

A Corpus-based Investigation of the Near-synonymy of “ziyâd and farâvân” Size Adjectives according to Behavioral Profile Approach

Vol. 12, No. 1, Tome 61
pp. 523-550
April & May 2021

Mehrzed Mansouri¹ , Samad Moghimi Sarani² , & Amir Saeid Moloodi³ 

Abstract

Investigations show that in Persian no significant research has been done on lexical relationships of size adjectives. In this paper, near-synonymy relationship of “ziyâd and farâvân” size adjectives and their inflectional forms has been studied based on Behavioral Profile Analysis approach of Divjack and Gries (2006) and the data of Hamshahri 2 corpus (Alahmad and et.al, 2009). Accordingly in this study, first, by using AntConc software (2011) the data related to “ziyâd, ziyâdi, ziyâdtar, ziyâdtarin, ziyâdtari, farâvân, farâvâni, farâvântar, farâvântarin, farâvântari” size adjectives were extracted from the corpus and then were annotated on the basis of the morphological, syntactic and semantic features. The result of behavioral profile approach which is acquired according to hierarchical cluster analysis indicated that the highest level of near-synonymy belongs to “ziyâdi, farâvâni”, “ziyâdtari, farâvântari”, “ziyâdtar, farâvântar”, “ziyâd, farâvân” and “ziyâdtarin, farâvântarin” groups and the lowest level of near-synonymy is between “ziyâdi, farâvâni, ziyâdtari, farâvântari” and “ziyâd, farâvân, ziyâdtarin, farâvântarin, ziyâdtar, farâvântar” groups. Finally, annotation analysis based on Z-scores has specified all meaningful and meaningless ID-tags of size adjectives.

Keywords: Near-synonymy, Size adjectives, Behavioral profile

Received: 18 March 2019

Received in revised form: 22 May 2019

Accepted: 4 July 2019

1. Associate Professor, Department of Foreign Languages and Linguistics, Faculty of Literature and Humanities, University of Shiraz, Shiraz, Iran;
ORCID ID: <https://orcid.org/0000-0001-6510-911X>
2. Corresponding author, PhD Candidate, Department of Foreign Languages and Linguistics, Faculty of Literature and Humanities, University of Shiraz, Shiraz, Iran;
Email: moghimimath@pnu.ac.ir; ORCID ID: <https://orcid.org/0000-0002-0672-7011>
3. Assistant Professor, Department of Foreign Languages and Linguistics, Faculty of Literature and Humanities, University of Shiraz, Shiraz, Iran;
ORCID ID: <https://orcid.org/0000-0003-4682-3160>

1. Introduction

Persian is one of many languages throughout the world which have a class of words whose grammar is clearly different from the others, and can be identified as adjectives. Adjectives, also, can perform some semantic tasks. For example, one major task of adjective is to show the property of things. For example, *Ali's house is big*, this sentence could referred to *Ali's house*, that it has a property of being big.

Although traditional grammarians stressed on the formal features, today, linguists typically classify adjectives according to semantic features. However, one lexical field that has received a lot of attention both in general and in corpus is adjective size, presumably because it includes two pairs of canonical antonyms and many studies have already focused on size adjectives. For example, Charles and Miller (1989) showed that *big* and *little* as well as *large* and *small* tend to occur in one and the same sentence in the Brown corpus with a probability that is much larger than chance would predict and concluded (Gries and Otani, 2010). In Persian, however, no attention has been paid to the size adjectives. Therefore, this is a motivation for us, to investigate near-synonymy of *ziyâd* and *farâvân* size adjectives and their inflectional forms according to behavioral profile analysis approach.

The body of this paper, generally, structured as follow: In section 2, we examine briefly background study of adjectives, in particular, size adjectives, synonymy and the application of behavioral profile. In section 3, we investigate behavioral profile approach. Section 4 deals with methodology of research according to corpus-based data, process of annotation and random data selection by using Kerjcie & Morgan table (1970). In section 5, we analyze statistical information acquired from the process of annotation by using SPSS software. And finally in section 6, we present the results of data.

2. Methodology

Divjak and Gries (2006) identified the following four steps for the behavioral profile approach: 1. The retrieval of (a representative random sample of) all instances of the lemmas from a corpus. 2. Manual analysis and annotation of many properties of each match in the concordance of lemmas. 3. The conversion of data into a co-occurrence table. 4. The evaluation of the table by means of exploratory and other statistical techniques. In this study to do the above steps, we first used Antconc software to retrieve all matches of the lemmas *ziyâd* and *farâvân* plus their inflectional forms from 150 million-word corpus of Hamshahri 2 AleAhmad et al (2009). Then we annotated morphological, syntactic and semantic features to lexical forms in order to identify distributional features. And finally, we utilized SPSS software to evaluate the data and acquire Hierarchical agglomerative cluster charts.

3. Discussion and Results

No considerable study has been done on lexical relationships of size adjectives in Persian and only some grammarians and linguists such as Anvari and Givi (2008), Assi (2006) and etc simply have categorized size adjectives. Consequently, it is difficult to compare the achievements of this study with previous studies on near-synonymy.

In this article, generally, determining near-synonymy is performed according to the observation and analyzing similarities and differences of distributional behaviors of size adjectives and therefore is different from previous researches such as Abdolkarimi (2014, 2016) and etc. Our findings also, support Gries and Otani (2010) assumption on theoretical implication of behavioral profile analysis and most corpus-based works that believe formal differences reflect functional differences. For example, formal tag levels such as active, passive, plural, singular and etc have significant role in

near-synonymy of "ziyâd and farâvân" inflectional forms. The findings of this study, moreover, clear that size adjectives (base form) or along with "tar" inflectional suffix are always dependent of a noun or a causative or predicative verb but size adjectives plus "tarin" superlative suffix can only be dependent of a noun not a verb, because no occurrence is observed for "tarin" as dependent of verbs.

4. Conclusion

In this study, we explored the semantic field of *size* by analyzing the distributional behavior of lexical forms "ziyâd, ziyâdi, ziyâdtar, ziyâdtarin, ziyâdtari, farâvân, farâvâni, farâvântar, farâvântarin, farâvântari" according to behavioral profile analysis and also three variable with 61 ID tags level were annotated to find near-synonymy relationship between size adjectives inflectional forms. The analysis of hierarchical agglomerative cluster charts all ID tags signified that mentioned lexical forms in terms of tag level made sub-groups which have near-synonymy between their own members sub-group. These sub-groups include "ziyâdi, farâvâni", "ziyâdtari, farâvântari", "ziyâdtar, farâvântar", "ziyâdi, farâvân", "ziyâdtarin, farâvântarin", "ziyâdi, farâvâni, ziyâdtari, farâvântari", "ziyâd, farâvân, ziyâdtarin, farâvântarin" and "ziyâdtar, farâvântar, ziyâd, farâvân, ziyâdtarin, farâvântarin".

بررسی پیکره بنیاد هم معنایی تقریبی صفت‌های اندازه «زیاد و فراوان» و صورت‌های تصریفی آن‌ها براساس رویکرد نمای رفتاری

مهرزاد منصوری^۱، صمد مقیمی سارانی^{۲*}، امیرسعید مولودی^۳

۱. دانشیار گروه زبان‌های خارجی و زبان‌شناسی دانشگاه شیراز، شیراز، ایران.

۲. دانشجوی دکتری زبان‌شناسی دانشگاه شیراز، شیراز، ایران.

۳. استادیار گروه زبان‌های خارجی و زبان‌شناسی دانشگاه شیراز، شیراز، ایران.

