

Semantic Classification of Persian Basic Words Instruction

Amin Navidi¹, Hayat Ameri², Zahra Abolhassani Chime³

Received: 2/10/2023

Accepted: 8/4/2024

Abstract

An attempt has been made in this article to classify Persian basic words from the semantic point of view, with the aim of vocabulary teaching. The list of words that are the basis of this research include 445 words in elementary level and 889 words in intermediate level (Navidi et al., 1400); since the mentioned list is the only Persian basic words list whose corpus include the basic words of the four books of Persian language teaching. The result of this article is a list of basic words in elementary and intermediate level, being classified in 28 groups.

Keywords: *Persian language teaching, vocabulary teaching, basic words, semantic classification.*

¹ Ph.D. Student of Department of General Linguistics, Science and Research Branch, Islamic Azad University, Tehran, Iran. Email: aminnavidi@gmail.com
ORCID:0000-0001-9442-3378

² Corresponding author, Linguistics assistant professor, Tarbiat Modarres University, Tehran, Iran. Email: h.ameri@modares.ac.ir ORCID:0000-0002-2627-3862

³ Linguistics Assistant Professor, SAMT Organization. Email: abolhassani@samt.ac.ir



1. Introduction

It's for years that languages are being taught on the basis of their most frequent vocabularies. Since in all linguistic interactions such as introducing ourselves or somebody else, talking about occupation or major, we deal with a subject, the semantic classification of these vocabularies could be so sufficient in educational content production, testing and language teaching. If an instructor or a writer knows well about the specific vocabularies that language learners in elementary and intermediate level require in addressing or asking for an address, he can then adjust the level of his instruction and content production with the students' linguistic level; as well as utilizing those words in his classes and books. Therefore, the language learners would learn the language faster and easier and will act more successfully in this process. In this research, we will aim to prepare a semantic classification for the basic words being inferred for the study of (Navidi & Ameri, 2021) which includes 445 words in elementary level and 889 words in intermediate level.

Regarding the categorization and vocabulary field, an attempt has been made to find a common definition and borderline for a group of words having at least specific links in one component, so that to be able to categorize them in the same category and domain. The words of different categories typically are similar in one or more component and this makes it impossible to have a certain and clear demarcation. Thereafter, a word can be classified in different categories due to its variant components.

Research Questions

- 1- To which semantic domain do the Persian basic vocabularies belong for the Persian learners in elementary level?
- 2- To which semantic domain do the Persian basic vocabularies belong for the Persian learners in intermediate level?

2. Literature Review

Foreign researches

2-1 Kaedingn (1897)

The first frequency corpus was that of German language which was done by Kaedings. To gather this corpus, he used a data consisted of 11

million words and showed that 50% of German texts are consisted of 207 LEMAs.

2-2 Thorndike (1921)

In “Teacher’s Vocabulary book”, Thorndike presented a list of 10000 English basic words. He manually gathered this list from a corpus of spoken and written sources consisted of 625000 words.

2-3 Felber (1984)

Felber believes that the contents are the main components of thought and utilizes a quadruple division in order to demonstrate the phenomenon in the brain with the purpose of showing the relationship between contents and the language units.

2-4 Seger (1990)

Seger believes that the process of cognition produces contents and consequently leads to categorization. These contents are classified abstractly and methodically. He introduces four methods in this process.

2-5 Felbaum (1998)

Fellbaum placed a conceptual unit at the head of a group by introducing a hierarchical division, with the underneath concepts as its sub-categories. He then presented 25 categories for the nouns and fifteen categories for the verbs.

2-6 Capel (2012)

English Vocabulary Profile is a beneficial tool to help English instructors, since it has classified English vocabularies and grammatical rules on the basis of the common reference framework for European languages. In this project, the words and the rules are classified on the basis of their theme as well as their level. Totally there are 21 thematic categories for vocabularies.

Internal research

۲-7 Badrei (1352)

In his research third appendix, Badrei has presented a list of 19 categories with the aim of exploring the basic vocabularies of Iranian primary school students.

2-8 Imen (1357)

Imen has investigated a list of 4 million words (out of the children’s’ most read books and their speech) and accordingly presented a list of 2398 most frequent vocabularies for the children of 9-12 ages.

2-9 Tahririan (1994)

Tahririan (1994) has prepared a list of the Persian most frequent vocabularies through randomly gathering 5200 sentences of 2 to 35 words and thus he's investigated the Persian's most frequent vocabularies.

Page | 36

2-10 Sanaati (2009)

In her PHD thesis, Sanaati has proposed a list containing 73 categories as well as investigating in details the different categorizations and word-formation processes, as listed below.

2-11 Nemat Zade et al. (2011)

This research is a large national project which introduces 497 words in 24 categories. The words are mainly derived from some comprehension and production exams for students and some special teacher exams for 20000 students and 750 primary school teachers in 175 schools around the country.

2-12 Dadras (2012)

In his research, Dadras has classified the words of first grade in primary school and some vocabularies from the first to fifth grade and thus introduced 24 thematic categories.

2-13 Miller & Aghajanian-Stewart (2017)

31 semantic categories have been presented in the frequency dictionary of Persian prepared by Miller & Aghajanian-Stewart.

3. Methodology

We've combined 8 frequency list of Persian words and thus have prepared a new list. The important point of this research is that the above-mentioned list has regarded a combination of all the previous corpus and thus is a comprehensive corpus of Persian-speakers' oral and written basic vocabularies.

In the following, in order to increase the research accuracy and regarding the instructional experiments of educational books, we've paid great attention to the educational vocabularies of the following four books. Excel software has been utilized with the aim to derive a united list in this regard. All the vocabularies of the research have been organized in one Excel column and the repeated words have been omitted.

The usual way of semantic domains classification is to categorize the language extracted basic vocabularies from the view point of Semantics and then to present the words of each category. For this purpose, the findings of the previous researches regarding the semantic categorization have been utilized (English profiles, Miller & Agha Janinan-Stewart, Nemat Zade & et al and Dadras) and concluded a new categorization with a specific model in teaching Persian to non-Persian speakers which will be introduced later in the present study.

In order to define the vocabularies' semantic category, a special formula as well as an Excel table has been specifically designed for this purpose, with the help of which we could extract the words' categories on the basis of Dadras words list.

Since some words in Dadras research belonged to more than one category, we utilized two more formulas to investigate the different categories; these two formulas assist in classifying all the words' categories in one cell.

Having passed these steps, we prepared a list of the related categories of each vocabulary on the basis of the research resources, but there were some words belonging to no specific category. In order to cover this issue in the reviewed resources, we applied some manual editing and on the basis of the Categorization Theory and investigating each word's application in linguistic corpus, we could then determine each word's main semantic category, as well as defining a semantic category for the words belonging to no specific category. In this regard, the final list includes the special semantic categories of the present research in which each word is classified in its most prominent semantic category

4. Results

As was explained in the previous section, in order to determine the basic vocabularies' semantic categories in this research, we utilized the findings of the three researches lead by Dadras, Miller Agha Janinan-Stewart and English index. In these three researches, vocabularies have been classified semantically. It is worth-mentioning that in Dadras research, a word might belong to more than one semantic category. For instance, the word 'house' belongs to two categories of "house/living environment" and "sport/game/entertainment". In this research, we could on a new revision,

determine the vocabularies' semantic categories on the basis of the findings of the above-mentioned three researches and our own linguistic investigations regarding the categorization theory. In this semantic categorization, in case of existing a word belonging to more than one semantic field, we selected the nearest and the most related semantic category. Radial Network Theory has been utilized in this research with the aim to finalize the semantic categories; for example, the word 'water' belonged to three categories of 'behavior', 'sports/game/entertainment' and 'food', but since the closest meaning was related to 'food' field, the semantic category of 'food' was finally selected.

The list of vocabularies' semantic categories of each level is available at the following link.

https://s25.picofile.com/d/8455254826/bde406da-b6ab-4dba-9c61-d9a00a971ef8/word_list.pdf

References

- Abdolkarimi, S. (2015). Theoretical Frameworks in Semantics. *Linguistic Analysis*, 1: 57-79. [In Persian]
- Badrei, F. (1971). *Written Vocabulary of Primary School Students in Iran*. Tehran: Iran Language Academy. [In Persian]
- Dadras, M. (2012). Multifold Catagorization of Persian Core Vocabulary. *Journal of Curriculum Studies*, 27, 33-66. [In Persian]
- Doroudi, A., Baradaran Hashemi, H., Al Ahmad, A., Zare Bidaki, A., Habibian, A., Mehdikhani, F., Shakeri, A., Rahgozar, M. (2008). *Standard Benchmark for Persian Web Information Retrieval Research*, Tehran. Technical Report of the Database Research Group of Tehran University. No. DBRG-TR-138702. [In Persian]
- Gandomkar, R. (2018). Analyzing the Cognition of Persian Speakers in terms of Categorization: Taxonomic or Thematic Categorization. *Language Related Researches*, 60 (6): 515-538. [In Persian]
- Gharibi, A., Sahraee, R. (2013). Language Functions: Needs Analysis of Persian Language learners' Functional Needs. *Journal of Teaching Persian to Speakers of Other Languages*, 2(5), 37-53. [In Persian]
- Khodami, M., Karami Nouri, R. (2002). A Tool for the Assessment of Word Recognition in First Grade Farsi Books and Investigating its Validity and Reliability in Tehran. *Advances in Cognitive Sciences*, 16, 1-9. [In Persian]
- Navidi, A., AMERI, H., Abolhasani Chime, Z. (2021). Comparison of Vocabulary in the Textbooks of Teaching Persian to Non-Persian Speakers with Persian Basic Words, 10 (2), 213-235. [In Persian]
- Nemat Zade, Sh., Dadras, M., Dastjerdi Kazemi, M., Mansuri Zade, M. (2011). *Persian Basic Vocabulary from Iranian Children Language*. Tehran: Madrese Publications. [In Persian]

