



## Semantic Classification of Persian Basic Words Instruction

Amin Navidi<sup>1</sup>, Hayat Ameri<sup>2</sup>, Zahra Abolhassani Chime<sup>3</sup>

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.*

---

<sup>1</sup> Ph.D. Student of Department of General Linguistics, Science and Research Branch, Islamic Azad University, Tehran, Iran .Email: [aminnavidi@gmail.com](mailto:aminnavidi@gmail.com)  
ORCID:0000-0001-9442-3378

<sup>2</sup>Corresponding author, Linguistics assistant professor, Tarbiat Modarres University, Tehran, Iran. Email: [h.ameri@modares.ac.ir](mailto:h.ameri@modares.ac.ir) ORCID:0000-0002-2627-3862

<sup>3</sup>Linguistics Assistant Professor, SAMT Organization. Email: [abolhassani@samt.ac.ir](mailto: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.

### **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](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: Allame Tabatabai University. [In Persian]
- Sanaati, M. (2008). *Conceptology in Persian Word Formation*. Unpublished PHD Thesis. Allame Tabatabaei University. [In Persian]
- Tahririan, M. (1994) A look at the frequency of basic words and vocabulary in Persian; *Allameh Tabatabaei University Proceedings*, 80, 47-70.