دریافت: ۹۷/۱۲/۲۷ پذیرش: ۹۸/۰۴/۱۳

چکیده

بررسی‌ها نشان می‌دهد که در زبان فارسی پژوهش مهمی درخصوص روابط واژگانی صفت‌های اندازه انجام نشده است. از همین رو، در این پژوهش رابطه هم معنایی تقریبی صفت‌های اندازه «زیاد و فراوان» و صورت‌های تصریفی آن‌ها بر اساس رویکرد تحلیل نمای رفتاری دیویاک و گرایس (2006) و داده‌های پیکره همشهری ۲ آل‌احمد و همکاران (2009) مورد بررسی قرار گرفته است. در همین راستا، در ابتدا با استفاده از نرم‌افزار آنکانک (2011) داده‌های مرتبط با صفت‌های اندازه «زیاد، زیادی، زیادتر، زیادترین، زیادتری، فراوان، فراوانی، فراوانتر، فراوانترین، فراوانتری» از پیکره استخراج شدند. سپس، براساس ویژگی‌های ساخت‌واژی، نحوی و معنایی برچسب‌دهی شدند. نتایج رویکرد نمای رفتاری که براساس تحلیل خوش‌های سلسله‌مراتبی به دست آمده است نشان می‌دهد که بالاترین سطح هم معنایی تقریبی متعلق به گروه‌های «زیادی، فراوانی»، «زیادتری، فراوانتری»، «زیادتر، فراوانتر»، «زیاد، فراوان» و «زیادترین، فراوانترین، فراوانتری» است و پایین‌ترین سطح هم معنایی تقریبی بین اعضا با پایه غیرهمسان گروه‌های «زیادی، فراوانی، زیادتری، فراوانتری» و «زیاد، فراوان، زیادترین، فراوانترین، زیادتر، فراوانتر» است. درنهایت، تحلیل برچسب‌دهی بر پایه نمرات Z تمام برچسب‌های معنادار و غیرمعنادار صفت‌های اندازه را مشخص ساخت.

واژه‌های کلیدی: هم معنایی تقریبی، صفت‌های اندازه، نمای رفتاری.

۱. مقدمه

بر مبنای پژوهش‌هایی که در زبان فارسی انجام شده، صفت به عنوان یکی از اجزای اصلی کلام درنظر گرفته شده است. برخی از پژوهشگران زبان فارسی صفت را به اقسام بیانی، اشاره، پرسشی، شمارشی و غیره تقسیم‌بندی کرده‌اند. با این حال، امروزه زبان‌شناسان مقولهٔ صفت را براساس ویژگی‌های معنایی، نحوی، نقشی، کاربردشناصی و یا ترکیبی از معیارهای معنایی، نحوی و ساخت‌واژی دسته‌بندی می‌کنند. در همین راستا، یک زمینهٔ واژگانی که چه به صورت اصلی و یا فرعی در دسته‌بندی‌های مرتبط با صفت لحاظ شده مبحث مربوط به صفت‌های اندازهٔ است. به طور کلی، این دسته شامل واژگانی مانند بزرگ، وسیع، پهناور، کوچک، خرد، کم، فراوان، زیاد و غیره می‌شود. برای نمونه در جملهٔ «آهو فراوان‌ترین حیوان این منطقه از کشور است»، لفظ واژگانی «فراوان» یک صفت اندازه است. روابط واژگانی صفت‌های اندازه از جمله موضوعاتی است که در سالیان اخیر مورد توجه معناشناسان پیکرهای بوده است. آن‌ها ضمن استفاده از رویکردهای پیکربندیاد مانند تحلیل نمای رفتاری^۲ اعتقاد دارند که این مدل‌ها به خوبی از عهدهٔ تبیین روابط معنایی مانند هم‌معنایی تقریبی بر می‌آیند، زیرا این رویکردها تحلیلگر را وارد خواهد کرد که همهٔ نمونه‌های به دست آمده از پیکره را با پارامترهای یکسان گذگاری کند تا از صحت اطلاعات واردشده به تحلیل اطمینان حاصل کند و هیچ مسئلهٔ مبهمی را در پشت چارچوب تحلیل قرار ندهد. همچنین، این پژوهشگران ضمن انتقاد از رویکردهای پیکربندیاد قبلی بر این باورند که با استفاده از رویکردی مانند تحلیل نمای رفتاری می‌توان کارایی رویکردهای پیکربندیاد قبلی را از نظر دامنهٔ عناصر مورد بررسی، داده‌ها و روش و پیشینهٔ نظری پژوهش بهبود بخشید. علاوه بر این، آن‌ها بیان می‌کنند که به لحاظ عناصر مورد بررسی، مطالعات گذشته فقط صورت‌های اصلی کلمات را مورد بررسی قرار داده و به صورت‌های تصریفی کلمات توجیهی نشان نداده‌اند و مبنای قضاؤت آن‌ها در مورد کلمات هم‌معنای فقط براساس تحلیل‌های شمی بوده است، نه ویژگی‌های توزیعی این کلمات مانند ساخت‌واژی، نحوی، معنایی و غیره که در بافت‌های زبانی مشاهده می‌شود.

در همین راستا، تا آنجایی که نگارندگان بررسی کرده‌اند در پژوهش‌های مرتبط با هم‌معنایی در زبان فارسی فقط صورت اصلی کلمات مورد بررسی قرار گرفته و از مدل‌های پیکربندیاد استفاده نشده است. بر همین اساس، این پژوهش در نظر دارد که با استفاده از

رویکرد تحلیل نمای رفتاری دیویاک و گرایس^۳ (2006) و داده‌های پیکره همشهری ۲ آل‌احمد و همکاران (2009) رابطه هم‌معنایی تقریبی^۴ صفت‌های اندازه «زیاد و فراوان» و صورت‌های تصریفی آن‌ها را مورد بررسی قرار دهد تا توجیه دقیق‌تری برای هم‌معنایی تقریبی صفت‌های اندازه فوق به دست آورد و میزان هم‌معنایی تقریبی بین آن‌ها را مشخص کند.

۲. پیشینه تحقیق

همان‌طور که در مقدمه اشاره شد دستورنویسان زبان فارسی مانند (انوری و گیوی، ۱۳۸۵) صفت را یکی از اجزای اصلی کلام می‌دانند که معمولاً برای توصیف اسامی مورد استفاده قرار می‌گیرد و حالت، مقدار، شماره و یا یکی از چگونگی‌های اسم را توصیف می‌کند. با این حال، در زبان فارسی با وجود تنوع گسترده مطالعات در حوزه صفت مانند (عبدالکریمی، ۱۳۹۳ و ۱۳۹۴)، پژوهش قابل توجهی درخصوص صفت‌های اندازه انجام نشده است و فقط برخی از پژوهشگران مانند عاصی و همکاران (۱۳۸۵) صفت‌های اندازه را به واسطه ویژگی‌های معنایی‌شان در طبقه مشخصی قرار داده‌اند. این در حالی است که در زبان انگلیسی مقوله صفت‌های اندازه مورد توجه پژوهشگران فراوانی مانند دیز^۵ (1964)، چارلز و میلر^۶ (1989)، جاستسون و کاتر^۷ (1991)، لی^۸ (۱۹۹۴) بایبر^۹ و همکاران (1998)، دیکسون^{۱۰} (2004)، جونز^{۱۱} و همکاران (2007)، گرایس و اُتانی^{۱۲} (2010) و غیره بوده است. برای نمونه، دیز (1964) با استفاده از یک مدل تداعی واژه نشان داد که صفت‌های اندازه «small، little» و «big， large» به لحاظ معنایی بسیار به یکدیگر شبیه هستند. علاوه بر این، وی جفت‌واژه‌هایی مانند «large， small» و «big， little» را بهترین و طبیعی‌ترین مثال برای نشان‌دادن رابطه واژگانی تضاد معرفی کرده است. بایبر و همکاران (1998) بیان داشتند که در مقایسه با «big， great»، صفت اندازه «large» با فراوانی ۴۰۸ وقوع در هر یک میلیون کلمه بسامد بیشتری در زبان انگلیسی دارد و برای هر دو جنس مؤنث و مذکور برای ارجاع به واژگانی مورد استفاده قرار می‌گیرد که به اندازه فیزیکی اشاره دارند. گرایس و اُتانی (2010) بیان کردند که براساس داده‌های پیکره‌بنیاد، صفت اندازه «tiny» نسبت به «tiniest» و «less» نسبت به «smallest» بالاترین سطح هم‌معنایی تقریبی را دارد و همچنین صفت‌های اندازه «large، small» و «big， little» بالاترین سطح تضاد را نسبت به جفت خود نشان می‌دهند. در همین زمینه، بررسی‌ها در زبان فارسی نشان داده است که از رویکرد نمای