- Safavi, K. (2003). Applied Linguistics. Tehran: Hamshahri Publications. [In Persian]
- Safavi, K. (2014). An Introduction to Semantics. Tehran: Soormehr Publications. [In Persian]
- Saffar Moghadam, A. (2007). Persian Language 1: Basic Lessons. Tehran: Persian Language and Literature Development Council. [In Persian]
- Sahraee, R., Gharibi, A., Maleklou, D., Sadeghi, S., Shahbaz, M., & Soltani, M. (2019). Mina 1. Tehran: Saadi Foundation, Azfa Publication. [In Persian]
- Sahraee, R., Marsous, F., & Maleklou, D. (2019). First Step. Tehran: Saadi Foundation, Azfa Publication. [In Persian]
- Sahraee, R., Ahmadighader, S., Marsous, F., Banafshe, L. (2017). Persian Vocabulary in Use: Elementary & Pre-Intermediate. Tehran: Allameh Tabatabaie University. [In Persian]
- Sanaati, M. (2008). Conceptology in Persian Word Formation. Unpublished PHD Thesis. Allameh Tabatabaie University. [In Persian]
- Tahririan, M. (1994) A look at the frequency of basic words and vocabulary in Persian; Allameh Tabatabaie University Proceedings, 80, 47-70.





فصلنامه

سال ۲۱، شماره ۸۶، زمستان ۱۴۰۳، ص ۱۳۵-۱۵۷

مقاله پژوهشی



DOI: <https://doi.org/10.2634/Lire.21.86.135>

طبقه‌بندی معنایی واژگان پایه آموزش زبان فارسی

امین نویدی^۱؛ دکتر حیات عامری^{۲*}؛ دکتر زهرا ابوالحسنی چیمه^۳

پذیرش مقاله: ۱۴۰۳/۵/۲۲

دریافت مقاله: ۱۴۰۲/۱۱/۲۰

چکیده

در این پژوهش تلاش شده‌است که واژگان پایه زبان فارسی به منظور آموزش آن از نظر معنایی نیز طبقه‌بندی شود. فهرستی که بر مبنای آن این طبقه‌بندی صورت گرفته، شامل ۴۴۵ واژه سطح مقدماتی و ۸۸۹ واژه سطح میانی در پژوهش (نویدی و دیگران، ۱۴۰۰) است. این فهرست علاوه بر هشت پژوهش پیکره‌بنیاد، که داده‌های تولیدی، نوشتاری و گفتاری فارسی‌زبانان را جمع‌آوری کرده، تنها فهرستی از واژگان پایه زبان فارسی است که در آن واژگان چهار کتاب آموزش زبان فارسی نیز موجود است. نتیجه این مقاله فهرستی از واژگان پایه در سطح مقدماتی و میانی است که در ۲۸ گروه طبقه‌بندی شده‌است.

کلیدواژه‌ها: آموزش زبان فارسی، آموزش واژه، واژه‌های پایه، طبقه‌بندی معنایی.

۱۳۵



فصلنامه پژوهش‌های ادبی سال ۲۱، شماره ۸۶، زمستان ۱۴۰۳

۱. دانشجوی دکتری گروه زبان‌شناسی، واحد علوم تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

aminnavidi@gmail.com ORCID: 0000-0001-9442-3378

۲. دانشیار گروه زبان‌شناسی، دانشگاه تربیت مدرس، تهران، ایران - نویسنده مسئول

h.ameri@modares.ac.ir ORCID: 0000-0002-2627-3862

۳. دانشیار گروه زبان‌شناسی، پژوهشگاه سمت، تهران، ایران

abolhassani@samt.ac.ir



Copyright© 2023, the Authors | Publishing Rights, ASPL. This openaccess article is published under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License which permits Share (copy and redistribute the material in any medium or format) and Adapt (remix, transform, and build upon the material) under the Attribution-NonCommercial terms

۱. مقدمه

سالهاست که زبانها بر اساس پرکاربردترین و پربسامدترین واژه‌ها آموزش داده می‌شود. از آنجا که در تمامی تعاملات زبانی مانند معرفی خود یا دیگری، صحبت درباره شغل یا رشته تحصیلی ما با یک موضوع سر و کار داریم، طبقه‌بندی موضوعی این واژگان می‌تواند در تولید محتوای آموزشی، آزمون‌سازی و تدریس بسیار کارآمد و مفید باشد. اگر معلم یا نویسنده‌ای بداند، زبان‌آموز سطح پایه و میانی برای پرسیدن نشانی یا آدرس دادن از چه واژه‌هایی باید استفاده کند، می‌تواند آن واژه‌های مورد نظر را هم در کلاس و هم در کتابش به کار ببرد و سطح تدریس و محتوای تولیدیش را مناسب زبان‌آموزان تنظیم کند؛ بدین ترتیب زبان‌آموزان، سریعتر و راحت‌تر زبان را یاد می‌گیرند و در این فرایند موفقتر عمل می‌کنند. کودکی که در حال فراگیری^۱ زبان مادری است با بزرگسالی که در حال یادگیری^۲ زبان خارجی است در فهمیدن و انتقال اطلاعات به مفاهیم مشترکی نیاز دارند که با واژه‌های مشخص و معینی بیان می‌شود. در واقع زبان‌آموز چیزی را می‌آموزد که نیاز دارد و وقت آموختن آن است (غریبی و صحرایی، ۱۳۹۲: ۱۶). می‌توان این واژه‌ها را به شکل تخصصی نیز حوزه‌بندی کرد؛ مثلاً واژه‌های مهم و پربسامد در حوزه ادبیات، هنر، فیزیک و یا اقتصاد را تهیه کرد. فردی که می‌خواهد به زبان فارسی درباره اقتصاد مطالعه کند، باید بداند که پرکاربردترین و پربسامدترین واژه‌های اقتصادی کدام است و او با دانستن کدام واژه‌ها می‌تواند متنی اقتصادی را زودتر و راحت‌تر بخواند؛ بفهمد و درباره آن صحبت کند. در این پژوهش فهرست واژگان پایه به دست آمده از پژوهش (نویدی و دیگران، ۱۴۰۰)، که شامل ۴۴۵ واژه سطح مقدماتی و ۸۸۹ واژه سطح میانی است در یک تقسیم‌بندی معنایی قرار می‌گیرد.

در بحث مقوله‌بندی و حوزه واژگانی، تلاش می‌شود حد و مرز و تعریف مشترکی برای گروهی از واژه‌ها به دست داده شود که دست‌کم در یک مؤلفه پیوندهای مشخصی داشته باشند تا بتوان آنها را در حوزه و مقوله‌ای مشترک قرار داد. واژه‌ها معمولاً در یک یا چند مؤلفه با واژه‌هایی از مقوله‌های دیگر مشترک هستند و این مرزبندی قطعی و روشن میان مقوله‌ها را ناممکن می‌کند؛ بدین ترتیب یک واژه می‌تواند با توجه به گزینش مؤلفه‌های گوناگون، عضوی از مقوله‌های مختلف و متفاوت باشد؛ مثلاً چاقو، که ابزاری در آشپزخانه است، می‌تواند در حوزه معنایی قضایی و

جرم‌شناسی، اسلحه سرد و آلت مجرمانه باشد. بسیاری از مقولات با هم وجوه مشترک دارند و یک واژه می‌تواند در یک یا چند طبقه‌بندی مختلف قرار بگیرد.

سؤالات پژوهش

- واژه‌های پایه برای فارسی‌آموزان خارجی سطح مقدماتی به کدام حوزه‌های معنایی تعلق دارد؟

- واژه‌های پایه برای فارسی‌آموزان خارجی سطح میانی به کدام حوزه‌های معنایی تعلق دارد؟

فرضیه‌های تحقیق

در این پژوهش با توجه به ماهیت سؤالات پژوهش، نمی‌توان فرضیه‌های صفر^۳ برای پژوهش مطرح کرد؛ اما به طور خلاصه هدف پژوهش در تناسب با سؤالات اول و دوم این است که طبقه‌بندی معنایی واژه‌های پایه در سطح مبتدی و میانی ارائه شود.

۲. مبانی نظری

۲-۱ طبقه‌بندی معنایی واژه

طبقه‌بندی واژه‌ها می‌تواند از دیدگاه‌های متفاوتی صورت بگیرد؛ مثل طبقه‌بندی نوع دستوری واژه که شامل اسم، ضمیر، فعل، صفت، قید، حرف اضافه و حرف ربط می‌شود. البته طبقه‌بندی‌های مختلفی از دیدگاه دستوری هست که (کامیار ۱۳۸۶) موارد بسیاری از آنها را در آثار اندیشمندان ذکر کرده است. طبقه‌بندی دیگر، نقش دستوری است؛ یعنی جایگاهی که واژه در محور همنشینی پیدا می‌کند. این طبقه‌بندی به حوزه نحو و نقش نحوی واژه مربوط می‌شود که از بحث این پژوهش خارج است. طبقه‌بندی دیگر، که در این پژوهش مبنای کار قرار گرفته، طبقه‌بندی معنایی و حوزه‌ها و مقوله‌های معنایی است (خدای و کرمی‌نوری، ۱۳۸۱، ۳). گاهی یک حوزه معنایی نسبت به حوزه معنایی دیگر، پایه‌تر و از لحاظ بسامدی بیشتر است که سبب می‌شود به لحاظ یادگیری مقدم باشد؛ مثلاً واژه‌های مادی و ملموسی که به نیازهای جسمی مربوط است مثل «آب» در واژه‌های پایه محک وب (درودی و همکاران، ۱۳۸۷)، واژه شماره ۱۷۰ با ۴۹۶۶۶ بار تکرار است؛ ولی واژه «مقدس»، که به حوزه معنایی ناملموس و غیرمادی تعلق دارد و به عقاید و باورها مربوط است، واژه شماره ۹۷۴ با بسامد ۹۳۲۷۱ است.

