چنین چیزی محدود به جهان امروز نیست، و ما از رهگذر بررسی متن‌های به‌جای‌مانده از گذشته، درمی‌یابیم که در گذشته نیز وام‌گیری زبان‌ها از یکدیگر - دست کم در جاهایی از جهان که جایگاه آمدوشدها و جابه‌جایی قدرت در میان مردمان گوناگون بوده - بسیار روی داده است. قلمرو تمدنی-فرهنگی ایران از این دست جاها است. و از همین رو است که در فارسی و دیگر زبان‌های این پهنه، نشان‌های گاه چشم‌گیری از زبان‌های دیگر دیده می‌شود. بررسی چنین نشان‌هایی افزون بر آن که زبان‌شناسان و دانشوران ادبیات را در شناخت دگرگونی‌های زبان فارسی و دیگر زبان‌های این قلمرو و عامل‌های آن به کار می‌آید، پژوهشگران دیگر زمینه‌های علمی را نیز در فهم تاریخ و دگرگونی‌های اجتماعی، سیاسی، اقتصادی و جابه‌جایی‌ها و دگرگونی‌های قومی این قلمرو به کار می‌آید.

در این مقاله، شرف‌نامهٔ بدلیسی (زین‌پس شرف‌نامه) - نوشتهٔ شرف خان بدلیسی در ۱۰۰۵ هجری قمری - که از متن‌های مهم زبان فارسی در دورهٔ صفویه، مرتبط با یکی از بزرگ‌ترین تیره‌های ایرانیان یعنی کُردها است^۵ بررسی شده، تا از یک سوی، بر پایهٔ آگاهی از کُرد بودن نویسندهٔ کتاب، اثر زبان کردی بر متن شرف‌نامه را دریابیم، و از سوی دیگر، با توجه به این واقعیت که او در هر دو دولت ایران و عثمانی - که زبان ترکی در آنها جایگاه مهمی داشته، «امیر» و «امیرالأمرا» («میرمیران») بوده - دریابیم که چه اثرهایی از زبان‌های ترکی بر متن شرف‌نامه وجود دارد. همچنین، با علم به این که دولت عثمانی با مسیحیان در خاور و مرکز اروپا و قفقاز در برخورد و دادوستد بوده، کوشش شده است اثر زبان‌های آنها در متن شرف‌نامه نیز شناسایی شود. بنابراین، دو پرسش کانونی مقاله عبارتند از:

الف. چه اثرهایی از نوع نگاه شرف‌خان به «کُردها» و زبان‌های کُردی و ترکی، در شرف‌نامه دیده می‌شود؟

ب. چه عنصرهایی از زبان‌های مسیحیان که عثمانی‌ها همواره با آنها در برخورد بودند، در شرف‌نامه وجود دارد؟

^۱ borrowing

^۲ phonetic-phonological

^۳ morphological

^۴ syntactic

^۵ کردها نقششان در دولت صفویه مهم بوده است، آن‌چنان که در کتاب‌هایی همچون *تاریخ عالم‌آرای عباسی*، نوشتهٔ اسکندر بیگ ترکمان، نقش آنها در جنگ‌های دولت صفویه برجسته است (ترکمان، ۱۳۸۷؛ محمدی و همکاران، ۱۳۹۳).

۲. پیشینه پژوهش

پژوهش‌هایی که تا کنون درباره شرف‌نامه بدلیسی انجام شده را می‌توان به دو دسته بخش کرد: دسته نخست پژوهش‌هایی را دربرمی‌گیرد که درباره شرف‌نامه هستند، و با موضوع پژوهش کنونی، به گونه‌ای، پیوند دارند، و پژوهش‌هایی که پیوند نزدیکی با موضوع پژوهش کنونی ندارند؛ با این همه، به پرسش‌هایی درباره آن کتاب، یا بر پایه آن، پاسخ گفته‌اند.

در دسته نخست، می‌توان به آقامجیدی و همکاران (۱۴۰۱) درباره سبک و محتوای زبانی، ادبی و فکری - محتوایی شرف‌نامه اشاره کرد که پژوهش سودمندی است و در آن به وجود واژه‌های کُردی و ترکی در شرف‌نامه نیز پرداخته شده؛ با این همه، اشاره آن به این که «شرف‌نامه از نظر فکری، دربردارنده رویکردی ستایش‌آمیز به کردها است»، سخن دقیقی نیست. امینی (در دست انتشار) نیز در مقاله‌ای «مفهوم ایران را در شرف‌نامه بدلیسی» بررسی کرده که موضوع آن در جاهایی با موضوع پژوهش کنونی هم‌راستایی دارد، چه این که در آن پژوهش نیز، بر پایه شرف‌نامه بدلیسی، هویت ایرانی کردها برجسته شده است.

در دسته دوم، می‌توان به پژوهش‌هایی همچون سلطانی^۱ (۱۹۹۸)، محمدی و پارسا (۱۴۰۰) و حاتم و همکاران (۱۳۹۵) اشاره کرد که در شناخت نسخه‌های گوناگون شرف‌نامه، «مآخذ فارسی» این کتاب و آشنایی با «نگاره‌های» نسخه نام‌ور به شرف‌نامه بودلیان کارهای سودمندی هستند. به این پژوهش‌ها می‌توان محمدی و فیضی (۱۳۹۹)، جمشیدیان (۱۳۹۰) و امینی (۱۴۰۵) را هم افزود که خواندن آنها برای آشنایی با «رویکرد تاریخ‌نگاری و تاریخ‌نگاری» شرف‌خان و ویژگی‌های متنی-زبانی شرف‌نامه سودمند است. با این همه، با نگاه به پژوهش‌های پیشین، درمی‌یابیم که تا کنونی پژوهشی با عنوان یا هدف پژوهش کنونی انجام نشده است.

۳. روش پژوهش

برای انجام این پژوهش، متن کامل شرف‌نامه بدلیسی مطالعه شد و پاره‌گفتارهای کُردی، ترکی و دیگر زبان‌های موردنظر پژوهش از آن استخراج و سپس بررسی شده است. با هدف پاسخ گفتن دقیق‌تر به پرسش‌های پژوهش، به برخی متن‌های دیگری که در دوره نگارش شرف‌نامه نوشته شده‌اند نیز مراجعه شده است. بنابراین، از دید روش، این پژوهش بر بررسی متنی، توصیف و تحلیل استوار است. در این مقاله، از آن نسخه‌ای از شرف‌نامه بهره گرفته شده که ولادیمیر ولیامینوف زرنوف^۲ آن را تصحیح و در ایران / انتشارات / اساطیر - در سال ۱۴۰۲ - منتشر کرده است. شرف‌نامه دو بخش دارد که در یک جلد منتشر شده است. از این رو، در ارجاع‌هایی که به این کتاب در این مقاله داده شده است، در درون دو کمان، نخست صفحه کتاب آمده و به دنبال آن «یک» یا «دو» که به ترتیب نشانگر جلد‌های «اول» و «دویم» کتاب هستند.

۴. تحلیل داده‌ها

۴-۱ «کُرد» و زبان کُردی در شرف‌نامه

جغرافیای زندگی کردها، آنها را در پیوند با مردمان بسیاری قرار داده که همین، در بسیاری جاها، گونه‌ای دوزبانگی / چندزبانگی و آمیختگی فرهنگی را بر سرنوشت آنها رقم زده است. این پیوند زبانی-فرهنگی در شرف‌نامه هم آشکار است. برای نمونه، شرف‌خان در جایی می‌گوید:

^۱ A. Soltani

^۲ V. V. Zernov

«... و مغنیان نغمه‌پرداز و مطربان شیرین زبان خوش آواز و رامشگران دلکش الحان سازنواز بیوسون اکراد و قاعده اعراب و طریقه فرس و قانون عجم در مقام خوانندگی و سازندگی درآمده آواز نشاط و ندای انبساط بدایره فلک زحل رسانیدند» (۴۰۴ یک).

بر پایه این پاره‌گفتار، می‌توان گفت در محیط زندگی شرف‌خان، هم «یوسون» («رسم/ شیوه») «آکراد» پذیرفته بوده، هم «قاعده اعراب»، هم «طریقه فرس و قانون عجم» و احتمالاً «یوسون» «اتراک» و دیگر مردمانی که «آکراد» با آنها در دادوستد فرهنگی-زبانی بوده‌اند. با این همه، از همان بخش آغازین *شرف‌نامه* که در آن نویسنده به تبار کردها می‌پردازد، هویت و تبار ایرانی آنها برجسته است؛ آن‌چنان که نخستین استناد او در این باره به داستان ضحاک در *شاه‌نامه فردوسی* است. در پیوند با این داستان، شرف‌خان چگونگی شکل‌گیری زبان کردها را هم به نبود درازمدت ارتباط میان آنها و «مردمان» دیگر مربوط دانسته و موصوف شدنشان به «کرد» را برآمده از «وفور شجاعت و تهور» در آنها دانسته است:

«... و چون مدت مدید و عهد بعید از اختلاط مردمان و تردد بلدان معرض و متوحش بودند برای خود لسان و زبانی پیدا کرده در جنکل و جبال در میانه بیشه و قلال آثار عمارت و زراعت و آبادانی کردند و بعضی از ایشان صاحب اموال و اغنام کشته بصحاری و بیابان رفتند و بروایتی از وفور شجاعت و تهور که لازمه ذات این طایفه است ملقب بکرد گشتند ...» (۱۳ یک)

با این همه، شرف‌خان، این را هم می‌گوید که برخی «حکما»، کردها را طایفه‌ای از جن، یا حاصل ازدواج «دیو» و «انسان» می‌دانند: «و بقول بعضی از حکما، [الاکراد طایفه من الجن کشف الله عنهم الغطا و بروایت برخی از مورخان دیو با انسان ازدواج کرده طایفه اکراد از ایشان پیدا شده ...» (۱۳ یک). سپس، شرف‌خان باز به زبان کردها اشاره کرده: «طایفه اکراد چهار قسم است و زبان و آداب ایشان مغایر یکدیگر است اول کرمانج دویم لر سیم کلهر چهارم کوران» (۱۳ یک). در این اشاره او، از کرد دانستن لرها که بگذریم، نکته درستی وجود دارد و آن «مغایر» بودن، یعنی گونه‌گونه بودن «زبان و آداب» کردها است که تا به امروز نیز دیده می‌شود.

در *شرف‌نامه*، محدوده «ولایت کردستان» هم بیان شده است:

«... و ابتدای ولایت کردستان از هرمز که بر ساحل دریای هند واقع شد و از آنجا بر خط مستقیم کشیده می‌آید تا در ولایت ملاطیه و مرعش منتهی می‌گردد و در جانب شمالی این خط ولایت فارس و عراق عجم و آذربایجان و ارمن است و بر طرف جنوبی دیاربکر و موصل و عراق عرب اما شعبات او از اقصای ولایت مشرق تا بینهایت دیار مغرب رسیده ...» (۱۳-۱۴ یک).

این‌که شرف‌خان ابتدای «ولایت کردستان» را «هرمز» می‌داند - و می‌توان گفت که منظور او از هرمز «هرمز بَری» در حدود میناب امروزی است و نه «هرمز بَحری» یعنی «جزیره هرمز» - از آن رو است که او لرها را نیز کُرد می‌داند، وگرنه در همه جغرافیایی که او ترسیم کرده «کردها» زندگی نمی‌کنند؛ مگر آن‌که «کُرد» را نه در مفهوم قومی آن، بلکه در مفهوم «رمه‌دار و کوچ‌نشین» به کار برده باشد که از معناهای این واژه هستند (سعیدی و زرگری‌نژاد، ۱۳۹۶).

اشاره‌ها و صفتهایی که شرف‌خان درباره کردها به کار بسته، نشان می‌دهد که او هویتی ویژه برای آنها قائل است. برای نمونه، در جایی، از زبان کسی، می‌گوید: «... ما طایفه اکرادیم و قاعده تراکمه در نزد مردمان ما مستحسن و مقبول نیست، [ترک آن اولیست ...» (۳۸۱ یک). این اشاره نشان‌دهنده مرزگذاری هویتی و آداب و رسومی میان «آکراد» و «تراکمه» در زمان نگارش *شرف‌نامه* است. شرف‌خان در جاهای زیادی نیز از «داد مردی دادن» «کردان گرد» و «دلیر» و «پرخاشجوی» و «تندخو»

و «خون‌خوار» بودن آنها سخن گفته که این سه صفت آخر نیز - به وارون آن چه در فارسی امروز داریم - گویا معنای مثبت دارند نه منفی:

«... از هر دو جانب کشش و کوشش بسیار کرده کردان کرد داد مردی و مردانگی دادند اما فایده نکرد» (۳۹۵ یک).

«... بدین دستور کردان پرخاشجوی و دلیران تندخوی دل از جان و جهان برداشته ...» (۳۹۹ یک).

«... امیر با کردان خون خوار خنجر کذار روی توجه بتسخیر قلعه و حصار آورد ...» (۴۰۳ یک).

گاه نیز دیده می‌شود که شرف‌خان‌گُردها را به مثابه انسان‌هایی صاف و صادق توصیف کرده است:

«آن کرد صافی صادق در جواب گفت ...» (۴۲۳ یک).

با توجه به اشاره‌هایی از گونهٔ زیر (دربارهٔ برخی اولاد «حکام اردلان») درمی‌یابیم که صاحب *شرف‌نامه* کوشیده است آن چه می‌نویسد موثق باشد:

«اما چون احوال اینجماعت که مذکور شد نزد راقم حروف محقق نبود و از کسی معتمدالقول نیز چیزی نشنیده بود که بر آن اعتماد توان کرد لاجرم بمسامحه و اغماض از آن درگذشت و شروع در احوال جمعی ازین طبقه نمود که بکرات از ثقات قدسی سمات احوالات ایشان استماع افتاده و بتواتر بصحت رسیده و بعضی را معاینه مشاهده نموده چه قرار با خامهٔ دوزبان درین نسخه بی‌سامان آنست که اصل قصه و داستان را بلا زیاده و نقصان در عبارت مختصر رقم نماید و از اقوال مختلفه بیمزه که سبب استنکار ارباب عقول گردد معرض و مستحوش باشد ...» (ص ۸۳)

بر این پایه، گفتنی است که در روایتی که شرف‌خان از پیشینه و زندگی گُردها به دست داده است، تعصب و یک‌سونگری دیده نمی‌شود، و او در شناساندن گُردها، آن چه را نوشته، خود خوانده یا شنیده یا بدان باور دارد. در این روایت، با آن که شرف‌خان خود گُرد بوده از انتقاد از گُردها یا بیان صفتهایی که دست کم امروزه ممکن است ناپسند به دید آیند، دربارهٔ آنها پرهیز نکرده است؛ برای نمونه، می‌گوید:

«... و اکثر این طایفه شجیع و متهور و سخی و متکبر باشند چنانچه از کمال تهور و شجاعت و کثرت مردانگی و غیرت اسم دزدی و قطاع الطریقی بر خود می‌نهند و درین وادی سربازی کرده خود را بکشتن میدهند و دست‌کدایی بجهت یک نان بدونان و لثیمان دراز نمی‌کنند ... بمقتضای من تفکر فی العواقب لم یشجع در اکثر امور دنیوی و شغل مهمات و معاملات آن بی‌فکر و بی‌تأمل اند ...» (۱۴ یک).

یا باز جای دیگر دربارهٔ «اُکُراد» می‌گوید:

«... و طوایف اُکُراد متابعت و مطاوعت همدیگر نمی‌کنند و اتفاق ندارند چنانچه جناب فضایل مآبی مولانا سعدالدین ... در صفت اُکُراد می‌گوید هر یک بدعوای انفراد رایت استبداد برافراشته‌اند و در قلال جبال باستقلال مجبول کشته بغیر از کلمهٔ توحید در هیچ امور اتفاق ندارند و سبب نفاق این طایفه را چنین روایت می‌کنند که چون صیت و صدای نبوت محمدی ... در اطراف و اکناف عالم غلغله افکند خواقین جهان و سلاطین عالیشان را داعیه آن شد که حلقهٔ بندگی و مطاوعت آن سرور را در کوش کنند ... اغوزخان که در آن زمان از عظمای سلاطین ترکستان بود از اعیان کرد بغداد نام کریه منظر دیوپیکر زشت چهره سیه‌چرده را بطریق رسالت باستان اقبال آشیان خواجه کونین و سید ثقلین ... ارسال نموده ... چون ایلچی کریه منظر بنظر سعادت اثر حضرت خیرالبشر درآمد از هیات و ماهیت او متنفر و منزجر کشته از عشایر و قبایل او سوال فرمود گفت از طایفهٔ اُکُرادم آنحضرت فرمود که حق سبحانه و تعالی این طایفه را موفق باتفاق نکرداند والا عالمی در دست ایشان تباه خواهد شد دیگر از آن روز دولت عظمی و سلطنت کبری میسر این طایفه نشده ... و چون در میانه طایفه (اُکُراد) فرمان‌فرمایی نافذالحکم نیست اکثر سفاک و بی‌باک و خونریز می‌باشند

چنانچه باندک جرایمی فساد بسیار میکنند و دیت نفس کامله بدختری یا اسبی یا دو سه راس چاروا معمول شده و دیت سقط دست و پا و چشم و دندان چندان معتبر نیست ...» (۱۶-۱۷ یک)

پیداست که برخی از این گفته‌ها پایه تاریخی ندارند. دست کم، با توجه به حدودی که شرف‌خان خود برای سرزمین‌های کُردنشین برشمرده، «اعیان کرد» نمی‌توانسته‌اند در «ترکستان» و نزد «اغوزخان» بوده باشند که او بخواهد از میان آنها یکی را «بطریق رسالت» به سویی بفرستد. در *شرف‌نامه* در مورد‌های چشم‌گیر دیگری نیز نویسنده نگاه انتقادی به کُردها دارد یا از آنها به‌نیکی یاد نمی‌کند؛ آن‌چنان که از کُردها به عنوان «صحرائشینان و دیوساران بدنهاد» یاد کرده که «که هرگز در خواب و بیداری روی آبادانی ندیده بودند»:

«... چون صحرائشینان اکراد و دیوساران بدنهاد که هرگز در خواب و بیداری روی آبادانی ندیده بودند داخله قصبه خوی شدند هر یک خود را بمثابة کودرز و کیو و سام نریمان تصور کرده شاه طهماسب ما را در مقابل عسکر روم نهاده است ... محصل کلام چنان عجب و غرور پیدا کرده که بالاتر از آن تصور نتوان کرد مشهور است که چند نفر از متعینان بدکان حلوایی رفته حلوای بسیار تناول کردند در محل رفتن چون استاد حلوایی از ایشان قیمت حلوا طلب نمود گفتند که شاه این شهر را با حلوا بما عنایت کرده است این مثل در میانه خلق بترکی مشهور شد که شهر بزم حلوا بزم وهم» (۳۱۱-۳۱۲ یک).

در جاهای دیگری نیز صفت «دیوسار» را شرف‌خان درباره «اکراد» به کار برده است:

«چون دو سه سال برین منوال گذشت جماعت بسیار از اکراد دیوسار بر سر خود جمع نمود ... و رونده و مترددین از اوضاع ناهموار اکراد متزجر کشته جمعیت ایشان باعث تفرقی خلق شد.» (۳۲۴ یک)

«سلیمان بیک بن قلی بیک بن سلیمان ... موازی سیزده هزار پیاده و سوار از اکراد دیوسار جمع نمود تاخت بولایت زرزا برده آنجا را نهب و غارت کرد ...» (۲۷۸ یک)

«دیونهاد» از دیگر صفت‌هایی است که شرف‌خان، در چندجا، درباره «اکراد» به کار برده است:

«سلطان سلیمان بواسطه آنکه از رفتن او میانه اکراد دیونهاد فتنه و فساد پیدا نشود سنجاق دکر از توابع حلب بدو ارزانی داشته ...» (۲۲۴ یک).

در جایی هم شرف‌خان از «اکراد» به «طایفه پرفتنه و فساد» یاد کرده است:

«... حسین پاشا نیز عرض کرد که مادام که حکومت اکراد بجان فولاد بیک تفویض نشود کسی بضبط و صیانت آن طایفه پرفتنه و فساد قادر نیست» (۲۲۴ یک)

و در جایی هم از «سخنان بیهوده و لاف و کزاف» «طایفه اکراد» سخن گفته است:

«با وجود آنکه شرف‌خان را از علوم رمل و نجوم بهره تمام بوده ... اما بواسطه سخنان بیهوده و لاف کزاف طایفه اکراد ضبط خود نتوانست کرد ...» (۴۳۲ یک).

از نکته‌های درخور درنگ دیگر *شرف‌نامه*، اشاره‌ها به زبان کُردی است؛ اشاره‌هایی که برخیشان در زبان‌شناسی «زبان‌های ایرانی»^۱ و آگاهی به ریشه واژه‌ها و نام‌های کُردی مهم هستند. در این کتاب، «روزکی» که نام یک «عشیرت» و «امارت» کُرد است با املاهای گوناگونی دیده می‌شود؛ آن‌چنان که گاه در یک جمله دو املاء گونه‌گون دارد:

«... شخصی که متصدی امر حکومت و متکفل مهام امارت روزکی بود فوت شده منقطع النسل کشت و عشیرت روزکی از اینجا یکی در یکدیگر افتاده کردن باطاعت هم ننهادند ...» (۳۵۹ یک).

^۱ درباره «زبان‌های ایرانی»، نگاه کنید به آرانسکی (۱۳۷۹) و ویندفور / Windfuhr (۲۰۰۹).

«... پردلان روزکی را بقصد شبیخون بکنار رودخانه زلم فرستاد ... در خلال این احوال آقایان روجکی واقف شده ...» (۳۸۵ یک). توضیح خود شرف‌خان دربارهٔ چنین گونه‌گونی‌های املایی درخور درنگ است؛ آن‌جا که «در بیان احوال عشیرت روزکی و سبب وجه تسمیه ایشان» می‌گوید:

«... روزکی لفظ دریست و بعضی املاء آنرا بجیم و شین هم نوشته‌اند ... اما روزکی در اصل عبارت از یک روز است و کی که در آخر لفظ روز واقع شده کاف و یای وحدت است مثل خواجکی و پردکی و امثال آن و بعضی از فصحای بلاغت انتما برانند که املای جیم قاعده فصحاء عرب باشد که هرگاه ژا در لفظی دری باشد بجیم درج میکنند و شین بمقتضای طبع بلغای کردی باشد چه از ثقات روات منضبط است و در سلک صحاح اخبار و آثار منخرط که عشیرت روزکی در یکرروز از بیست و چهار قبیله اکراد در موضع طاب من اعمال ناحیه خویت جمع کشته ...» (۳۵۸ یک).

آن‌چه شرف‌خان در این‌جا گفته است را می‌توان نخستین بیان مستدل تفاوت‌های زبان‌های ایرانی باختری شمالی^۱ و زبان‌های ایرانی باختری جنوبی^۲ دانست؛ امری که در سدهٔ بیستم نگاه کسروی تبریزی (۱۳۹۶: ۴۱-۴۰) را به خود جلب کرد:

«... آنچه در شمال بیشتر «شین» بوده، در جنوب «سین» می‌شده؛ چنان‌که اکنون نیز آنچه در شمال «شمیران» و «شمیرم» هست، در جنوب به جای آن «سیمیران» و «سمیرم» می‌باشد ... نیز آنچه در شمال «گاف» بوده، در جنوب «جیم» می‌شده؛ چنان‌که هنوز هم در شمال «گهرام‌دز» و در جنوب پاسخدهٔ آن «جهرم» می‌باشد. نیز آنچه در شمال «زاء» بوده در جنوب «دال» می‌شده؛ چنان‌که هنوز هم به جای «دانم» و «داماد» که از زبان جنوب است، در شمال «زانم» (در کردی) و «زوما» (در سمنانی) گفته می‌شود. اینها چیزهایی است که ما یافته‌ایم و باشد که چیزهای دیگری نیز بوده.»

پاره‌گفتار بالا گواه آگاهی کسروی تبریزی به گونه‌گونی‌هایی نظام‌مند دو گروه زبانی «شمال» و «جنوب» است که با توجه به نمونه‌های به‌دست‌داده‌شده و اشاره‌های او می‌توان آنها را با زیرشاخه‌های «شمالی» و «جنوبی» زبان‌های ایرانی باختری یکی دانست. پس از کسروی تبریزی، کسانی همچون مکنزی^۳ (۱۹۶۱) به گونه‌گونی‌های قاعده‌مند دو زیرشاخهٔ نام‌برده پرداختند. اشارهٔ شرف‌خان به تفاوت‌های آوایی «دری»/فارسی با کُردی و تبیینی که برای آن به دست داده، گذشته از آن‌که گواه آگاهی او به پیوند زبان‌های فارسی و کُردی است، گویای آن است که خوانش علمی متن‌های معتبر فارسی، ما را به دانشی که در گذشته در زمینهٔ زبان‌شناسی درقلمرو تمدنی - فرهنگی ایران تولید شده آگاه می‌کند، تا درمی‌یابیم که فقط پژوهشگران امروزی باخترزمین در این زمینه کار نکرده‌اند.

در شرف‌نامه با شماری واژهٔ کُردی هم روبه‌رو می‌شویم. در دو مورد، عدد «دوازده» به صورت «دوانزده» آمده که به «دوانزه» در گونه‌هایی از کُردی نزدیک است (ن. ک. مردوخ روحانی، ۱۳۹۲):

«... و دوانزده هزار فلوری از مال خود بخزینہ پادشاهی عاید می‌سازم.» (۱۳۵ یک)

«... و در شهر نیشابور دوانزده روز شمار کشتکان نمودند سوای عورات و اطفال هزار هزار و هفتصد و چهل و هفت هزار در

قلم آمد.» (۷-۸ دو)

به همین سان، در چند جا هم دیده می‌شود که شرف‌خان «پاییز» را به صورت «پایز» نوشته که در کردی نیز همین‌گونه است: «و در هنگام پاییز ... متوجه قشلاق تبریز کشته» (۱۶۸ دو)

¹ Northwestern Iranian languages

² Southwestern Iranian languages

³ D. N. Mackenzie

«عسکر قزلباش از شدت کرما مضطرب گشته ... قرار بدان دادند که مهم محاصره بفصل پاییز و زمستان انداخته ...» (۱۷۷دو)

این که چرا شرف‌خان این واژه‌ها را به صورت کُردی آنها در *شرف‌نامه* آورده دانسته نیست؛ شاید گواهی بر «تداخل^۱» (ن. ک. ضیاء حسینی، ۱۴۰۳) زبان کُردی او در نوشتار فارسی‌اش باشد.

در چند مورد هم پس از اشاره به واژه‌ای، گفته شده است که برابر آن در کُردی چیست:

«و ستونی چند در آنجا نصب کرده‌اند که مردم آن دیار تشخیص نکرده‌اند که چه درختست و بعضی می‌گویند درختیست که ترکان ایت بورنی و کردان شیلان مینامند.» (۲۱۱یک)

«... ماک خلف المشهور بجف سرخ در اصطلاح اکراد یعنی خلف چشم سرخ ...» (۱۵۴یک)

«... آخر از کثرت استعمال طایفه اکراد که سرخ را سهر می‌گویند بسهران اشتها دارند ...» (۲۷۲یک)

در چند مورد هم می‌بینیم که شرف‌خان در توضیحی که به دست داده، به تلفظ کُردی نام‌های خاص اشاره کرده است. چنین اشاره‌هایی در ریشه‌شناسی نام تیره‌ها و طایفه‌های کُرد راه‌گشا می‌نماید:

«و چون طایفه اکراد عزالدین را عزیزین می‌خوانند حکام آنجا بعزرائی مشهورند.» (۱۹۱یک)

«و در السنه و افواه مذکور که اسم خیزان در اوایل سحرخیزان بوده ... آخر از کثرت استعمال اکراد که ایشان همواره اسم را بتصغیر می‌خوانند چنانچه شمس‌الدین را *شمو* و عزالدین را *عزو* و جمشید را *جمو* و ابدال را *ابدو* درین اسم اسقاط لفظ سحر کرده *خیزان* گفته‌اند ...» (۲۱۰یک)

«حکام حکاری که از نسل شمس‌الدین‌اند و باصطلاح اکراد ایشانرا *شمو* و حکام عمادیه که از نسل بهاء‌الدین‌اند ایشانرا *به‌دین* و حکام کلیس که از نسل منتشاند ایشانرا *مند* می‌خوانند.» (۲۲۰-۲۲۱یک)

در چند مورد هم در *شرف‌نامه* به کُردی اشاره شده، بی‌آنکه پاره‌گفتاری از این زبان آمده باشد:

«... محمود عمادان شخصی ... در دیوان حاضر بود بلفظ کُردی توجه بطرف شمس‌الدین بیک کرده گفت بعد از آنکه ولایت موروثی و اوجاق روزکی از ید تصرف ما برود زندگانی ما بچه کار می‌آید ...» (۴۴۱یک)

«و بکر بیک روزبهرانی ... از مقدمات روزکی واقف شده بزبان کُردی گفت که زنهار بقول جهلای اکراد عمل نمایی ...» (۴۴۱یک)

«... امیر شمس‌الدین از سخن او بغایت خشناک و اعراضی گشته ... گفت ای کرد نادان این سخنان یاه و هذیان بود که بر روی من کفتی ... و عمر زبان استکانت بصنوف اعتذار کشاده گفت ... بنده ترک ادب نکرده بلکه از محض دولت خواهی و خیراندیشی کلمه چند باداء کردانه و وضع صادقانه بموقف عرض رسانیده ...» (۳۸۳یک).