رفتاری فقط در دو اثر گلشایی (2016) برای بررسی چندمعنایی فعل سبک «خوردن» و کاویانی (۱۳۹۶) برای بررسی چندمعنایی پیشوندهای نفی «بی- ضد- غیر- و نا» استفاده شده است. در طرف مقابل، در سایر زبان‌ها رویکرد نمای رفتاری مورد توجه محققانی همچون دیویاک و گرایس (2006)، گرایس و دیویاک (2009)، گرایس (2010)، گرایس و اُتانی (2010)، موزر^{۱۳} (2014)، کوزنتسووا^{۱۴} (2015) و غیره بوده است. برای نمونه، گرایس و دیویاک (2009) ضمن تبیین چارچوب نظری رویکرد نمای رفتاری و ارائه مثال‌هایی از مطالعات صورت‌گرفته با این رویکرد بیان داشتند که معناشناسان شناختی برخورد یکسان و عادلانه‌ای با انواع روابط واژگانی نداشته‌اند و همواره به چندمعنایی توجه بیشتری نشان داده‌اند، اما در رویکرد نمای رفتاری سعی شده است به روابط واژگانی نگاه عادلانه‌تری شود. موزر (2014) با استفاده از رویکرد نمای رفتاری بیان کرده است که صفت‌های اندازه «big, large» به لحاظ ویژگی‌های نحوی، هم‌معنایی تقریبی بیشتری نسبت به ویژگی‌های معنایی از خود نشان می‌دهند.

۳. چارچوب نظری تحقیق

چارچوب نظری این تحقیق بر مبنای رویکرد تحلیل نمای رفتاری است. در این رویکرد دو مفهوم اساسی به نام برچسب‌های شناسایی^{۱۵} و نمای رفتاری وجود دارد. برچسب‌های شناسایی بهوسیله آنکینز^{۱۶} (1987) و نمای رفتاری بهوسیله هنکس^{۱۷} (1996) معرفی شد. با وجود این، هیچ یک از آن‌ها تحلیل جامعی از اهمیت برچسب‌های شناسایی و نمای رفتاری ارائه نکرده‌اند، تا اینکه دیویاک و گرایس (2006) در رویکرد تحلیل نمای رفتاری از دو مفهوم فوق برای شکل‌دهی ساختار مدل خود استفاده کردند. این رویکرد برای بررسی روابط واژگانی در چارچوب معناشناسی شناختی طراحی شده است. به عقیده دیویاک و گرایس (2006) رویکرد تحلیل نمای رفتاری از چهار مرحله زیر تشکیل می‌شود:

در مرحله اول، صورت‌های مختلف تصریفی سروژدها^{۱۸} از یک پیکره زبانی بازیابی می‌شوند. این داده‌ها در قالب جمله و یا پاره‌گفتار هستند. مرحله دوم، شامل برچسب‌دهی^{۱۹} به داده‌ها و تحلیل آن‌ها می‌شود. در این شیوه به تمامی صورت‌های مختلف تصریفی سروژدها با توجه به مقوله دستوری آن‌ها، انواع برچسب‌های نحوی، معنایی، ساخت‌واژی و غیره اختصاص می‌یابد. البته هیچ محدودیتی در انتخاب برچسب‌ها وجود ندارد و پژوهشگر می‌تواند مناسب با متغیر مورد تحقیق برچسب‌های مورد نظر خود را انتخاب کند. در مرحله سوم، از اطلاعات

جدول‌های هموقوعی^{۲۰} برای تحلیل کمی داده‌ها استفاده می‌شود. مرحله سوم شامل دو زیرمرحله است: ۱) در ابتدا، اطلاعات موجود در جدول‌های هموقوعی به صورت بسامد مطلق آورده شود. ۲) سپس، بسامد مطلق به بسامد نسبی تغییر پیدا می‌کند. در مرحله چهارم، اطلاعات بسامدی جدول‌های هموقوعی با استفاده از تحلیل خوش‌ای سلسه‌مراتبی تجمعی^{۲۱} و آزمون آماری Z مورد بررسی قرار می‌گیرد. این دو راهکار آماری کمک می‌کند تا میزان (عدم) شباهت برچسب‌ها و متغیرهای پژوهش مشخص شود (Gries & Divjak, 2009).

۱-۳. برچسب‌های شناسایی

در مرحله دوم رویکرد نمای رفتاری از برچسب‌های شناسایی استفاده می‌شود. در این بخش ابعاد نظری این برچسب‌ها مورد بررسی قرار می‌گیرد. بیشتر برچسب‌های مورد استفاده در این مقاله براساس گرایس (2010)، گرایس آتانی (2010) و موذر (2014) انتخاب شده‌اند.

۱-۱-۳. برچسب‌های ساخت‌واژی

مبنای انتخاب برچسب‌های ساخت‌واژی این پژوهش بر پایه رفتار موصوف و فعل جمله‌ای بوده که صورت‌های واژگانی در آن حضور داشته‌اند. در جدول ۱، تعداد ۹ برچسب ساخت‌واژی برای این پژوهش تعیین شده است که در مجموع دارای ۲۹ سطح برچسب هستند.

جدول ۱: برچسب‌های ساخت‌واژی

Table 1: Morphological ID-tags

نوع برچسب	برچسب	سطوح برچسب
ساخت‌واژی	زمان	حال، آینده، گذشته
	جهت	معلوم، مجہول
	شمار	مفرد، جمع
	شخص	اول شخص، دوم شخص، سوم شخص
	درجۀ صفت	مشدد، وافر، ملایم
	درونداد	اسم، صفت، فعل، بن مضارع، بن ماضی، قید
	برونداد	اسم مرکب، صفت مرکب، اسم مصدر
	نوع نشانه تصrifی	تر تفضیلی، ترین عالی، واژه‌بست ربطی، ی نکره
	نوع کلمه به لحاظ ساخت	ساده، اشتراقی، مرکب



در جدول ۱، متغیر زمان دارای سه سطح برچسب «حال، گذشته و آینده» است. متغیر «جهت» دارای دو سطح برچسب معلوم و مجھول است. «معلوم» مانند، فرزانه کتاب را آورد. و «مجھول» مانند، کتاب آورده شد. متغیر «شخص» براساس شناسه فعل جمله و یا پاره‌گفتاری که صفات‌های اندازه در آن حضور دارند، مشخص می‌شود و شامل اول شخص، دوم شخص و سوم شخص می‌شود. متغیر «شمار»، مربوط به وضعیت مفرد یا جمع‌بودن موصوف صفات‌های اندازه می‌شود. برای نمونه در عبارت «تلاش/ تلاش‌های فراوان»، که موصوف (تلاش) به ترتیب مفرد و جمع‌بودن صفت (فراوان) را نشان می‌دهد. متغیر «درجه صفت»، توصیف‌کننده قبیل است که قبل از صفات‌های اندازه قرار می‌گیرد و دارای درجات مشدد، وافر و ملایم است. برای نمونه، داده‌های بسیار فراوانی مشاهده کردیم (مشدد)، غذا زیاد فراوان نیست (وافر)، احمد نسبتاً زیاد دوید (ملایم) (ماهوتیان، ۱۳۷۸). متغیر «درونداد»، به مقوله پایه جزء ترکیب‌شونده با صفات‌های اندازه در ترکیب‌های مرکب اشاره دارد. برای نمونه، صفت زیاده‌خواه از ترکیب زیاد+ه+بن مضارع (خواستن) ساخته شده است. متغیر «برونداد»، به مقوله آن لفظ واژگانی اشاره دارد که از ترکیب صفت اندازه (زیاد و فراوان) با جزء ترکیب‌شونده به دست آید. برای نمونه، واژه مرکب زیاده‌خواه یک صفت مرکب است. متغیر «نوع ساخت» دارای سه سطح برچسب ساده (مانند فراوان)، اشتاقاچی (مانند فراوانی) و مرکب (مانند زیاده‌خواه) است.