مقوله‌ها در حوزه علوم شناختی، طبقه‌بندی مشخص و روشنی ندارد و نمی‌توان مدعی شد که مثلاً یک واژه یا گروهی از واژه‌های مشخص در یک مقوله قرار دارد و به مقوله یا مقوله‌های دیگر مرتبط نیست. خود مقوله‌ها نیز مرز مشخصی ندارد و با مقولات دیگر وجوه مشترکی دارد.

معلوم کردن مقوله‌ای مشخص برای هر واژه کار مشکلی است. بسیاری از واژه‌ها در مرز بین مقوله‌های مختلف قرار دارند و می‌توانند با تغییر یک یا چند ویژگی یا همنشینی با واژه‌های مختلف بخشی از یک حوزه خاص به‌شمار آیند؛ بدین ترتیب بسیاری از واژه‌ها را نمی‌توان با قطعیت در یک طبقه یا حوزه به‌شمار آورد.

برای مشخص کردن یک حوزه واژگانی معمولاً یک یا چند واژه مرکزی وجود دارد که هسته و مرکز آن حوزه به حساب می‌آید. اگر از این مرکز فاصله بگیریم و به حاشیه و اطراف حرکت کنیم، واژه‌ها گویا از مشخصه‌های آن حوزه فاصله می‌گیرند و مشخصه‌های حوزه‌ای دیگر را نشان می‌دهند. در این فضای بینایی نمی‌توان به یک حوزه مشخص و مرزبندی‌شده قائل بود؛ اما می‌توان آن را پیوستاری در نظر گرفت که از یک حوزه به حوزه یا حوزه‌های دیگر حرکت می‌کند. بنابراین واژه‌های مختلف با حرکت روی پیوستاری از حوزه‌های معنایی، همواره در حال نزدیک شدن یا دور شدن از حوزه‌ای به حوزه‌ای دیگر هستند.

مثال آشنا در این مورد مقوله پرندگان است. توانایی پروازکردن از ویژگی‌های مهم پرنده بودن است؛ هر چند شرط کافی برای پرنده‌بودن نباشد؛ بدین ترتیب خفاش در این ویژگی به مقوله پرنده‌بودن خیلی نزدیکتر است تا شترمرغ یا بوقلمون؛ اما طبق ویژگی‌های دیگر پرنده‌بودن مثل تخم‌گذاری، پرداشتن یا مقدارداشتن، شترمرغ و بوقلمون را با وجود پرواز نکردن درون مقوله پرندگان به‌شمار می‌آوریم. خفاش به دلیل نداشتن این مؤلفه‌ها و بچه‌زا بودن و هم‌چنین پوست و پوزه داشتن، پستاندار به‌شمار می‌رود؛ هرچند پرواز کند؛ اما در هر یک از مقوله‌های پرندگان و پستانداران شترمرغ یا خفاش گونه‌هایی نیستند که نمایندگان مشخص و واضحی برای مقوله خود باشند؛ بدین ترتیب می‌توان مقوله‌ها را پیوستاری دانست که با توجه به مؤلفه‌های تشکیل‌دهنده آن مقوله دارای مرکزیت^۱ هستند؛ مثلاً کبوتر بهتر و بیشتر از شترمرغ، مقوله پرندگان را نمایندگی می‌کند و گاو، مؤلفه‌های پستاندار بودن را بهتر از خفاش نشان می‌دهد؛ برای همین با

شنیدن مقوله‌های پرندگان و پستانداران اولین حیوانی که به ذهنمان می‌رسد، کبوتر و گاو است. چیزی را که به محض شنیدن یک مقوله به ذهن متبادر می‌شود، سرنمون^۵ می‌نامند. سرنمونها در هر جامعه زبانی و جغرافیایی فرق دارد؛ مثلاً ممکن است در ایران، کبوتر سرنمون پرنده باشد ولی در انگلیس سینه‌سرخ یا در استرالیا و برزیل طوطی. پرندگان مختلف در میانه و مرکز طیف پرندگان قرار می‌گیرند؛ بدین ترتیب واژه‌هایی که در دو سر این طیف قرار دارند و از مرکز فاصله می‌گیرند، ممکن است با تمرکز روی یکی از عناصر چندگانه مقوله‌بندی براحتی وارد مقوله‌های دیگر شوند؛ مثلاً «کاهو» می‌تواند هم در مقوله غذا باشد و هم در مقوله گیاهان یا واژه «سرد» با توجه به متنی که در آن به کار رفته‌است، می‌تواند متعلق به حوزه فیزیک و دما باشد یا روانشناسی و روابط اجتماعی. در این پژوهش آثاری بررسی می‌شود که برای هر واژه، چندین مقوله معنایی را مشخص کرده‌اند؛ اما در نهایت برای آسانی کار در نتایج پژوهش برای هر واژه فقط اصلی‌ترین و پرکاربردترین مقوله را انتخاب می‌کنیم.

تفاوت دیگر در شناخت مقوله‌ها بین فرهنگهای مختلف نوع مقوله‌بندی است. برخی فرهنگها مقوله‌بندی را بر اساس موضوع و برخی دیگر بر حسب رده دسته‌بندی می‌کنند. مقوله‌بندی رده‌ای بر اساس شباهت بین ویژگیهای چیزها و مقوله‌بندی موضوعی بر اساس روابط مختلف بین چیزها انجام می‌گیرد؛ مثلاً بین «مرغ و سنجاب و درخت» دو واژه مرغ و سنجاب از نظر رده در یک دسته هستند و سنجاب و درخت از نظر موضوع (گندمکار ۱۳۹۷) نشان داده است که ایرانیان به طبقه‌بندی موضوعی (رابطه‌ای) مقوله‌ها گرایش دارند.

مفهوم دیگری که در همین رابطه باید بررسی کرد، هسته معنایی^۶ است. گاهی واژه‌ای در گذر زمان و به دلیل نوع معنایی که دارد، گسترش معنایی می‌یابد و به مفاهیم مشابه هم دلالت می‌کند؛ مثلاً واژه «سرد» در «غذای سرد»، «رفتار سرد» و «اسلحه سرد». در همه این عبارات، معنی واژه حول محور و مفهوم اصلی دلزدگی، کمبود انرژی و اشتیاق و گرما است. مثال دیگر در زمینه مفاهیم حمل و نقل است. زمانی که انسانها از اسب به عنوان وسیله نقلیه استفاده می‌کردند، کلماتی را به کار می‌بردند که بعدها از همانها برای نامیدن اجزای لوازم جدید هم بهره بردند؛ مثل «زین اسب» و «زین دوچرخه» یا «رکاب اسب» و «رکاب ماشین».

یکی دیگر از چالشهای مقوله‌بندی واژه‌ها در ارتباط با واژه‌هایی است که همانندی‌هایی دارند. در ادامه به معرفی مهمترین همانندی واژه‌ها می‌پردازیم که موجب چالش برای یافتن مقوله واژه می‌شود.

چند معنایی^۷ و چندحوزه بودن، که در سطح واژه و تک‌واژ اتفاق می‌افتد، یکی از چالشهای مهم مقوله‌بندی است؛ مثلاً پسوند «ین» سه کارکرد دارد: (صفوی ۱۳۸۲)

۱. کارکردی ساختاری که می‌تواند به تغییر مقوله نوع کلمه منجر شود؛ مانند «چوب» و «چوبین».

۲. کارکرد معنایی که سبب شکل‌گیری چندمعنایی می‌شود؛ مانند «سوم» و «سومین» یا «آتش» و «آتشین».

۳. کارکرد هنری که در متون ادبی زیبایی‌آفرینی می‌کند؛ مانند «کدام» و «کدامین». باید توجه کرد که حضور یک واژه در چند حوزه واژگانی مثل «کاهو»، که هم غذا است هم گیاه (و درباره آن صحبت کردیم) به چندمعنایی ارتباطی ندارد. نوع دیگری از چندمعنایی می‌تواند کاربردی باشد؛ مثلاً در واژه «روی»:

- آن کتاب روی میز است.
- ساعت روی دیوار زنگ زد.
- من دارم روی رساله‌ام کار می‌کنم.

همچنین چندمعنایی می‌تواند به دلیل کاربردهای مجازی، استعاری، تحولات تاریخی و ادبی یا کاربردهای ویژه اجتماعی باشد. این واژه‌ها معنای مرکزی و محوری دارند و در فرهنگ‌ها هم مدخلی واحد را به خود اختصاص می‌دهند؛ مثلاً واژه «ساعت» را در نظر بگیرید که می‌تواند به مفهوم زمان اشاره کند یا وقتی مشخص یا وسیله‌ای برای سنجش زمان:

- او ساعت یازده آمد.
- الان ساعت مناسبی نیست.
- ساعت روی میز قرمز است.

یکی دیگر از چالشهای مقوله‌بندی، مفهوم هم‌نویسه^۸ است. هم‌نویسه به واژه‌هایی گفته می‌شود که به شکل تصادفی در نوشتن مشابه هستند و میان آنها هیچ تشابه لفظی و معنایی دیگری وجود ندارد. این همان پدیده‌ای است که در سنت مطالعات ادبی

«جناس خَط» نامیده می‌شود (صفوی ۱۳۹۳)؛ مثل «سر» که با سه تلفظ مختلف سه معنای متفاوت دارد؛ /sar/، /ser/ و /sor/:

- سر من درد می‌کند.
- من روی برفها سر خوردم.
- کسی از سر این کار خبر نشد.
- یا واژه گل با دو تلفظ متفاوت /gol/ و /gel/:
- بعد از باران زمین گل شد.
- این گُلها خودشان اینجا رشد کردند.