در موردهایی هم، در *شرف‌نامه*، به سروده شدن شعر به زبان کُردی یا شاعران کُرد اشاره شده است:

«یعقوب بیک بن محمد بیک ... طبع نظم داشته ... اکثر اشعارش بزبان کردیست بلکه در آن فن دیوانی دارد.» (۲۴۰یک)

«شعرای اکراد قضایای احوال او را ... نظم کشیده ...» (۲۸۱یک)

این دست اشاره‌ها در شناخت پیشینه ادبیات نوشتاری کُردی^۲ اهمیت دارد. در برخی جاها نیز با توجه به بافتی که درباره آن سخن گفته شده است، درمی‌یابیم که گفته خاصی می‌تواند کُردی باشد، همانند «هی بنقاره» در پاره‌گفتار زیر:

«... مادر پیر دولتیار در آن معرکه کیرودار بر اسب بی زین سوار گشته در عقب کریختکان افتاده فریاد میکرد *هی بنقاره* یعنی *اول نقارهای اینجماعت را بستانید ...*» (۲۶-۳۲۵یک).

¹ interference^۲ در این زمینه، نگاه کنید به بلو (۱۳۹۴).

۲-۴ زبان‌های تُرکی و مغولی در شرف‌نامه

تُرکی شناخته‌شده‌ترین زبان خانوادهٔ زبان‌های تُرکی تبار^۱ است. این زبان در هر دو دولت ایرانِ عصر صفویه و عثمانی کارکرد مهم و محوری داشته است. صاحب شرف‌نامه از امیران دودمان‌های صفویه و عثمانی بوده است. از این رو، وجود پاره‌گفتارهای تُرکی در شرف‌نامه، همانند دیگر کتاب‌های نوشته‌شده در زمان نگارش این کتاب، برای نمونه، دستور شهریاران (ن. ک. امینی، ۱۴۰۱)، دور از انتظار نیست. در واقع امر هم در شرف‌نامه واژه‌های تُرکی پُرشماری دیده می‌شود که برخیشان، همچون «بیلاق» (۲۶۳ یک)، قشلاق (۲۶۴ یک)، «قرق» (۲۷ دو)، «اردو» (۱۴ دو)، «تفنگچی» (۴۳۲ یک)، «قراول» (۱۹۹ یک)، «یورش» (۴۵۵ یک)، «چاوش» (۱۳۳ یک)، «چاکر» (۳۷۵ یک)، «قزل‌باش» (۱۶۴ یک)، «یکچری»^۲ (۴۲۳ یک) و ... در فارسی امروز نیز کاربرد دارند، یا برای گویشوران آن شناخته هستند. شماری از این واژه‌ها هم در فارسی امروز کاربرد ندارند (ن. ک. جدول ۱).

بخش دیگری از پاره‌گفتارهای تُرکی شرف‌نامه، نام‌های خاص تُرکی آن هستند. برخی از این نام‌های خاص، جای‌نام^۳ هستند یعنی به جای مشخصی اشاره دارند: «... در دمور قاپی نام محلی ... (۱۸۶ یک)؛ «ناحیهٔ سوماقلی» (۲۷۳ یک)؛ «چون بموضع رابیه بولاق رسید» (۲۸۴)؛ «از راه اسکندر بولاغی» (۴۱۴ یک)؛ «و صحرای موش بین الاتراک بموش اواسی اشتهار دارد» (۳۵۳ یک)؛ «کوک کنبدی» (۱۸۲ دو)؛ «التون کوپری» (۴۴۰ یک)؛ «قلعهٔ اینه کول و یوند حصار و قلعهٔ ینکی شهر» (۱۱ دو)؛ «قلعهٔ قرقلی یکیکه‌سی» (۲۳ دو)؛ «در بیلاق المه قولاغی» (۱۳۵)؛ «بالفعل باسکی جامع اشتهار دارد» (۱۰۳ دو)؛ «بیلاق کوزلدرهٔ سلطانیه» (۱۷۲)؛ «و از آنجا بر راه التون کوپری متوجه تبریز شد» (۱۸۵)؛ ... برخی نام‌های خاص شرف‌نامه هم نام افراد هستند: «قزل ارسلان» (۳۶۴ یک)، «اورکمز بیگ» (۲۲۰ یک)؛ «الغ بیگ» (۲۹۰ یک)؛ «بکتاش قولی بیگ استاجلو» (۲۹۲ یک)؛ ... بخشی از برخی از نام‌های خاص در شرف‌نامه نیز تُرکی است: قره مصطفی پاشای (۱۸۸ یک)؛ سلیمان بیگ بیجن اوغلی (۱۴۶ یک)؛ امیر ارتق (۱۷۸ یک)؛ اوزن حسن بایندوری (۲۵۰)؛ بوداق بیگ بن عمر بیگ (۲۵۱)؛ در استنبول در یدی قلعه محبوس شده» (۲۴۹) ...

جدول (۱)

شماری از واژه‌های تُرکی شرف‌نامه

ترکی	فارسی	ترکی	فارسی	ترکی	فارسی	ترکی	فارسی	ترکی	فارسی
اقچه	سکهٔ نقره	یزک ^۴	جلودار	یورت	جایگاه/	الکاء	سرزمین	قپو آغاسی	رئیس خواجهگان
(۸۸ یک)	عثمانی	(۱۸۷ دو)		(۴۳ یک)	خانه	(۶۴ یک)		(۲۳۵ دو)	حرم‌سرا
آغا ^۵	عنوانی برای	التون ^۶	زر	ایلخی	رمة اسب	ایلچی	سفیر	اویماق	قبیله/ طایفه
(۹۹ یک)	بزرگان	(۱۳۱ یک)		(۲۹۲ یک)		(۳۷۴ یک)		(۱۶۴ یک)	

^۱ Turkic languages دربارهٔ این خانوادهٔ زبانی، نگاه کنید به یوهانسون و کساتو/ Johanson and Csattó ۲۰۲۲؛ منگس/ Menges، ۱۹۹۴.

^۲ = ینچیری (۴۲۳ یک)

^۳ toponym

^۴ برای نمونه: «یزک عسکر شاهی» (۱۸۷ دو).

^۵ برای نمونه: «و با وجود آنکه پدر و پسر او را با چهل نفر از آغایان ملکیشی بقتل آورده‌ایم» (۱۶۷ یک)

^۶ برای نمونه: «دویست هزار التون سرخ سکه سلطانی» (۱۳۱ یک).

ترکی	فارسی	ترکی	فارسی	ترکی	فارسی	ترکی	فارسی	ترکی	فارسی
کول ^۱	دریاچه	طغرا	فرمان	ایلی ^۲	فرمان-	یوسون ^۳	رسم/ شیوه	یراق	ساز و برگ
(دو۶۷)				(دو۷۷)	برداری	(۴۰۴)		(دو۱۰)	نظامی
یساول	ملازم	سنجاق ^۴	ولایت	بیک	امیر/	قیوجی‌باشی	رئیس	یوزباشی	رئیس گروهی
(دو۲۰)		(۱۴۲ یک)		(۱۶۵ یک)	بزرگ	(۲۳۰ یک)	دربانان	(۴۴۵ یک)	صدنفره
ارپه لیق ^۵	گونه‌ای تیول	یرتاولی ^۶		تواچی	جارچی	ایشک‌آقاسی	رئیس	بیکلربیکی ^۷	مقام بیکلربیگ
		(دو۱۸۰)		(۴۲۸ یک)		(دو۱۸۱)	تشریفات	(۲۲۹ یک)	
							دربار		
قورچی	سرباز	طوی ^۸	جشن	اوجاق ^۹	میراث/				
(۳۱۳ یک)		(۴۳۴ یک)		(۱۰۸ یک)	دودمان				

دسته دیگری از نام‌های خاص *شرف‌نامه* هم مواردی همچون نام ترکی-مغولی سال‌ها را شامل می‌شود که نشان‌دهنده کاربرد

«گاه‌شماری دوازده حیوانی^{۱۰}» در زمان نگارش این کتاب است:

«در شهر سینه ست و تسعین و سبعمایه موافق فروردین ماه جلالی و موافق ایت ییل ...» (۳۷۲ یک).

«فی سینه احدی و سبعین و سبعمایه درین سال موافق ایت ییل امیر تیمور کورکان در خطه کش که حالا بشهرسبز مشهور

است قدم بر سریر سلطنت نهاد ... و اطلاق لفظ کورکانی بر وی بواسطه دامادی میرحسین است ...» (۵۹ دو)

«و جلوس همایونش در دوشنبه نوزدهم رجب همین سال موافق پیچین ییل ...» (۱۶۸ دو).

«این فقیر خاکسار ... در قصبه کرهرود من اعمال قم عراق در تاریخ بیستم ذی القعدة سینه تسع و اربعین و تسعمایه موافق

توشقان ییل تولد یافت ...» (۴۴۹ یک)

^۱ برای نمونه: «کنار کول وان گرفته بعدلجواز آمد ...» (دو۶۷)

^۲ برای نمونه: «راه برادری و ایلی» (دو۷۷)

^۳ برای نمونه: «و میرزا سلطان حسین قریب ده سال در بیابانهای خوارزم بیوسون قازاق بسر میرد». (دو۱۳۸)

^۴ برای نمونه: «سنجاق حسنکیفا» (۱۴۲ یک)

^۵ برای نمونه: «تا موازی دویست هزار اچه از قرایا ... بطریق ارپه لیق و افراز بدو ارزانی داشته ...» (۳۰۷ یک).

^۶ برای نمونه: «قمش اوغلان اوزبک بر سبیل یرتاولی ایلغار بولایت ری آورده ...» (دو۱۸۰).

^۷ برای نمونه: «بیکلربیکی» (دو۱۷۹)

^۸ برای نمونه: «... دختر محمد بیک حزوی را برای خود خواستکاری کرده هفت شبانه روز طوی عظیم ترتیب داده ...» (۴۳۴ یک).

^۹ برای نمونه: «اوجاق موروثی او را از تحت تصرف قزلباشیه بیرون آورده ...» (۱۶۷ یک)؛ در *شرف‌نامه*، «اوجاق‌لق» («موروثی») که بر پایه «اوجاق» ساخته شده نیز دیده می‌شود، برای نمونه: «سلاطین ماضی ناحیه سکمن آباد خوی را بطریق اوجاقلق بدو ارزانی داشتند» (۳۱۰ یک). این واژه به صورت «اوجاقلق» نیز در متن آمده است: «حکم همایون بقید اوجاقلق نافذ شده (۲۵۸ یک).

^{۱۰} درباره این گاه‌شماری که خاستگاه‌های گوناگون چینی، مغولی، ترکی، ایرانی و غیره برایش برمی‌شمارند و روایت‌های گونه‌گونی درباره چگونگی پدیداری آن وجود دارد، کاشغری (۱۳۷۶) - صاحب «دیوان لغات الترک» - در مدخل «بُرس» توضیح مفصّلی داده است. فراهی (۱۳۶۱) - صاحب «نصاب الصبّان» - نیز دوازده حیوان این گاه‌شماری را به شعر نام برده است:

«موش و بقر و پلنگ و خرگوش شمار
زین چار چو بگذری نهنگ آید و مار»

«آنگاه باسب و گوسفند است حساب
همدونه و مرغ و سگ و خوک آخر کار»

مصحح «نصاب البیان»، در بخش «ملحقات» نسخه‌ای که تصحیح کرده، شعر دَر بُردارنده نام ترکی این دوازده حیوان را هم از نسخه‌های دیگر کتاب یادشده آورده است (فراهی، ۱۳۶۱: ۷۷):

«سیچقان و اود و پارس و توشقان و لوی پس
ئیلان و یونت و قوی است نامهای سال»

«پیچی و پس تخاقوی و ایت است بعد از آن
تنگوز تو ای پسر صاحب کمال»

برخی از نام‌های خاص شرف‌نامه ساختی ترکی دارند. برای نمونه، در پاره‌گفتارهای زیر «ایقوت اوغلی»، «خالق ویردی»، «قوتور دره‌سی»، «چولاق خالد»، «بیمشچی حسن آغا»، «صارو قپلان»، «یازیچی اوغلی» و «دلو بهاء‌الدین» که سه نمونه نخست مضاف و مضاف‌الیه و شش مورد دیگر موصوف و صفت هستند، ساختشان ترکی است؛ و در آنها به وارون فارسی، «مضاف» و «صفت» پیش از «مضاف‌الیه» و «موصوف» آمده است:

طایفه قزلباش بسرداری ایقوت اوغلی» (۲۵۶ یک)

«خالق ویردی بیک» (۳۱۱ یک)

«ناحیه قتور دره‌سی» (۳۱۳ یک)

«... پسرش خالد بیک بملازمت شاه اسمعیل صفوی مبادرت نمود در یکی از معارک ازو آثار مردانگی ... بظهور آمده حتی یک دست او از مفاصل جدا کشته شاه اسمعیل دستی از طلا ساخته بجای دستش نصب کرده موسوم بچولاق خالد کردانید ...» (۳۲۹ یک).

«حسن آغای قپوجی‌باشی المشهور بیمشچی حسن آغا» (۲۳۰ یک)

«قاسم بیک که از کمال بهادری و شجاعت و فرط تهور و شهامت بصارو قپلان مشهور بود ...» (۱۴۳ دو)

«دیگر شیخ محمد کلیبولیست که بیازیچی اوغلی مشهور است» (۱۰۴ دو)

«سلطان سلیمان خان او را موسوم بدلو بهاء‌الدین کردانیده ...» (۱۹۷ یک)

«حمزه سلطان المشهور بقازق حمزه» (۱۷۸ دو)

ساخت‌هایی که «اوغلی» (به معنای «پسر/ زاده» و «ویردی/ وردی» (به معنای «داده») شکل گرفته‌اند، در ترکی بسیار متعارف هستند، آن‌چنان که در شرف‌نامه نیز نام‌های زیاد دیگری با این ساخت دیده می‌شود: بیجن اوغلی (۱۴۶ یک)، سلیمان بیک بیژن اوغلی (۳۸۸ یک)، ...؛ حق ویردی ترکمان (۱۸۱ دو)؛ الله ویردی (۱۷۰ یک)، شاهوردی (۵۳ یک)،

در موردهایی نیز دیده می‌شود که در پاره‌گفتاری فارسی، پاره‌گفتاری ترکی آمده است. برای نمونه، در بیت «نجم ثانی که نباشد بدو کونش ثانی *» و در دگر جا بود الله تعالی اعلم» (۱۴۶ دو)، واژه «کونش» («گونش») واژه‌ای ترکی به معنای «خورشید» است. در دو پاره‌گفتار زیر نیز «چوق شکر خدایه که استدوکمزی بولدوق^۱» و «دوه کوزی^۲» ترکی هستند:

«... کسی دیگر از حاکمان اکیل در قید حیات نماند مگر عورت پیر بدر که حامله بود و شب و روز چشم طایفه مرداسی بر آن وضع حمل بود ... تا روزی که موعد وضع حمل بود بطریق عادت بدر خانه آمده ... شخصی بیرون آمده بوضع رمز بزبان ترکی گفت که چوق شکر خدایه که استدوکمزی بولدوق بنابر آن نام آن پسر سعادت اثر بامیر بولدوق اشتهار یافت و حاکمان اکیل ببلدوقانی ملقب شدند» (۱۷۹ یک).

«و جمله سنگهایش سنک سوداست و بعضی را سنک سیاه که ترکان او را دوه کوزی میخوانند مانند شان غسل سوراخهای او را پر کرده صلابت پیدا کرده ...» (۳۵۷ یک).

در شرف‌نامه، واژه‌های ترکی هم گاه در دو معنا به کار رفته‌اند. برای نمونه، همان‌گونه که در جدول ۱ دیدیم، «کول»، در جاهایی به معنای «دریاچه» به کار رفته است^۳. اما در جاهای دیگری، این واژه - که به شکل «کل» هم در شرف‌نامه دیده می‌شود - به معنای آبگیرهای کوچک‌تر نیز آمده است:

^۱ به معنی «خدا را بسیار سپاس که آن‌چه را می‌خواستیم گرفتیم».

^۲ «دوه کوزی» به معنی «چشم شتر».

^۳ واژه «دریاچه» نیز در متن آمده است: «دریاچه مشهور ببولانق» (۳۵۵ یک).

«و تخمینا هزار و پانصد ذراع میان کوه بزمین فرو رفته کول آب عظیم پیدا شده ... و آب حوض بغایت صاف و سرد است» (۳۵۷یک).

«پیشگاه لشکر خود را که زمین زراعتگاه ارزن بود شب آب بسته کل عظیم شده ...» (۴۳۳یک).

اشاره به شعرگویی به زبان ترکی، یا نقل شعرهای ترکی در *شرف‌نامه* هم می‌تواند گواه علم شرف‌خان در این زبان و جایگاه آن در دوره و محل زندگی او باشد. دربارهٔ سلطان سلیمان عثمانی می‌گوید:

«گاهی زبان معجز بیان و لسان کوهر افشان باسلوب نفیس و طرز سلیس بگفتن اشعار آبدار و نظم و نثر بدایع نکار اشتغال نموده این دو بیت فارسی و ترکی از نتیج طبع وقاد آن خسرو قیصر نژاد است!»: [نظم:] باز از آتش دل دیده پر آبست مرا * کار این چشمه ز سر چشمه خرابست مرا * های و هودن فارغ اول عالمده سلطانلق بودر * پندنی کوش ایلکل مورک سلیمانلق بودر^۱ ...» (۲۳۲دو).

در جای دیگری نیز شرف‌خان به سروده شدن شعری ترکی اشاره می‌کند:

«در این اثنا محمود اوغلی شاعر که مداح سلیمان بیک بود در غزلی ترکی این بیت را در سلک نظم آورده بنزد حسن بیک فرستاد نظم شها اول بدلسیک کردی مطیع اولمز سلیمانه^۲» (۳۸۹یک).

جدول (۲)

شماری از واژه‌های مغولی شرف‌نامه

مغولی	فارسی	مغولی	فارسی	مغولی	فارسی
سبیه (۲۸۶دو)	سنگر	تالان (۱۴۴دو)	تاراج	یاسامیشی (۱۸دو)	آراستگی/ تدبیر
یساق (۲۷۵دو)	سیاست	منقلا ^۳ (۱۸۳دو)	پیش‌رو لشکر	بیلاک ^۴ (۱۶۵دو)	هدیه
سیورغال ^۵ (۳۷۷یک)	تیول	یاسا ^۶	مجازات/ قانون	یرلیغ ^۷ / یارلیغ (۲۳۵، ۴۴یک)	فرمان
الوس (۴۶یک)	قوم/ قبیله	ایلغار ^۸ (۵۱یک)	تاختن/ هجوم	تمغا (۹۱یک)	مهر/ نشان
یرغو (۲۸یک)	استنطاق/ بازپرسی	تنسوق/ تنسیق ^۹ (۵۸، ۱۹۰یک)	چیز نادر	تکامیشی ^{۱۰} (۹۰دو)	تکاپوی شتابان

^۱ به معنی «از های و هوی فارغ باش که سلطان بودن در عالم همین است/ پند کلید واقعی سلطنت در جهان است».

^۲ به معنی «شاهان آن کُردِ بدلیسی مطیع سلیمان نمی‌شود».

^۳ برای نمونه: «بر سبیل منقلای روانهٔ آذربایجان ساخت و خود متعاقب ایشان روان گردیده». (۱۸۳دو)

^۴ برای نمونه: «با بیلاکات و تحف و هدایا بدرگاه شاهی آمده». (۱۶۵دو)

^۵ برای نمونه: «و هم در این سال سلطان احمد بن سلطان اویس که اردبیل در وجه سیورغال او مقرر بود برادرش سلطان حسین خروج کرده ...» (۶۵دو).

^۶ برای نمونه: «... بیاسا رسیدند» (۱۶۵یک)

^۷ برای نمونه: «یرلیغ همایون موکد بلعنت(نامه)» (۳۷۳یک).

^۸ برای نمونه: «اما چون اسبان این طایفه جمام بود هرچند ایلغار کردند بکرد ایشان نرسیده چاروا بسیار هم ضایع شد» (۵۱یک).

^۹ برای نمونه: «با تحف و تنسوقات فراوان» (خاتمه: ۳۵)؛ «تنسیق امور امارت» (۱۹۰یک)

^{۱۰} برای نمونه: «... پسر خود میرزا جوکی را با امراء کبار برسم تکامیشی نامزد فرمود و ایشان تا ارضوم از قفای مخالفان رفته اثری ندیدند» (۹۰دو).

وجود واژه‌های مغولی^۱ در زبان‌های ایرانی، با توجه به برخوردهایی که دست کم از هزار سال پیش ایرانیان با مغولان داشته‌اند امری طبیعی است. از آن‌جا که بخش دوم *شرف‌نامه* به تاریخ «ایران» و «توران» و «عثمانی» اختصاص دارد که بخش بزرگی از آن هم‌زمان با قدرت یافتن مغولان در این سرزمین‌ها است، وجود واژه‌های مغولی در *شرف‌نامه* هم امری طبیعی است (ن. ک. جدول ۲). بسیاری از واژه‌هایی که ترکی پنداشته می‌شوند در اصل مغولی هستند. از همین رو، در واژه‌نامه‌های فارسی، چنین واژه‌هایی که بسیاریشان از رهگذر زبان ترکی وارد زبان فارسی شده‌اند، «ترکی-مغولی» معرفی می‌شوند.^۲

در *شرف‌نامه* گاهی به خاستگاه مغولی واژه‌ها اشاره شده است:

«فی سنه اربع و اربعین و سبعمایه در این سال دولت امیر شیخ ابواسحاق بن امیر شاه محمود اینجو روی در ترقی آورد ...

و اینجو در اصطلاح طایفه مغول ضابط اموال پادشاه را میکوبند» (۴۴دو).

دربارهٔ رواج «جاو»/ «چاو» یا نخستین پول کاغذی در ایران هم شرف‌خان روایتی از زمان گیخاتو خان آورده است که خاستگاه مغولی واژه را هم نشان می‌دهد:

«عزالدین مظفر بن محمد بن عمید ... بعرض صدر جهان رسانید که مصلحت وقت در آنست که بدستور ولایت خطای و مملکت قآن جاو در ممالک محروسهٔ خانی بعوض زر و نقره روان سازیم و من بعد بیع و شرا بدان جاو نمایند تا جنس سیم و زر بتمام واصل خزینة عامره شود ... بنابر آنکه در آن زمان صدر جهان را زیاده از پانصد تومان قرض بیوجه بهم رسیده بود ... سخن عزالدین مظفر مرضی طبع او کشته ... و باتفاق پولاد چنکسانک که ایلچی قآن بود این خیال محال را بعرض کیخاتون رسانید و پادشاه حکم نمود که ... در هر شهر جاو‌خانها بنیاد نهند و عمال جهت تمشیت آن کار تعیین نمودند در هر بلده و شهر مبلغی کلی در مؤنت بنای جاو‌خانه مصروف شد و این جاو کاغذ پاره مربع مستطیل بود که چند کلمه بخط خطایی بر آن نوشته و بر هر دو طرف آن کلمه صِبَّه لا اله الا الله محمد رسول الله و لفظ ایرنجی دورجی خاقان خطای بآن عبارت پادشاه را ملقب گردانیده بود مثبت ساخته و دایرة در میان کاغذ کشیده بودند و وزنش از نیمدرم تا ده درم و ایضا سطر چند بر آن کاغذ مسطور بود مضمون آنکه بتاریخ ثلث و تسعین و ستمایه پادشاه جهان این جاو مبارک را در ممالک روان گردانید تغییر و تبدیل کننده با زن و فرزند و پیوند خویش بیاسا خواهد رسید ...» (۱۴-۱۵دو).

با توجه به آن‌چه دربارهٔ حکومت مغولان در *شرف‌نامه* گفته شده، وجود نام‌های مغولی در آن نیز امری طبیعی است که در پاره‌گفتار اخیر نیز نمونه‌هایی از آنها را می‌بینیم.

۳-۴ زبان‌های مسیحیان در شرف‌نامه

بخش چشم‌گیری از *شرف‌نامه* به شرح تاریخ دودمان عثمانی اختصاص دارد که همواره با اروپاییان و مسیحیان در دادوستد یا جنگ بود. از همین رو، در این کتاب اشاره‌های چشم‌گیری هم به اروپا، اروپایی‌ها و مسیحیان وجود دارد که از آنها با عنوان‌هایی همچون «فرنچ» (۱۹۲دو)، «فرنکستان» (۱۲۶دو)، «فرنک» (۵۶یک)، «فرنکان لعین» (۶۳یک)، «کفره کرجستان و ارامنه» (۲۶۳یک)، «کفره فرنک» (۱۰۰دو)، «کفره» (۳۵۵یک)، «کرجی» (۷۹دو) و «ارمن» (۳۵۱یک) یاد شده است. دادوستدها و جنگ‌های عثمانی‌ها و اروپایی‌ها/مسیحیان، سوئے زبانی هم داشته است. بنابراین، در *شرف‌نامه* با واژه‌های برآمده از زبان‌های اروپایی‌ها/مسیحیان نیز برخورد می‌کنیم که شرف‌خان در این‌جا نیز گاه به خاستگاه یا ریشهٔ آنها اشاره کرده است. یکی از این

^۱ زبان مغولی/ Mongolian language از زبان‌های «خانوادهٔ زبان‌های مغولی»/ Mongolic Language Family است. ویژگی‌هایی همچون «هماهنگی واکه‌ای»/ vowel harmony و «ساخت‌واژهٔ پیوندی»/ agglutinative morphology مغولی را با زبان‌های ترکی تبار همانند کرده است. از دید واژگانی نیز اشتراک‌های زیادی در میان زبان‌های این دو خانواده دیده می‌شود. برای آشنایی، با خانوادهٔ زبان‌های مغولی، نگاه کنید به یانهونن/ Janhunnen (۲۰۰۳).

^۲ برای واژه‌های «تُرکی» و «مغولی»، در فارسی نو، نگاه کنید به دوئرفر/ Doerfer (۱۹۷۵).

واژه‌ها «قرال» است. این واژه که اکنون در ترکی استانبولی به صورت «kral» در معنای «پادشاه» دیده می‌شود، در شرف‌نامه بارها به کار رفته است. برای نمونه:

«فی سنه ثلث و خمسین و ثمانمائه درین سال یانقو قرال و قرال انکروس ویوده افلاق را با خود متفق ساخته بنیکبولی آمدند ... و یانقو قرال فرار کرده کافران پراکنده شدند.» (۱۰۰دو)

«فی سنه احدی و ستین و ثمانمائه درین سال محمد خانرا با قرال انکروس محاربه و مجادله صعب دست داده قرال زخمدار کشته بدرک اسفل رفت.» (۱۰۸دو)

«قرال المان و نیمچه باراده ولایت انکروس بیرون آمده ...» (۱۷۵دو)

«ایلچیان قرال المان و نمچه آمده بموعد سه سال قبول جزیه و خراج نموده» (۱۸۰دو)؛

«قرال اسپانیه» (۱۸۸، ۱۸۹دو)؛

«قرال بغداد» (۱۹۱دو)؛

«فرندوش قرال» (۱۹۳دو)؛

«قرال بچ و اسپانیه و قرال ریم بابا و دوقه افرنج و قرال اردل» (۳۰۵دو)

با توجه به پاره‌گفتارهای بالا، درمی‌یابیم که «قرال» در شرف‌نامه درباره پادشاهان «فرنگ» («المان و نمچه»، «اسپانیه»، ...) به کار برده می‌شود و از این رو می‌تواند از یکی از زبان‌های اروپایی وارد زبان ترکی شده باشد. همچنین موردهای همچون «یانقو قرال» و «قرال انکروس» نشان می‌دهد که این عنوان همانند «شاه» فارسی هم پیش از نام می‌آید (مانند «شاه عباس») و هم پس از آن (مانند «نادرشاه»). شرف خان درباره «قرال» و تعلق آن به زبان‌های نامسلمانان («اصطلاح کفره») توضیح داده است: «فی سنه سبع و خمسین و سبعمائه درین سال قرال انکروس که قرال در اصطلاح کفره عبارت از پادشاه است باتفاق امرای لاز و طایفه سرف بعزم مقابله و مقاتله سلطان اورخان متوجه اناتولی شد» (۴۹دو).