۲-۱-۳. برچسب‌های نحوی

مبناًی انتخاب برچسب‌های نحوی این پژوهش، براساس ساختار صوری جمله و پاره‌گفتاری است که صفات‌های اندازه در آن‌ها حضور دارند. در جدول ۲، برچسب‌های نحوی دارای ۵ متغیر هستند که در مجموع، شامل ۱۳ سطح برچسب می‌شود.

جدول ۲: برچسب‌های نحوی

Table 2: Syntactic ID-tags

نوع برچسب	برچسب یا متغیر	سطوح برچسب یا متغیر
نوع جمله	خبری، پرسشی، امری	
	پایه، پیرو	نوع بند

نوع برچسب	برچسب یا متغیر	سطوح برچسب یا متغیر
نحوی	کارکرد صفت	اسنادی، وصفی
	نقش نحوی صفت	مسند، تمیز
	مقوله نحوی هسته	فاعل، مفعول مستقیم، مفعول غیرمستقیم، فعل

در جدول ۲، متغیر «نوع جمله» شامل سه سطح جمله خبری، پرسشی و امری می‌شود. متغیر «نوع بند» شامل دو سطح پایه و پیرو است. متغیر بعدی «کارکرد صفت» است که شامل دو سطح اسنادی و وصفی می‌شود. کارکرد اسنادی صفت زمانی رخ می‌دهد که صفت بدون ساخت اضافه و پس از اسم بیاید و جمله‌ای که صفت در آن حضور دارد افعالی نظری است، بود، شد و غیره داشته باشد. مانند، سال پیش برداشت محصول فراوان بود. کارکرد وصفی زمانی به‌وقوع می‌پیوندد که صفت با ساخت اضافه و پس از اسم بیاید. مانند، همهٔ تیم‌ها تلاش زیادی کردند (همان). همچنین، بعد از بررسی داده‌ها در پیکرهٔ همشهری ۲، مشخص شد که می‌توان برای صفت‌های «زیاد و فراوان» نقش‌های نحوی مانند مسند و تمیز لحاظ کرد. درنهایت، متغیر «مقوله نحوی هسته» در چهار سطح برچسب فاعل، مفعول مستقیم، مفعول غیرمستقیم و فعل برچسب‌گذاری شد تا وضعیت هستهٔ صفت‌های اندازه «زیاد و فراوان» مورد بررسی قرار گیرد.

۳-۱-۳. برچسب‌های معنایی

در جدول ۳، تعداد ۵ برچسب برای بررسی روابط معنایی صفت‌های اندازه انتخاب شده است که درمجموع، دارای ۱۹ سطح برچسب است.

جدول ۳: برچسب‌های معنایی

Table 3: Semantic ID-tags

نوع برچسب	برچسب یا متغیر	سطوح برچسب یا متغیر
معنایی	نوع موجودیت (شیء)	انتزاعی، عینی
	موجودیت (شیء)	جاندار (انسان و غیرانسان)، بی‌جان
	توصیف‌گر اندازه	استعاری، تحت‌اللفظی

نوع برچسب	برچسب یا متغیر	سطح برچسب یا متغیر
مفهوم معنایی موصوف	اسم عام، اسم مکان، اسم خویشاوندی، اسم شغل، اسم زمان، اسم گزاره‌ای، اسم جمع	اسم عارف
نمود واژگانی افعال	کنشی ^{۲۲} ، استا ^{۲۳} ، لحظه‌ای ^{۲۴} ، پایا ^{۲۵}	

ولین برچسب جدول ۳ «نوع موجودیت» است که شامل دو سطح برچسب انتزاعی و عینی است. متغیر انتزاعی، به مفاهیمی مانند عشق، تنفس، دوستی، حسادت و غیره گفته می‌شود که در جهان بیرون غیرملموس هستند. در مقابل، متغیر عینی به موجودیت‌هایی مانند میز، لیوان، کاغذ و غیره گفته می‌شود که در جهان بیرون قابل‌لمس هستند. متغیر بعدی «موجودیت» است که به دو سطح کلی جاندار و بی‌جان تقسیم‌بندی می‌شود. سطح برچسب جانداری، خود دارای دو سطح انسان و غیرانسان است و دلالت بر مجموعه انسان‌ها و حیوانات و هر چیز جانداری دارد. برای نمونه: زن، چیتا، معلم، درخت و غیره. البته جانداربودن و یا نبودن تا حدی فرهنگ-محور است؛ یعنی بستگی به فرهنگ هر کشور هم دارد(Brown & Miller, 2013). همچنین، برچسب غیرجاندار بر موجوداتی که ویژگی حیات ندارند مانند میز، صندلی، کفش و غیره دلالت دارد. متغیر بعدی «توصیفگر اندازه» است که براساس مبانی نظریه استعاره مفهومی لیکاف و جانسون^{۲۶} (1980)، لیکاف (1993، 1987)، لیکاف و ترنر^{۲۷} (1989) برای آن سطوح برچسب‌های استعاری و تحت‌اللفظی مشخص شده است. برای متغیر بعدی یعنی «مفهوم معنایی موصوف» سطوح برچسب‌گذاری فراوانی می‌توان درنظر گرفت. حال با توجه به داده‌های این تحقیق سطوح اسامی عام، مکان، خویشاوندی، شغل، زمان، گزاره‌ای و جمع درنظر گرفته شده است. درنهایت، آخرین متغیر «نمود واژگانی افعال» دارای چهار طبقه اصلی کنشی، لحظه‌ای، پایا و ایستاست که در مجموع، این چهار طبقه شامل ده طبقهٔ فرعی می‌شوند (Croft, 2012).

۴. روش بررسی داده‌ها

در این پژوهش، از پیکره ۱۵۰ میلیون کلمه‌ای همشهری ۲ آل‌احمد و همکاران (2009) برای جمع‌آوری داده‌ها استفاده شده است. متن‌های موجود در این پیکره شامل انواع ژانرهای خبری، سیاسی، اقتصادی، ورزشی و غیره می‌شود. همچنین، با درنظر گرفتن محدوده زمانی آسناد پیکره همشهری، وضعیت تحلیل داده‌ها در این پژوهش بر مبنای استفاده از داده‌های

همزمانی است و شامل داده‌های درزمانی نمی‌شود.