اگر در پیکره مورد مطالعه و به شکل متنی شروع به شمارش واژه‌ها کنیم در حالت مطلوب باید واژه‌های هم‌نویسه، یکسان شمرده و طبقه‌بندی نشود؛ اما چنین کاری در عمل دشوار است و در گزارش، بسیاری از پیکره‌های موجود واژه‌های هم‌نویسه جدا نشده‌است. این تفکیک را یا به روشهای دستی باید انجام داد که زمانبر و هزینه‌بر است و یا با روشهای جدید آماری^۹ و یادگیری ماشینی^{۱۰} که مستلزم داشتن مقدار زیادی متن است.

این نکته نیز مهم است که واژه‌های هم‌آوا،^{۱۱} که در مطالعات ادبی «جناس لفظ» نامیده می‌شود، گرچه مفهومی نزدیک به واژه‌های هم‌نویسه دارد، در مطالعات پیکره‌بنیاد متن‌محور نمی‌تواند مشکلی ایجاد کند؛ چرا که این واژه‌ها هر چند تلفظ یکسان دارد در نوشتار متفاوت است؛ برای مثال واژه‌هایی مانند «خوار» و «خار» را در نظر بگیرید.

چالش سوم در مقوله‌بندی مفهوم همانمی یا هم‌آوا-هم‌نویسی^{۱۲} است که در ادبیات به آن «جناس تام» می‌گویند. این واژه‌ها بدون هیچ ارتباط معنایی صرفاً به یک شکل نوشته و تلفظ می‌شوند. این واژه‌ها در فرهنگها نیز مدخلی جداگانه دارد و باید جدا شمرده و طبقه‌بندی شود. واژه «شانه» نمونه خوبی از این اتفاق است.

- او با شانه چوبی، موهایش را شانه می‌زند.
- کوله‌پشتی سنگین، شانه مرا خسته کرد.
- آنها یک شانه تخم‌مرغ خریدند.

در آخر باید گفت که بعضی از پژوهشگران، میان حوزه واژگانی و حوزه معنایی تمایز قائل هستند؛ مانند لاینز^{۱۳} (لاینز، ۱۹۷۷) که معتقد است حوزه واژگانی تنها شامل

واژه‌ها است؛ اما حوزه معنایی شامل مفاهیمی است که به شکل ترکیبی و اصطلاحی ساخته شده‌است؛ برای مثال واژه‌های «خنده»، «لبخند»، «تبسم» و «قهقهه» در حوزه واژگانی شادی است؛ اما ترکیبات «گل از گل شکفتن»، «ریسه رفتن» و «قند تو دل آب شدن» در حوزه معنایی شادی است.

(لیپکا، ۲۰۱۰) هم این تمایز را پذیرفته‌است؛ اما به جای اصطلاح حوزه واژگانی، حوزه واژگان و به جای حوزه معنایی حوزه واژگانی را به کار می‌برد. در تعریف او حوزه واژگانی شامل هر دو نوع حوزه واژه و معنا است. فرضیه واژگانی^{۱۴} یکی دیگر از مفاهیم مرتبط با حوزه معنایی است. زبان چطور می‌تواند پدیده‌های جهان را در قالبی بریزد و چگونه می‌تواند با دنیای بیرون ارتباط برقرار کند؟^{۱۵} (گودارد و ویرژیستسکا ۱۹۹۴) با ارائه این فرضیه معتقد بودند که مفاهیم نخستین در تمامی زبانهای جهان واژگانی شده‌اند. به باور اینان، مفاهیم مهم و جهانی فارغ از فرهنگ و جغرافیا در قالب واژگان درآمده، و به کمک واژه‌ای خاص یا دست‌کم به کمک یک عبارت بیان شده‌اند؛ هم‌چنین تریر^{۱۶} در مطالعات خود درباره حوزه‌های واژگانی،^{۱۷} که تحت عنوان نظریه حوزه واژگانی مطرح کرد به ارتباط واژه‌ها و مختصات ذهنی پرداخت. او بر این باور بود که واژه‌ها را نمی‌توان به شکل تنها بدون ارتباط با شبکه واژه‌هایی بررسی کرد که با هم در ارتباط معنایی هستند. مرزبندی واژه‌های مرتبط هم به همین شکل باید بررسی شود (عبدالکریمی، ۱۳۹۴: ۶۶). دانش شناختی انسان از طریق همین حوزه‌های کوچک و چینش آنها در کنار هم شکل می‌گیرد؛ مانند موزاییک‌هایی که با کنار هم قرار گرفتن، الگوها را شکل می‌دهد.

به حوزه معنایی و ترکیبات اصطلاحی مرتبط با آن از آنجا که از سطح واژه فراتر رفته در این پژوهش پرداخته نشده است.

۳. ادبیات و پیشینه تحقیق

۳-۱ پژوهش‌های خارجی

۳-۱-۱ کدینگ، (۱۸۹۷)

اولین فرهنگ بسامدی را می‌توان کاری نامید که کدینگ^{۱۸} برای زبان آلمانی انجام داده است. او برای تهیه این فرهنگ از پیکره‌ای شامل یازده میلیون واژه استفاده کرد و نشان داد که پنجاه درصد از متنهای زبان آلمانی فقط از ۲۰۷ فرم واژه (لما) تشکیل شده‌است.

۲-۱-۳ ثورندایک (۱۹۲۱)

ثورندایک^{۱۹} (۱۹۲۱) در «کتاب واژه معلم»، فهرستی از ده هزار واژه پایه انگلیسی ارائه کرد. او این فهرست را به شکل دستی از پیکره‌ای شامل ۶۲۵ هزار واژه از ادبیات کودکان، سه میلیون واژه از انجیل و کتابهای کلاسیک انگلیسی، ۳۰۰ هزار واژه از کتابهای سطح مقدماتی مدارس، پنجاه هزار واژه از کتابهای آشپزی، خیاطی، کشاورزی و تجارت، ۹۰ هزار واژه از روزنامه‌ها و ۵۰۰ هزار واژه از نامه‌ها به دست آورده بود. در مجموع ۴۱ منبع مختلف برای این کتاب استفاده شده‌است. او برای هر واژه مشخص کرده است که چند منبع از این ۴۱ منبع از آن واژه استفاده کرده‌اند.

۳-۱-۳ فلبر (۱۹۸۴)

فلبر می‌گوید که مفاهیم واحدهای تشکیل دهنده تفکر است. او برای نشان دادن پدیده‌ها در ذهن از یک تقسیم بندی چهار گانه کمک می‌گیرد، که رابطه واحدهای زبانی و مفاهیم را نشان می‌دهد.

۱. شیئی یا موجود که با مقوله دستوری اسم نشان داده می‌شود.
 ۲. کیفیت که به وسیله صفت یا اسم بیان می‌شود.
 ۳. عمل که توسط فعل یا اسم بازنمایی می‌شود.
- مکان یا موقعیت که با اسم، قید، حرف اضافه و یا حرف ربط نشان داده می‌شود.

۴-۱-۳ سگر (۱۹۹۰)

سگر بر این باور است که فرایند شناخت، مفاهیم را تولید می‌کند و نهایتاً به طبقه‌بندی چیزها منجر می‌شود. این مفاهیم به شکلی انتزاعی و روشمند صورت‌بندی می‌شود. او در این فرایند، چهار روش را به شرح زیر معرفی می‌کند:

- مفاهیم را می‌توان یک طبقه *class* به‌شمار آورد و آن را مفاهیم طبقه‌ای نامید؛ مثل «قرمز» که به طبقه «رنگها» تعلق دارد یا «اسکی» که در طبقه «ورزشها» قرار می‌گیرد.

- مفاهیم می‌تواند به مقولاتی تقسیم شود که به کمک برخی ویژگیها *property* مشخص می‌شود؛ مثل ویژگی «ساخته شده توسط انسان» که می‌تواند اشیا را به «مصنوعی» و «طبیعی» تقسیم کند.

- مفاهیم را می‌شود به دلیل داشتن رابطه با هم در یک گروه قرار داد؛ مثل «مداد» و

«پاک کن» که از وسایل «لوازم التحریر» هستند.
- مفاهیمی که می‌تواند از تعامل مقولات مختلف با هم به وجود بیاید؛ مثل درونداد، ذخیره، بازیابی و فرایند، نتیجه و واکنش، ترکیب.

۵-۱-۳ فلبوم (۱۹۹۸)

فلبوم با معرفی یک تقسیم‌بندی سلسله مراتبی، یک واحد مفهومی را در رأس یک گروه قرار داد و همه مفاهیمی را که زیر آن قرار می‌گیرند در واقع زیر شمولهای آن به‌شمار آورد. او یک تقسیم‌بندی بیست و پنج‌گانه برای نامها به این شرح پیشنهاد می‌کند: عمل و فعالیت، مصنوع، بدن، ارتباط، احساس و عاطفه، گروه و دسته بندی، انگیزه و علت، پدیده طبیعی، رستنیها و گیاهان، فرایند، رابطه، حالت، زمان، حیوان و جاندار، ویژگی، شناخت و آگاهی، رویداد و اتفاق، خوردنی، مکان، شیء طبیعی، فرد و انسان، مالکیت، کمیت و مقدار، شکل، ماده و یک تقسیم‌بندی پانزده‌گانه برای فعل به این شرح: حرکت، تماس، رقابت، شناخت، ایجاد، آب و هوا، رفتار مربوط به بدن، ایستا، ادراک، ارتباط، تغییر، مصرف، احساس، مالکیت، رفتارها و روابط اجتماعی.