در بالا، کاربرد واژه «ویوده» را نیز می‌بینیم. این واژه، در شرف‌نامه، به مثابه مقام و لقب به کار برده شده است. برای نمونه، گفته می‌شود: «... درین سال حاکم افلاق که بدیرالغوا ویوده اشتها داشت ...» (۸۷دو)؛ یا «استیفان ویوده حاکم آنجا فرار کرد» (۱۲۰دو). معنای این واژه را نیز شرف‌خان در یکی از کاربردهای آن توضیح داده است:

«... اسفندیار بیک او را در کشتی نشانده نزد ویوده افلاق فرستاد و ویوده در اصطلاح رومیه داروغه و صوباشی را می‌گویند و بواسطه تخفیف حاکمان کافر را که خراج قبول نموده‌اند ویوده می‌نامند و نام اصلی او مرجو بود» (۷۶دو).

بر پایه این توضیح، مسلمانان پس از آن که «حاکمان کافر» گردن به اطاعت پایین می‌آوردند و به آنها «خراج» می‌دانند، با عنوان «ویوده» به آنها اشاره می‌کردند تا آنها را در حد «داروغه» فروکاسته باشند. در «کتاب‌الثانی فی اللغات الرومی» فرهنگ نصیری که در حدود ۱۰۹۵ هجری قمری / ۱۷۰۰ میلادی نوشته شده و واژه‌های «ترکی جغتایی، رومی، قزلباشی، روسی و قلماقی» در آن فهرست شده، «ویوده پاشا» به عنوان «حاکم قرا الوس» تعریف شده است (نصیری، ۱۳۹۳: ۲۷۱/مدخل «ویوده پاشا»). مصححان این فرهنگ در پانوشتی گفته‌اند «ویوده» (BoñBoda) کلمه ایست روسی به معنی حاکم (نصیری، ۱۳۹۳: ۲۷۱/مدخل «ویوده»). منظور از «اللغات الرومی» در این فرهنگ نیز واژه‌های ترکی نام‌ور به «عثمانی» است.

واژه دیگری که شرف‌خان درباره آن توضیح داده «بان» است که با توجه به پاره‌گفتار زیر، درمی‌یابیم در «اصطلاح کفره» فرنگ، به معنای «وزیر» است:

«... در خلال این احوال سلطان محمد خان از مغنيسا^۱ رسید باتفاق پدر بر سر کفار فجار رفته در مابین ایشان جدال و قتال عظیم واقع شده چنانچه بان له و بان چه و چند بان دیگر در آن معرکه خونخوار ب خاک بوار افتاد و بان در اصطلاح کفره فرنک عبارت از وزیر است.» (۱۰۰دو)

«ویره» کردن هم به نظر می‌رسد که واژه‌ای مربوط به زبان‌های اروپایی باشد که خود شرف‌خان آن را به «امان طلبیدن» معنا کرده است:

«کافران ویره کردند یعنی امان طلبیدند که قلعه تسلیم نمایند و خودشان در کشتی و سفینه نشسته با اهل و عیال روانه اوطان و دیار فرنکستان شوند طایفه یکچیری و قپو خلقی بصلح و ویره مصطفی پاشا راضی نکشته تیغ بیدریغ بفرنگان آخته جمله را از پای درآورده زنان و پسران و دختران ایشانرا اسیر کرده اموال و اسباب آنجماعت را نهب و غارت کردند ...» (۲۴۱دو)

«قپودان» از دیگر واژه‌های اروپایی به‌کاررفته در شرف‌نامه است. این واژه دگرگون‌شده واژه «کاپیتان» است که وارد ترکی عثمانی و از آن رهگذر، فارسی شده است:

«فی سنه ست و اربعین و تسعمایه ... سلطان سلیمان خان سلیمان پاشای وزیر را باراده تسخیر هندوستان از راه دریای عمان روانه ساخت و خیرالدین پاشا باندره نام شخصی از قپودان افرنج در روی دریا دوچار یکدیگر شده ...» (۱۹۲دو)

«قپودان» در عنوان‌های مقام‌های نظامی عثمانی هم به کار رفته است: «علی پاشای قپودان (۲۳۶یک)». همان‌گونه که «اندره» از سویی و «انکروس»^۲، «اسپانیه» و «نمچه» و «بغدان»^۳ و «افلاق»^۴، «اردل»^۵، «بیچ»^۶، از سوی دیگر، در پاره‌گفتارهای اخیر نشان می‌دهد، بخش دیگری از عنصرهای زبان‌های اروپایی شرف‌نامه، نام‌های خاص مربوط به این زبان‌ها است. از آن میان، برخی نام‌ها همچون «اسپانیه» («اسپانیا») تلفظشان کم‌وبیش با صورت امروزشان در فارسی متفاوت است، و برخی همچون «نمچه» و «بغدان» و «بیچ» با واژه یک‌سره متفاوت دیگری یعنی «اتریش» و «مولداوی» / «ترانسیلوانیا» و «وین» جایگزین شده‌اند. «نمچه» در متن‌های فارسی به صورت «نمسه/ نمسا» هم آمده است.

همان‌گونه که مواردی همچون «تاویت نام شخص کرجی» (۱۹۱یک)، «کستندیل کرجی» (۷۹دو)؛ «سماعون کرجی» (۱۹۴یک) نشان می‌دهد، برخی از نام‌های خاص موجود در شرف‌نامه به زبان گرجی مربوط هستند که با توجه به هم‌جواری سرزمین‌های گرجیان با ایران و عثمانی ورود آنها به فارسی و ترکی امری طبیعی است. پاره‌گفتارهایی همچون زیر نیز وجود نام‌های ارمنی و گرجی در شرف‌نامه را نشان می‌دهد:

«دلکلوکیا و ... قلعه جریس و قلعه ایدنیک و قلعه سلیک ... را از تصرف کفره کرجستان و ارامنه بیرون آورده ...» (۲۶۳یک)

^۱ نام شهر و استانی در باختر ترکیه امروزی که در ترکی استانبولی، Manisa نوشته می‌شود.

^۲ «انکروس» / «انگروس» / engerus در واژه‌نامه برخط «kelime.com» ترکیه به عنوان «Eski Osmanlıların Macarlara verdikleri isimdi» یعنی «نام کهنی که عثمانی‌ها به مجارها داده بودند» تعریف شده است (https://B2n.ir/yz1331). از این رو، آن را می‌توان برابر «مجارستان» دانست.

^۳ درباره این واژه که به صورت «بُغدان» تلفظ می‌شود، نگاه کنید به خسروشاهی (۱۳۹۹).

^۴ «افلاق»، یا «افلاق»، نام تاریخی سرزمینی در جنوب رومانی میان رشته کوه‌های کارپات و رود دانوب» (رضا، ۱۳۹۸).

^۵ Erdély که نام مجاری و تاریخی «ترانسیلوانیا» / Transylvania در باختر رومانی امروز است (اورت-هیث / Everet-Heath، ۲۰۰۵: مدخل Transylvania).

^۶ این واژه که در ترکی عثمانی «بیچ» تلفظ می‌شود و در ترکی استانبولی به صورت Beç نوشته می‌شود، برآمده از واژه مجاری Bécs است که نام شهر وین در این زبان است (اورت-هیث / Everet-Heath، ۲۰۰۵: مدخل Vienna).

افزون بر نام‌های خاص، واژه‌هایی از زبان‌های گرجی و ارمنی در *شرف‌نامه* به چشم می‌خورد، از جمله واژه گرجی «ازناور» به معنای «شریف و بزرگ قوم»:

«قرايوسف کستندیل را با برادرش و سیصد نفر ازناوران کرجی بدوزخ فرستاد.» (۷۹دو)

«... مصطفی پاشای سردار با لشکر بیشمار بعزم تسخیر کرجستان و شیروان در حرکت آمده منوچهر ولد غرغره کرجی حاکم آخصه

بقدم اطلاعات پیش آمده با او جمع کثیر ازناوران کرجی بشرف اسلام مشرف کشته ...» (۲۵۷دو)

واژه «کوتان» در پاره‌گفتار زیر نیز شاید خاستگاه یا ریشه‌ای در زبان‌های ارمنی و گرجی داشته باشد:

«... و موش در اصطلاح ارامنه دمان را می‌گویند و از کثرت دمان اشجار مثمره در آنجا کمتر است. و موش صحرای خوب و علف

زار مرغوب دارد ... رعایای آنجا کاو و کاومیش و گوسفند بسیار نگاه می‌دارند چنانچه هر جفت گاو که عبارت از کوتانست بیست و

چهار کاو و کاومیش می‌بندند.» (۳۵۳یک)

«موش»، به نوشته شرف‌خان، از ناحیه‌های «بدلیس» و «بلده قدیم البناست» که با توجه به آن‌چه در پاره‌گفتار بالا آمده، در «اصطلاح» («زبان») ارمنی‌ها معنای «دمان» می‌دهد که این واژه گویا در کُردی معنای «مه» دارد (ن. ک. مردوخ روحانی، ۱۳۹۲: مدخل «مه meh»). واژه «کوتان» در این پاره‌گفتار نیز در هر دو زبان ارمنی و گرجی، معنای «شخم زدن» دارد. این معنا با موضوع پاره‌گفتاری که این واژه در آن به کار رفته پیوند دارد. با توجه به آن‌که «موش» منطقه‌ای بیش‌تر ارمنی‌نشین بوده - چنانچه شرف‌خان در وصف بیش‌تر این ناحیه از «قرايای ارامنه [که] قریب بصد خانوار در آن مرغزار پهلوی یکدیگر افتاده ...»، «قرايای اوقاف و املاک مع جزیه و خراج و چهار هزار نفر کفره [که] ...» و «حکام ارامنه» «قبل از ظهور اسلام» سخن گفته است (۳۵۵-۳۵۳یک) - ارمنی بودن این واژه محتمل است. با این همه، جستجوهای نگارنده نشان می‌دهد در «کُردی سورانی» هم واژه «کوتان» در معنای «جای نگهداری گوسفندان» کاربرد دارد که این معنا هم با موضوع پاره‌گفتار یاشده پیوند دارد. از این رو، «کوتان» را باید از آن دسته واژه‌های *شرف‌نامه* برشمرد که تعیین قاطع معنا و خاستگاهشان، دست کم بر پایه جستجوهای نگارنده این پژوهش، شدنی نیست.

۵. نتیجه‌گیری

از هر دو زبان کُردی و ترکی اثرهایی در *شرف‌نامه* دیده می‌شود. این اثرها به شکل کاربرد واژه‌های کُردی یا واژه‌ها و پاره‌گفتارهای ترکی در متن پدیدار شده است. نوع نگاه صاحب *شرف‌نامه* به «کُردها» و «زبان کُردی»، از یک سو نشان‌دهنده نگاه عینی و بی‌کسونگرانه به این تیره و زبان ایرانی است که خود هم برآمده از یکی و گویش‌ور دیگری بوده، و از سوی دیگر، گواه دانش زبان‌شناسی درخور توجه او است، آن‌چنان که همان‌گونه که در بالا آمد، شرح شرف‌خان درباره علت نگارش گونه‌گون «روزکی»، «روزکی» و «روجکی» را - تا آن‌جا که این نگارنده می‌داند - می‌توان نخستین بیان نظام‌مند تمایزهای شاخه‌های شمالی و جنوبی زبان‌های ایرانی باختری دانست.

افزون بر واژه‌ها، شعرها و جمله‌های ترکی، برخی پاره‌گفتارهای فارسی *شرف‌نامه* نیز ساختی ترکی دارند که این برآمده از برخورد دیرین زبان‌های ترکی و فارسی است. در بررسی دقیق‌تر درمی‌یابیم که بخش چشم‌گیری از واژه‌های ترکی-مغولی *شرف‌نامه* به امور جنگی و حکومتی مربوط هستند که این را می‌توان پیامد ترک - مغول بودن دودمان‌هایی دانست که از چندین سده پیش از زمان نگارش *شرف‌نامه* بر «توران» و «ایران» و جز آن حکمرانی کرده بودند.

شماری از واژه‌های شرف‌نامه هم برآمده از زبان‌های مربوط به مردمان گوناگون مسیحی هستند که دولت عثمانی با آنها در برخورد بوده است. این واژه‌ها، یا نام‌های خاص افراد هستند، یا جای‌نام، یا عنوان‌های مقام‌ها، و گاه واژه‌های دیگری که به مقتضای برخوردهای عثمانیان با مسیحیان از زبان‌های آنها وارد ترکی عثمانی و سپس نوشتار شرف‌نامه شده است. از آن‌جا که بیش‌تر برخوردهای عثمانی‌ها با مسیحیان - دست کم در دوره نگارش شرف‌نامه - در بالکان، خاور، مرکز اروپا و منطقه قفقاز بوده، این دست واژه‌های شرف‌نامه، بیش‌تر، برآمده از زبان‌های اسلاوی^۱ (روسی، اوکراینی، چکی، صربی، ...)، زبان‌های ارمنی و گرجی و آلمانی/ اتریشی بوده‌اند.

سپاس‌گزاری

از سرکار خانم دکتر پونه مصطفوی، دانشیار پژوهشگاه میراث فرهنگی و گردشگری، و سرکار خانم سوسن نزهت‌زاده، کارشناس پژوهشی پژوهشگاه میراث فرهنگی و گردشگری، برای راهنمایی نگارنده درباره برگردان فارسی برخی پاره‌گفتارهای ترکی بررسی - شده در این مقاله، بسیار سپاس‌گزاری می‌شود.

منابع

- آقامجیدی، لیلای، سلطانی کوهینانی، حسن، شوهانی، علیرضا، شیری، قهرمان. (۱۴۰۱). «سبک‌شناسی ساختاری و محتوایی شرف‌نامه بدلیسی اثر شرف خان بدلیسی»، *سبک‌شناسی نظم و نثر فارسی (بهار ادب)*، ۱۵، ۲۶۷-۲۴۹. DOI:10.22034/bahareadab.2022.15.6421
- آرانسکی، ای. م. (۱۳۷۹). *مقدمه فقه اللغة ایرانی*، ترجمه کریم کشاورز، تهران: انتشارات پیام.
- امینی، رضا. (۱۴۰۱). «واکاوی ورود واژه‌های ترکی به زبان فارسی»، *مطالعه موردی «دستور شهریاران»*، *زبان‌شناسی و گویش‌های ایرانی*، ۷(۱)، ۳۹-۲۱. <https://doi.org/10.22099/jill.2022.44446.1324>
- امینی رضا. (۱۴۰۵). «بررسی متنی - زبانی شرف‌نامه بدلیسی در مقایسه با فارسی امروز»، *زبان فارسی و گویش‌های ایرانی*، منتشرشده به صورت برخط: <https://B2n.ir/yy7730>
- امینی، رضا. (منتشرنشده)، «مفهوم ایران در شرف‌نامه: تاریخ مفصل کردستان».
- بدلیسی، شرف خان بن شمس‌الدین. (۱۴۰۲). *شرف‌نامه تاریخ مفصل کردستان*، به اهتمام ولادیمیر ویلیامینوف زرنوف، تهران: انتشارات اساطیر.
- بلو، جویس. (۱۳۹۴). «ادبیات نوشتاری کردی»، در *تاریخ ادبیات فارسی: ادبیات شفاهی زبان‌های ایرانی، کردی، پشتو، بلوچی، آسی، فارسی و تاجیکی*، زیر نظر دکتر احسان یارشاطر، فیلیپ کرینبروک، اولریخ مارزلف (ویراستاران انگلیسی) و احمد رضا قائم‌مقامی، غلامحسین کریمی دوستان، تهران: انتشارات سخن.
- ترکمان، اسکندریگ. (۱۳۸۳). *تاریخ عالم آرای عباسی*، تهران: امیرکبیر.
- جمشیدیان، سجاد. (۱۳۹۰). «شرف‌نامه بدلیسی و ظهور تاریخ‌نگاری کردستان»، *کتاب ماه تاریخ و جغرافیا*، ۱۵۸، ۱۰۵-۱۰۲.
- حاتم، غلامعلی، قادرزاد، مهدی، آقایی، عبدالله. (۱۳۹۵). «معرفی و مطالعه نگاره‌های شرف‌نامه امیر شرفخان بدلیسی (شرف‌نامه بودلیان)»، *نگره*، ۳۸، ۴۹-۶۰. <https://doi.org/10.22070/negareh.2016.359>
- سعیدی، مدرس، زرگری‌نژاد، غلامحسین. (۱۳۹۶). «تطور و تکوین معنایی واژه کُرد در منابع دوره اسلامی (قرن اول تا نهم هجری)»، *تاریخ ایران*، ۲۱(۷۹/۵)، ۹۱-۱۱۶. <https://B2n.ir/fw7244>
- فراهی، ابونصر. (۱۳۶۱). *نصاب الصبّیان*، با مقدمه و ملحقات و فهرس لغات و تصحیح و تعریب، به اهتمام دکتر محمدجواد مشکور، تهران: اشرفی.
- کاشغری، محمود بن حسین بن محمد. (۱۳۷۵). *نام‌ها و صف‌ها و ضمیرها و پسوندهای دیوان لغات‌الترک*، ترجمه و تنظیم و ترتیب الفبائی سید محمد دبیرسیاقی. تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.

¹ Slavic Languages

- کسروی تبریزی، احمد (۱۳۹۶). *آذری یا زبان باستان آذربایجان*، به کوشش عقاب علی‌احمدی، تهران: نشر هزار کرمان.
- محمدی، ذکراالله، سورنی، برومند، عزیزیان، زهره. (۱۳۹۳). «بازتاب عملکرد کردها در «تاریخ عالم آرای عباسی»»، *پیام بهارستان*، ۲۴، ۳۱۱-۳۲۶. <https://B2n.ir/qu5237>
- محمدی، ذکراالله، فیضی، کیومرث. (۱۳۹۹). «تحلیل رویکرد تاریخ‌نگاری و تاریخ‌نگاری شرف‌خان بدلیسی در شرف‌نامه»، *تاریخ‌نگاری و تاریخ‌نگاری*، ۱۱۰، ۲۶۹-۲۹۵. <https://doi.org/10.22051/hph.2021.28380.1393>
- محمدی، یداله، پارسا، سید احمد. (۱۴۰۰). «معرفی مآخذ اشعار فارسی کتاب شرف‌نامه بدلیسی»، *پژوهشنامه ادبیات کردی*، ۷(۲)، ۱۰۹-۱۳۳. <https://dor.isc.ac/dor/20.1001.1.26453657.1400.7.2.8.1>
- مدرسی، یحیی (۱۳۹۳). درآمدی بر جامعه‌شناسی زبان، ویراست دوم، تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- مردوخ روحانی، ماجد. (۱۳۹۲). *فرهنگ فارسی - کُردی دانشگاه کردستان*، سنندج: دانشگاه کردستان.
- نصیری، محمدرضا و عبدالجلیل. (۱۳۹۳). *فرهنگ نصیری، ترکی جغتایی، رومی، قزلباشی، روسی، و قلماقی به فارسی*، به کوشش حسن جواد و دکتر ویلم فلور، با همکاری مصطفی کاجالین، تهران/ تبریز: کتابخانه، موزه و مرکز اسناد مجلس شورای اسلامی/ انتشارات آیدین.
- Aghamajidi, Leila, Soltani Kouhbanani, Hassan, Shohani, Alireza, Shiri, Qahraman (2022). "Structural and Content Stylistics of Badlisi Sharafnameh by Sharafkhan Badlisi", *Journal of the stylistic of Persian poem and prose*, 15(number 75), 249-267. DOI :10.22034/bahareadab.2022.15.6421 [in Persian]
- Amini, R. (2022). "An Analysis of the Entrance of Turkish Words into Persian: The Case Study of "Dastur-e Shahriaran". *Journal of Linguistics & Iranian Languages*, 7(1), 21-39. <https://doi.org/10.22099/jill.2022.44446.1324> [in Persian]
- Amini, R. (unpublished). "The Concept of Iran in Sharafname-ye Bedlisi: A Comprehensive History of Kurdistan". [in Persian]
- Amini, R. (in Press). "A Textual Review of Sharafname-ye Bedlisi in Comparison to Modern Standard Persian", *Persian Language and Iranian Languages*. [in Persian] <https://B2n.ir/yy7730>
- Bedlisi, Sharaf Khan ibn Shams-al-din (2022). *Sharafname The Comprehensive History of Kurdistan*, Vladimir Vélyaminov Zernov (ed.), Tehran: Asatir Publications. [in Persian]
- Doerfer, Gerhard. (1975). *Türkische und Mongolische Elemente im Neupersischen*, Franz Steiner Verlag GMBH Wiesbaden.
- Farahi, Abunaser (1982). *Nesab-al-Sebyan*, Mohamad Javad Mashkur (ed.), Tehran: Ashrafi. [in Persian]
- Hatam, Gholam Ali, Qadernezhad, Mehdi, Aghaie, Abdollah (2016). "The study of paintings of Amir Sharaf Khan Bidlisi's Sharafnameh", *Negareh Journal*, 11(38), 48-60. <https://doi.org/10.22070/negareh.2016.359> [in Persian]
- Jamshidian, Sajad (2010). "Bedlisi's Sharafname and the Emergence of Kurdish Historiography", *KETAB-E MAH-E TĀRIKH VA JOGHRAFIĀ*, 158, 102-105. [in Persian] <https://B2n.ir/qk9942>
- Janhun, Juha. (2003). *The Mongolic Languages*, London: Routledge.
- Johanson, Lars, Csátó, Éva Á. (2022). *The Turkic Languages*, second edition, London: Routledge.
- Mackenzie, D. N. (1961). "The Origins of Kurdish", *Transactions of the Philological Society*, 86, 68-86.
- Menges, Karl H. (1994). *The Turkic Languages and People, An Introduction to Turkic Studies*, 2nd revised edition, Göttingen: Harrassowitz Verlag.
- Mohammadi, Z. and Feizi, K. (2020). "An Analysis of Sharafkhan Badlisi's Approach to Historical Perspective and Historiography in Sharafnameh". *Historical Perspective & Historiography*, 30(25), 269-295. <https://doi.org/10.22051/hph.2021.28380.1393> [in Persian]
- Mohamadi, Z., Surni, B., Azizian, Z. (2013). "The Refraction of thKurds in "The History of Ālamāray-e Abasi"", *PAYAM-E BAHARESTAN*, 24, 311-326.[in Persian] <https://B2n.ir/qu5237>
- Mohamadi, Y. and Parsa, S. A. (2022). "Introducing the Sources of Persian Poems in Bidlisi's Sharafnameh". *Journal of Kurdish Literature*, 7(2), 109-133. <https://dor.isc.ac/dor/20.1001.1.26453657.1400.7.2.8.1> [in Persian]
- Oranski, I. M. (2000). *Introduction to Iranian Philology*, translated into Persian by Karim Keshavarz, Tehran: Payam Publications. [in Persian]
- Saidi, Modares, Zargarinezhad, Gholam Hosein (2017). "The Evolution and Development of the Meaning of the Term "Kurd" in the Sources of the Islamic Period (1th to 9th Century AH). *Tārīkh-i Īrān*", 10(1), 91-116. <https://dor.isc.ac/dor/20.1001.1.20087357.1396.10.1.5.9> [in Persian]

-
- Soltani, Anwar. (1998). "The Sharafnama of Bitlisi: manuscript copies, translations and appendixes", Presented to the *International Conference on the Sharafnama in Berlin*, 1-3 May, 1998: <https://B2n.ir/xg5041>.
- Torkaman, Eskandarbeyg (2022). *TARIKH-E ĀLAMĀRĀY-E ABĀSI*, Tehran: Amirkabir. [in Persian]
- Windfuhr, Gernot. (2009). *The Iranian Languages*, London: Routledge.
- Ziahosseiny, S. M. (2024). *Contrastive Analysis & English Error Analysis*, 3rd edition, Tehran: Rahnama Press.

Extended Abstract

Lenition in Arabic Loanwords Borrowed into Russian Language; the Optimality Theory Framework

Maryam Moradi 

Associate Professor of Department of Russian Language, Faculty of Persian Literature and Foreign Languages, Allameh Tabataba'i University, Tehran, Iran.
m.moradi@atu.ac.ir

Sakineh Navidi-Baghi 

Assistant Professor of Linguistics, Shiraz University, Shiraz, Iran.
s.navidi@saadi.shirazu.ac.ir

Abstract

One of the ways new words enter the world's languages is borrowing. Through this phenomenon, a word from the donor language enters the recipient language due to language contact. Due to differences in phonotactics and the various phonological rules that govern languages, loanwords undergo changes during this transfer to adapt to the structure of the recipient language. Arabic loanwords, likewise, must undergo the filtering of the Russian phonological system in order to be incorporated, so as to conform to its phonotactic and phonological rules. In this research, the process of lenition in Arabic loanwords in Russian language has been examined within the framework of Optimality Theory. The data for this descriptive-analytical study have been extracted from the "Dictionary of Foreign Words in Russian" compiled by Spirkin and his colleagues. The results of the study show that out of 149 loanwords examined, which were extracted through complete sampling, 29 words have undergone various types of lenition processes: spirantization, voicing, degemination, and deletion. In a general overview, the Russian language employs two strategies: deleting the foreign consonant or replacing it with the closest consonant available in Russian in terms of place of articulation.

Keywords:

Russian language
Arabic language
Borrowing
Lenition
Optimality Theory

How to cite this article:

Moradi, M. Navidi-Baghi, S. (2026). Lenition in Arabic Loanwords Borrowed into Russian Language; the Optimality Theory Framework. *Journal of Linguistics & Iranian Languages*, 11(2), 67-88. doi: <https://doi.org/10.22099/jlil.2026.56534.1491>



COPYRIGHTS ©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publisher.

Introduction

* Corresponding author

Over time, languages inevitably undergo changes due to the various economic, political, and cultural interactions among countries, and these changes affect their lexical systems. Language contact can lead to the transfer of linguistic elements from the donor language into the recipient language. Loanwords undergo modifications owing to differences in phonotactics and the diverse phonological systems governing languages. Arabic loanwords, too, must pass through the filter of the Russian phonological system in order to enter Russian and conform to its phonological rules.

In this study, the phonological processes of lenition during the borrowing of lexical items from Arabic into Russian have been examined. Lenition is a type of sound change that affects consonants and makes them more similar to vowels. The greater the constriction of a consonant, the stronger it is, and the lesser the constriction, the weaker that consonant is (Naghzguye Kohan, 2025/1404: 245).

The present study seeks to answer the following questions:

1. What lenition changes have Arabic loanwords undergone in the phenomenon of borrowing in Russian?
2. What strategies has the Russian language adopted in encountering Arabic loanwords?

Materials & Methods

In this descriptive-analytical study, Arabic loanwords entering the Russian language were extracted from the *Dictionary of Foreign Words in Russian*, 7th edition, published in 1984, authored by Spirkin and colleagues. They were then compared with the original words in Arabic, and the phonetic changes made to the words were carefully recorded. The types of lenition processes, including spirantization, voicing, degemination, and deletion, were identified in the changes that occurred in 29 loanwords and were described and analyzed within the framework of Optimality Theory.

Research Background: Al-Kadimi (2010) studied the semantics and etymology of Arabic loanwords entering Russian; Bakhtiyarova and Fatkullina (2015) examined loanwords in the lexical system of Russian; Moradi and Ghorbani-Madavani (2020) investigated the category of semantic specification in Arabic loanwords in Russian; Reztsova (2011) provided a thematic classification of Arabic loanwords in Russian; Bakhtiyarova (2016) explored the phonetic aspect of Arabic loanwords entering Russian and phonetic assimilation in Arabic loanwords entering Bashkir; Moradi (2022) studied types of deletion processes in Arabic loanwords in Russian; and Kambouziya and Hashemi (2015/1394) examined the phonological adaptation of vowels in Arabic loanwords in Persian.

This study investigates lenition processes in Arabic loanwords in Russian within the framework of Optimality Theory.

Results & Discussion

This section introduces a number of the loanwords examined in which various lenition processes—including spirantization, voicing, degemination, and deletion—are attested. Subsequently, employing Optimality Theory, the study analyzes and explains the cause of the occurrence of lenition processes.

A) Spirantization (5 loanwords): This process results from the effort-minimization constraint LAZY. The faithfulness constraint IDENT[voice] opposes this constraint and prevents changes in voicing. Opposing this process is the faithfulness constraint IDENT[cont], which prevents any change in the [continuant] feature—this is the anti-spirantization constraint. For example, the word "الكيمياء" (al-kimiya) is examined in Tableau (1). The constraint ranking is as follows:

1) LAZY >> IDENT[voice] >> IDENT[cont]

Tableau (1)

Input:	LAZY	IDENT[voice]	IDENT[cont]
a. [alkím'ija]	*!		
b. [alxím'ija]			*
c. [algím'ija]		*!	

B) Voicing (one loanword): The word "السُّمُوت" (as-sumu:t) became "āzīmut" in Russian. In this process, the voiceless fricative /s/ in Arabic changed into the voiced fricative /z/—that is, the voicing of /s/ between two vowels and its change to /z/. Since producing a voiced consonant requires less articulatory effort, we are dealing here with the effort-minimization constraint LAZY. On the other hand, the faithfulness constraint of voicing works against the effort-minimization constraint.