در این تحقیق، در ابتدا داده‌های موجود در پیکره همشهری ۲ با استفاده از نرم‌افزار آنکانک^{۲۸} (2011) استخراج شد. آنکانک یک نرم‌افزار واژه‌نمای^{۲۹} نسل سوم است که برای استخراج و پردازش داده‌ها از پیکره، مورد استفاده قرار می‌گیرد (McEnery & Hardie, 2012). در ادامه، برای انتخاب تصادفی داده‌ها از جدول کرجسی و مورگان^{۳۰} (1970) استفاده شد. در این جدول براساس حجم جمعیت، حجم نمونه مناسب برای هر صفت اندازه مشخص شد. با احتساب این روش نمونه‌گیری، حجم نمونه داده‌های صورت‌های واژگانی «زیاد، زیادتر، زیادترین، زیادتری، فراوان، فراوانی، فراوانتر، فراوانترین، فراوان‌تری» به‌شرح زیر درنظر گرفته شد:

جدول ۴: حجم جمعیت و حجم نمونه پژوهش

Table 4: The research population and sample size

صفت اندازه	حجم نمونه	صفت اندازه	حجم نمونه	صفت اندازه	حجم نمونه	صفت اندازه
زیاد	۳۷۵	فراآن	۳۸۰	زیاد	۳۲۲۸۵	۱۳۰۳۴
زیادی	۳۷۵	فراآنی	۳۸۲	زیادی	۶۱۸۲۸	۱۲۵۰۱
زیادتر	۲۸	فراآن‌تر	۲۵۶	زیادتر	۷۰۰	۳۰
زیادتری	۱۹	فراآن‌تری	۱۳۶	زیادتری	۲۱۵	۲۰
زیادترین	۶	فراآن‌ترین	۶	زیادترین	۶	۶
جمع کل	۸۰۳	جمع کل	۱۱۶۰	جمع کل	۹۰۰۳۹	۲۵۶۴۱

در رابطه با انتخاب صورت‌های واژگانی مهم است که بیان شود «سی» موجود در صورت‌های «زیادی و فراوانی»، سی نکره است که آن را از موصوف خود گرفته است. برای نمونه، «احمد تلاش زیادی کرد» که در اصل «احمد تلاشی زیاد کرد» بوده است. در مرحله بعدی، داده‌های انتخاب شده براساس ویژگی‌های نحوی، معنایی و ساخت‌واثری برچسب‌دهی شدند. در رابطه با نحوه برچسب‌دهی، دیویاک (2010) بیان می‌کند که در زبان‌شناسی پیکره‌ای برچسب‌دهی دستی^{۳۱} بهترین شیوه برچسب‌دهی است. در این شیوه، انتخاب، اعمال و تحلیل برچسب‌ها به‌وسیله پژوهشگر انجام می‌شود؛ یعنی پژوهشگر متناسب با متغیر پژوهش، برچسب‌هایی را در سطح اجزای کلام و یا در سطح جمله انتخاب می‌کند. سپس، آن

برچسب‌ها را بر روی داده‌ها اعمال می‌کند و درنهایت، داده‌ها را با توجه به برچسب‌های زده‌شده مورد تجزیه و تحلیل قرار می‌دهد. در این پژوهش، برای تحلیل خوش‌های سلسله‌مراتبی و آزمون آماری Z از نرم‌افزار آماری اس.پی.اس. استفاده شد.

۵. تحلیل داده‌ها

در تحلیل داده‌های این پژوهش، وضعیت هم‌معنایی تقریبی صفت‌های اندازه «زیاد، فراوان» و صورت‌های تصریفی آن‌ها بر مبنای برچسب‌های ساخت‌واژی، نحوی و معنایی مورد بررسی قرار می‌گیرد.

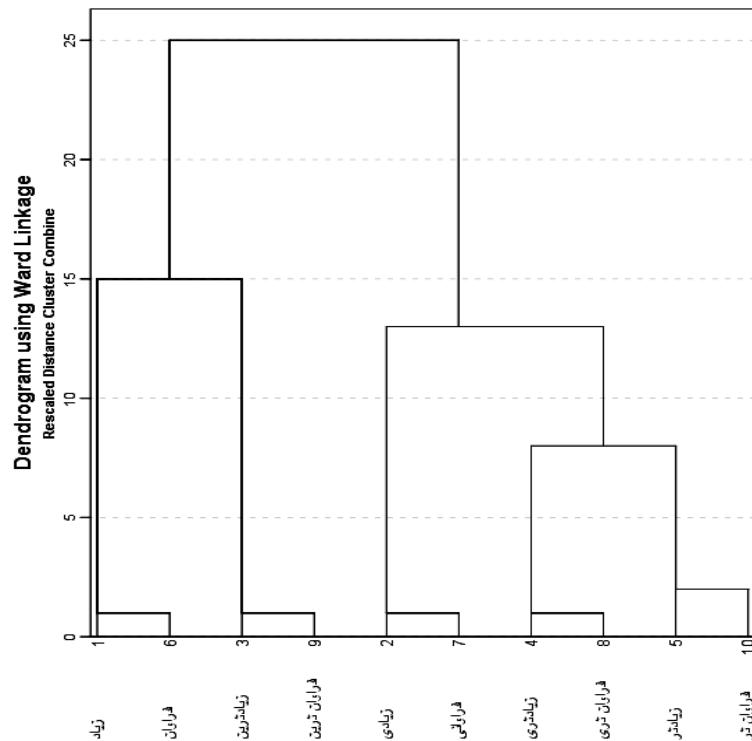
۱-۵. تحلیل برچسب‌های ساخت‌واژی

نمودار ۱، تحلیل خوش‌های سلسله‌مراتبی برچسب‌های ساخت‌واژی را نشان می‌دهد. در این نمودار، علاوه بر سطوح برچسب‌ها که در بخش پایینی نمودار مشخص شده است، دو دسته اعداد در محورهای افقی و عمودی قرار دارند که اعداد نمودار افقی شماره خوش و اعداد نمودار عمودی میزان ارتفاع را مشخص می‌کنند. به‌طور کلی در این پژوهش، رابطه هم‌معنایی تقریبی به دو صورت زیر سنجدیده می‌شود:

(۱) هر نمودار سلسله‌مراتبی گروه‌هایی را تشکیل می‌دهد که اعضای هر گروه با یکدیگر رابطه هم‌معنایی تقریبی دارند. برای نمونه، نمودار ۱ از گروه‌های «زیاد/۱، فراوان/۶»، «زیادترین/۳، فراوان‌ترین/۹»، «زیادی/۲، فراوانی/۷»، «زیادتری/۴، فراوان‌تری/۸» و «زیادتر/۵، فراوان‌تر/۱۰» تشکیل شده است.

(۲) در نمودارهای سلسله‌مراتبی، هر چه ادغام خوش‌های در ارتفاع کمتری صورت بگیرد میزان هم‌معنایی تقریبی (شباهت) بین آن‌ها بیشتر است و بالعکس. برای نمونه در نمودار ۱، نقطه اتصال خوش‌های گروه‌های «زیادتر/۵، فراوان‌تر/۱۰» و «زیادتری/۴، فراوان‌تری/۸» در ارتفاع بین (۵ تا ۱۰) است. در حالی که، نقطه اتصال خوش گروه «زیادی/۲، فراوانی/۷» با اعضای دو گروه قبلی در ارتفاع بین (۱۰ تا ۱۵) است و این بدان معناست که به‌دلیل ارتفاع کمتر، میزان هم‌معنایی تقریبی بین اعضا با پایه غیرهمسان گروه‌های «زیادتر/۵، فراوان‌تر/۱۰» و «زیادتری/۴، فراوان‌تری/۸» بیشتر از میزان هم‌معنایی تقریبی بین گروه‌های

«زیادی/۲، فراوانی/۷» و «زیادتر/۵، فراوان تر/۱۰، زیادتری/۴، فراوان تری/۸» است. همچنین در سمت چپ نمودار ۱، خوشة گروه «زیاد/۱، فراوان/۶» و «زیادترین/۳، فراوان ترین/۹» در ارتفاع ۱۵ با یکدیگر برخورد می کنند که نشان از هم معنایی تقریبی بسیار کم بین آن هاست. علاوه بر این، صفت های اندازه «زیاد/۱، فراوان/۶، زیادترین/۳، فراوان ترین/۹» پایین ترین سطح هم معنایی تقریبی را نسبت به صورت های «زیادی/۲، فراوانی/۷، زیادتری/۴، فراوان تری/۸، زیادتر/۵، فراوان تر/۱۰» نشان می دهند، زیرا خوشه های آن ها در بالاترین ارتفاع نمودار یعنی ۲۵ با یکدیگر برخورد می کنند.