۶-۱-۳ کیپل (۲۰۱۲)

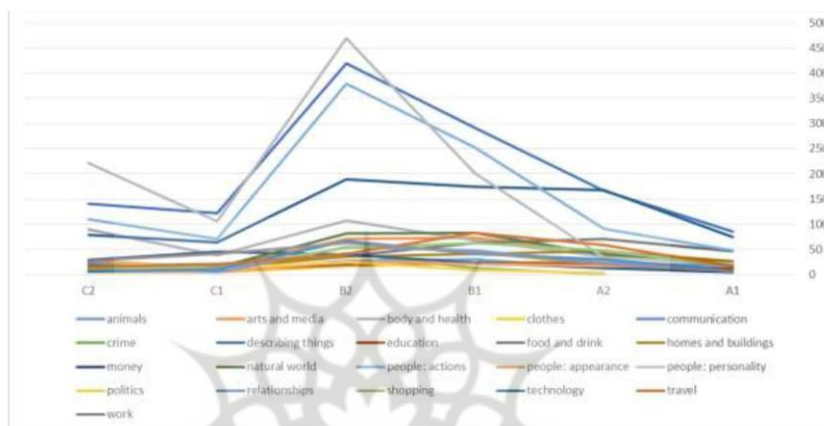
نمایه انگلیسی،^{۲۰} طرحی برای کمک به معلمان آموزش زبان انگلیسی است که در آن بر اساس چارچوب مرجع مشترک زبانهای اروپایی، واژه‌ها و قواعد دستوری زبان انگلیسی را سطح‌بندی کرده است. در این طرح، واژه‌ها علاوه بر دسته‌بندی بر اساس سطح بر اساس موضوع نیز دسته‌بندی شده‌است (کیپل، ۲۰۱۲). این طرح با حمایت دانشگاه کمبریج^{۲۱} و انتشارات این دانشگاه، دانشگاه بدفوردشایر^{۲۲}، بریتیش کانسیل^{۲۳} و انگلیش یوکی^{۲۴} به ثمر رسیده است. داده‌ها بر اساس آنچه زبان‌آموزان می‌دانند (و نه آنچه باید بدانند) تدوین شده (کیپل، ۲۰۱۰) و مبنای آن پیکره مرجع زبان انگلیسی و پیکره زبان‌آموز بوده است.

در این طرح ۲۱ دسته موضوعی برای واژه‌ها وجود دارد:

- | | | |
|------------------|-------------------------|-----------------|
| ۱. حیوانات | ۸. آموزش | ۱۵. مردم: شخصیت |
| ۲. هنرها و رسانه | ۹. غذا و نوشیدنی | ۱۶. سیاست |
| ۳. بدن و سلامتی | ۱۰. خانه‌ها و ساختمانها | ۱۷. ارتباطها |
| ۴. لباسها | ۱۱. پول | ۱۸. خرید |

- | | | |
|----------------|--------------------|------------|
| ۵. ارتباطات | ۱۲. دنیای طبیعت | ۱۹. فناوری |
| ۶. جنایت | ۱۳. مردم: فعالیتها | ۲۰. سفر |
| ۷. توضیح چیزها | ۱۴. مردم: ظاهر | ۲۱. کار |

شکل ۱. نمودار واژه‌های موضوعات مختلف بر اساس سطح زبانی (منبع: نمایه انگلیسی)



۳-۲ پژوهشهای داخلی

۳-۲-۱ بدره‌ای (۱۳۵۰)

شاید اولین کار درباره فهرست واژه‌های پایه زبان فارسی را بتوان کار فریدون بدره‌ای در سال ۱۳۵۰ دانست (این اثر را چاپخانه فردوسی در اسفند ۱۳۵۲ به چاپ رساند).

بدره‌ای در پیوست سوم پژوهش خود با هدف استخراج واژه‌های پایه کودکان دبستانی ایرانی، این فهرستهای موضوعی را نیز ارائه می‌کند:

۱. گیاهان
۲. جانوران و آنچه به آنها مربوط است.
۳. اندامها، حسها، اجزا و کارهای مربوط به آنها
۴. پیشه‌وران: پیشه‌ها و کارها و هر آنچه بدان مربوط است
۵. گورها و عناصر طبیعی و پدیده‌های
۶. جاییها، سازمانها، اداره‌ها و هر آنچه به ساختمان و ورزشها و بازیها
۷. افزارها و هر آنچه به آنها مربوط است.
۸. گاهشماری و واژه‌های مربوط به زمان
۹. دین
۱۰. اعداد
۱۱. پیشه‌وران: پیشه‌ها و کارها و هر آنچه بدان مربوط است
۱۲. رنگها و بویها
۱۳. جاییها، سازمانها، اداره‌ها و هر آنچه به ساختمان و ورزشها و بازیها
۱۴. اندازه‌ها و وزنها

۸. بافتنی‌ها و پوشیدنیها
 ۹. خوراکها و مزه‌ها و آنچه بدانها مربوط است.
 ۱۰. خویشاوندی و خانواده

۲-۲-۳ ایمن (۱۳۵۷)

ایمن در پژوهش فهرست پربسامدترین واژه‌های خردسالان ۹ تا ۱۲ سال، ۲۳۹۸ واژه پربسامد را از مجموع چهار میلیون واژه (با تحلیل پرخواننده‌ترین کتابها و نوشته‌های خردسالان و گفتار آنان) تهیه کرده است

۳-۲-۳ تحریریان (۱۳۷۳)

تحریریان (۱۳۷۳) در مقاله‌نگاهی به فراوانی واژه‌ها و واژه‌های پایه در فارسی به بررسی فهرست واژه‌های پربسامد فارسی پرداخته است. او برای این کار ۵۲۰۰ جمله ۲ تا ۳۵ کلمه‌ای را به شکل تصادفی از منابع زیر جمع‌آوری کرده است:

۱. روزنامه‌های کثیرالانتشار (۱۳۰۰ جمله: ۲۵ درصد)
 ۲. کتابهای علوم انسانی و اجتماعی (۱۳۰۰ جمله، ۲۵ درصد)
 ۳. تدریس کلاسهای دروس معارف اسلامی یا علوم تربیتی (۱۳۰۰ جمله، ۲۵ درصد)
 ۴. دو برنامه «صبح جمعه با شما» و «صدای صبح اصفهان» از رادیو (۶۵۰ جمله، ۱۲.۵ درصد)
- صحبت‌های دوستانه کودکان و نوجوانان (۶۵۰ جمله، ۱۲.۵ درصد)

۴-۲-۳ صنعتی (۱۳۸۷)

خانم صنعتی در رساله دکتری خود ضمن بررسی مفصل انواع تقسیم‌بندیها و فرایندهای ساخت واژه، یک طبقه بندی ۷۳ گانه برای مفاهیم پیشنهاد کرده‌است که فهرستوار در پی می‌آید:

۱. آمیختگی	۲۰. حرکت	۳۸. شباهت	۵۶. مذهب
۲. ابزار	۲۱. حواس پنجگانه	۳۹. شغل	۵۷. مراسم
۳. اجتماع	۲۲. خوراک	۴۰. شکل	۵۸. مرحله
۴. احتمال	۲۳. خویشاوندی	۴۱. شیء	۵۹. مرکزیت
۵. ارتباط	۲۴. دانش	۴۲. ظرف	۶۰. مسیر
۶. ارزش	۲۵. درمان	۴۳. عارضه	۶۱. مقیاس
۷. انتزاع	۲۶. دستگاه	۴۴. عامل	۶۲. مکان

۸. اندام	۲۷. دما	۴۵. عامل مبتلا	۶۳. منشا
۹. ایستایی	۲۸. رتبه	۴۶. علت	۶۴. نسبت
۱۰. بازی	۲۹. رده	۴۷. عمل	۶۵. نظام
۱۱. بافت	۳۰. رطوبت	۴۸. قابلیت	۶۶. نقلیه
۱۲. پدیده طبیعی	۳۱. رقابت	۴۹. کاربرد	۶۷. نوشت افزار
۱۳. پوشش	۳۲. رنگ	۵۰. کمیت	۶۸. نوشته
۱۴. تغییر	۳۳. روش	۵۱. گرایش	۶۹. نهاد اجتماعی
۱۵. توالی	۳۴. زمان	۵۲. ماده	۷۰. نیرو
۱۶. جاندار غیر انسان	۳۵. زیورآلات	۵۳. مالکیت	۷۱. وزن
۱۷. جزء سازنده	۳۶. سطح	۵۴. مجموعه	۷۲. وضعیت روحی - رفتاری
۱۸. جزء و کل	۳۷. سن	۵۵. محصول	۷۳. وضعیت صوری
۱۹. جنسیت			

۵-۲-۳ نعمت‌زاده و همکاران (۱۳۹۰)

این پژوهش، طرحی بزرگ و ملی است که از سال ۱۳۸۰ در سازمان پژوهش و برنامه‌ریزی آموزشی وزارت آموزش و پرورش آغاز شد. هدف این طرح، یافتن واژه‌ها و عباراتی بود که کودکان دبستانی ایرانی می‌شناسند. این واژه‌ها استخراج شده از آزمونهای ادراکی، تولیدی (برای دانش‌آموزان) و آزمونهای مخصوص معلمان از ۲۰ هزار دانش‌آموز و ۷۵۰ معلم ابتدایی در ۱۷۵ مدرسه سراسر کشور است که شامل ۴۹۷ واژه در ۲۴ موضوع مختلف است و شامل موارد زیر می‌شود:

۱. بدن ۲. پوشاک/آرایش ۳. تربیت / تحصیل ۴. جانوران ۵. خانواده ۶. خانه / محیط زندگی ۷. خوراک ۸. دین / اخلاق ۹. رابطه اجتماعی ارتباط ۱۰. رخداد ۱۱. رفتار ۱۲. رنگ ۱۳. مسافرت/ حکل و نقل ۱۴. سلامتی / بهداشت ۱۵. طبیعت ۱۶. کار / مهارت ۱۷. کمیت ۱۸. کیفیت ۱۹. گیاهان ۲۰. مفاهیم ۲۱. مواد / عناصر ۲۲. ورزش / بازی / تفریح ۲۳. وسایل / فناوری ۲۴. هنر / ادبیات.