2) LAZY >> IDENT[voice]

Tableau (2)

Input: /ʔassumu:t/	LAZY	IDENT[voice]
a. [asimut]	*!	
b. [azimut]		*

C) Degemination (8 loanwords): Among the examples of degemination in loanwords, [nabob] has been examined through two methods: the Obligatory Contour Principle (OCP) and the effort-minimization principle, of which only the latter is mentioned here. An effort-minimization constraint LAZY causes the removal of gemination, and opposing it is the gemination faithfulness constraint, defined as IDENT[cont]. Additionally, the constraints ONSET and *COMPLEX are involved in selecting the optimal candidate.

3) LAZY >> ONSET >> *COMPLEX [onset,coda], IDENT[GEM]

Tableau (3)

Input: /nuw.wāb/	LAZY	ONSET	*COMPLEX _{onset}	*COMPLEX _{coda}	IDENT[cont]
a. [na.bób]					*
b. [nab.ób]		*!			*
c. [na.bbób]	*!		*		
d. [nab.bób]	*!				
e. [nabb.ób]	*!	*		*	

D) Deletion (15 loanwords): Various types of initial, medial, and final deletion were observed in these loanwords, of which only one case of medial deletion is examined here. In the word "تعلب" (/θaʕ.lab/), since the pharyngeal consonant "ع" /ʕ/ does not exist in Russian, it was deleted from the word. Candidates created by replacing the absent consonant with adjacent Russian consonants are also penalized by the NOCODA constraint. Since vowel raising itself is in line with the principle of effort-minimization, the candidate [sālab] is penalized by the LAZY constraint, although it satisfies the vowel-lowering faithfulness constraint IDENT[low].

4) LAZY >> NOCODA >> IDENT[low] >> MAX-IO

Tableau (4)

Input: /θaʕ.lab/	LAZY	NOCODA	IDENT[low]	MAX-IO
a. [sā.lib]			*	*
b. [sā.lab]	*!			*
c. [sax.lab]	*!	*		
d. [sag.lab]	*!	*		

Conclusion

This study examined various types of lenition processes in Arabic loanwords in Russian within the Optimality Theory framework. The results showed that out of 149 loanwords examined, 29 underwent various lenition processes: spirantization (5 cases), voicing (1 case), degemination (8 cases), and deletion (15 cases). The changes in the data were described and explained using various markedness and faithfulness constraints of Optimality Theory. An examination of these constraints shows that the effort-minimization constraint LAZY plays the most important role in consonant lenition in loanwords, taking priority over faithfulness constraints. However, if the output resulting from effort-minimization is homonymous with a frequently used Russian word, that output is penalized by the anti-homonymy constraint ANTI-IDENT, and another non-homonymous candidate is selected as optimal. Overall, the results of the data analysis within Optimality Theory show that Russian, when encountering foreign consonants in Arabic loanwords, uses two strategies: first, replacing them with the closest Russian consonant; second, deleting the foreign consonant, provided the resulting word does not violate Russian's linguistic and phonological constraints.

References

- Al-Kadimi, Mahmood G. Ch. (2010). Arabic borrowing in the Russian Language. *Voronezh State University Bulletin. Series of Linguistics and Intercultural Communication*, No. 1, 75-77. [In Russian]
- Bakhtiyarova, Alfira N. (2016). On the issues of phonetic development of Arabisms in the Bashkir language, *Vestnik of the Bashkir University*. 1 (4). 763-768. [In Russian].
- Bakhtiyarova, Alfira N. & Flyuza G. Fatkullina (2015). Arabic loanwords in the lexical system of the Russian language. *Journal of "Philological Sciences"*. 2. 6124-6128. [In Russian]. <https://fundamental-research.ru/ru/article/view?id=38632> (Date of access: 28.10.2024).
- Kord Zafaranlu Kambuziya, Aliyeh & Eftekharsadat Hashemi (2015/1394) Phonological Adaptation of Vowels of Arabic loanwords in Persian: *Optimality Approach*. *LRR* 2015; 6 (6) :229-257. [In Persian]. URL: <http://lrr.modares.ac.ir/article-14-9562-fa.html>
- Moradi, Maryam (2022). The analysis of phonetic elision: aphaeresis elision, syncope elision and apocope elision in Arabisms in Russian language in the phonetic process. *Bulletin of Udmurt University. Series History and Philology*. 32 (2). 206-214. [In Russian].
<https://doi.org/10.35634/2412-9534-2022-32-2-198-206>
- Moradi, Maryam & zohreh Ghorbani-Madavani (2020). Analysis of the semantic narrowing of Arabisms included in the Russian language. *Journal of "Theory of Language and Intercultural Communication" of Kursk State University*, 1 (36). 206-214. [In Russian].
- Naghzguye Kohan, Mehrdad (2025/1404). *Historical Linguistics*. Tehran: Agah Publication. [In Persian].
- Reztsova, Vera V. (2011). Lexical-thematic classification of Arabisms in the Russian language. Material presented at the III Russian students' electronic scientific conference "Student Scientific Forum 2011, Astrakhan State University, Astrakhan, Russia. *Fundamental research*. 2 (27). 6124-6128. [In Russian] <https://scienceforum.ru/2011/article/2011001521>
- Spirkin, Alexander G., Igor A. Akchurin, Regina S. Karpinskaya, Ivan V. Lekhin, Slava M. Lokshina, Fedor N. Petrov & Lev S. Dhaumiyani (1984). *Slovar' inostrannyh slov [Dictionary of foreign words]*. Moscow: Russian Language. [In Russian].

زبان‌شناسی و گویش‌های ایرانی

سال ۱۱، شماره ۲، پیاپی ۲۱ (تابستان ۱۴۰۵) شماره صفحات: ۶۷ - ۸۸

فرایند واجی نرم‌شدگی در وام‌واژه‌های عربی دخیل در زبان روسی؛ چارچوب نظریه بهینگی

مریم مرادی^۱، سکینه نویدی باغی^{۲*}

دانشیار آموزش زبان روسی، گروه زبان روسی، دانشکده ادبیات فارسی و زبان‌های خارجی، دانشگاه علامه طباطبائی (ره)، تهران، ایران.
استادیار زبان‌شناسی، دانشکده ادبیات و علوم انسانی، دانشگاه شیراز، شیراز، ایران.

چکیده

یکی از فرایندهای ورود واژه‌های جدید در زبان‌های دنیا پدیده وام‌گیری است. در طی این پدیده، بر اثر تماس زبانی، واژه‌ای از زبان وام‌دهنده وارد زبان پذیرا می‌شود. به دلیل تفاوت‌های واج‌آرایی و قواعد واج‌شناختی گوناگونی که بر زبان‌ها حاکم است، وام‌واژه‌ها در طی این انتقال دچار تغییراتی می‌شوند تا با ساختار زبان پذیرا هماهنگ شوند. وام‌واژه‌های عربی نیز برای آن که وارد زبان روسی شوند لازم است از صافی نظام واجی این زبان عبور کنند تا با قواعد واج‌شناختی آن هماهنگ شوند. در این پژوهش، فرایند نرم‌شدگی در وام‌واژه‌های عربی موجود در زبان روسی در چارچوب نظریه بهینگی بررسی شده است. در این پژوهش توصیفی-تحلیلی، جهت استخراج وام‌واژه‌های عربی دخیل در زبان روسی از فرهنگ لغات بیگانه روسی چاپ سال ۱۹۸۴ (ویرایش هفتم) به تألیف اسپیرکین^۴ و همکاران استفاده شده است. نتایج پژوهش نشان می‌دهد که از ۱۴۹ وام‌واژه مورد بررسی که به روش نمونه‌گیری کامل استخراج شدند، تعداد ۲۹ واژه دستخوش انواع فرایندهای نرم‌شدگی سایشی‌شدگی، واکدارشدگی، تشدیدزدایی و حذف شده‌اند. در یک نگاه کلی، زبان روسی در مواجهه با وام‌واژه‌ها از دو راهبرد حذف همخوان بیگانه یا جایگزینی آن با نزدیکترین همخوان موجود در زبان روسی از نظر جایگاه تولید، بهره برده است که برخی از این تغییرات نمونه‌هایی از فرایند نرم‌شدگی هستند.

واژه‌های کلیدی:

زبان روسی
زبان عربی
وام‌گیری^۱
نرم‌شدگی^۲
نظریه بهینگی^۳

تاریخچه مقاله: دریافت: ۱۰ تیرماه ۱۴۰۵

بازنگری: ۱۴ شهریورماه ۱۴۰۵

پذیرش: ۹ مهرماه ۱۴۰۵

* آدرس ایمیل نویسنده مسئول: s.navidi@saadi.shirazu.ac.ir

مرادی، مریم و نویدی باغی، سکینه (۱۴۰۵). فرایند واجی نرم‌شدگی در وام‌واژه‌های عربی دخیل در زبان روسی؛ چارچوب نظریه بهینگی. زبان‌شناسی و گویش‌های ایرانی (۳)، ۶۷-۸۸. <https://doi.org/10.22099/jlil.2026.56534.1491>

حق نشر ۲۰۲۶-نویسنده(گان). این یک مقاله با دسترسی آزاد است که تحت شرایط مجوز Creative Commons Attribution (CC BY-NC 4.0) منتشر شده است، استفاده، توزیع و تکثیر نامحدود را در هر رسانه‌ای مجاز می‌داند، مشروط بر اینکه به نویسندگان و منبع اصلی اشاره شود. هیچ اجازه‌ای از نویسندگان یا ناشر لازم نیست.

¹ borrowing² lenition³ Optimality Theory⁴ A. G. Spirkin

۱. مقدمه

وام‌گیری یا قرض‌گیری زبانی عبارت است از روندی که طی آن یک زبان یا گویش برخی از عناصر زبانی را از زبان یا گویشی دیگر وام می‌گیرد و در خود جای می‌دهد (آرلاتو^۱، ۱۹۸۱: ۲۱۱). تماس زبانی می‌تواند موجب ورود عناصر زبانی از زبان وام‌دهنده به زبان وام‌گیرنده یا پذیرا شود. این وام‌گیری می‌تواند به دو صورت جوهری^۲ و انگاره‌ای^۳ رخ دهد. در اولی خود عنصر زبانی وارد زبان پذیرا می‌شود و در دومی انگاره یا قاعده‌ای انتزاعی (مانند جمع مکسر) وام گرفته می‌شود (نغزگوی کهن، ۱۴۰۴: ۳۲ و ۸۵). زبان‌ها در گذر زمان به دلیل ارتباطات گوناگون اقتصادی، سیاسی و فرهنگی کشورها خواه ناخواه دستخوش تغییراتی می‌گردند و این تغییرات، نظام واژگانی آنها را تحت تأثیر قرار می‌دهد. در اثر پدیده وام‌گیری، ذخیره واژگانی زبان پذیرا افزایش می‌یابد. یک واقعیت بسیار مهم درباره وام‌گیری آن است که واژه‌های قرضی، با نظام واجی (یا صوتی) زبان پذیرا همگون یا منطبق می‌شوند. این نکته به معنی آن است که پس از آن که یک واژه به طور کامل در زبان دیگری جذب شد، مانند واژه‌های معمولی آن زبان تلفظ می‌گردد و مشمول کلیه قواعد زبان وام‌گیرنده نیز می‌شود. اصواتی که در زبان وام‌دهنده وجود دارد اما در زبان وام‌گیرنده به کار نمی‌رود، یا حذف می‌گردد یا با اصوات مشابه بومی جایگزین می‌شود (آرلاتو، ۱۹۸۱: ۲۱۱).

در این پژوهش فرایندهای واج‌شناختی نرم‌شدگی یا تضعیف در طی وام‌گیری واژگانی از زبان عربی به زبان روسی مورد توجه و بررسی قرار گرفته است. نرم‌شدگی نوعی تغییر آوایی است که تنها بر همخوان‌ها اثر می‌گذارد و آن‌ها را هرچه بیشتر شبیه واکه‌ها می‌کند. هرچه بست یک همخوان بیشتر باشد، سخت‌تر یا قوی‌تر است و هرچه بست کم‌تر باشد آن همخوان نرم‌تر یا ضعیف‌تر است (نغزگوی کهن، ۱۴۰۴: ۲۴۵).

پژوهش حاضر در پی پاسخ به پرسش‌های زیر است:

- ۱- وام‌واژه‌های عربی در پدیده وام‌گیری در زبان روسی، دستخوش چه دگرگونی‌های نرم‌شدگی شده‌اند؟
- ۲- زبان روسی در برخورد با وام‌واژه‌های عربی چه راهبردهایی را اتخاذ کرده است؟

۱-۱. مقایسه واج‌های زبان‌های روسی و عربی

زبان روسی و عربی هر یک به خانواده‌های زبانی متفاوتی تعلق دارند. زبان عربی مهم‌ترین زبان زیرگروه زبان‌های سامی است که جزیی از خانواده زبان‌های افریقایی-آسیایی است (نغزگوی کهن، ۱۴۰۴: ۳۲۸). زبان روسی به خانواده زبان‌های هندواروپایی متعلق است و یکی از زبان‌های اسلاو شرقی^۴ از شاخه زبان‌های اسلاوی این خانواده زبانی به شمار می‌رود و در کنار زبان‌های اوکراینی و بلاروسی اعضای گروه اسلاو شرقی را تشکیل می‌دهد (نغزگوی کهن، ۱۴۰۴: ۳۱۵). بیشتر وام‌واژه‌های عربی هنگام ورود به زبان روسی، دستخوش دگرگونی‌های آوایی مختلف شده‌اند و تنها تعداد کم شماری از آنها بدون تغییر وارد زبان روسی شده‌اند.

پیش از بررسی فرایند نرم‌شدگی در وام‌واژه‌های عربی موجود در زبان روسی، ابتدا لازم است مروری بر فهرست واکه‌ها و همخوان‌های این دو زبان داشته باشیم و ساخت هجایی آن‌ها را با هم مقایسه کنیم. جدول (۱) فهرست واج‌های زبان روسی را نشان می‌دهد:

¹ A. Arlotto

² matter borrowing

³ pattern borrowing

⁴ East Slavic languages

جدول ۱.

همخوان‌ها و واکه‌ها در زبان روسی^۱

همخوان ها										واکه ها				
روش تولید		گرفته				رسا				گرد	غیر گرد	روش تولید		
		بی واک			واکدار							جایگاه تولید		
					غلط	کناری	چسبشوی	لرزشی	سایشی				انفجاری	سایشی
جایگاه تولید		نرم	к'		х'	г'					й'	پیشین	افراشته	
کامی (پس زبانی/پیشین)		سخت	к		х	г					з'			
الثری (میان زبانی)		نرم		ч'	ш'			р'				پیشین	افزاده	
		سخت			ш		ж	р						
دنداننی (پیش زبانی/پیشین)		نرم	т'		с'	д'	з'		н'	л'		پیشین	افراشته	
		سخت	т	ц	с	д	з		н	л				
لبی		نرم	п'		ф'	б'	в'		м'			پیشین	افزاده	
		سخت	п		ф	б	в		м					

همچنین، واکه‌ها و همخوان‌های زبان عربی در جدول (۲) آمده است:

^۱ هر یک از خانه‌های جدول ۱ و ۲ حاوی حرف روسی/عربی و علامت آوانگاری آن است که باید بین دو خط مورب قرار بگیرد (برای مثال، /b/). با این وجود، برای ساده‌تر شدن جدول از خطوط مورب استفاده نشده است.

جدول ۲.

واکه‌ها و همخوان‌های زبان عربی^۱

همخوان‌ها									واکه‌ها				
جایگاه تولید	گرفته				رسا				گرد	غیر گرد	کشش واکه	جایگاه زبان	
	بی‌واک		واکدار										
	انسدادی	سایشی	انسدادی	انسایشی	سایشی	لرزشی	خیشومی	کناری					غلت
دو لبی			بb				مm		وw		i	کوتاه	پیشین افراشته
لبی-داندانی		فf								i:		بلند	
دندانی		ثθ			ذð ظðʔ								
لثوی	تt طtʔ	سسs صsʕ	ددd ضdʕ		ززz	رr	نn	للl			a	کوتاه	مرکزی افتاده
لثوی-کامی		شʃ		جdʒ							a:	بلند	پسین افراشته
کامی									یj				
نرم‌کامی	کk								وw	u		کوتاه	
ملازی	قq	خχ			غɣ					u:		بلند	
حلقی		حħ			عʕ								
چاکنایی	ءʔ	هh											

یکی از عوامل تغییر وام‌واژه‌ها در زبان پذیرا، تفاوت الگوهای ساختار هجایی در دو زبان وام‌دهنده و پذیرا است. در زبان عربی ساختار هجایی به صورت CV(C) بازنمایی می‌شود؛ یعنی تنها یک واج در آغاز و یک واکه ساده در هسته الزامی است که با هم یک هجای باز CV را تشکیل می‌دهند. همچنین، وجود یک همخوان در پایانه اختیاری بوده و هجای بسته CVC را تشکیل می‌دهد. در زبان روسی تنها وجود واکه به عنوان هسته هجا الزامی است و وجود همخوان(های) آغاز و پایانه برای ساخت هجای زبان روسی الزامی نیست. آغاز و پایانه، هر یک می‌توانند تهی باشند یا به صورت یک همخوان یا خوشه همخوانی وجود داشته باشند. البته خوشه‌های همخوانی از برخی ملاحظات واج‌شناختی چون اصل توالی رسایی^۲ تبعیت می‌کنند (کرد زعفرانلو کامبوزیا و همکاران، ۱۳۹۳: ۴۱). بنابراین، به دلیل این که در زبان روسی ملاحظات واج‌شناختی بیشتری وجود دارد و وجود آغاز هجا

^۱ افزودن اندیس [ʕ] در بالای برخی همخوان‌ها به معنای حلقی‌شدگی آنهاست. برای مثال در تلفظ «ط» [tʕ] علاوه بر ایجاد بست تولیدی بین تیغه زبان و لثه بالا، ریشه زبان به سمت دیواره حلق گرایش پیدا می‌کند.

^۲ Sonority Hierarchy Principle

نیز الزامی نیست، و از طرفی، برخی واج‌های زبان عربی در این زبان موجود نیست، شاهد تغییراتی در وام‌واژه‌های عربی در این زبان هستیم که بخشی از آنها دستخوش فرایند واجی نرم‌شدگی شده‌اند.

۲. پیشینه پژوهش

در زمینه وام‌واژه‌های عربی موجود در زبان روسی، بیشتر پژوهش‌ها در حوزه واژه‌شناسی و معناشناسی انجام گرفته است. محمود قاضی چلیوب القدیمی^۱ (۲۰۱۰) معناشناسی و ریشه‌یابی وام‌واژه‌های عربی دخیل در زبان روسی را بررسی نموده است. بختیارووا و فاتکولینا^۲ (۲۰۱۵) به بررسی وام‌واژه‌ها در نظام لغوی زبان روسی پرداخته‌اند. مقوله تخصیص معنایی در وام‌واژه‌های عربی زبان روسی نیز توسط مرادی و قربانی مادوانی (۲۰۲۰) بررسی شده است. سیرازیف^۳ (۲۰۰۲) جنبه واژگانی کلمات عربی زبان تاتاری معاصر را بررسی کرده است. طبقه‌بندی موضوعی وام‌واژه‌های عربی در زبان روسی توسط رزتسووا^۴ (۲۰۱۱) بررسی شده است. جنبه آواشناختی وام‌واژه‌های عربی دخیل در زبان روسی نیز در پژوهش بختیارووا (۲۰۱۶) مورد توجه قرار گرفته و همگونی آوایی در وام‌واژه‌های عربی که وارد زبان باشقیری شده‌اند، به بحث گذاشته شده است. مرادی (۲۰۲۲) نیز انواع حذف (آغازین، میانی و پایانی) در وام‌واژه‌های عربی زبان روسی را بررسی کرده است. تحولات آوایی و معنایی وام‌واژه‌های عربی در فارسی توسط عامری (۱۳۹۵) مورد پژوهش قرار گرفته است و کرد زعفرانلو کامبوزیا و هاشمی (۱۳۹۴)، انطباق واجی واژه‌های وام‌واژه‌های زبان عربی را در زبان فارسی بررسی نموده‌اند. ایمانی و آهنگر (۲۰۲۴) تعدادی از واژه‌های گونه‌های عربی استاندارد و عربی لبنانی را با تکیه بر نظریه آواشناسی زایشی و به ویژه فرایندهای نرم‌شدگی و سخت‌شدگی را بررسی نموده‌اند. ولی‌پور و جمال‌زاد (۱۴۰۱) ویژگی‌های صوت‌شناختی همخوان‌های سایشی در زبان روسی در بیان دانشجویان فارسی‌زبان ایرانی و مقایسه آن با زبان معیار را مورد پژوهش قرار داده‌اند. پژوهش حاضر، اولین تلاش در حوزه بررسی انواع فرایندهای نرم‌شدگی در وام‌واژه‌های عربی زبان روسی می‌باشد.

۳. چارچوب نظری پژوهش

الف) نظریه بهینگی

نظریه بهینگی توسط پرنس^۵ و اسمولنسکی^۶ (۱۹۹۳) ارائه شد. در این نظریه محدودیت‌بنیاد، داده‌ها به صورت درون‌داد عرضه می‌شوند و پس از ارزیابی آن‌ها توسط انواع محدودیت‌های پایایی^۷ (یا وفاداری) و نشان‌داری^۸، در نهایت گزینه بهینه به عنوان برون‌داد انتخاب می‌شود. محدودیت‌های پایایی بر همانندی درون‌داد و برون‌داد، و محدودیت‌های نشان‌داری بر خوش‌ساختی گزینه برون‌داد نظارت می‌کنند. در این فرایند، ابتدا زایشگر^۹ از یک درون‌داد دریافتی خاص تعدادی گزینه برون‌دادی رقیب را تولید می‌کند. زایشگر به عنوان یک سازوکار جهانی، باید گزینه‌های متنوعی را تولید کند که با هر زبانی با هر ویژگی مطابقت

¹ M. G. Ch. Al-Kadimi

² A. N. Bakhtiarova & F. G. Fatkulina

³ I. I. Siraziyev

⁴ V. V. Reztsova

⁵ A. Prince

⁶ P. Smolensky

⁷ faithfulness constraints

⁸ markedness constraints

⁹ generator

کنند. سپس ارزیاب^۱ سلسله مراتب محدودیت‌های خاص یک زبان را بر مجموعه گزینه‌ها اعمال می‌کند تا هماهنگ‌ترین گزینه مشخص شود. این هماهنگ‌ترین گزینه همان برون‌داد بهینه است (جم، ۱۳۸۸؛ به نقل از مک‌کارتی^۲، ۲۰۰۲).

در این نظریه گزینه‌ها در یک تابلو ارائه می‌شوند. در ردیف افقی ستون‌های انواع محدودیت‌ها به ترتیب اولویت قرار دارند و در ردیف عمودی انواع گزینه‌های تولید شده قرار دارند. اگر گزینه‌ای یک محدودیت را نقض کند یک ستاره دریافت می‌کند (*). که به معنای جریمه نقض محدودیت مورد نظر است. اگر این تخطی موجب حذف گزینه شود، یک علامت تعجب هم پیش از ستاره قرار می‌گیرد (!*)، که به معنای تخطی مهلک است. اگر گزینه مورد نظر از دو جهت از محدودیتی تخطی کرده باشد، دو ستاره دریافت می‌کند. گزینه بهینه نیز با این علامت (EF) نشان داده می‌شود.

رتبه‌بندی محدودیت‌های نشاننداری و پایایی در هر نمونه مورد بررسی، براساس تقدم اثربخشی آنها بر روی وام‌واژه مورد نظر انجام می‌شود. با توجه به برون‌داد حاصل، رتبه‌بندی‌ها و تقدم و تأخر محدودیت‌ها می‌تواند متفاوت باشد. نوع برون‌داد موجب می‌شود که رتبه‌بندی‌ها مشابه هم نباشند. همچنین، محدودیت‌هایی که تاثیری در انتخاب گزینه بهینه ندارند با زمینه خاکستری نشان داده شده‌اند. در برخی موارد، تقدم و تأخر دو محدودیت تاثیری در برون‌داد ندارد که در این صورت بین دو محدودیت کاما قرار داده می‌شود.

ب) فرایند واجی نرم‌شدگی

نرم‌شدگی نوعی تغییر عمده آوایی است که تنها بر همخوان‌ها اثر می‌گذارد و آنها را هر چه بیشتر شبیه واکه‌ها می‌کند. همخوان‌های نرم‌تر شباهت بیشتری به واکه‌ها دارند زیرا در زمان تولید آنها جریان هوا با آزادی بیشتری از جایگاه تولید عبور می‌کند (نغزگوی کهن، ۱۳۹۳: ۱۲۵). طبق نظر مدرسی قوامی (۱۳۹۴: ۲۲۶)، «آواها در نرم‌شدگی واکدار می‌شوند، درجه گرفتگی آنها در مجرای گفتار کاهش می‌یابد، از دیرش آنها کاسته می‌شود، یا حذف می‌شوند».

کرشنر^۳ (۱۹۹۸) معتقد است که عامل رخداد فرایند نرم‌شدگی یا تضعیف، کاهش «کوشش تولیدی» است. برای نمونه، در فارسی، «واو» در واژه‌ای چون «دویدن» زمانی که صورت «بدو» از آن تلفظ می‌شود، شکلی نرم‌تر به خود می‌گیرد. صرامی (۱۴۰۲: ۴۳) نیز، نرم‌شدگی یا تضعیف را کاهش قدرت آواها تعریف می‌کند که به لحاظ واج‌شناسی به طور معمول در بافت پایانه و یا میان دو واکه رخ می‌دهد. تضعیف یا نرم‌شدگی میان دو واکه را تضعیف میان‌واکی^۴ می‌گویند. کرشنر (۲۰۰۱) نرم‌شدگی در

^۱ evaluator

^۲ J. McCarthy

^۳ R. Kirchner

^۴ intervocalic

بافت میان‌واکه‌ای را معمول‌تر از بافت پایانه می‌داند (جم، ۱۳۸۸: ۱۸۷). کرشنر (۱۹۹۸) انواع فرایندهای نرم‌شدگی را این‌گونه معرفی کرده است:

- ۱- سایشی‌شدگی^۱: عبارت است از تبدیل یک همخوان انسدادی یا انسایشی به همخوان سایشی. تبدیل همخوان [b] به [v] در واژه فارسی باستان «nibištan» و تلفظ آن در فارسی معاصر به صورت /neveštan/ نوعی سایشی‌شدگی است. همچنین تبدیل همخوان انفجاری ملازی /G/ به /ʃ/ در واژه ی «باقی» سایشی‌شدگی به شمار می‌رود (مدرسی قوامی، ۱۳۹۴: ۱۴۳).
- ۲- واکدارشدگی^۲: تبدیل یک همخوان بیواک به جفت واکدار خود. مانند واکدارشدگی /t/ فارسی باستان در واژه‌ای مانند «پدر» و تبدیل شدن آن به /d/ در فارسی جدید /pedær/ (مدرسی قوامی، ۱۳۹۴: ۲۶۰).
- ۳- تشدیدزدایی^۳: تشدیدزدایی فرایندی است که طی آن یکی از دو جزء همخوان مشدد حذف می‌شود: [t] → /tt/ برای مثال، حذف /d/ در ترکیب «قند» + «دان» /gand+dan/ و تلفظ شدن آن به صورت [Gandan] (مدرسی قوامی، ۱۳۹۴: ۷۱). لازم به ذکر است که، کرشنر فرایند تشدیدزدایی را به عنوان کوتاه شدن یک همخوان کشیده تعریف کرده است: /t:/ → [t] (جم، ۱۳۸۸: ۱۸۷).
- ۴- زنشی‌شدگی^۴: تغییر از یک همخوان انسدادی به یک همخوان زنشی؛ مانند [r] → /t/.
- ۵- ناسوده‌شدگی^۵: تبدیل همخوان‌ها به یک همخوان ناسوده؛ برای نمونه، در فارسی، /v/ در «دویدن»، به [w] تبدیل می‌شود و به صورت «بدو» تلفظ می‌شود و شکلی نرم‌تر به خود می‌گیرد.
- ۶- غیردهانی‌شدگی^۶: تبدیل یک همخوان دهانی به یک همخوان چاکنایی؛ مانند [h] → /s/.
- ۷- حذف^۷: حذف یک آوا که کاملترین نوع تضعیف است؛ فرایند هم‌زمانی یا درزمانی تلفظ نشدن یک همخوان، واکه یا هجا. در فارسی محاوره‌ای، همخوان [r] در «صبر کن» [sabr kon] اغلب تلفظ نمی‌شود و این ترکیب با حذف [r] به صورت [sab kon] تلفظ می‌شود.