نمودار ۱: ساختار سلسله مراتبی تجمعی برچسب های ساخت واژی

Chart 1: Hierarchical agglomerative structure of Morphological ID-tags

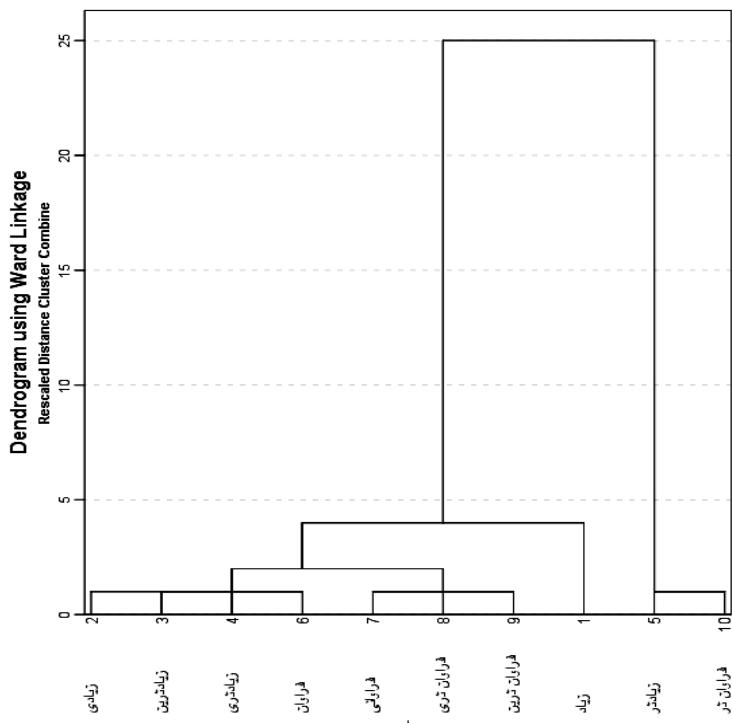
داده‌های آماری نحوه خوشه‌بندی نمودار ۱ را تأیید می‌کند. برای نمونه، تمام گروه‌های این نمودار در سطوح برچسب‌های «حال، گذشته، آینده، معلوم، مجہول، اول شخص، دوم شخص، سوم شخص» بسامد نسبی تقریباً مشابهی را نشان می‌دهند. با این وجود، برخی از صورت‌های واژگانی در بعضی از سطوح بسامد‌های نسبتاً متفاوتی را نشان می‌دهند که به‌دلیل نداشتن بسامد قابل توجه تأثیر مهمی را در نحوه خوشه‌بندی نمودار نمی‌گذارند. برای نمونه، صفت اندازه «زیاد» در سطوح برچسب‌های «بن ماضی و بن مضارع» به‌ترتیب بسامد ۲ درصد و ۴ درصد را نشان می‌دهد. در حالی است که سایر صورت‌های تصویری این پژوهش، هیچ وقوعی را برای ساخت واژه‌های مرکب در پیکره همشهری نداشته‌اند. همچنین، با بررسی همه صورت‌های واژگانی مشخص شد که آن‌ها ترجیح می‌دهند در جمله‌ای حضور پیدا کنند که بیشتر سطوح «حال، معلوم، مفرد، سوم شخص، مشدد» را داشته باشد و این بدان معناست که آن‌ها کمتر تمایل دارند در جملاتی شرکت کنند که سطوح برچسب‌های «گذشته، آینده، مجہول، جمع، اول شخص، دوم شخص، وافر و ملایم» را داشته باشد.

از طرف دیگر، در این پژوهش برای اطمینان از (عدم) شباهت سطوح برچسب‌های ساخت‌واژی از آزمون Z استفاده شده است تا تفاوت‌های بین صورت‌های تصویری به‌درستی مشخص شود. در این آزمون، نمره استاندارد بالاتر از ۱.۹۶ و پایین‌تر از ۰.۵ در سطح ۰.۰۵ معنادار به‌حساب آورده می‌شود. بر این اساس، آزمون Z برچسب‌های ساخت‌واژی نشان می‌دهد که صفت‌های اندازه «زیاد و فراوان» و صورت‌های تصویری آن‌ها در سطوح برچسب‌های «مفرد، جمع، اول شخص، دوم شخص، سوم شخص، مشدد، ملایم، اسم مصدر، تر تفضیلی، ترین عالی و ساده» دارای نمره استانداردی بین ۱.۹۵ تا ۱.۴۸ هستند. بنابراین، بر پایه آزمون Z می‌توان بیان داشت که برچسب‌های فوق نه تنها تفاوت معناداری را نشان نمی‌دهند، بلکه عامل افزایش شباهت‌های توزیعی بین صفت‌های اندازه این پژوهش هستند. همچنین، برخی از سطوح برچسب‌ها نمرات معناداری را نشان می‌دهند. برای نمونه، صفت اندازه «زیاد» در سطوح «بن مضارع، اسم مرکب، صفت مرکب و مرکب» نمره استاندارد (۲.۸۴) را به‌دست آورد که این نمره تفاوت معناداری را بین «زیاد» با صورت‌های «زیادتر، فراوان‌تر، زیادتر، فراوان‌تری» ایجاد کرد و موجب جدایی بیشتر این صورت‌های تصویری در نمودار شده است. با وجود این، در نمودار ۱، صفت اندازه «زیاد» به‌دلیل بسامد‌های

مشابه در بیشتر سطوح دیگر برچسب‌ها، در کنار صفت اندازه «فراوان» قرار گرفته است.

۲.۵. تحلیل برچسب‌های نحوی

تحلیل برچسب‌های نحوی در نمودار ۲ نشان داده شده است. این نمودار، به ترتیب از چپ به راست شامل گروه‌های «زیادی/۲، زیادترین/۳، زیادتری/۴، فراوان/۶»، «فراوانی/۷، فراوانانتری/۸، فراوانانترین/۹» و «زیادتر/۵، فراوانانتر/۱۰» است. در گروه‌های «زیادی/۲، زیادترین/۳، زیادتری/۴، فراوان/۶» و «زیادتر/۵، فراوانانتر/۱۰» بین اعضا با پایه غیریکسان هم‌معنایی تقریبی و با پایه یکسان شباهت توزیعی وجود دارد. همچنین در گروه «فراوانی/۷، فراوانانتری/۸، فراوانانترین/۹» فقط رابطه شباهت توزیعی بین اعضا مشاهده می‌شود. از طرف دیگر، نمودار ۲ نشان می‌دهد که خوشه برآمده از صفت اندازه «زیاد/۱» و خوشه‌های گروه‌های «زیادی/۲، زیادترین/۳، زیادتری/۴، فراوان/۶» و «فراوانی/۷، فراوانانتری/۸، فراوانانترین/۹» در ارتفاع بین (۰ تا ۵) با یکدیگر برخورد می‌کنند که بیانگر هم‌معنایی تقریبی بسیار زیاد بین پایه‌های غیریکسان این صفت‌های اندازه است. این در حالی است که خوشه برآمده از صفت‌های اندازه «زیاد/۱، زیادی/۲، زیادترین/۳، زیادتری/۴، فراوان/۶، فراوانی/۷، فراوانانتری/۸، فراوانانترین/۹» در ارتفاع (۲۵) با خوشه گروه «زیادتر/۵، فراوانانتر/۱۰» ترکیب می‌شود که دلالت بر پایین‌ترین سطح هم‌معنایی تقریبی بین پایه‌های غیریکسان آن‌هاست.