۶-۲-۳ دادرس (۱۳۹۱)

دادرس در پژوهش خود با هدف ارتقای کیفیت آموزش در نظام آموزشی مدارس، واژه‌های پایه اول دبستان و در مواردی اول تا پنجم را در قالب بیست و چهار طبقه موضوعی آورده است:

۱. بدن	۹. رابطه اجتماعی / ارتباط	۱۷. کمیت
۲. پوشاک / آرایش	۱۰. رخداد	۱۸. کیفیت
۳. تربیت / تحصیل	۱۱. رفتار	۱۹. گیاهان
۴. جانوران	۱۲. رنگ	۲۰. مفاهیم
۵. خانواده	۱۳. سفر / حمل و نقل	۲۱. مواد / عناصر
۶. خانه / محیط زندگی	۱۴. سلامتی / بهداشت	۲۲. ورزش / بازی / تفریح
۷. خوراک	۱۵. طبیعت	۲۳. وسیله / فناوری
۸. دین / اخلاق	۱۶. کار / مهارت	۲۴. هنر / ادبیات

در پژوهش دادرس، یک واژه ممکن است در چند دسته آمده باشد. هم‌چنین برخی از واژه‌های سطوح بالاتر از پایه اول ذکر نشده‌است. این واژه‌ها با در نظر گرفتن واژه‌های تکراری ۵۱۹ مورد است و با حذف موارد تکراری ۲۲۶ واژه یکتا. مشخص است که بیشترین نوع واژه در این پژوهش نام است و تأکید آن بر پایه اول بوده است.

۷-۲-۳ میلر و آقاجانیان - استوارت (۲۰۱۷)

در فرهنگ بسامدی زبان فارسی میلر و آقاجانیان - استوارت ۳۱ طبقه‌بندی معنایی ارائه شده است:

۱. حیوانات	۱۲. نامهای خانوادگی	۲۲. ورزشها
۲. بدن	۱۳. افعال مرکب	۲۳. زمان
۳. لباس	۱۴. نامهای مردان	۲۴. روزهای هفته
۴. رنگها	۱۵. مواد	۲۵. ماههای فارسی
۵. کشورها	۱۶. ملیتها	۲۶. ماههای دری
۶. وسایل الکترونیک	۱۷. طبیعت	۲۷. ماههای اسلامی
۷. احساسات	۱۸. سیایت	۲۸. ماههای فرانسوی
۸. خانواده	۱۹. شغلها	۲۹. حمل و نقل
۹. نامهای زنان	۲۰. دین	۳۰. جنگ
۱۰. غذا	۲۱. افعال ساده	۳۱. آب و هوا
۱۱. سلامتی		

۴. روش اجرای پژوهش

در این بخش با ترکیب هشت فهرست بسامدی زبان فارسی، که از پیکره‌های مختلفی تهیه شده‌است یک فهرست جدید تهیه کردیم که حاصل اجماع آنها است. این فهرستها

گاهی متون گفتاری را هم در برمی‌گیرد؛ گاهی شامل طیف گسترده‌ای از متون می‌شود و گاهی شامل نوع مشخصی از متون (مانند متون مطبوعاتی)؛ اما نکته مهم درباره پژوهش این است که ترکیبی از داده‌های همه این فهرستهای پیکره‌بنیاد را در نظر گرفته است که سبب می‌شود فهرستی جامع از واژه‌های کاربردی در گفتار و نوشتار فارسی‌زبانان ایجاد شود.

جدول زیر، توصیف کلی این هشت پژوهش را نشان می‌دهد:

توصیف پیکره‌های مورد استفاده در پژوهش

شماره	نام پیکره	سال تهیه	نوع متن‌ها	تعداد واژه	فهرست واژه‌ها
۱	حسنی	۱۳۸۴	مجله، کتاب و روزنامه	یک میلیون	۸۴۳۸
۲	محک وب	۱۳۸۷	وب‌سایت‌های ایرانی	< ۱۰۰ میلیون	< ۸۰۰۰
۳	همشهری	۲۰۰۹	روزنامه همشهری	۱۵۰ میلیون	< ۸۰۰۰
۴	نعمت‌زاده و همکاران	۱۳۹۰	آزمون‌ازکودکان دبستانی	-	۴۹۷
۵	بیجن‌خان	۱۳۹۰	منابع گفتاری و نوشتاری	۱۰ میلیون	۱۸۶۸
۶	میلر و آقاجانیان	۲۰۱۷	منابع گفتاری و نوشتاری	۱۵۶ میلیون	۵۰۰۰
۷	صحرايي و همکاران	۱۳۹۶	روزنامه‌ها	۱.۲ میلیون	< ۸۰۰۰
۸	عاصی	۱۳۹۸	منابع نوشتاری	۵۰ میلیون	۲۷۹۱

از میان این هشت پژوهش در دو پژوهش نعمت‌زاده و همکاران و عاصی، عدد بسامد واژه‌ها بیان نشده است (چون هدف این دو پژوهش متفاوت بوده است)؛ اما در دیگر پژوهش‌ها این عدد یا رتبه واژه در میان واژه‌های پربسامد ذکر شده است.

هم‌چنین در سه پژوهش اول (حسنی، همشهری و محک‌وب) اصل واژه‌ها ذکر شده بود و نه ریشه واژه؛ مثلاً ممکن است دو واژه «رفتیم» و «می‌رویم» هر دو در این فهرست‌ها باشد، اما چون هدف در اینجا آموزش زبان است و در آموزش زبان بیش از اصل واژه‌ها با ریشه واژه‌ها (لما) کار داریم، نیاز بود که ابتدا ریشه واژه‌ها در این پژوهش‌ها استخراج شود. برای این کار، واژه‌ها تک‌به‌تک ریشه‌یابی، و سپس بسامد

ریشه‌های مشترک با هم ترکیب شد و فهرستی از ریشه‌های غیرتکراری پربسامد به دست آمد.

در پژوهش برای استخراج فهرست واحد از نرم‌افزار اکسل، و برای رسم نمودارها به تناسب نوع نمودار از اکسل یا گوگل‌شیت استفاده شد. برای دو پژوهش نعمت‌زاده و همکاران و عاصی، کل واژه‌های بیان‌شده در نظر گرفته شد (چون بین این واژه‌ها اولییتی نبود) و برای پژوهشهای دیگر هزار واژه اول در نظر گرفته شد. واژه‌های ذکرشده در این پژوهشها در یک ستون در نرم‌افزار اکسل قرار داده، و سپس واژه‌های تکراری حذف شد.

در ادامه برای افزایش دقت پژوهش و در نظر گرفتن تجربه تدریس مؤلفان کتابهای آموزشی، واژه‌های آموزشی چهار کتاب زیر نیز در نظر گرفته شد:

آموزش کاربردی واژه (صحرايي و همکاران، ۱۳۹۶): این کتاب، کتابی سه جلدی با هدف

آموزش واژه‌های زبان فارسی است که از سطح مقدماتی تا فوق میانی (مطابق استاندارد بنیاد سعدی) را پوشش می‌دهد. در این پژوهش از جلد اول و دوم این کتاب استفاده شده‌است که سطوح مقدماتی و پیش‌میانی را پوشش می‌دهد.

آموزش فارسی (صفارمقدم، ۱۳۸۶): این کتاب چهار جلدی همه سطوح زبان فارسی را پوشش می‌دهد و ما فقط از جلد اول آن استفاده کرده‌ایم که مخصوص سطح مقدماتی است.

مینا (صحرايي و همکاران، ۱۳۹۸): این کتاب اصلی‌ترین کتاب آموزش زبان فارسی بنیاد سعدی است که در چهار جلد از سطح مقدماتی تا فوق‌میانی را پوشش می‌دهد. در این پژوهش از جلد اول این کتاب استفاده شده که با سطح مقدماتی و پیش‌میانی متناسب است.

گام اول (صحرايي و همکاران، ۱۳۹۸): این کتاب اولین کتاب سیر آموزشی بنیاد سعدی است و با هدف آموزش زبان به سطح نوآموز (مطابق استاندارد بنیاد سعدی) طراحی شده است. این سطح مناسب زبان‌آموزانی است که تازه یادگیری زبان فارسی را آغاز کرده‌اند و القبا را هم نمی‌دانند. این کتاب به زبانهای واسطه مختلفی تدوین شده است که محتوای آنها یکی است و زبان واسطه فقط در ترجمه واژه‌ها و عبارات مهم به کار رفته است. در اینجا از نسخه انگلیسی کتاب استفاده کرده‌ایم.