ج) سلسه مراتب نرمی یا سختی واجها

فرایند نرم‌شدگی در مقابل فرایند سخت‌شدگی^۸ قرار دارد که طی آن یک همخوان به سمت بازتر شدن و کاهش گرفتگی حرکت می‌کند؛ برای مثال تبدیل همخوان انسدادی /b/ به همخوان سایشی /v/ با گرفتگی کمتر نمونه‌ای از فرایند نرم‌شدگی محسوب می‌شود. برخلاف فرایند نرم‌شدگی، در فرایند سخت‌شدگی همخوان‌ها به همخوانی سخت‌تر با درجه گرفتگی بیشتر تبدیل می‌شوند. برای میزان نرم یا سخت بودن همخوان‌ها سلسله مراتب‌هایی ارائه شده است.

ثمره (۱۳۷۸: ۲۶) در بخشی با عنوان «شدت»^۹ توضیح می‌دهد که فشار و نیروی ماهیچه‌ای که در تولید آوای بیواک به کار می‌رود بیشتر از نیرویی است که در تولید آوای واکدار صرف می‌شود. وی در این خصوص به مقایسه دو آوای S و Z می‌پردازد.

^۱ Spirantization/ fricativization

^۲ voicing

^۳ degemination

^۴ flapping

^۵ reduction to approximant

^۶ debuccalisation

^۷ elision

^۸ fortition

^۹ intensity

فشار هوایی که از ششها به بیرون منتقل می‌شود در مورد S و Z یکسان است منتهی هنگام تولید Z مقداری از آن صرف به ارتعاش درآوردن تارآواها می‌گردد و بقیه آن به مصرف ایجاد سایش می‌رسد. ولی در مورد S چون هیچ‌گونه فعالیتی در حنجره وجود ندارد، بنابراین تمام نیرو و فشار هوا برای ایجاد سایش به کار می‌رود. به همین علت برای دعوت کسی به سکوت از صدای S (که آن را هیس می‌گوییم) استفاده می‌کنید و نه از صدای Z. از این جهت آواهای بیواک را «سخت» و آواهای واکدار را «نرم» می‌نامند.

طبق نظر هایمن^۱ (۱۳۶۸: ۲۵۷)، همخوان های لبی سخت تر یا قویتر از دندانی/لثوی ها و دندانی/لثوی ها سخت تر از نرم‌کامی‌ها هستند. همچنین، مشددها از انسدادی‌های بیواک غیرمشدد سخت‌ترند و دسته اخیر از انسدادی‌های واکدار سخت‌ترند. انسدادی‌های واکدار سخت‌تر از سایشی‌های واکدارند و بالاخره سایشی‌های واکدار از رساهای^۲ واکدار سخت‌ترند. از نظر ولی‌پور (۱۳۹۳: ۲۷-۲۶) نیز، ترتیب اصوات از نظر مصرف انرژی به شرح زیر است: واکه‌ها > خیشومی‌ها > همخوان‌های واکدار > همخوان‌های بیواک.

د) فرایند نرم‌شدگی در نظریه بهینگی

در خصوص چرایی انتخاب نظریه بهینگی جهت تحلیل فرایند واجی نرم‌شدگی، کرشنر (۲۰۰۶) استدلال می‌کند که رویکردهای صورت‌گرا قادر به تبیین شرایط وقوع طبیعی فرایند نرم‌شدگی نیستند. برای مثال، فرایندهای تشدیدزدایی، چاکنایی شدگی و حذف در واج‌شناسی خودواحد به صورت حذف یا قطع پیوند^۳ تحلیل می‌شوند. همچنین، رویکردی دیگر پیشنهاد شده که در آن نرم‌شدگی را «ارتقای رسایی»^۴ در نظر گرفته‌اند. با توجه به سلسله مراتب رسایی که در آن رسایی سایشی‌ها کمتر از خیشومی‌ها است^۵، بنابراین طبق پیش‌بینی این رویکرد، فرایند نرم‌شدگی باید موجب تبدیل سایشی‌ها به خیشومی‌ها بشود، که در عمل این اتفاق نمی‌افتد. بنابراین، کرشنر (۲۰۰۶) برخلاف رویکردهای صورت‌گرا، فرایند نرم‌شدگی را به عنوان کاهش «کوشش تولیدی»^۶ در نظر می‌گیرد و این اصل را به عنوان یک محدودیت صوری کم‌کوشی در چارچوب نظریه بهینگی معرفی می‌کند

^۱ L. M. Hyman

^۲ sonorants

^۳ delinking

^۴ Sonority promotion

^۵ کرشنر به نقل از دل و المدلویی (۱۹۸۵، Dell & Elmedlaoui) درجه بندی زیر را برای رسایی واجها به دست داده است: انسدادی > سایشی بی‌واک > سایشی واکدار > خیشومی > روان > واکه افراشته/ غلت > واکه افتاده

^۶ articulatory effort

(جم، ۱۳۸۸: ۱۸۸؛ به نقل از کرشنر، ۲۰۰۶). بر همین اساس، محرک کاهش کوشش را در قالب یک محدودیت صوری کم‌کوشی که «LAZY» نامیده می‌شود به صورت زیر معرفی می‌کند:

LAZY: واحدهای واجی با کم‌کوشی تولید می‌شوند.

محدودیت کم‌کوشی در تقابل با محدودیت‌های وفاداری یا پایایی^۱ و سخت‌شدگی^۲ (تقویت) قرار دارد. محدودیت سخت‌شدگی به صداها قدرت می‌بخشد و آن‌ها را به لحاظ شنیداری و ادراکی متمایز می‌سازد. فرایند نرم‌شدگی به لحاظ سهولت تولید گوینده‌مدار است و فرایند سخت‌شدگی به دلیل ابهام‌زدایی و آسان‌تر شدن درک گفتار، شنونده‌مدار است (جم، ۱۳۸۸: ۱۸۹). علاوه بر سخت‌شدگی، محدودیت پایایی نیز که موجب یکسانی درونداد و برون‌داد می‌شود و هرگونه تغییر در برون‌داد را جریمه می‌کند، در تقابل با فرایند نرم‌شدگی قرار می‌گیرد. این محدودیت با عبارت «IDENT»^۳ که به معنی یکسانی است، نمایش داده می‌شود و می‌تواند درباره مشخصه‌های مختلفی مانند رسایی یا پیوسته بودن رخ دهد و از تبدیل آن به چیزی غیر از درونداد جلوگیری کند. البته در رده‌بندی محدودیت‌ها، محدودیت کم‌کوشی «LAZY» نسبت به «IDENT» در اولویت قرار دارد (کرشنر، ۱۹۹۸: ۲۶).

۳-۱. روش‌شناسی

در میان دگرگونی‌های آوایی که در وام‌واژه‌های عربی زبان روسی به وجود آمده، فرایند واجی نرم‌شدگی بسیار متداول است. در این پژوهش توصیفی-تحلیلی که به روش کتابخانه‌ای انجام شده، جهت استخراج وام‌واژه‌های عربی دخیل در زبان روسی از فرهنگ لغات بیگانه روسی (اسپیرکین و همکاران، ۱۹۸۴) استفاده شده است. در مجموع ۱۴۹ وام‌واژه از این فرهنگ لغت به روش نمونه‌گیری کامل استخراج شد. پس از استخراج وام‌واژه‌ها و مقایسه آن‌ها با واژه اصلی موجود در زبان عربی، تغییرات آوایی صورت گرفته بر روی واژه‌ها با دقت ثبت شد. سپس انواع فرایندهای نرم‌شدگی، از جمله سایشی‌شدگی، واکدارشدگی، تشدیدزدایی و حذف در تغییرات صورت گرفته شناسایی شد (۲۹ وام‌واژه) و در چارچوب نظریهٔ بهینگی مورد توصیف و تحلیل قرار گرفت.

^۱ faithfulness

^۲ fortition

^۳ identity

یافته‌های این پژوهش می‌تواند در دروس آواشناسی دانشجویان ایرانی رشته زبان روسی و نیز در بررسی فرایندهای نرم‌شدگی در وام‌واژه‌ها برای دانشجویان زبان‌شناسی مؤثر و مفید واقع شود.

۴. تحلیل داده‌ها

در این بخش، به معرفی برخی از وام‌واژه‌های بررسی شده که انواع فرایندهای نرم‌شدگی سایشی‌شدگی، واکدارشدگی، تشدیدزدایی و حذف در آن‌ها مشاهده شده می‌پردازیم و سپس، با استفاده از نظریه بهینگی، به تحلیل و تبیین علت وقوع فرایندهای نرم‌شدگی می‌پردازیم.

۴-۱. فرایند سایشی‌شدگی^۱

برخی از وام‌واژه‌های عربی موجود در زبان روسی که دستخوش فرایند سایشی‌شدگی شده‌اند عبارتند از:

جدول ۳.

فرایند سایشی‌شدگی در وام‌واژه‌های عربی دخیل در زبان روسی

ردیف	واژه عربی	تلفظ در عربی	واژه روسی	تلفظ در روسی	آوای قبل از فرایند	آوای پس از فرایند
۱	الکیمیا	/ʔalkimija:ʔ/	алхі́мия	[alxímʲija]	[k]	[x]
۲	دُکان	/dukka:n/	духа́н	[duxán]	[k]	[x]
۳	القبه	/ʔalqubbah/	алько́в	[alʹkóv]	[b]	[v]
۴	سُمّاق	/suma:q/	сума́х	[sumáx]	[q]	[x]
۵	سِگَه	/sɪkkah/	цехи́н	[tsixiˈn]	[kk]	[x]

از آن‌جا که تولید همخوان‌های انسدادی به کوشش بیشتری نسبت به همخوان‌های سایشی نیاز دارند، بنابراین، فرایند سایشی‌شدگی حاصل محدودیت کم‌کوشی LAZY است. از سوی دیگر، محدودیت پایایی IDENT[voice] در مقابل این محدودیت قرار دارد و مانع تغییر مشخصه واکداری می‌شود. در مقابل این فرایند، محدودیت پایایی IDENT[cont] قرار دارد که مانع هرگونه تغییر در مشخصه [پیوسته^۲] می‌شود و محدودیت ضدسایشی‌شدگی قلمداد می‌شود. لازم به ذکر است که همخوان‌های انسدادی [-پیوسته] و همخوان‌های سایشی [+پیوسته] هستند (جم، ۱۳۸۸: ۱۹۸-۲۰۴).

بنابراین، رتبه‌بندی (۱) فرایند سایشی‌شدگی و واک‌رفتگی را به ترتیب اولویت نشان می‌دهد:

1) LAZY >> IDENT[voice] >> IDENT[cont]

^۱ spirantization

^۲ Continuant

نکته قابل توجه این که، در زبان روسی جفت واکدار همخوان /x/ یعنی همخوان /x/ موجود نیست و نمی‌تواند یکی از گزینه‌ها در تابلوی بهینگی باشد. با این توضیحات، با توجه به محدودیت‌های فوق، تابلوی (۱) برای فرایند سایشی‌شدگی قابل ترسیم است.

تابلوی (۱)

Input:	LAZY	IDENT[voice]	IDENT[cont]
a. [alkím'ija]	!* b. [alxím'ija]		*
c. [algím'ija]		*!	
a. [dukán]	!* b. [duxán]		*
c. [dugán]		*!	
a. [ʔalkub]	!* b. [ʔalkuf]	!* c. [ʔalkóv]	* *
a. [suma:k]	*! b. [suma:x]		*
c. [suma:g]		*!	

لازم به ذکر است که، در زبان روسی همخوان «خ» /x/ به صورت نرم‌کامی تلفظ می‌شود اما همخوان ملازی، حلقی و چاکنایی در این زبان وجود ندارد. بنابراین در تبدیل /k/ به /x/ تغییری در جایگاه تولید مشاهده نمی‌شود و محدودیت IDENT[place] رعایت شده است، اما شیوه تولید از انسدادی به سایشی تغییر کرده است. به عبارت دقیق‌تر، همخوان غیرپیوسته k به همخوان پیوسته x تبدیل شده و از محدودیت IDENT[cont] تخطی شده است. در واژه «القبّه» نیز بر اثر اصل کم‌کوشی یک همخوان انسدادی به همخوان سایشی تبدیل می‌شود. جایگزین اول همخوان لبی سایشی /β/ است که در زبان روسی وجود ندارد. سپس، نوبت به جایگاه تولید بعدی (لبی-دندانی) می‌رسد که دو همخوان سایشی /v/ و /f/ هستند که یکی واکدار و دیگری بیواک است. به دلیل محدودیت IDENT[voice] همخوان بیواک /f/ جریمه می‌شود و همخوان واکدار /v/ انتخاب می‌شود. از آنجا که محدودیت پایایی واکداری پیش از محدودیت پایایی پیوسته نبودن (IDENT[-cont]) قرار دارد، همخوان /v/ در نهایت به عنوان وام‌واژه در زبان روسی مورد استفاده قرار می‌گیرد.

در وام واژه «سماق» /suma:q/ صورت سایشی شده همخوان انسدادی ملازی /q/ همخوان سایشی ملازی /x/ است. زبان روسی فاقد همخوانهای ملازی است و باید یک همخوان بومی را جانشین این همخوان بیگانه کند یا آن را حذف کند.^۱ در نگاه اول به نظر می‌رسد که باید همخوان بومی انسدادی نرم‌کامی /k/ جانشین این واج شود، اما گزینه بهینه، همخوان نرم تر /x/

^۱ ذکر این نکته لازم است که همخوان‌های تبدیل شده همواره بیگانه نیستند و گاهی یک همخوان آشنا در زبان روسی مانند /s/ نیز برای سهولت تلفظ دستخوش فرایند نرم‌شدگی شده است.

است که یک همخوان سایشی نرم‌کامی است و بر اثر سایشی شدگی ایجاد می‌شود. در فرایند سایشی شدگی شاهد تقدم محدودیت پایایی واکداری IDENT[voice] بر محدودیت پایایی پیوسته بودن IDENT[cont] هستیم و به همین دلیل همانطور که در تابلوی (۱) مشاهده می‌شود جفت‌های واکدار به عنوان گزینه بهینه انتخاب نمی‌شوند.

۴-۲. فرایند واکدارشدگی

در بررسی تغییرات واجی وام‌واژه‌های عربی پس از ورود به زبان روسی، در داده‌های مورد بررسی تنها در یک واژه فرایند واکدارشدگی رخ داده است^۱ (جدول ۴).

جدول ۴.

فرایند واکدارشدگی در وام‌واژه‌های عربی دخیل در زبان روسی

ردیف	واژه عربی	تلفظ در عربی	واژه روسی	تلفظ در روسی	آوای قبل	آوای پس از
۱	السَّمت	/ʔassumu:t/	áзимут	[áz'imut]	/s/	/z/

واژه «السَّمت» (یا «السَّمت») کوتاه شده ترکیب «رأس السَّمت» است که در زبان روسی تبدیل به «áзимут» [áz'imut] شده است. در این فرایند، همخوان سایشی بی واک /s/ در زبان عربی به سایشی واکدار /z/ تبدیل شده است؛ یعنی واکدارشدگی /s/ میان دو واکه و تبدیل آن به /z/. از آنجا که برای تولید همخوان واکدار کوشش تولیدی کمتری صورت می‌پذیرد در اینجا با محدودیت کم‌کوشی LAZY روبرو هستیم. از طرفی محدودیت پایایی واکداری بر خلاف محدودیت کم‌کوشی عمل می‌کند. از آنجا که محدودیت کم‌کوشی در این میان پیروز می‌شود. رتبه‌بندی محدودیت‌ها به صورت زیر خواهد بود:

2) LAZY >> IDENT[voice]

با توجه به این توضیحات تابلوی (۲) می‌تواند بیانگر دلیل وقوع فرایند نرم‌شدگی در واژه «السَّمت» باشد:

تابلوی (۲)

Input: /ʔassumu:t/	LAZY	IDENT[voice]
a. [asimut]	!* [asimut]	
b. [azimut]		*

شایان ذکر است که، همخوان تبدیل شده در این نمونه یک همخوان سایشی است و محدودیت پایایی پیوسته بودن IDENT[cont] در انتخاب گزینه بهینه تاثیری ندارد. اگر همخوان تبدیل شده یک همخوان انسدادی بیواک مانند /t/ و

^۱ ذکر این نکته ضروری است که در وام‌واژه‌هایی مانند «حَرَم» /garém/ «гарém» و «موهاریه» /mağár/ «морáп» به ترتیب شاهد تبدیل همخوان حلقی سایشی بی واک «ح» /h/ به همخوان انسدادی نرم‌کامی واکدار «گ» /g/ و تبدیل همخوان چاکنایی سایشی بی واک «ه» /h/ به همخوان /g/ هستیم، اما از ذکر این نمونه‌ها به عنوان فرایند واکدارشدگی خودداری شده؛ زیرا دو همخوان جایگاه تولید متفاوتی دارند و همخوان /g/ جفت واکدار همخوانهای اولیه محسوب نمی‌شود. از سوی دیگر، تبدیل همخوان حلقی یا چاکنایی به همخوان دهانی در خلاف مسیر «غیردهانی شدگی» که یکی از فرایندهای نرم‌شدگی است رخ داده، و همخوان سایشی نیز بر خلاف فرایند سایشی شدگی به یک همخوان انسدادی تبدیل شده است. یادآور می‌شویم که زبان روسی فاقد همخوان‌های حلقی و چاکنایی است و در این نمونه‌ها نزدیکترین همخوان موجود در زبان روسی جانشین همخوان بیگانه شده است.

همخوان جانشین جفت واکدار آن یعنی /d/ باشد، آنگاه محدودیت پایایی پیوسته بودن بر محدودیت پایایی واکداری تقدم دارد و باعث حذف گزینه سایشی بیواک یعنی /s/ می شود.

۳-۴. فرایند تشدیدزدایی

در نمونه‌های استخراج شده هشت وام واژه عربی دخیل در زبان روسی دستخوش فرایند تشدیدزدایی شده اند که فهرست آن ها در جدول ۵ آمده است.

جدول ۵.

فرایند تشدیدزدایی در وام واژه های عربی دخیل در زبان روسی

ردیف	واژه عربی	تلفظ در عربی	واژه روسی	تلفظ در روسی	آوای قبل از فرایند	آوای پس از فرایند
۱	صَفَّه	/sʕuffah/	сóфа	/sófa/	/ff/	/f/
۲	زَبَّادَان	/zabba:da:n/	цибетін	/tsibʔitʔín/	/bb/	/b/
۳	نُوب	/nuwwa:b/	набóб	/nabób/	/ww/	/b/
۴	القُبَّة	/ʔalqubbah/	алькóв	/alʔkóv/	/bb/	/v/
۵	السُّمُوت	/ʔssumu:t/	ázимут	/ázimut/	/ss/	/z/
۶	سِكَّة	/sikkah/	цехін	/tsixiʔn/	/kk/	/x/
۷	دُكَّان	/dukka:n/	духán	/dukhaʔn/	/kk/	/x/
۸	سَيِّد	/sajjid/	се́йд, се́йид	/seid/, /seyid/	/jj/	/j/

از نمونه های فوق تشدیدزدایی در وام‌واژه [nabob] را با دو روش مختلف بررسی می کنیم. اول بر اساس اصل مرز اجباری و دوم کشش همخوان که کرشنر آن را مطرح کرده است.

الف) اصل مرز اجباری یکی از اصول مطرح در واج‌شناسی خودواحد^۱ است. براساس این اصل در یک لایه خودواحد، مجاورت دو خودواحد یکسان مجاز نیست. اساساً اگر توالی واحدهای مستقل یکسان در یک اشتقاق وجود داشته باشد، اصل مرز اجباری مانند صافی عمل کرده، آن دو را یکی می کند (کرد زعفرانلو کامبوزیا، ۱۳۸۵: ۵۳). با در نظر گرفتن محدودیت‌های ساختاری

^۱ autosegmental phonology

هجا به لحاظ رویکرد بهینگی و نظر به بی‌نشان‌ترین (و پرتکرار و معمول‌ترین) ساخت هجا که CV است، بایستی به چهار نکته زیر توجه نمود:

- ۱- تمام هجاها باید آغازه داشته باشند (محدودیت ONSET)؛
 - ۲- تمام هجاها باید یک هسته یا مرکز داشته باشند (محدودیت NUCLEUS)؛
 - ۳- هجاها نباید پایانه داشته باشند: (محدودیت NO-CODA)
 - ۴- هیچ دو عنصر مشابهی نباید در مجاورت یکدیگر باشند مطابق محدودیت اصل مرز اجباری یا «OCP».
 - ۵- عناصر ساختی هجا (چه در آغازه، چه در هسته و چه در پایانه) نباید پیچیده باشند: «*COPMLEX».
- البته، تشدیدزدایی در پایانه هجا با محدودیت CodaGemination که با CODAGEM* نشان داده می‌شود نیز بیان می‌شود که وجود همخوان مشدد در پایانه را جریمه می‌کند.
- در تقابل با محدودیت‌هایی که سبب تشدیدزدایی می‌شوند، محدودیت پایایی تشدید IDENT[GEM] نیز وجود دارد که تغییر پرونداد نسبت به درونداد را جریمه می‌کند. با توجه به مطالب بالا، رتبه‌بندی محدودیت‌ها در فرایند تشدیدزدایی به صورت زیر خواهد بود:

3) OCP >> ONSET >> *COMPLEX[onset,coda] , IDENT[GEM]

بر این اساس، می‌توان انواع گزینه‌ها را در تابلوی (۳) بررسی کرد:

تابلوی (۳).

تشدیدزدایی همخوان /b/ بر اساس اصل مرز اجباری

Input:	OCP	ONSET	*COMPLEX _{onset}	*COMPLEX _{coda}	IDENT _[GEM]
/nuwwa:b/					
a. [na.bób]					*
b. [nab.ób]		*!			*
c. [na.bbób]	*!		*		
d. [nab.bób]	*!				
e. [nabb.ób]	*!	*		*	

ب) از نگاه دوم، طبق نظر کرشنر (1998، 1) که تشدیدزدایی را حذف کشش همخوان قلمداد کرده است، یک محدودیت کم‌کوشی LAZY موجب حذف کشش شده و در تقابل با آن، محدودیت پایایی کشش به صورت IDENT[cont] تعریف می‌شود. علاوه بر این، محدودیت‌های ONSET و *COMPLEX* نیز در فرایند انتخاب گزینه بهینه دخیل هستند که رتبه‌بندی آنها به صورت زیر است:

4) LAZY >> ONSET >> *COMPLEX[onset,coda] , IDENT[GEM]

تابلوی (۴).

تشدیدزدایی همخوان /b/ بر اساس اصل کم‌کوشی

Input:	LAZY	ONSET	*COMPLEX _{onset}	*COMPLEX _{coda}	IDENT[cont]
/nuw.wāb/					
a. [na.bób]					*

b. [nab.ób]	!	*	
c. [na.bbób]	!	*	
d. [nab.bób]	!		
e. [nabb.ób]	!	*	*

همان طور که در تابلوی (۳) و (۴) مشخص است، گزینه [nabb.ob] به دلیل تخطی از سه محدودیت کم‌کوشی، آغاز هجا، و تجمع همخوان در پایانه دارای بیشتری تخطی است. در این میان، گزینه [na.bob] خوش‌ساخت‌ترین گزینه در زبان روسی تلقی شده و به عنوان گزینه منتخب در این زبان مورد استفاده قرار گرفته است.

بنابراین، در صورتی که فرایند تشدیدزدایی را حاصل از اصل مرز اجباری قلمداد کنیم تابلوی (۵) و اگر بخواهیم طبق تعریف کرشنر، تشدیدزدایی را حاصل حذف کشش همخوان بدانیم، تابلوی (۶) به خوبی از عهده تبیین پدیده تشدیدزدایی بر می‌آیند.

۴-۴. فرایند حذف

در فرایند نرم‌شدگی، ضعیف شدن همخوان آن‌قدر می‌تواند ادامه پیدا کند تا به حذف همخوان منجر شود (نغزگوی کهن، ۱۴۰۴: ۲۴۶). در واژگان مورد بررسی انواع فرایند حذف در جایگاه‌های مختلف وام‌واژه‌ها مشاهده شد:

جدول ۷.

فرایند حذف در وام‌واژه‌های عربی دخیل در زبان روسی

ردیف	کلمه عربی	تلفظ عربی	کلمه روسی	تلفظ روسی	واج حذف شده	جایگاه واج حذف شده
۱	عاده	/ʕa:dat/	Ада́т	/adat/	ع /ʕ/	آغازی
۲	علماء	/ʕu:lama:ʔ/	уле́мы	/ul'émí/	ع /ʕ/، ء /ʔ/	آغازی و پایانی
۳	طرحه	tʰarħah/	tára	/tara/	ح /ħ/، ه /h/	میانی و پایانی
۴	مرباط	/mura:bitʕ/	марабу́	/marabu/	ط /tʕ/	پایانی
۵	صاحب	/sʕa:ħib/	саи́б	/saíb/	ح /ħ/	میانی
۶	رعیه	/raʕijah/	раи́я	/rajia/	ع /ʕ/، ه /h/	میانی و پایانی
۷	رَهج الغار	/rahdzalɣa:r/	реальгар	/realgar/	ه /h/، ج /dʒ/	میانی
8	تَعَلَب	/θaʕlab/	сале́п	/salep/	ع /ʕ/	میانی
9	سَيِّد	/sajjid/	сеи́д, ейи́д	/seyid/, /seid/	ی /j/	میانی
10	سَيِّل	/sajl/	сель	//sel	ی /j/	میانی
11	بُورَق	/bu:raq/	бура́	/bura/	ق /q/	پایانی
12	حناء	/ħana:ʔ/	хна	/xna/	أ /a/، ء /ʔ/	میانی و پایانی
۱۳	القبه	/ʔalqubbah/	алько́в	[al'kóv]	أ /a/، ه /h/	پایانی
۱۴	صَفَّة	/sʕuffah/	со́фа	/sófa/	ه /h/	پایانی
15	ذره	/ðarrah/	ду́рра	/durrá/	ه /h/	پایانی

در ادامه هر یک از انواع حذف را در چارچوب نظریه بهینگی بررسی می‌کنیم.

الف) فرایند حذف آغازی

فرایند حذف آغازی^۱، مانند حذف همخوان واکدار حلقی سایشی /ع/ /ʕ/[¹] در آغاز هجای واژه «عاده» [adát] «علماء» [ul'émí] «ولعلمی» همان‌گونه که پیش‌تر در بخش ساخت هجایی در زبان روسی و عربی اشاره شد، وجود همخوان آغازی در زبان عربی الزامی است، اما در زبان روسی ضروری نیست. بنابراین، هر جا که وام‌واژه با همخوانی آغاز شود که در زبان روسی موجود نیست، شاهد حذف همخوان آغازی در آغاز هجا هستیم. برای مثال، همخوان حلقی «ع» /ع/ در زبان روسی وجود ندارد؛ به همین دلیل در وام‌واژه‌های «عاده» و «علماء» شاهد حذف همخوان آغازی هستیم. در این حالت محدودیت ONSET وجود آغازی برای هجا را الزامی می‌داند، از اولویت برخوردار نبوده و تقریباً در انتهای رتبه‌بندی قرار می‌گیرد. فرایند حذف در راستای پیروی از اصل کم‌کوشی و رعایت محدودیت LAZY رخ می‌دهد. با این وجود، اگر برون‌داد حاصله با واژه پربسامدی در زبان روسی رابطه همنامی^۲ داشته باشد، با محدودیت «ضد همنامی»، ANTI-IDENT جریمه می‌شود^۳. در مقابل، محدودیت پایایی ضد حذف MAX-IO نیز وجود دارد که حذف عناصر واجی از واژه را جریمه می‌کند. در ادامه تعریف این محدودیت‌ها را ارائه می‌دهیم:

MAX-IO محدودیت ضد حذف: هر واحد زنجیری که در درون‌داد دیده می‌شود، باید در برون‌داد نیز ظاهر شود.

ANTI-IDENT محدودیت ضد همنامی: برون‌داد نباید با واژه‌ای که در زبان وام‌گیرنده وجود دارد هم‌آوا باشد.

رتبه‌بندی محدودیت‌ها برای وام‌واژه «عاده» به صورت زیر است:

>> ONSET , NOCODA , MAX-IO >> IDENT[cont] 5) ANTI-IDENT >> LAZY

تابلوی (۵)

Input: /ʕadát /	ANTI- IDENT	LAZY	ONSET	NOCODA	MAX- IO	IDENT[cont]
a. [adá]	*!		*		**	
b. [adát]			*	*	*	
c. [gadát]		*!		*		
d. [xadát]		*!		*		*

طبق تابلوی (۵) گزینه [adát] محدودیت NOCODA را نقض کرده و نسبت به گزینه [adá] باید از برتری کمتری برخوردار باشد. با این حال، به دلیل این که واژه [adá] (به معنای «جهنم» در حالت مضاف‌الیه‌ی)، از قبل در زبان روسی موجود است، با محدودیت ANTI-IDENT جریمه می‌شود. از سوی دیگر، از لحاظ جایگاه تولید، همخوان‌های مجاور همخوان حلقی

^۱aphaeresis
^۲homonymy

^۳ این محدودیت در مقاله کراس‌وایت (Crosswhite, ۱۹۹۷) معرفی شده است.