نمودار ۲: ساختار سلسله‌مراتبی تجمعی برچسب‌های نحوی

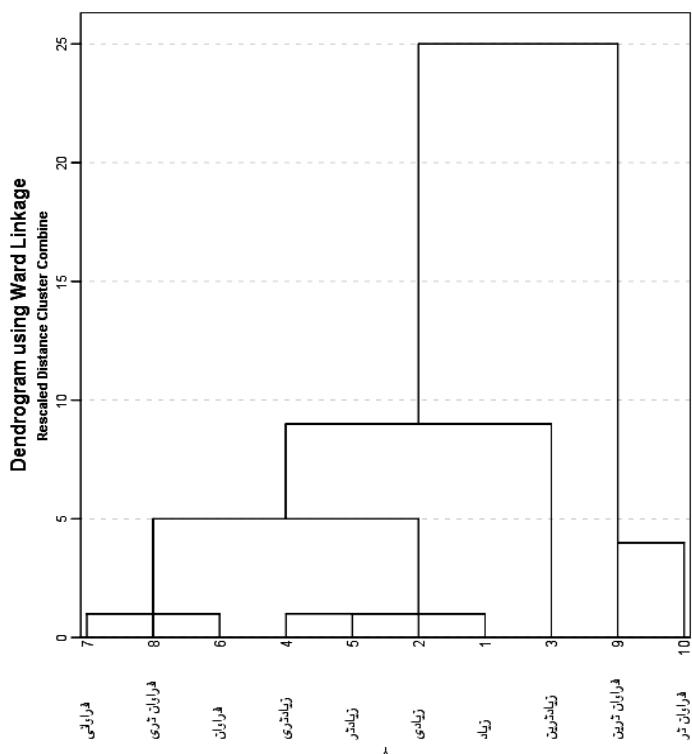
Chart 2: Hierarchical agglomerative structure of Syntactic ID-tags

تحلیل بسامد مطلق برچسب‌های نحوی در راستای گروه‌بندی‌های نمودار ۲ قرار دارد. برای نمونه، صفت‌های «زیادی، زیادترین، زیادتری، فراوان» به دلیل شباهت نسبی در سطوح برچسب‌های «خبری، پرسشی، امری، پایه، پیرو» در اولین گروه نمودار ۲ قرار می‌گیرند و یا شباهت نسبی بسامدها در سطوح «خبری، پرسشی، امری، پایه، استنادی، وصفی، مستند» باعث شده است که صورت‌های «فراوانی، فراوان‌تری، فراوان‌ترین» در گروه دوم کنار یکدیگر قرار گیرند. همچنین، عامل تشکیل گروه «زیادتر، فراوان‌تر» شباهت نسبی در سطوح «خبری، پرسشی، امری، پایه، پیرو، استنادی، وصفی، مستند، فاعل» است.

بررسی نمره‌های استاندارد برچسب‌های نحوی نشان داد که سطوح برچسب‌های «استنادی و وصفی» در همه صورت‌های واژگانی دارای نمرات استانداردی بین ۱.۹۳ تا ۱.۹۴ هستند. از این‌رو، صورت‌های تصویری فوق تفاوت معناداری را در این سطوح نشان نمی‌دهند. با وجود این، صورت تصویری «فراوان‌تر» در سطوح برچسب‌های «خبری، امری، پایه، پیرو» با نمرات استانداردی بالاتر از ۲ تفاوت معناداری را با سایر صورت‌ها ایجاد کرده است، اما این صورت تصویری در نمودار سلسله‌مراتبی به دلیل نزدیکی نسبی نمره‌های Z در سطوح برچسب‌های «استنادی، وصفی، مستند، فاعل» در کنار صورت تصویری «زیادتر» قرار گرفته است.

۳-۵. تحلیل برچسب‌های معنایی

نمودار ۳ تحلیل سلسله‌مراتبی برچسب‌های معنایی را نشان می‌دهد. این نمودار براساس ادغام سه گروه «فراوانی/۷، فراوان‌تری/۸، فراوان/۶»، «زیادتری/۴، زیادتر/۵، زیادی/۲، زیاد/۱» و «فراوان‌ترین/۹، فراوان‌تر/۱۰» تشکیل شده است. در این گروه‌ها، به دلیل کنار هم قرار گرفتن صفت‌های اندازه با پایه‌های یکسان فقط شباهت توزیعی بین اعضای هر گروه وجود دارد. علاوه بر این، خوشة گروه «فراوانی/۷، فراوان‌تری/۸، فراوان/۶» در ابتداء و در ارتفاع (۵) با گروه «زیادتری/۴، زیادتر/۵، زیادی/۲، زیاد/۱» ادغام شده که نشان از شدت هم‌معنایی تقریبی فراوان بین آن‌هاست و سپس در ارتفاع بین ۵ تا ۱۰ با «زیادترین/۳» ترکیب شده که سطح هم‌معنایی تقریبی کمتری را بین صفت‌های اندازه فوق نشان می‌دهد. درنهایت، خوشة برآمده از گروه «فراوان‌ترین/۹، فراوان‌تر/۱۰» در ارتفاع ۲۵ با خوشة برآمده از صفت‌های اندازه «فراوانی/۷، فراوان‌تری/۸، فراوان/۶، زیادتری/۴، زیادتر/۵، زیادی/۲، زیاد/۱، زیادترین/۳» برخورد می‌کند که دلالت بر پایین‌ترین سطح هم‌معنایی تقریبی بر پایه برچسب‌های معنایی است.



نمودار ۳: ساختار سلسله مراتبی برچسب‌های معنایی

Chart 3: Hierarchical agglomerative structure of Semantic ID-tags

بررسی اطلاعات آماری برچسب‌های معنایی نشان می‌دهد که صورت‌های «زیادی»، «زیادتر»، «فرادی»، «فراد تری» ترجیح می‌دهند که غالباً موصوف آن‌ها «انتزاعی» باشد، در بافت‌های «استخاری» حضور داشته باشد، مقوله معنایی موصوف آن‌ها «اسم عام و یا گزاره‌ای» باشد و زمانی که فعل جمله‌ای را توصیف می‌کند، آن فعل دارای نمود واژگانی «فرایندی (کنشی) و یا ایست» باشد. این در حالی است که صورت «فراد ترین» ترجیح می‌دهد که غالباً موصوف آن «عینی» باشد، در بافت‌های «تحت‌اللفظی» حضور داشته باشد و مقوله معنایی موصوف آن «اسم عام» باشد. همچنین، صورت‌های «زیادترین»، «زیادتری»، «زیاد»،

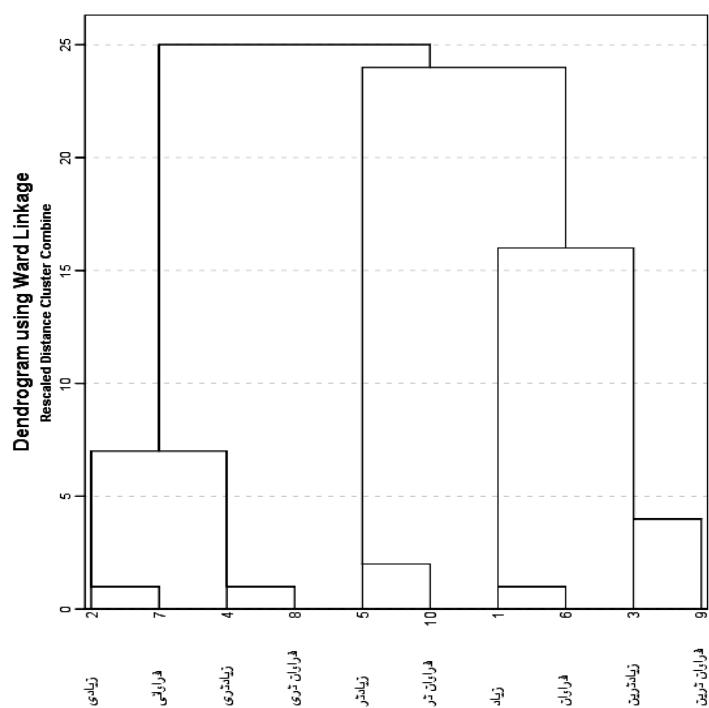
فراوان‌تر» ترجیح می‌دهند که غالباً موصوف آن‌ها «انتزاعی» باشد در بافت‌های «تحتاللفظی» حضور داشته باشد و مقوله معنایی موصوف آن «اسم عام، اسم گزاره‌ای و یا اسم مکان» باشد. علاوه بر این، براساس اطلاعات آماری می‌توان نتیجه گرفت که صورت‌های «زیاد، زیادی، زیادتر، زیادترین، فراوان، فراوانی، فراوان‌تر، فراوان‌ترین، فراوان‌تری» هیچ تمایلی ندارند که موصوف آن‌ها اگر اسم بود مقوله معنایی «اسم خاص و اسم خویشاوندی» داشته باشد و اگر موصوف آن‌ها فعل بود، نمود واژگانی «پایا» را داشته باشد.