واژه‌های این چهار کتاب اولویت‌بندی نداشت و کل آنها برای پژوهش در نظر گرفته شد. جدول ذیل خلاصه‌ای از اطلاعات این کتابها را نشان می‌دهد که در این پژوهش به کار رفته است:

توصیف کتابهای آموزشی مورد استفاده در این پژوهش

شماره	نام کتاب	تعداد جلد/ جلد استفاده‌شده	سال انتشار	سطوح زبانی	تعداد واژه
۱	گام اول	۱/۱	۱۳۹۸	نوآموز	۳۷۱
۲	مینا	۱/۴	۱۳۹۸	مقدماتی - پیش‌میانی	۸۵۳
۳	آموزش کاربردی واژه	۲/۳	۱۳۹۶	مقدماتی - پیش‌میانی	۱۵۴۱
۴	آموزش فارسی	۱/۴	۱۳۸۶	مقدماتی	۲۹۹۵

با ترکیب واژه‌های کتابهای آموزشی و پیکره‌ها به فهرستی رسیدیم که حداقل در یکی از این پژوهشها یا کتابهای آموزشی ذکر شده بود.

نگارنده براساس سابقه تدریس خود و هم‌چنین هنگام پژوهش به این نتیجه رسیده که طبقه‌بندی معنایی برای تدریس زبان به شکل صفر و یکی نگاه نمی‌شود؛ یعنی نمی‌توان گفت که یک طبقه معنایی (مثلا حیوانات) در یک سطح آموزش داده شود و طبقه معنایی دیگر در سطحی دیگر؛ چرا که در گام اول برای حفظ جذابیت تدریس به زبان‌آموز لازم است موضوعها و در نتیجه طبقه‌های معنایی مختلفی در هر سطح به او تدریس شود و در گام دوم برای اینکه زبان‌آموز خیلی زود بتواند ثمره آموزش خود را مشاهده کند، لازم است با بخشهای مختلف زبان (و از جمله طبقه‌های معنایی مختلف) آشنا شود و یا به آنها تسلط یابد.

بنابراین در هر سطح زبانی می‌توان همه طبقه‌های معنایی را دخیل کرد؛ اما واژه‌هایی از آن طبقه معنایی تدریس شود که با آن سطح در ارتباط است. نمودار واژه‌های هر طبقه معنایی، شکلی زنگوله‌ای دارد (مانند خود واژه‌های پایه که در بخش میانی تقریباً دوبرابر سطح مقدماتی است)؛ یعنی مثلاً برای حوزه معنایی «حیوانات» در سطح میانی بیشتر از دو سطح مقدماتی و پیشرفته واژه آموزش داده می‌شود.

راهکار استخراج حوزه‌های معنایی این است که واژه‌های پایه استخراج شده را از نظر طبقه معنایی طبقه‌بندی، و سپس واژه‌های هر فهرست را ارائه کنیم. به این منظور از پژوهشهای قبلی استفاده کرده‌ایم که وارد طبقه‌بندی معنایی واژه‌ها نیز شده‌اند (پژوهشهای نمایه انگلیسی، میلر و آقاجانیان- استوارت، نعمت‌زاده و همکاران و دادرس) و بر اساس آنها به طبقه‌بندی جدید با الگوی خاص آموزش زبان فارسی فارسی به غیر فارسی زبانان رسیده‌ایم که در ادامه آن را معرفی و ارائه می‌کنیم. برای یافتن طبقه معنایی هر واژه از اکسل و فرمولی استفاده شد که مناسب این منظور نوشته شده است برای مثال برای استخراج موضوع هر واژه طبق فهرست دادرس از فرمول زیر استفاده کردیم:

=IFERROR(INDEX(!دادرسB:B;MATCH(\$A:\$A!دادرسA:A;0));"")

اما علاوه بر این فرمول با توجه به اینکه در پژوهش دادرس، بسیاری از واژه‌ها بیش از یک موضوع داشت برای اینکه همه موضوعها بررسی شود از دو فرمول زیر نیز استفاده شد. این دو فرمول کمک می‌کند که کل موضوعات هر واژه در یک خانه قرار گیرد:

=IF(OR(A2<>A1);B2; C1 & "،" & B2)
=A2<>A3

برای درک بهتر این فرمولها فرض کنید که همه واژه‌ها در ستون اول جدول (A) قرار دارد و به ترتیب حروف الفبا مرتب شده‌است. موضوع هر واژه نیز در ستون دوم (B) قرار گرفته است. با قرار دادن فرمول اول در ستون سوم، این موضوعها با هم ترکیب می‌شود و با قرار دادن فرمول دوم در ستون چهارم، می‌توان آخرین سطر مربوط به هر واژه را نمایش داد؛ به این ترتیب مثلاً می‌توانیم بفهمیم که اسب به سه طبقه موضوعی جانوران، سفر/ حمل و نقل و ورزش/ بازی/ تفریح تعلق گرفته است. برای اینکه فهرستی از واژه‌های یکتا و موضوع هر کدام داشته باشیم، کافی است ستون آخر را فقط بر اساس اندازه‌های درست فیلتر کنیم.

چالش چندموضوعی در دو طبقه‌بندی موضوعی دیگر وجود نداشت؛ اما در مورد نمایه انگلیسی، واژه‌ها به زبان انگلیسی بود. برای ترجمه واژه‌ها از انگلیسی به فارسی از فرمول زیر در گوگل شیت (که مشابه اکسل است) استفاده شد:

GOOGLETRANSLATE(A2,"en","fa")=

در این فرمول، خانه مدنظر (مثلاً A2) از زبان مبدأ (انگلیسی با علامت اختصاری en)

به زبان مقصد (فارسی با علامت اختصاری fa) ترجمه می‌شود. این فرمول کمک می‌کند که کل واژه‌های یک ستون یکجا توسط موتور ترجمه گوگل ترجمه شود. پس از این مراحل، فهرستی از موضوعات مرتبط به هر واژه براساس منابع مورد بررسی داشتیم؛ اما بسیاری از واژه‌ها به هیچ حوزه معنایی متعلق نبود. به دلیل برطرف کردن این ضعف در منابع مورد بررسی، ویرایش دستی را مبنای کار خود قرار دادیم و بر اساس نظریه مقوله بندی و مطالعه نمونه‌های کاربرد هر واژه در پیکره‌های زبانی، که متن کاملشان در دسترس بود (مانند پیکره همشهری و صحرائی و همکاران)، اصلی‌ترین طبقه معنایی هر واژه را مشخص کردیم و برای واژه‌هایی که به هیچ طبقه‌ای تعلق نداشت نیز یک طبقه معنایی مشخص کردیم؛ بدین ترتیب فهرست نهایی شامل طبقه‌های معنایی مختص این پژوهش است و هر واژه در آن در مهمترین طبقه معنایی مرتبط به آن قرار دارد.

۵. تجزیه و تحلیل داده‌ها

۵-۱ طبقه‌بندی معنایی واژه‌های پایه

همان طور که در فصل قبل توضیح داده شد برای استخراج طبقه معنایی واژه‌های پایه این پژوهش از سه پژوهش دادرش، میلرآقاجانیان- استوارت و نمایه انگلیسی استفاده شد. در این سه پژوهش، واژه‌ها طبقه‌بندی معنایی شده‌است. توضیح اینکه در پژوهش دادرش ممکن است یک واژه به چند طبقه معنایی متعلق باشد؛ برای مثال واژه «خانه» به دو طبقه «خانه/ محیط زندگی» و «ورزش/ بازی/ تفریح» تعلق دارد. در اینجا طبق نتایج این سه پژوهش و هم‌چنین بررسی‌های زبانی بر اساس مقوله بندی در ویرایشی جدید برای هر واژه یک طبقه معنایی مشخص شده‌است. در این طبقه‌بندی معنایی در مواردی که یک واژه به چند طبقه معنایی تعلق داشته است، نزدیکترین و مرتبط‌ترین طبقه معنایی را گزارش کرده‌ایم. در این بررسی از نظریه شبکه شعاعی برای تصمیم‌گیری نهایی و مشخص کردن طبقه معنایی واژه‌ها بهره برده‌ایم؛ مثلاً واژه «آب» به سه دسته «رفتار»، «ورزش/ بازی/ تفریح» و «غذا» تعلق داشت؛ اما چون نزدیکترین شعاع معنایی آن به هسته «غذا» بوده‌است، همین مورد را گزارش کرده‌ایم.

۵-۲ فهرست طبقات معنایی به تفکیک واژه‌های هر سطح

تقسیم‌بندی افعال ساده برای نمونه اینجا آورده می‌شود. فهرست کامل واژگان در لینک

زیر در دسترس است:

https://s25.picofile.com/d/8455254826/bde406da-b6ab-4dba-9c61-d9a00a971ef8/word_list.pdf

افعال ساده

واژه‌های سطح میانی

ایستادن، برداشتن، بردن، بودن، توانستن، خواستن، دادن، داشتن، دیدن، ریختن، زدن، شدن، شنیدن، کردن، گذاشتن، گرفتن، گفتن

واژه‌های سطح میانی

افتادن، انداختن، آوردن، بخشیدن، پذیرفتن، پیچیدن، چیدن، ساختن، سوختن، شکستن، شناختن، کشیدن، گذاشتن، نگاه کردن، یافتن

۶. نتیجه‌گیری و پیشنهادها

در این بررسی و بر اساس نظریه مقوله بندی و شبکه شعاعی و با منابعی که از پژوهشهای مبتنی بر طبقه‌بندی معنایی واژه‌ها داشتیم، نشان دادیم که تقریباً از همه حوزه‌های معنایی در سطح مقدماتی، واژه‌هایی وجود دارد. هر چند در برخی حوزه‌های معنایی تعداد واژه‌ها کمتر و در برخی حوزه‌ها تعداد واژه‌ها بیشتر است در نتایج نهایی پژوهش برای همه واژه‌ها اصلی‌ترین طبقه معنایی آنها گزارش شده است.