«ع» /ʕ/ همخوان‌های نرم‌کامی /g/ و /k/ و /x/ هستند. اما از آنجا که در زبان روسی وجود آغازۀ الزامی نیست، طبق اصل کم‌کوشی، گزینه‌ای انتخاب شده است که کمترین کوشش در تولید آن به کار گرفته شود.

ب) فرایند حذف میانی:

فرایند حذف میانی^۱، مانند حذف همخوان میانی در واژه‌های «رعیۀ» «ра́йя» [rájja]، «رهج الغار» «реальга́р» [real'gár]، «صاحب» «са́иб» [saíb]، «ثعلب» «сáлеп» [sál'ep]، «سید» «се́йд» [seíd]، «سَل» «се́ль» [sel']، با دقت در واژه‌هایی که در آنها فرایند حذف رخ داده است درمی‌یابیم که در همه این فرایندها همخوانی بیگانه که در زبان روسی موجود نیست، از وام‌واژه حذف شده است. برای مثال، همخوان /h/ و /dʒ/ در «رهج الغار»، همخوان /ع/ /ʕ/ در «ثعلب». همچنین، حذف واکه از واژه ی «حناء» «хна» /xna/ نیز دیده می‌شود. در اینجا دو واژه «صاحب»، «ثعلب» و «حناء» را مورد بررسی قرار می‌دهیم.

در وام‌گیری واژه «صاحب»، اتفاقی که افتاده قابل توجه است. این واژه بر اثر فرایند نرم‌شدگی حذف به صورت [saib]، و همچنین به دو صورت [saxib]، [sagib] در زبان روسی موجود است، که در جهت عکس فرایند غیردهانی‌شدگی (تبدیل واج‌های دهانی به حلقی) صورت گرفته‌اند. یعنی علاوه بر فرایند حذف، راهبرد تبدیل همخوان بیگانه به نزدیکترین همخوان موجود در زبان نیز مورد استفاده قرار گرفته است. در تابلوی (۶) این گزینه‌ها بررسی شده‌اند:

6) Lazy >> MAX-IO >> ONSET >> IDENT[voice, cont]

تابلوی (۶)

Input:	LAZY	MAX-IO	ONSET	IDENT[voice]	IDENT[cont]
/s ^ʕ a.hib /					
a. saib		*	*	*	*
b. sa.xib	*				
c. sa.gib	*			*	*
d. sa.kib	**!				*

همان‌طور که در تابلوی (۶) مشخص است، گزینه (a) اصل کم‌کوشی را رعایت کرده است و به همین دلیل از محدودیت ضد حذف MAX-IO و محدودیت آغازۀ هجا ONSET که در رتبه بعدی قرار دارند تخطی کرده است. گزینه‌های (b) و (c) هم در جهت عکس فرایند نرم‌شدگی (غیردهانی‌شدگی) تولید شده‌اند. که بررسی آنها خارج از حیطه این پژوهش است. اما از آنجا که هر دو در زبان روسی مورد استفاده قرار می‌گیرند به عنوان گزینه‌ای که تخطی مهلک داشته معرفی نشده‌اند. در گزینه

^۱ syncope

(d) همخوان /k/ یک همخوان انسدادی بیواک است و از دو جهت غیرسایشی و غیرواکدار بودن اصل کم‌کوشی را نقض کرده است و دو ستاره دریافت کرده است.

در واژه «ثعلب»، از آنجا که همخوان حلقی «ع» /ʕ/ در زبان روسی وجود ندارد، مطابق راهبرد دوم، این همخوان از واژه حذف شده است. گزینه‌هایی که با راهبرد جایگزینی همخوان غایب با همخوانهای مجاور موجود در زبان روسی ایجاد می‌شوند نیز در تابلوی (۷) مد نظر قرار گرفته‌اند.

7) Lazy >> NOCODA >> IDENT[low] >> MAX-IO

تابلوی (۷)

Input:	LAZY	NOCODA	IDENT[low]	MAX-IO
/θaʕ.lab/				
a. [sā.lib]			*	*
b. [sā.lab]	*!			*
c. [sax.lab]	*!	*		
d. [sag.lab]	*!	*		

طبق محدودیت کم‌کوشی LAZY، گزینه (a) گزینه بهینه است. از طرفی گزینه‌های (c) و (d) هم تا حدی در جهت نرم‌شدگی ایجاد شده‌اند، اما بر خلاف واژه «صاحب» که همخوان مورد نظر در آغاز بود، در واژه «ثعلب» همخوان حلقی در پایانه هجا قرار دارد. بنابراین، اگر از راهبرد جایگزینی همخوان حلقی یا همخوانهای مجاور موجود در زبان روسی استفاده شود و پایانه حذف نشود، از محدودیت NOCODA نیز تخطی کرده است. بنابراین، گزینه بهینه همان (a) است نه دو گزینه دیگر. از طرف دیگر، در گزینه ب، واکه افتاده /a/ در وام‌واژه بدون تغییر قرار گرفته و تبدیل به واکه افراشته /i/ نشده. از آنجا که ارتقای واکه، خود در راستای اصل کم‌کوشی است، گزینه [sālab] نیز با این که محدودیت پایایی افتادگی واکه IDENT[low] را رعایت کرده است، با محدودیت LAZY جریمه می‌شود.

فرایند حذف واکه از واژه «حناء»:

زبان روسی در مواجهه با وام‌واژه‌هایی که با همخوان‌های بیگانه (ملازی، حلقی، چاکنایی و...) آغاز می‌شوند، معمولاً از راهبرد حذف همخوان حلقی استفاده می‌کند، زیرا در این زبان وجود آغاز الزامی نیست؛ مانند حذف همخوان حلقی /ʕ/ در واژه‌های «عادة» و «علماء». اما در مواجهه با وام‌واژه «حناء»، از آنجا که در زبان روسی واژه [ana] به عنوان یک ضمیر پرکاربرد وجود دارد، برای اجتناب از وام‌واژه همنام، آغاز هجا یعنی واج حلقی /h/ حذف نمی‌شود. بلکه از راهبرد جایگزینی واج بیگانه با نزدیکترین واج در زبان وام‌گیرنده استفاده می‌شود و وام‌واژه به [xna] تبدیل می‌شود. در نظریه بهینگی، محدودیت ANTI-IDENT واژه‌های مشابه با واژه‌های از قبل موجود در زبان را جریمه می‌کند.

لازم به ذکر است که، در بعضی ساخت‌های واژگانی زبان روسی واکه هجای اول حذف شده و واژه دوهجایی به صورت یک هجا با دو همخوان در آغاز تلفظ می‌شود. این واکه حذف شده «مصوت فرّار» نامیده می‌شود؛ مانند تلفظ [rota] به صورت

[ɾta]؛ که در آن واکه /o/ مصوت فرّار است. در واژه «حناء» نیز واکه /a/ مصوت فرّار است و از واژه حذف شده است. در نظریهٔ بهینگی، عامل حذف واکه میانی را محدودیت SYNC می‌نامند.

ANTI-IDENT: برون‌داد نباید با واژه ای که در زبان وام گیرنده وجود دارد هم آوا باشد.

SYNC: عامل حذف واکه میانی

بر اساس توضیحات بالا و داده‌های موجود در تابلوی (۸) رتبه‌بندی محدودیت‌ها برای واژه «حناء» به صورت زیر است:

8) NUCLEUS >> ONSET >> NOCODA >> SYNC >> ANTI-IDENT >> IDENT[voice],
IDENT[cont] >> LAZY >> MAX-IO, *COMPLEX_{ONSET}

تابلوی (۸)

Input:	NUCL	ONSE	NOCOD	SYNC	ANTI-IDENT	IDENT[voice]	IDENT[cont]	LAZY	MAX-IO	*COMPL EX _{onset}
/hanāʔ/	EUS	T	A							
a. [a.nā]		*!			*				*	
b. [x.nā]								*	*	*
c. [x.nā]	*!							*	*	
d. [xnāx]			*!							
e. [gnā]						*!	*		*	*
f. [xa.nā]				*!				*		
g. [ga.nā]				*!		*	*			

در تابلوی (۸)، گزینه a به دلیل نداشتن آغاز، گزینه c به دلیل هجای بدون هسته، گزینه d به علت حضور پایانه جریمه شده‌اند. در گزینه e و g واج /g/ به عنوان واج جایگزین در واژه به کار رفته که به دلیل تخطی از محدودیت یکسان نبودن مشخصه بی‌واک و پیوسته جریمه شده‌اند.

ج) فرایند حذف پایانی:

فرایند حذف پایانی^۱، مانند واژه‌های «ذرة» [dur:á] «دُرّ» [dur:á]، «رعيّة» [rájja] «رایّ» [rájja]، «بورق» [bura] «بورق» [bura]، «مُرابط» [marabú] «مارابُ» [marabú]. در داده‌های مورد بررسی، فرایند حذف همخوان پایانی در تعدادی از داده‌ها از جمله دو واژه «بورق» و «مُرابط» مشاهده شد. هر دو همخوان حذف شده /G/ و /tʰ/ در زبان روسی وجود ندارند و این زبان از راهبرد حذف همخوان بیگانه یا غایب در وام‌گیری استفاده کرده است. در تابلوی (۹) این گزینه‌ها مورد بررسی قرار گرفته‌اند:

9) LAZY >> NOCODA >> MAX-IO >> IDENT[voice,cont]

تابلوی (۹)

Input:/mu.ra.bitʰ/	LAZY	NOCODA	MAX-IO	IDENT[voice]	IDENT[cont]
a. [ma.ra.bú]			*	*	*
b. [ma.ra.bút]	*!	*			
Input: /bu.rag/					
a. [bu.ra]			*	*	*
b. [bu.rak]	**!	*			*
c. [bu.rag]	*!	*		*	*

apocope^۱

همانطور که پیش‌تر هم اشاره شد، هجای بهینه در اکثر زبان‌ها، از جمله زبان روسی، هجای باز (CV) است. بنابراین، گزینه‌ای که پایانه نداشته باشد در اولویت قرار دارد. گزینه‌هایی که پایانه دارند هم از محدودیت کم‌کوشی LAZY تخطی کرده‌اند و هم از محدودیت NOCODA. همچنین در گزینه دوم وام‌واژه /bu.raɡ/، همخوان /k/ به دو دلیل انسدادی و بیواک بودن دو بار از محدودیت LAZY تخطی کرده است.

در تابلوی (۱۰) واژه «طرحه» مورد بررسی قرار می‌گیرد:

10) LAZY >> ONSET >> NOCODA >> MAX-IO

تابلوی (۱۰)

Input: /tʰar.ħat/, /tʰar.ħah/	LAZY	ONSET	NOCODA	MAX-IO
a. [tar.xat]	*		**!	
b. [tar.xa]	*		*!	*
c. [tar.at]		*!	*	*
d. [ta.ra]				*
e. [tar.a]		*!	*	*
f. [ta.rat]			*!	*

واژه «طرحه» در زبان عربی به دو صورت کامل و همراه با وقف تلفظ می‌شود. در حالت اول، با تایی تأنیث و به صورت [tʰar.ħat] و صورت دوم با حالت وقف و به صورت [tʰar.ħah]. همانطور که پیش‌تر هم بیان شد، زبان روسی فاقد همخوان‌های حلقی و چاکنایی است و در مواجهه با وام‌واژه‌هایی که دارای چنین همخوان‌هایی هستند، یکی از این دو نوع راهبرد را برمی‌گزیند: ۱- حذف همخوان حلقی / چاکنایی؛ ۲- جایگزینی آن با نزدیکترین همخوانی که در روسی وجود دارد. در واژه «طرحه» نیز همخوان بیگانه «ط» /tʰ/ وجود دارد که با همخوان مشابه /t/ در زبان روسی جایگزین می‌شود؛ همچنین در دو صورت ارائه شده این وام‌واژه، دو همخوان حلقی /ħ/ و همخوان چاکنایی /h/ وجود دارد که هر دوی آنها حذف می‌شوند. با توجه به حذف /t/ پایانی در صورت اول می‌توان اینطور فرض کرد که صورت پایه این وام‌واژه همان [tar.ħah] (در حالت وقف) در نظر گرفته شده است. طبق داده‌های جدول، گزینه [ta.ra] با رعایت محدودیت NOCODA و محدودیت کم‌کوشی (LAZY) به عنوان گزینه بهینه انتخاب شده است.

نکته قابل توجه این که، در واژه «مرابط» همخوان بیگانه /tʰ/ در پایانه هجا قرار دارد و بر اثر محدودیت NOCODA حذف می‌شود، اما در واژه «طرحه» در آغاز قرار داشته و به دلیل محدودیت ONSET که وجود آغاز را برای هجا الزامی می‌داند، به جای حذف، تبدیل به همخوان مشابه موجود در زبان روسی یعنی /t/ می‌شود (راهبرد جایگزینی به جای راهبرد حذف).

۵. نتیجه گیری

در این پژوهش انواع فرایندهای نرم‌شدگی در وام‌واژه‌های عربی موجود در زبان روسی مورد بررسی قرار گرفت. نتایج نشان داد که از ۱۴۹ وام‌واژه بررسی شده ۲۹ واژه دستخوش انواع فرایندهای نرم‌شدگی شده‌اند که عبارتند از: سایشی‌شدگی (۵ مورد)، واکنارشدگی (۱ مورد)، تشدیدزدایی (۸ مورد) و حذف (۱۵ مورد). تغییرات به وجود آمده در داده‌های مورد نظر با انواع محدودیت‌های نشاننداری و پایایی نظریه بهینگی توصیف و تبیین شدند. بررسی انواع این محدودیت‌ها نشان می‌دهد که محدودیت

کم‌کوشی LAZY مهمترین نقش را در پدیده نرم‌شدگی همخوان‌ها در وام‌واژه‌ها ایفا می‌کند؛ به طوری که نسبت به محدودیت‌های وفاداری در اولویت قرار دارد. با این وجود، در صورتی که واژه برون‌داد حاصل از کم‌کوشی با یکی از واژه‌های پرکاربرد موجود در زبان روسی همنام باشد، آن برون‌داد با محدودیت ضد همنامی ANTI-IDENT جریمه می‌شود و گزینه دیگر غیرهمنام به عنوان گزینه بهینه انتخاب می‌شود. به طور کلی، نتایج تحلیل داده‌ها در چارچوب نظریه بهینگی نشان می‌دهد که زبان روسی در مواجهه با همخوان‌های بیگانه موجود در وام‌واژه‌های عربی از دو راهبرد استفاده می‌کند: اول، جایگزینی آن با نزدیکترین همخوان موجود در زبان روسی؛ دوم، حذف همخوان بیگانه، در صورتی که واژه به دست آمده محدودیت‌های زبانی و واج‌شناختی زبان روسی را نقض نکند.

منابع

- آرلاتو، آنتونی (۱۹۸۱). در آمدی بر زبان‌شناسی تاریخی. ترجمه: یحیی مدرسی (۱۳۸۳). تهران: انتشارات سمت.
- ثمره، یدالله (۱۳۷۸). *آواشناسی زبان فارسی: آواها و ساخت آوایی هجا*. ویراسته ی رضا نیلی پور. امید طبیب زاده. تهران: مرکز نشر دانشگاهی.
- جم، بشیر (۱۳۸۸). *نظریه بهینگی و کاربرد آن در تبیین فرایندهای واجی زبان فارسی*. رساله دکتری. تهران: دانشگاه تربیت مدرس.
- صرامی، ولی الله (۱۴۰۲). «تحول فرایند تضعیف در گویش خلخال (درو) از ایرانی باستان تا به امروز: بر پایه ی نظریه ی بهینگی». *فصلنامه پازند*. سال ۱۷. شماره ۶۲. صص. ۳۹-۵۴.
- عامری، حیات (۱۳۹۵). «وام واژه های عربی در فارسی، بررسی تحولات آوایی و معنایی». *دوفصلنامه مطالعات تطبیقی فارسی-عربی، سال ۱*. شماره ۲. صص ۸۷-۹۷.
- کرد زعفرانلو کامبوزیا، عالیہ (۱۳۸۵). *واج‌شناسی: رویکردهای قاعده بنیاد*. تهران: سمت.
- کرد زعفرانلو کامبوزیا، عالیہ و افتخارسادات هاشمی (۱۳۹۴). «انطباق واجی واکه های وام واژه های زبان عربی در فارسی. رویکرد بهینگی». *دوماهنامه جستارهای زبانی*. د. ۶. ش. ۶. صص. ۲۲۹-۲۵۷.
- کرد زعفرانلو کامبوزیا، عالیہ، فردوس آقاگل‌زاده، و سکینه نویدی باغی (۱۳۹۳). «بررسی قانون مجاورت هجا در واژه های بسیط زبان فارسی». *مجله زبانشناسی و گویش های خراسان*. دوره ۶، شماره ۱۰، صص ۲۵-۴۹.
- مدرسی قوامی، گلناز (۱۳۹۴). *فرهنگ توصیفی آواشناسی و واج شناسی*. تهران: انتشارات علمی.
- نغزگوی کهن، مهرداد (۱۳۹۳). *فرهنگ توصیفی زبان‌شناسی تاریخی*. تهران: انتشارات علمی.
- نغزگوی کهن، مهرداد (۱۴۰۴). *زبان‌شناسی تاریخی*. تهران: انتشارات آگه.
- ولی پور، علیرضا (۱۳۹۳). *آواشناسی زبان روسی*. تهران: انتشارات سمت.
- ولی پور، علیرضا و محمد جمال زاد (۱۴۰۱). «بررسی ویژگی‌های صوت‌شناختی همخوان‌های سایشی در زبان روسی در بیان دانشجویان فارسی زبان ایرانی و مقایسه آن با زبان معیار». *پژوهش های زبانشناختی در زبان های خارجی*. دوره ۱۲، شماره ۳، صص. ۲۳۸-۲۵۶.
- هایمن، لاری م. (۱۹۷۵). *نظام آوایی زبان. نظریه و تحلیل*. ترجمه: یدالله ثمره (۱۳۶۸). تهران: انتشارات فرهنگ معاصر.
- Al-Kadimi, Mahmood G. Ch. (2010). Arabic borrowing in the Russian Language. *Voronezh State University Bulletin. Series of Linguistics and Intercultural Communication*, No. 1, 75-77. [In Russian]
- Bakhtiyarova, Alfira N. (2016). On the issues of phonetic development of Arabisms in the Bashkir language, *Vestnik of the Bashkir University*. 1 (4). 763-768. [In Russian].
- Bakhtiyarova, Alfira N. & Flyuza G. Fatkullina (2015). Arabic loanwords in the lexical system of the Russian language. *Journal of "Philological Sciences"*. 2. 6124-6128. [In Russian]. <https://fundamental-research.ru/ru/article/view?id=38632> (Date of access: 28.10.2024).
- Crosswhite, Katherine (1997). Avoidance of Homophony in Trigrad Bulgarian Vowel Reduction. Paper presented at *Southwestern Workshop on Optimality Theory*, University of California, Los Angeles.
- Dell, François, & Mohamed Elmedlaoui (1985) "Syllabic Consonants and Syllabification in Imdlawn Tashlihyt Berber," *Journal of African Languages and Ling.* 7, 105-130.
- Kirchner, Robert (1998) *An Effort- Based Approach to Consonant Lenition*. Doctoral dissertation. Los Angeles: University of California.

- Kirchner, Robert (2004). "Consonant Lenition.". In: Bruce Hayes (Eds.), *PHONETICALLY Based Phonology*. Robert Kirchner, and Donca Steriade. Cambridge: Cambridge University Press, pp.313-345.
- McCarthy, John (2002) *A Thematic Guide to Optimality Theory*; Cambridge: Cambridge University Press.
- Moradi, Maryam (2022). The analysis of phonetic elision: aphaeresis elision, syncope elision and apocope elision in Arabisms in Russian language in the phonetic process. *Bulletin of Udmurt University. Series History and Philology*. 32 (2). 206-214. [In Russian].
- Moradi, Maryam & Zohreh Ghorbani-Madavani (2020). Analysis of the semantic narrowing of Arabisms included in the Russian language. *Journal of "Theory of Language and Intercultural Communication" of Kursk State University*, 1 (36). 206-214. [In Russian].
- Reztsova, Vera V. (2011). Lexical-thematic classification of Arabisms in the Russian language. Material presented at the III Russian students' electronic scientific conference "Student Scientific Forum 2011, Astrakhan State University, Astrakhan, Russia. *Fundamental research*. 2 (27). 6124-6128. [In Russian] <https://scienceforum.ru/2011/article/2011001521>
- Siraziyev, Ilgiz I. (2002). The Arabic layer of the vocabulary of the modern Tatar literary language, *abstract of the dissertation for the degree of PhD. In Philological Sciences*. Kazan: Kazan Federal University. [Unpublished doctoral thesis]. [In Russian]. <https://dspace.kpfu.ru/xmlui/handle/net/167177>
- Spirkin, Alexander G., Igor A. Akchurin, Regina S. Karpinskaya, Ivan V. Lekhin, Slava M. Lokshina, Fedor N. Petrov & Lev S. Dhaumiyan (1984). *Slovar' inostrannyh slov [Dictionary of foreign words]*. Moscow: Russian Language. [In Russian].

Extended Abstract

A Diachronic Study of the Morphology of the Infinitives /*hastīdan*/ and /*būdan*/ and a Critique of the Formation and Use of /*mī-bāšad*/ in Standard Written Persian

Roshanak Azari 

Assistant Professor, Department of Persian Language and Literature Education, Farhangian University, Tehran, Iran

Azari@cfu.ac.ir

Review History: Received: 2026-05-22 Revised: 2026-08-08 Accepted: 2026-08-23	Abstract In the Persian language, /<i>būdan</i>/ and /<i>hastīdan</i>/ are two of the most applied infinitives. Their conjugation is in the present and past systems, in different tenses, moods and aspects. The purpose of this descriptive-qualitative article is to study the morphological evolutions of the mentioned verbs with a diachronic approach from Ancient to Middle and New eras of the Persian language. This study indicates that according to the pattern [present stem + past stem suffix + infinitive suffix], the infinitive /<i>hastīdan</i>/ is more correct than /<i>hastan</i>/. Moreover, in Middle Persian, the suppletive stems were derived from two different roots; accordingly, the present and past stems had been from two different roots. This characteristic was observed in some Pahlavi verbs including /<i>būdan</i>/ and /<i>hastīdan</i>/. According to the Pahlavi texts, the conjugation of the verb /<i>būdan</i>/ was observed in the present system, with the stem /<i>bav-</i>/. Gradually, the future stem /<i>bāš-</i>/ had been replaced with it. As the future system is vanished in the New Persian, using /<i>ast</i>/ or /<i>hast</i>/ is preferred to /<i>mī-bāšad</i>/ in the speech and the writing.	Keywords: Suppletive stems /<i>hastīdan</i>/ Secondary infinitive /<i>bav-</i>/ /<i>mī-bāšad</i>/
How to cite this article: Azari, R. (2026). A Diachronic Study of the Morphology of the Infinitives /<i>hastīdan</i>/ and /<i>būdan</i>/ and a Critique of the Formation and Use of /<i>mī-bāšad</i>/ in Standard Written Persian. <i>Journal of Linguistics & Iranian Languages</i>, 11(2), 89-101. doi: https://doi.org/10.22099/jlil.2026.56205.1477		
 COPYRIGHTS ©2026 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publisher.		

Introduction

The verbs */būdan/* and */hastīdan/* are among the most frequent verbs in Persian. In addition to functioning as main verbs, they occur in copular and auxiliary constructions. Despite their semantic proximity, the inflected forms associated with these two verbs are not derived from a single stem. This difference raises questions about the historical relationship of */ast/* and */hast/* to */būdan/*, the formation of the infinitive */hastīdan/*, and the status of */mī-bāšad/* in standard written Persian. Answering these questions requires a distinction between root and stem and an examination of evidence from different periods of the Iranian languages.

This study examines the morphological and inflectional development of these verbs from the ancient period through Middle and New Persian. Its central question is how forms derived from different roots came to be associated within the Persian verbal system and whether their history can provide a basis for judging their present-day use. Alongside an examination of the morphology of the infinitives */hastīdan/* and */bāšīdan/*, the study presents evidence for the present-tense inflection of */būdan/* in Pahlavi. Finally, it clarifies the difference between the historical origin of */bāš/* and the morphological analysis of */mī-bāšad/* in contemporary Persian.

Research Method, Previous Studies, and Objectives

This study adopts a diachronic approach and a descriptive-qualitative method. Data for the ancient period were extracted from forms recorded in Avestan and Old Persian dictionaries. For the Middle period, verbal forms in the Indian Bundahišn, Selections of Zādspram, Pahlavi Rivāyat, and Zand ī Wahman Yasn were examined. Selected forms are presented in the article's tables alongside their person, number, tense and mood, and New Persian meaning to enable comparison of the stems and their paths of development. The discussion focuses on formation and inflection; an examination of the frequency and function of */mī-bāšad/* in a corpus of contemporary Persian falls outside its scope.

Previous studies have offered differing views on the infinitive corresponding to */ast/* and */hast/*. Some sources record */hastan/* or */astan/* as infinitives, whereas Abolqāsemi considers */hastīdan/* and */astīdan/* more justifiable morphologically, given the way an infinitive is formed from a past stem. Earlier grammatical and historical studies have also examined the function and development of the historical roots associated with */būdan/*.

The study aims to answer the following questions: (1) Why is the use of */hastīdan/* more appropriate than */hastan/*? (2) What is the difference between */ast/* and */mī-bāšad/*? Is the use of */mī-bāšad/* in contemporary Persian correct?

Findings and Discussion

In the ancient Iranian languages, verbal forms were built from various roots and stems across several tense and four systems. Avestan and Old Persian evidence shows that forms derived from the root */ah-/* were not confined to the present system, nor were forms derived from */bav-/* */bū-/* confined to the past system. For example, Old Persian records forms derived from */bav-/* with present meaning, such as */bavatiy/*, as well as forms with past meaning. In addition to present forms such as */astiy/*, other forms derived from */ah-/* occur in the perfect system. Thus,

a simple division that assigns one root exclusively to the present and the other exclusively to the past conflicts with the evidence from the ancient period.

As the verbal system developed in the Middle period, some of the systems and inflectional distinctions of the ancient period disappeared or merged. In this period, the relationship between the present and past stems of certain verbs cannot be explained solely by phonological changes within a single root. The concept of “suppletive stems” describes a situation in which synonymous stems originate from different roots and occur together in the inflection of a verb. This study examines the relationship between forms associated with */hastīdan/* and */būdan/* within this historical development.

One finding of the article is the occurrence of present-tense forms of */būdan/* in Pahlavi texts. Forms such as */bawēm/*, */bawē/*, */bawēd/*, and */bawēnd/* show that the stem */baw-/* was used in present-tense inflection. These forms are recorded in the Pahlavi Rivāyat, Selections of Zādspram, Bundahišn, and Zand ī Wahman Yasn. Forms such as */bawīd ēstēd/* also show the use of */baw-/* in forming a past stem by adding the morpheme */-īd/*. The existence of present-tense forms of */būdan/* in Pahlavi must therefore be taken into account when explaining the history of this verb. In infinitive formation, the infinitive suffix */-an/* is added to the past stem. If */hast/* is treated as a present stem, the past stem is first formed by adding */-īd/*, after which the infinitive suffix is attached. This process yields */hastīdan/*. On this basis, the article favors */hastīdan/* over */hastan/* from a morphological perspective.

An examination of */bāš/* also shows that, in the ancient Iranian period, it was a future stem associated with the root */bav-/* and that it survived in the imperative form */bāš/*. The infinitive */bāšīdan/*, formed by adding the elements of infinitive formation to this stem, is considered a secondary infinitive. Since */bāv + šy/* is itself a future stem, adding the past-stem-forming morpheme */-īd/* to it constitutes a secondary or artificial formation. From the article’s diachronic perspective, this history is one reason to favor */ast/* and */hast/* over */mī-bāšad/* in formal prose. Nevertheless, */mī-bāšad/* is a form used in contemporary Persian and can be segmented as */mī + bāš + -ad/* in a synchronic analysis. The historical origin of the stem */bāš/* does not, in itself, prevent its use in a construction with present-time meaning.

Conclusion

A comparative examination of ancient and Middle-period evidence shows that the inflectional relationship between */hastīdan/* and */būdan/* resulted from a development in several stages and cannot be reduced to a fixed opposition between two roots for the present and past. Forms recorded in Old Persian attest to the inflectional diversity of both roots; Pahlavi evidence also shows that */būdan/* was inflected in the present tense with the stem */bav-/*. Furthermore, on the basis of the pattern for forming an infinitive from a past stem, the article favors */hastīdan/* over */hastan/* in its morphological analysis and considers */bāšīdan/* a secondary infinitive.

Two kinds of judgment must be distinguished regarding */mī-bāšad/*. On the basis of the historical development of the stems, this study favors */ast/* and */hast/* in the present indicative of formal prose. This conclusion, however, does not deny the use of */mī-bāšad/* in present-day Persian. Establishing its place, frequency, and function in different varieties of standard Persian requires a separate synchronic, corpus-based study.