مهمترین شباهت واژه‌های پایه در سطح مقدماتی و میانی این است که توزیع طبقه‌بندی معنایی هر دوی آنها گوناگون و فراوان است و تقریباً از همه موضوعات در این دو سطح واژه‌هایی حضور دارد. اما تفاوت را می‌توان در نوع واژه‌ها جستجو کرد؛ مثلاً واژه‌های عینی بیشتر در سطح مقدماتی آمده‌است و واژه‌های انتزاعی در سطح میانی یا واژه‌های ساده‌تر و تک‌هجایی در سطح مقدماتی آمده‌است و واژه‌های پیچیده‌تر در سطح میانی.

برای تهیه متون آموزشی کارآمد باید تلاش شود که واژه‌های پرکاربردتر زودتر تدریس شود تا زبان‌آموز زودتر بتواند زبان را در موقعیتهای مختلف به کار گیرد و به این ترتیب انگیزه لازم را برای ادامه یادگیری بیابد؛ اگرچه در این مسیر سادگی واژه نیز اهمیت دارد. در این پژوهش نشان داده شد که برخلاف تصور رایج در همه سطح‌های



زبانی می‌توان واژه‌هایی از طبقه‌های معنایی مختلف به کار گرفت؛ اما در انتخاب واژه‌های آن طبقه معنایی باید دقت کرد.

از دیگر پیشنهادها در این حوزه می‌تواند تهیه فهرستها و فهرست موضوعی و تخصصی باشد؛ مانند برچسب‌گذاریهای موضوعی مثل سینما و ادبیات، اقتصاد، سیاست، ورزش و سرگرمی و ... که در واژگان پربسامد پیکره صحرائی و همکاران انجام شده است. لازم است که دانشگاه‌ها یا مراکزی که دانشجویان غیرایرانی را پذیرش می‌کنند، بررسی کنند و آمار فارسی آموزان را داشته باشند که برای مطالعه در چه حوزه یا ادامه تحصیل در چه رشته‌ای به ایران می‌آیند یا به یادگیری زبان فارسی اقدام می‌کنند و بر اساس آن و با توجه به نیاز مخاطب از همان حوزه‌ها به پیکره‌های موجود بیفزایند و واژگان تخصصی پایه و پربسامد برای آموزش هر رشته یا حوزه را استخراج کنند. کتابها و درسنامه‌ها با وجود چنین فهرستی از واژگان طبقه‌بندی شده، می‌توانند با تکیه بر مبنایی علمی و پژوهشی تدوین شوند و این موضوع به تهیه و تدوین و تولید محتوا محدود نیست بلکه در زمینه طراحی دوره‌های آموزشی، آزمونهای تعیین سطح و بسندگی و ارزشیابی مهارتهای مختلف زبان آموزان کاربرد دارد؛ مثلاً فارسی آموزانی که برای تحصیل در رشته پزشکی یا سینما به ایران آمده‌اند، می‌توانند واژگان پرکاربرد حوزه تخصصی خود را سریعتر بیاموزند و متون تخصصی را در سطح مورد نظر براحتی مطالعه کنند؛ بدین ترتیب می‌توان محتوایی تهیه کرد برای فارسی آموزی در سطح مقدماتی که مطالب سینمایی بخواند یا فارسی آموزی که مطالب مهم اقتصادی را با دانش زبانی در حد میانی مطالعه کند. زبان آموزان می‌توانند با دانش زبانی در حد پایه یا میانی در موضوع تخصصی مورد علاقه خود متونی را بخوانند و در حوزه‌ای که به آن علاقه دارند، ارتباط برقرار کند؛ البته به شرطی که قبلاً محتوای مورد نظر در سطح مناسب تهیه و تدوین شده باشد.

همان‌طور که می‌دانیم و در فرهنگهای واژگانی هم دیده می‌شود، گاهی برای یک مدخل واژگانی تا چندین معنا و حوزه معنایی ذکر می‌شود. مؤثرترین عامل در شناسایی و تشخیص معنای یک واژه، همنشین‌های آن است که در معنای آنها نقشی تعیین کننده دارد؛ به عنوان مثال واژه «ساعت» وقتی در کنار همنشین‌های خود قرار می‌گیرد، معنای مشخص و مورد نظر را می‌یابد؛ مثل «ساعت دیواری»، «ساعت سه» و «ساعت مناسب». بنابراین پیشنهاد دیگر این است که بر اساس پیکره‌های موجود، واژگان پرکاربرد با

بیشترین باهمایی‌های خود فهرست شود؛ بدین ترتیب هم سریعتر معانی متعدد یک واژه درک، و هم عبارتهای واژگانی و باهمایی‌ها آموخته می‌شود.

پی‌نوشتها

- | | |
|----------------------|--------------------------------|
| 1. acquisition | 13. Lyons |
| 2. learning | 14. lexicalization hypothesis |
| 3. Null hypothesis | 15. Goddard & Wierzbicka |
| 4. centrality | 16. Trier |
| 5. prototype | 17. Lexical fields |
| 6. core meaning | 18. Kaedings |
| 7. polysemy | 19. Thorndike |
| 8. homographic | 20. English Vocabulary Profile |
| 9. statistics | 21. Cambridge University |
| 10. machine learning | 22. Bedfordshire University |
| 11. homophony | 23. British Council |
| 12. homonymy | 24. English UK |

فهرست منابع

- بدره‌ای، فریدون؛ (۱۳۵۰) *واژگان نوشتاری کودکان دبستانی ایران*؛ تهران: فرهنگستان زبان ایران.
- خدامی، مریم؛ کرمی نوری، رضا؛ (۱۳۸۱) تهیه ابزار ارزیابی بازشناسی کلمه‌های کتاب فارسی پایه اول ابتدایی و بررسی روایی و پایایی آن در تهران؛ *تازه‌های علوم شناختی*، ش ۱۶، ۱-۹.
- تحریریان، م؛ (۱۳۷۳) *نگاهی به فراوانی واژه‌ها و واژگان پایه در فارسی*؛ مجموعه مقالات دانشگاه علامه طباطبائی، ۸۰، ۴۷-۷۰.
- دادرش، محمد؛ (۱۳۹۱) طبقه‌بندی چندگانه موضوعی واژگان پایه فارسی؛ *فصلنامه مطالعات برنامه درسی ایران*، ش ۲۷، ص ۳۳-۶۶.
- درودی ا؛ برادران‌هاشمی ه؛ و همکاران (۱۳۸۷) *محک استاندارد برای تحقیقات بازایی اطلاعات وب فارسی*؛ تهران: گزارش فنی گروه تحقیقاتی پایگاه داده‌ها دانشگاه تهران، شماره: DBRG-TR-138702.
- صحرائی، ر؛ احمدی‌قادر، ش؛ مرصوص، ف؛ بنفشه، ل؛ (۱۳۹۶) *آموزش کاربردی واژه برای فارسی‌آموزان سطح مقدماتی و پیش‌میان*؛ تهران: دانشگاه علامه طباطبائی.

صحرائی، ر.؛ غریبی، ا.؛ و همکاران (۱۳۹۸ الف) **مینا ۱: برای فارسی‌آموزان سطح A**؛ تهران: بنیاد سعدی، انتشارات آذفا.

صحرائی، ر.؛ مرصوص، ف.؛ ملکلو، د.؛ (۱۳۹۸ پ) **گام اول (در یادگیری زبان فارسی)**؛ تهران: بنیاد سعدی، انتشارات آذفا.

صفارمقدم، ا.؛ (۱۳۸۶) **زبان فارسی: درس‌های پایه**؛ تهران: شورای گسترش زبان و ادبیات فارسی.

صفوی، کوروش؛ (۱۳۸۲) **معنی‌شناسی کاربردی**؛ تهران: نشر همشهری.

صفوی، کوروش؛ (۱۳۹۳) **درآمدی بر معنی‌شناسی**؛ تهران: نشر سوره مهر.

صناعتی، مرضیه؛ (۱۳۸۷) **مفهوم‌شناسی در واژه‌سازی زبان فارسی**؛ [رساله دکتری منتشر نشده]، دانشگاه علامه طباطبائی.

عبدالکریمی، سپیده؛ (۱۳۹۴) چهارچوب‌های نظری در معنی‌شناسی؛ **زبان‌کاوی**، ش ۱: ص ۵۷-۷۹.

غریبی، افسانه؛ صحرائی، رضامراد؛ (۱۳۹۲) کارکردهای زبانی: واکاوی نیازهای زبانی کارکردی فارسی‌آموزان. **پژوهشنامه آموزش زبان فارسی به غیر فارسی‌زبانان**، ش ۳(۵)، ص ۵۳-۳۷. گندمکار، راحله؛ (۱۳۹۷) تحلیل شناخت سخنگویان فارسی‌زبان در قالب مقوله‌بندی؛ مقوله‌بندی رده‌ای یا موضوعی؛ **جستارهای زبانی**، ش ۶۰(۶)؛ ص ۵۱۵-۵۳۸.

نعمت‌زاده، شهین؛ دادرس، م.؛ و همکاران (۱۳۹۰) **واژه‌های پایه فارسی از زبان کودکان ایرانی**؛ تهران: موسسه فرهنگی مدرسه برهان (انتشارات مدرسه).

نویدی، امین؛ عامری، حیات؛ ابوالحسنی چیمه، زهرا؛ (۱۴۰۰) مقایسه واژه‌های آموزشی کتابهای آموزش زبان فارسی با واژه‌های پربسامد؛ **پژوهش‌نامه آموزش زبان فارسی به غیر فارسی‌زبانان**، ش ۱۰(۲)، ص ۲۱۳-۲۳۵.