References

- Abolghasemi, M. (1389). Historical grammar of Persian (8th ed.). Tehran: SAMT. [In Persian]
- Abolghasemi, M. (1376). A guide to the ancient languages of Iran (Vol. 2). Tehran: SAMT. [In Persian]
- Abolghasemi, M. (1373a). Stems of Persian Dari verbs. Tehran: Qoqnoos. [In Persian]
- Abolhasani Chimeh, Z. (1390). Types of lexical aspect in Persian verbs. *Persian Literature*, 1(6), 101–120. [In Persian]
- Aghamohseni, F. (1390). 'Būdan' as a copular and full verb in Old Iranian and Western Middle Iranian. *Literary Thoughts*, 2(7), 185–199. [In Persian]
- Anklesaria, B. T. (1957). *Zand ī Vohūman Yasn*. Bombay: Messrs. K. L. Bhargava.
- Anklesaria, B. T. (1964). *Zatsparam*. Bombay: Godrej Memorial Printing Press.
- Azari, R. (1400). A diachronic study of the formation of Persian active and passive participles and an analysis of its educational content in old and new secondary school textbooks. *Educational Research Quarterly*, 7(25), 36–48. [In Persian]
- Azari, R. (1401). Mood, aspect, and the structure of conditional sentences in Pahlavi: The present system. *Proceedings of the Fifth International Conference on Iranian Languages and Dialects*, Great Islamic Encyclopedia Center, 1–17. ISBN 978-622-5510-00-5. [In Persian]
- Bagheri, M. (1396). A history of the Persian language (22nd ed.). Tehran: Qatreh. [In Persian]
- Bartholomae, C. (1384). A history of Iranian phonemes (V. Dumanian, Trans.). Tehran: Institute for Humanities and Cultural Studies. [Persian translation]
- Bateni, M. R. (1382). A description of the grammatical structure of Persian (14th ed.). Tehran: Amir Kabir. [In Persian]
- Behzadi, R. (1368). *The Indian Bundahišn*. Tehran: Institute for Humanities and Cultural Studies. [In Persian]
- Dabir-Moghaddam, M. (1389). Theoretical linguistics: The emergence and development of generative grammar. Tehran: SAMT. [In Persian]
- Farshidvard, K. (1351). The verb /būdan/ and its development in Persian. *Vahid*, 10(7), 780–789. [In Persian]
- Jabbari, M. J. (1387). Constructed infinitives: A descriptive-synchronic approach. *Language and Linguistics*, 4(1), 41–61. [In Persian]
- Jackson, A. V. W. (1892). *Avesta grammar* (Part 1). Stuttgart: W. Kohlhammer.
- Kent, R. G. (1953). *Old Persian* (2nd ed.). New Haven: American Oriental Society.
- MacKenzie, D. N. (1379). A concise Pahlavi dictionary (M. Mirfakhraei, Trans.; 2nd ed.). Tehran: Institute for Humanities and Cultural Studies. [Persian translation]
- Mosaffa Jahromi, A. (1392). A generative view of the grammatical category of the infinitive in Persian: Approaches and solutions. *Language Research*, Alzahra University, winter special issue, 163–175. [In Persian]
- Mousavi, S. H. (2022). When past meets future in Persian: A construction grammar approach to futurity. *Australian Journal of Linguistics*, 42(1), 1–34. <https://doi.org/10.1080/07268602.2022.2073545>
- Najafi, A. (1390). *Foundations of linguistics and their application to Persian* (11th ed.). Tehran: Niloufar. [In Persian]
- Nyberg, H. S. (1974). *A manual of Pahlavi* (Vol. 2). Wiesbaden: Otto Harrassowitz.
- Rashed-Mohassel, M. T. (1364). An introduction to Avestan grammar. Tehran: Karyan. [In Persian]
- Rezaei, V., & Nisani, M. (1392). A syntactic and pragmatic study of copularized constructions in Persian based on Role and Reference Grammar. *Language Research*, Alzahra University, 4(8), 23–51. [In Persian]
- Sims-Williams, N. (2004). Ideographic writing I: Terminology and conventions. *Encyclopaedia Iranica*, 12(6), 617–618.
- Williams, A. V. (1990). *The Pahlavi Rivāyat*, accompanying the *Dādestān ī Dēnīg*. Copenhagen: The Royal Danish Academy of Sciences and Letters.
- Zarshenas, Z. (1387). *A handbook of Sogdian*. Tehran: Institute for Humanities and Cultural Studies. [In Persian]

زبان‌شناسی و گویش‌های ایرانی

سال ۱۱، شماره ۲، پیاپی ۲۱ (تابستان ۱۴۰۵) شماره صفحات: ۸۹ - ۱۰۱

بررسی ساخت‌واژی مصدرهای «هستیدن» و «بودن» و نقدی بر ساخت و کاربرد «می‌باشد» در زبان فارسی نوشتاری معیار: رویکرد در زمانی

روشنک آذری 

استادیار گروه آموزش زبان و ادبیات فارسی، دانشگاه فرهنگیان، تهران، ایران

چکیده

«بودن» و «هستیدن» در زبان فارسی، از پربسامدترین مصدرهایی هستند که در دستگاه حال و گذشته، در زمان‌ها، وجوه و نمودهای گوناگون صرف می‌شوند و نقش‌های متفاوتی چون فعل اصلی، ربطی و کمکی ایفا می‌کنند. هدف از نگارش این مقاله توصیفی- کیفی نقد صورت‌های صرفی بر اساس شواهد موجود در واژه‌نامه‌های زبان فارسی باستان، اوستایی و متون پهلوی، با هدف بررسی تحولات ساخت‌واژی این دو فعل از دوره باستان به میانه تا نو زبان فارسی است. این پژوهش نشان داد که کاربرد «هستیدن» بر اساس الگوی {ستاک حال + تکواژ ماضی‌ساز+ تکواژ مصدرساز}، درست‌تر از «هستن» است. علاوه بر این، در دوره میانه ستاک‌های مکمل یا ستاک‌های هم معنا، از شقاق یافته از دو ریشه کاملاً متفاوت بوده و ستاک حال و گذشته آنها از ریشه‌های مختلف سرچشمه می‌گرفته‌اند. این ویژگی برای «هستیدن» و «بودن» صادق است. همچنین در متون پهلوی، صرف فعل «بودن» بر دستگاه حال وجود داشته و ستاک حال آن «بَؤ-» /bav-/ بوده‌است. از آنجاکه «باش» /bāš-/ ستاک آینده از دوره باستان است و ساخت «می‌باشد» بر ستاک ثانویه صورت گرفته‌است؛ درست‌تر خواهد بود که در نوشتار به جای «می‌باشد»، «است» و «هست» به کار رود.

واژه‌های کلیدی:

ستاک ثانویه
ستاک‌های مکمل
می‌باشد
هستیدن
/bav-/

تاریخچه مقاله: دریافت: ۱ خردادماه ۱۴۰۵

بازنگری: ۱۵ مردادماه ۱۴۰۵

پذیرش: ۱ شهریورماه ۱۴۰۵

* آدرس ایمیل نویسنده مسئول: azari@cfu.ac.ir

آذری، روشنک (۱۴۰۵). بررسی ساخت‌واژی مصدرهای «هستیدن» و «بودن» و نقدی بر ساخت و کاربرد «می‌باشد» در زبان فارسی نوشتاری معیار: رویکرد در زمانی. زبان‌شناسی و گویش‌های ایران ۱۱(۲): ۸۹-۱۰۱. <https://doi.org/10.22099/jlil.2026.56205.1477>

حق نشر -۲۰۲۶ نویسنده(گان). این مقاله با دسترسی آزاد است که تحت شرایط مجوز Creative Commons Attribution (CC BY-NC 4.0) منتشر شده است، استفاده، توزیع و تکثیر نامحدود را در هر رسانه‌ای مجاز می‌داند، مشروط بر اینکه به نویسندگان و منبع اصلی اشاره شود. هیچ اجازه‌ای از نویسندگان یا ناشر لازم نیست.

۱. مقدمه

تغییرات در زمانی^۱ که فرآیندی مسلم و حتمی در همه زبان‌هاست، نظام آوایی، تکواژی و نحوی زبان را پیوسته تحت تاثیر و دگرگونی قرار می‌دهد و در هر جای این نظام بقایای این تغییرات مشهود است؛ از این رو برای توصیف همزمانی^۲ ساخت زبان، آگاهی از تغییرات در زمانی اجتناب‌ناپذیر است (دبیرمقدم؛ ۱۳۸۹: ۳۸). در میان زبان‌های هندواروپایی، زبان فارسی این ویژگی ممتاز را دارد که می‌توان تاریخ آن را بر اساس منابع و مدارک موجود، در تمامی ادوار تاریخی از قرن ششم پیش از میلاد تا زمان حاضر به طور پیوسته بررسی کرد (باقری، ۱۳۹۶: ۱۱).

زبان فارسی نو با پشت سر گذاشتن دو دوره تاریخی باستان و میانه، تغییر و تحول چشمگیری را در سه حیطه واج‌های زبان، نظام دستگاه صرفی-نحوی و واژگان، طی کرده‌است (باقری، ۱۳۹۶: ۱۹). در روی بُعد زمان، زبان فارسی به گونه‌های بیشماری تقسیم می‌شود که به صورت زنجیری از فارسی باستان تا امروز ادامه دارند. هر یک از گونه‌های کلی زبان فارسی، خود دارای گونه‌های فرعی زیاد است که در حیطه‌های گوناگون زبانی با هم تفاوت‌هایی دارند (باطنی، ۱۳۸۲: ۸).

«بودن» و «هستن» (هستیدن) در زبان فارسی، از پربسامدترین مصدرهایی هستند که در دستگاه حال و گذشته، در گونه گفتاری و نوشتاری زبان فارسی^۳، در زمان‌های گوناگون و در جوه و نمودهای متنوع صرف می‌شوند. صرف این دو مصدر در زبان فارسی نقش‌های متفاوتی در سطح جمله چون فعل اصلی، ربطی و کمکی ایفا می‌کند و در کتاب‌های دستور سنتی یا ساختگرا همواره بخشی به توضیح در مورد صرف این دو فعل اختصاص می‌یابد.

هدف از نگارش این مقاله بررسی ساختاری دو فعل «هستیدن» و «بودن» در صرف دستگاه حال و گذشته، با رویکرد در زمانی از فارسی باستان تا زبان‌های فارسی میانه و فارسی نو است تا بتوان به نتایج دقیق‌تری در مورد مصدر، ستاک و صرف این دو فعل دست یافت. در این مقاله به طور خاص به ساخت «می‌باشد» پرداخته می‌شود.

در این بررسی، بر اساس شواهد ضبط‌شده در متون اوستایی، فارسی باستان و پهلوی، صورت‌های مختلف دو فعل بررسی و تحولات صرفی آنها به بحث گذاشته می‌شود. تلاش بر این است تا بر پایه این شواهد، حلقه‌های تغییر و تحول در صرف و ساخت ستاک این دو فعل مشخص گردد. همچنین در نظر است در پژوهش حاضر به این پرسش‌ها پاسخ داده شود:

۱- چرا کاربرد «هستیدن» درست‌تر از «هستن» است؟

۲- تفاوت «است» با «می‌باشد» چیست؟ آیا کاربرد «می‌باشد» در زبان فارسی معاصر درست است؟

۲. پیشینه پژوهش

نظرات گوناگونی در مورد مصدر فعل «است» در مقالات و کتب گوناگون دستوری فارسی وجود دارد. نوابی در مقاله «بررسی فعل "است" در فارسی و پهلوی»، با استناد به مثال‌هایی از متون کهن فارسی نو و متون پهلوی، با در نظر گرفتن حیطه نحو و معنا، نتیجه می‌گیرد که «است» مانند همه فعل‌هایی که زمان گذشته آنها با تکواژ ما ضی ساز «-ست» ساخته می‌شود، بایستی زمان گذشته را برساند؛ چون: رست، گسست، خست بدین ترتیب در مثال‌هایی که «است» از نظر دستوری

^۱ Diachronic

^۲ Synchronic

^۳ در این مقاله صرفاً به گونه نوشتاری زبان فارسی معیار پرداخته می‌شود.

سوم شخص زمان حال را نشان می‌دهد، بایستی از مصدر «ایستادن/ایستادن» باشد. او «است» را کوتاه شده «ایستد/ایستد» *ēstēd* می‌داند (نوابی، ۱۳۷۹: ۲۸).

بنابر نظر ابوالقاسمی، «هستن» و «استن» مصدرهای نادرستی هستند که برای «هستم، هستی...» و «استم، استی...» فرض شده‌اند. از آنجاکه «ن» از ستاک گذشته مصدر می‌سازد، اگر قرار باشد برای «هستم» و «استم» مصدری فرض شود، باید آن را به صورت‌های «هستیدن» و «استیدن» فرض کرد؛ یعنی «هست» و «است» را ماده مضارع فرض کنیم و از آنها با «ید» ستاک گذشته بسازیم و سپس به آن نشانه مصدری «ن» بیافزاییم (ابوالقاسمی، ۱۳۷۳: ۸۱). باطنی در نقد «دستور پنج استاد» که دستوری سنتی است، به صرف فعل «استن» اشاره می‌کند (باطنی، ۱۳۷۰: ۶۸). کلباسی در توضیح فعل «پی بستی»^۱ نظیر «خوبم، خوبی، خوبست» واژه بست‌ها را صورت‌های تصریفی فعل «بودن» در زمان حال معرفی می‌کند و مصدر مجزایی برای آن ارائه نمی‌دهد (کلباسی، ۱۳۸۰: ۳۲). بنابر نظر فرشیدورد «بودن» از افعال معین و ناقص است و وابسته‌های آن از دو ریشه *ah-* (ام، ای، ... استم، استی...) و *bav-* در فارسی باستان آمده‌اند (فرشیدورد، ۱۳۵۱: ۷۸۱). وی مصدر «باشیدن» را نیز در کنار «بودن» می‌آورد و مثال‌هایی از متون کهن برای آن ذکر می‌کند: در تنها باشیدن وسواس غلبه کند (کیمیای سعادت)؛ مردی بود از عرب به بخارا با شیده (تاریخ بخارا) (فرشیدورد، ۱۳۵۱: ۷۸۸). ناتل خانلری در زبان پهلوی سه معین فعل *ah-*، *bu-* و *sta-* را معرفی می‌کند. نکته حایز اهمیت این است که وی کاربرد *sta-* (ایستادن) را برای ساخت ماضی نقلی و بعید می‌داند (ناتل خانلری، ۱۳۷۴ (۲): ۱۹۶).

از مقالات همایشی جدید می‌توان به «تکوین فعل "است" در زبان فارسی» اشاره کرد که فعل *ah-* را از فارسی باستان تا تبدیل به واژه بست بررسی می‌کند (شریفی، تابش‌فر، ۱۳۹۷). آقامحسینی در مقاله «بودن در نقش فعل ربطی و تام در ایرانی باستان و میانه غربی» با رویکرد ساختاری-معنایی از زبان‌های اوستایی و فارسی باستان تا پهلوی، جملاتی برای نقش ربطی و اصلی این فعل مثال آورده است (آقامحسینی، ۱۳۹۰). مقالاتی نیز بر اساس نظریه‌های زبان‌شناسی و رویکردهای گوناگون، فعل‌های خاص زبان فارسی نظیر «شدن» و «بودن» را بررسی و یا نمود این فعل‌ها را تحلیل می‌کنند (بهرامی‌خورشید؛ ورهاگن؛ گلفام، ۱۳۹۰ و ابوالحسنی‌چیمه، ۱۳۹۰).

در لغت‌نامه دهخدا مصدر «هستن» به معنی «وجود داشتن» و «زیستن» نوشته شده و مصدر «استن» مترادف با «هستن» آمده است. در فرهنگ معین دو معنا برای «هستن» دیده می‌شود: ۱- وجود داشتن، موجود بودن. ۲- وقوع داشتن، حاصل بودن. در همین فرهنگ «است» صورتی تصریفی از مصدر «استن» معرفی شده است. در فرهنگ عمید برای «هستن» معنای «بودن» و «وجود داشتن» ذکر گردیده است (دهخدا، ۱۳۷۷ (۱۳): ۱۱۴۴).

۳. چارچوب نظری پژوهش

زبان‌های جهان علیرغم وجوه اشتراکی که با هم دارند، از نظر ساختمان متفاوت هستند (باطنی، ۱۳۸۲: ۱۳). یکی از مزایای بررسی با دیدگاه علمی نسبت به دستور سنتی، توصیفی بودن زبان‌شناسی ساختاری است؛ بدین ترتیب مقولات خارجی بر زبان مورد مطالعه تحمیل و تجویز نمی‌شود (وحیدیان کامیار، ۱۳۸۵: ۵). برای مثال در زبان انگلیسی، معادل هر دو مصدر «بودن» و

^۱ Enclitic Verb نام دیگر آن «واژه‌بست» است.

«هستیدن» «to be» و یا در زبان فرانسه «être» است^۱؛ پس چرا در زبان فارسی برای صرف در دستگاه مضارع از مصدر «هستیدن» و برای دستگاه ماضی از مصدر «بودن» استفاده می‌شود؟ اگرچه هر سه زبان از شاخه زبان‌های هندواروپایی هستند و فعل‌های ربطی-کمکی^۲ در اکثر زبان‌های این سرشاخه وجود دارند؛ اما از آنجا که در بررسی علمی زبان‌ها، مطالعه هر زبان به قصد شناخت همان زبان صورت می‌گیرد؛ زبان‌شناسی تاریخی، با نشان دادن تحولات آوایی، صرفی و نحوی واژه‌ها در زبان فارسی بر اساس مستندات موجود در دوره‌های مختلف زبان، می‌تواند به طور قانونمند پاسخگوی این پرسش باشد. همچنین، در زبان‌شناسی تاریخی، بازسازی زبانی، امکان ارائه صورت‌هایی از مراحل پیشین زبان را فراهم می‌کند که شاهد مستقیمی از آنها در دست نیست؛ این صورت‌ها بر اساس شواهد زبانی موجود و روابط نظام‌مند میان داده‌ها بازسازی می‌شوند. صورت‌های بازسازی شده، که معمولاً با علامت ستاره (*) مشخص می‌شوند، فرضیه‌هایی مبتنی بر شواهد زبان‌شناختی درباره صورت‌های پیشین زبان به شمار می‌روند (لین^۳، ۲۰۱۸: ۶۹).

۳-۱. روش و پیکره پژوهش

روش این پژوهش توصیفی-کیفی است. ابتدا توضیح کلی در مورد شیوه‌های ساخت ستاک حال در دوره باستان تا میانه برای دو فعل «هستیدن» و «بودن» بیان می‌گردد. سپس بر اساس شواهد ضبط شده در متون اوستایی، فارسی باستان و پهلوی، صورت‌های مختلف صرف شده دو فعل می‌آید و تحولات صرفی به بحث گذاشته می‌شود. برای سهولت مشاهده و مقایسه، تعدادی صورت‌های تصریفی در جدول نمایش داده می‌شود. تلاش بر این است تا حلقه‌های تغییر و تحول در صرف و ساخت ستاک این دو فعل مشخص گردند (حیطه معنایی جای بحثی دیگر دارد). با شرح ستاک‌های مکمل و مصدر ثانوی در ادامه مقاله، پاسخ پرسش‌هایی که در بخش بیان مسئله مطرح شده‌است، به دست می‌آید.

پیکره پژوهش برای زبان‌های دوره باستان، مثال‌های مستخرج از واژه‌نامه‌های اوستایی و فارسی باستان و برای زبان پهلوی نمونه‌های موجود در کتاب‌های بندهش هندی (۱۳۶۸)، گزیده‌های زاد سپرم (۱۹۶۴)، روایت پهلوی (۱۹۹۰) و زند بهمن یسن (۱۹۵۷) بودند.

۳-۲. ساخت فعل در دوره باستان زبان فارسی

در میان زبان‌هایی که از دوره باستان زبان فارسی، از آن‌ها آثار مکتوبی بجا مانده که قابلیت بررسی آوایی، صرفی و نحوی یا معنایی داشته باشند، فارسی باستان و اوستایی هستند (ابوالقاسمی، ۱۳۷۳: ب: ۲). از سوی دیگر، زبان‌های ایرانی باستان و هندی باستان، هر دو منشعب از زبان هندواروپایی آغازین هستند و زبان‌های خواهر به شمار می‌آیند و واژه‌های مشترک فراوان در بین متون کهن ایرانیان و هندیان، مشاهده می‌گردد (سیفی، بختیاری، محمدی، ۱۳۹۲: ۷۲ و ۷۴). در زمینه صرف و نحو نیز شباهت به حدی است که در اوستا جملات کاملی یافت می‌شوند که به موجب تطابق قواعد آوایی بین دو زبان اوستایی و سنسکریت، قابل تبدیل به معادل‌های زبان خواهر یعنی سنسکریت هستند و لذا ترجمه آنها به کمک دستور و واژه‌نامه هندی کلمه به

^۱ با در نظر گرفتن حیطه‌های معنایی گسترده‌تر چون «وجود داشتن»، می‌توان معادل‌های بیشتری در زبان انگلیسی و فرانسه برای آن در نظر گرفت که طرح آن در این مقاله نمی‌گنجد.

^۲ Copular and Auxiliary verbs

^۳ J.K. Lin

کلمه ممکن می‌شود (بارتلمه،^۱ ۱۳۸۴: ۱). اگرچه زبان فارسی امروز ما دنباله فارسی باستان است؛ اما چون متون فارسی باستان محدود به کتیبه‌های دوره هخامنشیان می‌شود و شواهد موجود در آن برای بحث حاضر محدود است، در ادامه مقاله، تعدادی از مثال‌ها از زبان اوستایی خواهند بود.

۳-۳. ریشه

ریشه جزئی است که میان لغت‌های تشکیل دهنده یک خانواده لغوی مشترک است و معنی اصلی آن لغت‌ها را دربردارد (ابوالقاسمی، ۱۳۸۹: ۱۴۶). بررسی صرفی در دوره باستان زبان فارسی نشان می‌دهد که زمان‌های گوناگون دستوری و مشتق‌های اسمی و صفتی بر یکی از سه ریشه ضعیف (/a/ ندارد)، افزوده (قوی) که /a/ دارد و بالانده (قوی‌تر) که /ā/ دارد، ساخته می‌شده است. البته بسته به واژه موجود در ریشه و یا در صورت وجود همخوان خیشومی صرف آنها متفاوت است. مثلاً برای فعل «بردن»، صورت ضعیف آن /barə-/br-، صورت افزوده /bar- و صورت بالانده /bār- می‌شود. مثال دیگر: فعل «آمدن»، به صورت ضعیف /ja/ga- (حذف همخوان خیشومی m)، افزوده /gam- و صورت بالانده /gām- است (راشدمحصل، ۱۳۶۴: ۲۱). بدین ترتیب، واژه‌های «بار» و «گام» در زبان فارسی میانه و نو، اسم‌هایی هستند که از ریشه بالانده گرفته شده‌اند (آذری، ۱۴۰۰: ۴۰).

۳-۴. ماده/ستاک

در دوره باستان، ماده فعلی یا ستاک^۲ با افزودن جزئی که در اساس برای افزودن معنی جدید به ریشه بوده، ساخته می‌شده است. البته ساخت ستاک‌ها برای زمان‌های گوناگون بر روی ریشه‌های متفاوت بوده است. برای مثال، همه صیغه‌های التزامی و بعضی از صیغه‌های اخباری و امر بر ریشه افزوده (قوی) ساخته می‌شده است (راشدمحصل، ۱۳۶۴: ۲۱-۴۲)، ساخت زمان‌های دیگر، اغلب بر ریشه ضعیف دیده می‌شود (جکسون^۳، ۱۸۹۲: ۱۴۸-۱۴۷). در جدول شماره ۱ یک نمونه ساخت ماده/ستاک از ریشه در زبان اوستایی دیده می‌شود.

جدول ۱.

ساخت مضارع التزامی از ریشه «بردن»

زمان	ریشه	ماده ساز التزامی	شناسه سوم شخص مفرد	فعل صرف شده	معنا به فارسی نو
مضارع التزامی	bar-	-ā-	-ti	barā(i)ti	ببرد

۳-۵. تفاوت ریشه با ماده/ستاک

همان‌طور که در بخش پیشین نشان داده شد، در دوره باستان، برای صرف فعل و اسم، صورت‌های مختلف ریشه در اکثر مواقع، مستقیم به پسوندها یا شناسه‌های صرفی نمی‌چسبند، بلکه ابتدا باید از آن‌ها ستاک (ماده) ساخت. برای مثال، در ساخت ماده مضارع، بر اساس واجی که ریشه به آن مختوم می‌شده، ۹ روش مختلف ساخت ستاک (ماده) فعلی وجود داشته است.

^۱ C. Bartholomae

^۲ Stem

^۳ A.V. Jackson

بدین ترتیب «بن» در معنای ریشه است و با «ماده» (معادل واژه فارسی «ستاک» و «منشعب از ریشه») متفاوت است. زیرا همان‌گونه که در بخش ۳-۳ توضیح داده شد، ریشه (بن) تنها در دوره باستان زبان‌های ایرانی وجود داشته و از دوره میانه به بعد و به تبع آن، در فارسی نو، کاربرد «ریشه» از میان رفته است و فقط ماده‌های فعلی ساخته شده از آن کاربرد دارند؛ پس بهتر است در آموزش دستور زبان فارسی، لفظ دیگری جز «بن» (که به معنای ریشه است)، برای ماده ماضی یا مضارع انتخاب شود (آذری، ۱۴۰۰: ۴۰).

۳-۶. دستگاه‌های فعلی و ستاک‌های آن

دوره باستان زبان فارسی، چهار ستاک و بدین ترتیب چهار دستگاه فعلی حال، مطلق^۱، نقلی^۲ (تام) و آینده (فقط در اوستایی) وجود داشته است^۳. برای نشان دادن پیچیدگی و تنوع زیاد صرف فعل، ذکر این نکته کافیست که در هر یک از این دستگاه‌ها علاوه بر وجوه اخباری، التزامی، خواستاری^۴ (دعایی/تمنایی)، انشایی^۵ و امری؛ همچنین بر اساس جهت^۶ فعل، شناسه‌های صرفی گذرا^۷ با ناگذر^۸ متفاوت بوده است (کنت، ۱۹۵۱: ۷۵). از آنجا که زبان رو به سادگی در حرکت است، در دوره میانه، دستگاه فعلی مطلق (آئوریست) و نقلی (تام) ادغام شده و فقط دو دستگاه گذشته و حال به فارسی میانه و نو رسیده است (آموزگار، تفضلی، ۱۳۸۲: ۶۳). بعضی وجوه نیز از بین رفته‌اند؛ برای مثال ساخت وجه خواستاری در فارسی نو (معاصر) از بین رفته و کاربرد آن در وجه التزامی ادغام شده است (آذری، ۱۴۰۰: ۴).

۴. تحلیل داده‌ها

۴-۱. صرف ریشه‌های /ah-/ و /bav-/

۴-۱-۱. دوره باستان

از لحاظ معنایی در دوره باستان زبان‌های ایرانی، /ah-/ در مفهوم کلی «بودن» و «وجود داشتن» به کار رفته؛ درحالی‌که /bav-/bū- در مفهوم کلی «بودن»، «وجود داشتن» و «شدن» کاربرد داشته است (آقامحسینی، ۱۳۹۰: ۱۸۶). بر اساس شواهدی که در متون فارسی باستان و اوستایی وجود دارد، هر دو فعل در دستگاه‌های فعلی و زمان‌ها و وجوه گوناگون دوره باستان صرف می‌شده‌اند^۹. بدین ترتیب عبارت ذیل نمی‌تواند قابل تعمیم به همه موارد کاربرد این دو ریشه در دوره باستان باشد. در فارسی باستان بعضی فعل‌ها دو ریشه متفاوت دارند. از جمله فعل‌هایی که دارای دو ریشه متفاوتند و هر دسته از صیغه‌های آن از یک ریشه مشتق شده، فعل «بودن» است. /bav-/ شدن، بودن) برای صیغه‌های ماضی و /ah-/(استن، هستن) برای صیغه‌های مضارع است (ناتل خانلری، ۱۳۸۲: ۲/۹۲).

^۱ Aorist

^۲ Perfect

^۳ دستگاه فعلی آینده فقط در اوستایی بوده است. در فارسی باستان سه دستگاه فعلی داریم (کنت، ۱۹۵۳: ۷۰).

^۴ Optative

^۵ Injunctive

^۶ Voice

^۷ Active

^۸ Middle

^۹ برای نمونه‌های اوستایی بنگرید به بارتلمه، ۱۹۶۱: ۹۲۷؛ ذیل مدخل -bav و همان منبع ص. ۲۶۶، ذیل مدخل -ah؛ برای نمونه‌های فارسی باستان بنگرید به کنت، ۱۹۵۳: ۷۱ و ۷۲ و ۱۷۴ و ۲۰۰.

در جدول‌های شماره ۲ و ۳ صرف‌هایی که از دو ریشه مذکور در کتیبه‌های فارسی باستان ضبط شده‌اند، بر اساس واژه‌نامه کنت^۱ (۱۹۵۲) آمده است. توضیح آنکه imperfect زمانی بین ماضی ساده و نقلی است که در فارسی نو معادل ندارد^۲.

جدول ۲.

نمونه‌های تصریفی از ریشه /bav-/ در فارسی باستان

نمونه صرف	معنا	ریشه
/bavatiy/	می‌شود، است	
/bava ⁿ tiy	می‌شوند، هستند	ضعیف $\sqrt{b\bar{u}-}$
/abavam/	بودم، بوده‌ام ^۳	افزوده $\sqrt{bav-}$
/abava ⁿ	بودند، بوده‌اند ^۴	بالانده $\sqrt{*b\bar{a}v-}$
/Bavātiy/	باشم (وجه التزامی)	
/biyā/	باشی (وجه خواستاری)	

جدول ۳.

نمونه‌های تصریف ریشه /ah-/ در فارسی باستان

نمونه صرف	معنا	ریشه
a ^h miy	هستم	
a ^h mahy	هستیم	
ha ⁿ nty	هستند	
astiy	است	ضعیف $\sqrt{h-}$
āham	بودم (نقلی)	افزوده $\sqrt{ah-}$
āha ^t	بود (نقلی)	بالانده $\sqrt{\bar{a}h-}$
āha ⁿ	بودند (نقلی)	
ahaniy	باشم (وجه التزامی)	
āhy	باشی	
ahatiy	باشد	

۴-۱-۲. در دوره میانه

از فارسی میانه به بعد، برای اسامی ریشه وجود ندارد زیرا اسم صرف نمی‌شود و صورت ثابت پیدا می‌کند و نقش صرفی خود را با نقش‌نما نشان می‌دهد. برای افعال نیز فقط دو ماده مضارع و ماضی باقی می‌ماند و بیشتر پیچیدگی‌های ساخت ماده از ریشه در زبان‌های ایرانی میانه غربی از میان می‌رود (آذری، ۱۴۰۰: ۴۱). با ادغام شدن دستگاه‌های آئوریست و نقلی با هم و

^۱ R.G. Kent

^۲ در زبان فرانسه این زمان فعلی وجود دارد و بر اساس متن و نوع فعل در ترجمه فارسی نو ماضی ساده یا نقلی ترجمه می‌شود: J'étais (بودم/بوده‌ام).

^۳ imperfect

^۴ imperfect

حاصل شدن دستگاه گذشته در دوره میانه، تعدادی از ستاک‌ها از بین رفته یا ادغام شده‌اند. در تحول دوره ای زبان، دستگاه «نقلی» که در دوره باستان ستاک و شناسه‌های صرفی خود را داشته، در فارسی نو به زمان مضارع نقلی و ماضی نقلی، در دستگاه حال و گذشته تبدیل شده‌است.

بر اساس واژه‌نامه زبان پهلوی مکنزی، این اطلاعات در مورد فعل‌های مورد بحث (به جز مشتقات) وجود دارد (۱۳۷۹: ۵۴۴):

$ast [AYT < A 'yty] / =$ هست؛ موجود است

$/būdan, b(a/w)- [YHWWN-tn < A yhwon, \sqrt{hwh}] /$ = بودن؛ شدن

همانطور که مشاهده می‌گردد، برای «هست» در واژه‌نامه مکنزی مدخل مصدری نیامده‌است (نگارش حروف بزرگ نشان‌دهنده صورت واژه با هزوارش^۱ است. A کوتاه‌نوشت برای آرامی است).

در واژه‌نامه زبان پهلوی نیبرگ، ذیل مدخل */būtan/* این توضیحات وجود دارد: */baw-/* به عنوان مکمل فعل */h-/* برای ساخت مصدر؛ به عنوان فعل کمکی در ساخت زمان‌های حال و آینده؛ به عنوان ستاک سازنده گذشته ساده و نقلی فعل *h-* کاربرد دارد. در نقش فعل اصلی معانی «بودن، شدن، رخ دادن، به حقیقت پیوستن، برخاستن، به وجود آمدن، رشد کردن، نشأت گرفتن، وجود داشتن» برای آن آمده است (نیبرگ ۱۹۷۴: ۵۱).

اگرچه در دوره میانه صرف فعل «بودن» برای دستگاه گذشته و صرف فعل «هستن/هستیدن» برای دستگاه حال، تا حدی از یکدیگر متمایز می‌گردد، اما صرف فعل «بودن» برای زمان حال در متون گوناگون به چشم می‌خورد:

جدول ۴.

تصرف فعل */būdan/* در زمان حال- زبان پهلوی

معنا	شخص و شمار	مضارع اخباری	ستاک حال
هستم	اول ش. مفرد	<i>/bawēm/</i> ^۲	<i>/baw-/</i>
هستی	دوم ش. مفرد	<i>/bawē/</i> ^۳	
هست	سوم ش. مفرد	<i>/bawēd/</i> ^۴ / <i>/bēd/</i> ^۵	
هستید	دوم ش. جمع	<i>/bawēd/</i> ^۶	
هستند	سوم ش. جمع	<i>/bawēnd/</i> ^۷	

^۱ هزوارش به واژه‌های آرامی گفته می‌شود که در خط پهلوی نوشته می‌شدند، اما هنگام خواندن، برابر فارسی میانه آنها تلفظ می‌شد؛ به بیان ساده، صورت نوشتاری واژه آرامی بود، ولی خوانش آن فارسی میانه بود. سیمز ویلیامز هتروگرام‌های ایرانی میانه را «Aramaeograms» می‌نامد (Sims-Williams, 2004).

^۲ H.S. Nyberg

^۳ روایت پهلوی، ص. ۱۵؛ گزیده‌های زادسپر، ص. ۸۰.

^۴ روایت پهلوی، ص. ۱۹۴؛ بندهش، بندهای ۱ و ۱۳.

^۵ روایت پهلوی، ص. ۲ و ۸؛ گزیده‌های زادسپر، ص. ۴ و ۱۱.

^۶ روایت پهلوی، ص. ۶.

^۷ روایت پهلوی، ص. ۱۷؛ گزیده‌های زادسپر، ص. ۱۵۰.

^۸ روایت پهلوی، ص. ۳۹ و ۴۹؛ زند بهمن یسن، ص. ۲۲ و ۲۳ و ۲۹.

معنا	شخص و شمار	مضارع اخباری	ستاک حال
مضارع التزامی			
باشد	سوم ش. مفرد	<i>/bawād/</i> ^۱ <i>/bād/</i>	
باشند	سوم ش. جمع	<i>/bawānd/</i> ^۲	

در متون پهلوی امر فعل «بودن» به صورت */bāš/* نیز ضبط شده‌است که در بخش هفتم مقاله به آن پرداخته می‌شود. تحول ستاک گذشته فعل «بودن» را می‌توان در این دو نمونه ساخت ماضی نقلی نیز مشاهده کرد: */būd ēstēd/* و */bawīd ēstēd/* در مثال دوم، به ستاک حال */baw-/* تکواژ ماضی ساز */-īd/* اضافه شده و ستاک گذشته ساخته‌است. در دو نمونه از ساخت مصدر مجهول */bawīhistan/* (پدید آمدن، تشکیل شدن، ساخته شدن، شکل گرفتن) نیز کاربرد ستاک حال */baw-/* در زبان پهلوی مشهود است.

۳-۴. ستاک‌های مکمل^۴ در زبان سغدی

زبان سغدی از گروه شرقی و شاخه ایرانی خانواده زبان‌های هندواروپایی است. این زبان از نظر تنوع و حجم ادبیات مکتوب در دوره میانه، مهم‌ترین زبان ایرانی شرقی است (زرشناس، ۱۳۸۰: ۱۵۴). سغدی روزگاری مهم‌ترین زبان ایرانی آسیای مرکزی و زبان میانجی آسیای مرکزی در هزاره اول میلادی بوده‌است (زرشناس، ۱۳۸۰: ۲۲). مقدسی در قرن چهارم هجری، این زبان را گونه‌ای از زبان فارسی در سمرقند و بخارا معرفی می‌کند که از رواج افتاده‌است. این زبان به تدریج جای خود را به زبان‌های فارسی و ترکی داده‌است (زرشناس، ۱۳۸۷: ۱۴).

سغدی نه تنها از نظر زمان، وجه و نمود بزرگ‌ترین و متنوع‌ترین مجموعه ساختاری فعل را داراست؛ بلکه نوآوری‌هایی نیز در گونه‌های ستاک پدید آورده‌است (قریب، ۱۳۷۲: ۲). در زبان سغدی ستاک‌های مکمل یا ستاک‌های هم معنا و انشقاق یافته از دو ریشه کاملاً متفاوت وجود دارند. یعنی ستاک حال و گذشته آنها از ریشه‌های مختلف سرچشمه می‌گیرند. مشابه آن در فارسی فعل «بودن» در مضارع اخباری است که با مکمل آن، فعل «است»، صرف می‌شود (قریب، ۱۳۷۲: ۱۱). در پهلوی و دیگر زبان‌های ایرانی میانه غربی، ستاک حال و گذشته فعل، جز چند استثنا، از یک ریشه‌اند.

سماعی بودن ستاک‌های حال از دوره میانه به بعد، دو دلیل دارد: یکی این که با تحولات آوایی و تلفظی از ایرانی باستان به ایرانی میانه غربی، رابطه دستوری میان دو ستاک از بین رفته‌است؛ مثل: داد/ده؛ زاد/زای. دوم آن که بعضی ستاک‌ها از دو ریشه آمده‌اند؛ مثل: «آمد/آی؛ دید/بین؛ بود/است» (ابوالقاسمی، ۱۳۷۶: ۵۰).

^۱ گزیده‌های زادسپرم، ص. ۱۳۴.

^۲ روایت پهلوی، صص. ۹۱ و ۹۲؛ زند بهمن یسن، صص. ۷۱ و ۸۲.

^۳ گزیده‌های زادسپرم، صص. ۲۹ و ۶۱؛ ۱۵۹؛ زند بهمن یسن، ص. ۷۱.

^۴ Suppletive Stems

۴-۴. مصدر

۴-۴-۱. تکواژ سازنده مصدر

اختلاف نظرهای زیادی در مورد تعلق مصدر به فعل یا اسم در میان دستورنویسان و زبان‌شناسان وجود دارد. در علم زبان‌شناسی، بر اساس رویکرد کرافت^۱ (۱۹۹۱) نشان داده شده که مصدر از نظر داشتن کنش گزاره‌ای ارجاع مانند اسم است؛ اما از نظر طبقه معنایی به فعل شباهت دارد (رضایی، ۱۳۹۵: ۱). در دستور زایشی که مقوله‌های دستوری مرزهایی روشن دارند و نمی‌توان مصدر را متعلق به هر دو مقوله اسم و فعل دانست، شواهدی قوی برای فعلی بودن مصدرهای فارسی ارائه گردیده است؛ هر چند مقوله دستوری «مصدر» در دستور زبان فارسی نیازمند تأمل بیشتر است (مصفاچه‌رمی، ۱۳۹۲: ۱۷۴).

از نظر ساختاری، مصدر در فارسی دری از ماده ماضی با پسوند «-ن» ساخته می‌شود. این پسوند بازمانده */-an/* فارسی میانه است که برای ساختن مصدر از ماده ماضی به کار می‌رفته است. در ایرانی باستان مصدر از ریشه با پسوند */-tanai/* ساخته می‌شده است. بدین ترتیب */-an/* فارسی میانه، بازمانده */-tanai/* ایرانی باستان است (ابوالقاسمی، ۱۳۷۳: ۲۰). در کتاب «تاریخ زبان فارسی» دکتر ناتل خانلری (۱۳۷۴: ۲/۱۰۴-۱۰۱) طی بحثی بر اساس مطابقت نشانه‌های مصدرساز دوره باستان و میانه غربی زبان فارسی و استناد به هزوارش‌های فعلی دوره میانه، بر این باور است که تکواژ اشتقاقی مصدر ساز، «تن/دن» است. نظر به اینکه پسوند سازنده صفت مفعولی در دوره میانه به صورت */t/* یا */d/* و جزئی مرده از ستاک گذشته در آمده‌است (ابوالقاسمی، ۱۳۸۹: ۱۸۱)، نظر غالب دستورنویسان بر «-ن» است.^۲

همانطور که در بخش پیشینه مقاله آمد، بنابر نظر دکتر ابوالقاسمی، اگر «است» یا «هست» ستاک حال فرض شود^۳، ابتدا باید از آن ستاک گذشته (با */-īd/*) ساخته و سپس تکواژ مصدرساز به آن اضافه گردد. بدین ترتیب مصدر «هستیدن» به دست می‌آید.

۴-۴-۲. اگر مصدر «بودن» داریم پس «باشیدن» چیست؟

پیش از این گفته شد که در دوره باستان چهار دستگاه (نظام) فعلی وجود داشته و یکی از آنها دستگاه آینده بوده است. موارد کاربرد آینده در اوستا چندان فراوان نیست و تنها ساخت‌های اخباری و صفتی وجود دارد. به جای ساخت‌های دیگر آن معمولاً وجه التزامی می‌آید که در معنی آینده به کار می‌رود (راشدمحصل، ۱۳۶۴: ۴۹).

برای ساخت ستاک آینده از ریشه */bav-/* در ایرانی باستان، صورت بالانده ریشه به */-šy/-hy/* اضافه می‌گردد:

/bāv- + -šy + -a-/

پس اضافه کردن تکواژ ماضی‌ساز */-īd/* به آن ثانوی یا جعلی است؛ زیرا «باش» خود ستاک آینده است. بدین ترتیب «باشیدن» مصدر ثانوی محسوب می‌گردد (ابوالقاسمی، ۱۳۷۳: ۳۷). امر فعل «بودن»، به صورت «باش» نیز از ستاک آینده دوره باستان به فارسی نو رسیده‌است.

^۱ W. Croft

^۲ برای آگاهی بیشتر در مورد تفاوت‌های ساخت مصدرهای «بر ساخته»، «ثانوی»، «ساختگی» یا «جعلی» با مصدر اصلی، بنگرید به مقاله «مصدر بر ساخته: رویکردی توصیفی-هم‌زمانی» (جباری، ۱۳۸۷).

^۳ برای آگاهی بیشتر از صرف *-ah-* در شخص و شماره‌های مختلف مضارع اخباری زبان فارسی باستان، بنگرید به جدول شماره ۳ در همین مقاله.

۴-۵. آیا کاربرد «می‌باشد» و صفت فاعلی «باشنده» در زبان فارسی معیار درست است؟

بر اساس بحث ستاک‌های مکمل در بخش ۱-۲-۶ و توضیحاتی که در بخش ۲-۷ بیان شد، به نظر می‌رسد کاربرد فعل «است» و «هست» که بر ستاک حال «هستیدن» ساخته شده، بهتر از کاربرد فعل «می‌باشد» است؛ زیرا «می‌باشد» بر ستاک آینده فعل «بودن» ساخته شده است که این دستگاه در زبان فارسی نو وجود ندارد. «باش» ستاکي از دستگاه آینده در دوره باستان است و مصدر ثانوی (جعلی) «باشیدن» نیز بر پایه آن ساخته شده است. از این رو، از دیدگاه در زمانی و بر اساس سیر تاریخی بررسی شده در این پژوهش، کاربرد «است» در مضارع اخباری بر «می‌باشد» اولویت دارد. با این حال، در فارسی معاصر «می‌باشد» به کار می‌رود و ساخت صرفی «می + باش + ـد» در دستگاه همزمانی زبان قابل تحلیل است؛^۱ بدین ترتیب، یافته‌های تاریخی این پژوهش به معنای نفی کاربرد این صورت در فارسی معاصر نیست. بررسی جایگاه، بسامد و کارکرد «می‌باشد» در فارسی معیار مستلزم پژوهشی همزمانی و مبتنی بر پیکره فارسی نو است و از محدوده پژوهش حاضر خارج است.

۵. نتیجه‌گیری

پاسخ به پرسش‌های پژوهش:

- ۱- چرا کاربرد «هستیدن» درست‌تر از «هستن» است؟ پاسخ: از آنجاکه تکواژ */-an/* از ستاک گذشته مصدر می‌سازد، ابتدا باید ستاک گذشته، با افزودن */-īd/* به «است» یا «هست» که ستاک حال فرض شده است، ساخته شود. سپس تکواژ مصدر ساز یعنی «ن» به آن بیفزاییم. بدین ترتیب مصدر «هستیدن» به دست می‌آید.
- ۲- تفاوت «است» با «می‌باشد» چیست؟ آیا کاربرد «می‌باشد» در زبان فارسی معاصر درست است؟ پاسخ: در دوره باستان، بر اساس متون فارسی باستان و اوستایی، صرف دو ریشه *√bav-* و *√ah-* در دستگاه نقلی و ستاک ساخته شده از ریشه *√bav-* در دستگاه حال یا آینده تصریف داشته است. «باش» (*bāš-*) ستاک از دستگاه آینده و از ریشه *√bav-* در دوره باستان است و از نظر ساخت‌واژی مصدر ثانوی (جعلی) محسوب می‌شود؛ زیرا به ستاک (ماده) آینده، مجدداً تکواژ ماضی ساز */-īd/* اضافه شده و مصدر آن به صورت «باشیدن» درآمده است. از این رو، از دیدگاه در زمانی و بر اساس سیر تاریخی بررسی شده در این پژوهش، کاربرد «است» در مضارع اخباری بر «می‌باشد» اولویت دارد. با این حال، در فارسی معاصر «می‌باشد» به کار می‌رود و ساخت صرفی «می + باش + ـد» در دستگاه همزمانی زبان قابل تحلیل است؛ بنابراین، یافته‌های تاریخی این پژوهش به معنای نفی کاربرد این صورت در فارسی معاصر نیست. بررسی جایگاه، بسامد و کارکرد «می‌باشد» در فارسی معاصر مستلزم پژوهشی همزمانی و مبتنی بر پیکره فارسی نو است و از محدوده پژوهش حاضر خارج است.

همچنین در این پژوهش نشان داده شد که یکی از ویژگی‌های فعلی در زبان‌های دوره میانه و به ویژه در زبان سغدی ستاک‌های مکمل است. ستاک‌های مکمل یا ستاک‌های هم معنا، انشقاق یافته از دو ریشه کاملاً متفاوت هستند. یعنی ستاک حال و گذشته آنها از ریشه‌های مختلف سرچشمه می‌گیرند. در زبان پهلوی نیز چند فعل با این ویژگی وجود دارند که

^۱ در تحلیل صرفی صورت‌های فعلی، صورت آوایی و صورتی ستاک، مستقل از معنای تاریخی آن، در ساخت تصریفی فعل اهمیت دارد. از این رو، دلالت تاریخی ستاک «باش» بر آینده مانع از حضور آن در ساخت یک صورت فعلی با دلالت زمان حال نیست. برای نمونه، در ساخت آینده فارسی نیز ستاک گذشته فعل واژگانی در کنار صورت صرف شده فعل کمکی «خواستن» به کار می‌رود؛ مانند «خواهد ساخت». در این ساخت، دلالت بر آینده از کل ساخت حاصل می‌شود و نه لزوماً از ستاک فعل واژگانی (موسوی، ۲۰۲۲).

«هستیدن» و «بودن» جزو این گروه هستند. - در دوره میانه، بر اساس شواهد موجود در متون پهلوی، صرف فعل «بودن» بر دستگاه حال نیز وجود داشته است و ستاک حال /bav-/ بوده است.

پیشنهاد می‌شود در آموزش مقوله فعل، اصطلاح «ستاک» به جای «بن» به کار رود و دو مصدر «هستیدن» یا «استیدن» در تدریس افعال ربطی و جملات گذرا به مسند مورد توجه قرار گیرد. همچنین، با توجه به نتایج در زمانی این پژوهش، پیشنهاد می‌شود در نثر رسمی و معیار، به‌ویژه در رسانه‌ها، نشریات و کتاب‌ها، به کاربرد «است» و «هست» به جای «می‌باشد» اولویت داده شود. باین حال، از آنجا که در زبان‌شناسی توصیفی کاربرد اهل زبان مبنای توصیف است، بررسی جایگاه و کاربرد «می‌باشد» در فارسی معاصر نیازمند پژوهشی مستقل با رویکرد همزمانی و بر پایه داده‌های پیکره‌ای است.

منابع

- آذری، روشنک. (۱۴۰۰). بررسی در زمانی ساخت صفت‌های فاعلی و مفعولی زبان فارسی و تحلیل محتوای آموزشی آن در کتب متوسطه نظام قدیم و جدید، *فصلنامه آموزش پژوهی*، ۷(۲۵)، ۳۶-۴۸.
- آذری، روشنک. (۱۴۰۱). وجه، نوع و ساخت جملات شرطی زبان پهلوی: دستگاه مضارع. مجموعه مقالات پنجمین همایش بین‌المللی زبان‌ها و گویش‌های ایران، مرکز دائرةالمعارف بزرگ اسلامی، صص. ۱۷-۱، شابک: ۵-۵۵۱۰-۰۰-۶۲۲-۹۷۸.
- آقامحسنی، فیروزه. (۱۳۹۰). «بودن» در نقش فعل ربطی و تام در ایرانی باستان و میانه غربی. *اندیشه‌های ادبی*، ۲(۷)، صص. ۱۸۵-۱۹۹.
- آموزگار، ژاله؛ تفضلی، احمد. (۱۳۸۲). *زبان پهلوی: ادبیات و دستور آن*. چاپ چهارم. تهران: نشر معین.
- ابوالحسنی چیمه، زهرا. (۱۳۹۰). انواع نمود واژگانی در افعال فارسی، *مجله ادب فارسی*، ۱(۶)، صص. ۱۰۱-۱۲۰.
- ابوالقاسمی، محسن. (۱۳۷۳). الف. *ماده‌های فعل‌های فارسی دری*. تهران: انتشارات ققنوس.
- ابوالقاسمی، محسن. (۱۳۷۳). ب. *تاریخ مختصر زبان فارسی*. چاپ اول. قم: بنیاد اندیشه اسلامی.
- ابوالقاسمی، محسن. (۱۳۷۶). *راهنمای زبان‌های باستانی ایران*. جلد دوم. تهران: انتشارات سمت.
- ابوالقاسمی، محسن. (۱۳۸۹). *دستور تاریخی زبان فارسی*. چاپ هشتم. تهران: انتشارات سمت.
- بارتلمه، کریستین. (۱۳۸۴). *تاریخچه واج‌های ایرانی*. ترجمه واهه دومانیان. تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- باطنی، محمدرضا. (۱۳۷۰). *نگاهی تازه به دستور زبان*. چاپ چهارم. تهران: انتشارات آگاه.
- باطنی، محمدرضا. (۱۳۸۲). *توصیف ساختمان دستوری زبان فارسی*. چاپ چهاردهم. تهران: انتشارات امیرکبیر.
- باقری، مهری. (۱۳۹۶). *تاریخ زبان فارسی*. چاپ بیست و دوم. تهران: نشر قطره.
- بهرامی‌خورشید، سحر؛ ورهاگن، آری؛ گلفام، ارسلان. (۱۳۹۰). فضای معنایی فعل «شدن» در زبان فارسی: نگرشی شناختی. *جستارهای زبانی*، ۲(۴)، صص. ۱۴۵-۱۶۷.
- بهزادی، رقیه. (۱۳۶۸). *بندھش هندی*. تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- جباری، محمدجعفر. (۱۳۸۷). *مصدر بر ساخته: رویکردی توصیفی - هم‌زمانی. زبان و زبان‌شناسی*، ۴(۱)، صص. ۴۱-۶۱.
- دبیرمقدم، محمد. (۱۳۸۹). *زبان‌شناسی نظری (پیدایش و تکوین دستور زایشی)*. تهران: سمت.
- دهخدا، علی‌اکبر. (۱۳۷۷). *لغت نامه دهخدا، جلد سیزدهم*. تهران: مؤسسه لغت‌نامه دهخدا، دانشگاه تهران.
- راشدمحصل، محمدتقی. (۱۳۶۴). *درآمدی بر دستور زبان اوستایی*. تهران: انتشارات کاربان.
- رضایی، والی. (۱۳۹۵). *مصدر، اسم مصدر و حاصل مصدر از دیدگاه سرنمون رده‌شناختی*. *پژوهش‌های زبان‌شناسی*، ۸(۱)، صص. ۱-۱۶.
- رضایی، والی؛ نیسانی، مژگان. (۱۳۹۲). *بررسی نحوی و کاربرد شناختی ساخت‌های اسنادی شده در زبان فارسی بر اساس دستور نقش و ارجاع. زبان پژوهی دانشگاه الزهراء*، ۴(۸)، صص. ۲۳-۵۱.
- زرشناس، زهره. (۱۳۸۰). *جستارهایی در زبان‌های ایرانی میانه شرقی*. تهران: انتشارات نیلوفر.
- زرشناس، زهره. (۱۳۸۷). *دستنامه سغدی*. تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- سیفی، رویا؛ بختیاری، رحمان؛ محمدی اکمل، احمد. (۱۳۹۲). *پیشینه نفوذ واژه‌های ایرانی در زبان‌های هندی*. *فصلنامه مطالعات شبه قاره*، ۵(۱۶)، صص. ۷۱-۹۰.

- شریفی، شهلا؛ تابش‌فر، لیلا. (۱۳۹۷). تکوین فعل "است" در زبان فارسی. مجموعه مقالات نخستین همایش ملی تحقیقات ادبی دانشگاه علامه طباطبائی. اسفندماه. تهران.
- فرشیدورد، خسرو. (۱۳۵۱). فعل «بودن» و تحول آن در زبان فارسی. مجله وحید. دوره ۱۰ (۷). صص. ۷۸۹-۷۸۰.
- قریب، بدرالزمان. (۱۳۷۲). نظام فعل در زبان سغدی. مجله زبان‌شناسی. ۱۰ (۱). صص. ۵۲-۲.
- کلباسی، ایران. (۱۳۸۰). ساخت/اشتقاقی واژه در فارسی/امروز. چاپ دوم. تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- مصفاچهرمی، ابوالفضل. (۱۳۹۲). نگاه زایشی به مقوله دستوری مصدر در زبان فارسی (ره‌یافت‌ها و راه‌حل‌ها). زبان‌پژوهی دانشگاه الزهراء. ویژه نامه زمستان. صص. ۱۶۳-۱۷۵.
- مکنزی، دیوید. نیل. (۱۳۷۹). فرهنگ کوچک زبان پهلوی. ترجمه مهشید میرفرخایی. چاپ دوم. تهران: پژوهشگاه علوم انسانی و مطالعات فرهنگی.
- ناتل‌خانلری، پرویز. (۱۳۷۴). تاریخ زبان فارسی، جلد دوم، چاپ پنجم، تهران: فرهنگ نشر سیمیرغ.
- نجفی، ابوالحسن. (۱۳۹۰). مبانی زبان‌شناسی و کاربرد آن در زبان فارسی. چاپ یازدهم. تهران: انتشارات نیلوفر.
- نوابی، ماهیار. (۱۳۷۹). بررسی فعل «است» در فارسی و پهلوی. مجله دانشکده ادبیات و علوم انسانی دانشگاه فردوسی مشهد. ۱ (۳۳). صص. ۳۱-۲۳.

- Anklesaria, B. T. (1957). Zand ī Vohūman Yasn. Bombay: Messrs. K. L. Bhargava.
- Anklesaria, B. T. (1964). Zatsparam. Bombay: Godrej Memorial Printing Press.
- Bartholomae, C. (1961). Altiranisches wörterbuch. Berlin: Walter de Gruyter & Co.
- Croft, W. (1991). Syntactic categories and grammatical relations: The cognitive organization of information. University of Chicago Press.
- Dhbhar, E. Bamanji. (1913). The Pahlavi Riyāyat. Bombay: The Fort Printing Press.
- Jackson, A. V. W. (1892). Avesta Grammar (part 1). Stuttgart: W. Kohlhammer.
- Kent, R. G. (1953). Old Persian. 2nd Edition. New Haven: American Oriental Society.
- Lin, Jeremy K. (2018). "Interpreting the 'documents of language': Linguistic reconstruction as an approach to prehistoric culture." *Res: Anthropology and Aesthetics*, 69-70.
- Mousavi, S. H. (2022). When past meets future in Persian: A construction grammar approach to futurity. *Australian Journal of Linguistics*, 42(1), 1-34. <https://doi.org/10.1080/07268602.2022.2073545>
- Nyberg, H. S. (1974). A manual of Pahlavi. Volume II. Wiesbaden: Otto Harrassowitz.
- Sims-Williams, N. (2004). Ideographic writing I: Terminology and conventions. *Encyclopaedia Iranica*, 12(6), 617-618.
- Williams, A. V. (1990). The Pahlavi Rivāyat, Accompanying the Dādestān ī Dēnīg. Copenhagen: The Royal Danish Academy of Sciences and Letters.