تحلیل آزمون Z نشان می‌دهد که صورت‌های تصrifی «زیادی، زیاد، زیادتری، فراوان، فراوانی، فراوان‌ترین» در سطوح برچسب‌های «انتزاعی، عینی، غیرجاندار، انسان، غیرانسان، استعاره‌ای، تحتاللفظی، اسم عام، اسم جمع، اسم مکان، اسم گزاره‌ای» با نمرات استاندارد بین ۱.۶۶ تا ۱.۳۲ - موجب تفاوت معناداری نمی‌شوند. از این رو، آن‌ها بیشتر عامل گسترش هم-معنایی تقریبی بین صورت‌های واژگانی هستند. با وجود این، برخی از صورت‌های تصrifی مانند «فراوان‌ترین» در سطوح برچسب‌های «انتزاعی، عینی، اسم عام، اسم گزاره‌ای» به ترتیب با نمرات استاندارد ۲.۳۹، ۲.۳۹، ۲.۶۶ و ۱.۹۹ در سطوح برچسب «اسم مکان» با نمره استاندارد ۲.۳۹ تفاوت معناداری را ایجاد کرده‌اند و به همین دلیل، این دو صورت تصrifی در آخرین زیرگروه نمودار در کنار هم قرار گرفته‌اند. علاوه بر این، «زیاد» در سطوح برچسب‌های «ایستا، کنشی و لحظه‌ای» با نمرات استاندارد ۲.۱۶، ۲.۱۶ و ۲.۲۲ تفاوت معناداری را نسبت به سایر صورت‌های تصrifی از خود نشان می‌دهد. این تفاوت در سطوح برچسب‌های فوق باعث شده است که در نمودار ۳ صفت اندازه «زیاد» در کنار «فراوان» قرار نگیرد.

۴-۵. تحلیل مجموع برچسب‌ها

نمودار ۴ بر مبنای اعمال همه برچسب‌های ساختوازی، نحوی و معنایی شکل گرفته است. این نمودار شامل گروه‌های هم‌معنایی «زیادی/۲، فراوانی/۷، «زیادتری/۴، فراوان‌تری/۸»، «زیادتر/۵، فراوان‌تر/۱۰»، «زیاد/۱، فراوان/۱۰» و «زیادترین/۳، فراوان‌ترین/۹» است. اعضای هر کدام از این گروه‌ها به‌دلیل قرارگرفتن در پایین‌ترین ارتفاع نمودار یعنی بین (۰ تا ۵) بالاترین سطح هم‌معنایی تقریبی را نشان می‌دهند. از طرف دیگر، بررسی نقاط برخورد خوش‌ها در نمودار ۴ نشان می‌دهد که در ابتدا، خوشة برآمده از گروه «زیادی/۲، فراوانی/۷» در ارتفاع بین ۵ تا ۱۰ با خوشة گروه «زیادتری/۴، فراوان‌تری/۸» برخورد می‌کند که دلالت بر هم‌معنایی تقریبی نسبتاً زیاد بین پایه‌های

غیریکسان آن هاست. سپس، خوشهای گروههای «زیاد/۱، فراوان/۷» و «زیادترین/۳، فراوان ترین/۹» به دلیل هم معنایی تقریبی کم در ارتفاع بین (۱۵ تا ۲۰) با هم ادغام می شوند. در مرحله بعد، و در ارتفاع بین (۲۰ تا ۲۵) خوشهای برآمده از صفت‌های اندازه «زیاد/۱، فراوان/۶، زیادترین/۳، فراوان ترین/۹» با خوشهای تفضیلی «زیادتر/۵، فراوان تر/۱۰» ترکیب می شود. در آخرین مرحله، خوشهای سمت چپ و راست نمودار یعنی «زیادی/۲، فراوانی/۷، زیادتری/۴، فراوان تری/۸» و «زیاد/۱، فراوان/۶، زیادترین/۳، فراوان ترین/۹، زیادتر/۵، فراوان تر/۱۰» در ارتفاع (۲۵) با یکدیگر ادغام می شوند که این نقطه اتصال، پایین‌ترین سطح هم معنایی تقریبی را بین صفت‌های اندازه این تحقیق نشان می دهد.



نمودار ۴: ساختار سلسله‌مراتبی تجمعی همه برجسب‌ها

Chart 4: Hierarchical agglomerative structure of all ID-tags

۶. نتیجه

در این مقاله به منظور بررسی رابطه هم معنایی تقریبی صفت‌های اندازه «زیاد، فراوان» و صورت‌های تصریفی آنها از رویکرد تحلیل نمای رفتاری استفاده شده است. بر مبنای این رویکرد تعداد ۶۱ سطح برچسب ساخت‌واژی، نحوی و معنایی برای بررسی هر جمله درنظر گرفته شد که درمجموع کل جملات، شامل بیش از ۱۱۹۰۰ برچسب می‌شود. نتایج نمودارهای خوش‌ای سلسله‌مراتبی نشان می‌دهد که در سطوح برچسب‌های ساخت‌واژی اعضای گروه‌های «زیاد، فراوان»، «زیادترین، فراوان‌ترین»، «زیادی، فراوانی»، «زیادتری، فراوان‌تری»، «زیادتر، فراوان‌تر» بالاترین سطح رابطه هم معنایی تقریبی را نسبت به هم‌گروه خود دارند. همچنین، در سطح برچسب‌های نحوی، بالاترین سطح هم معنایی تقریبی متعلق به گروه‌های «زیادی، زیادترین، زیادتری، فراوان»، «فراوانی، فراوان‌تری، فراوان‌ترین» و «زیادتر، فراوان‌تر» است که در گروه اوّل بین «فراوان» و سه عضو دیگر رابطه هم معنایی تقریبی وجود دارد. در گروه دوم، صفت‌های اندازه «فراوانی، فراوان‌تری، فراوان‌ترین» دارای شباهت توزیعی هستند و درنهایت، در گروه چهارم بین «زیادتر، فراوان‌تر» رابطه هم معنایی تقریبی برقرار است. علاوه بر این، در سطح برچسب‌های معنایی بین اعضا با پایه غیریکسان گروه‌های «فراوانی، فراوان‌تری، فراوان» و «زیادتری، زیادی، زیاد» بالاترین سطح هم معنایی تقریبی و بین اعضای گروه‌های «فراوان‌ترین، فراوان‌تر» با «زیادتری، زیادتر، زیادی، زیاد، زیادترین» پایین‌ترین سطح هم معنایی تقریبی مشاهده می‌شود. درنهایت، تحلیل همه برچسب‌های این تحقیق، تشکیل گروه‌های «زیادی، فراوانی»، «فراوان‌تری، زیادتری»، «زیادتر، فراوان‌تر»، «زیاد، فراوان» و «زیادترین، فراوان‌ترین» را نشان داد که اعضای هر کدام از این گروه‌ها به صورت درون‌گروهی بالاترین سطح رابطه هم معنایی تقریبی را نسبت به هم‌گروه خود نشان می‌دهند. همچنین در این تحقیق به منظور مشخص نمودن برچسب‌هایی که تفاوت معناداری را موجب می‌شوند از نمرات استاندارد (آزمون \mathbb{Z}) استفاده شده است. بر همین اساس، تحلیل نمرات استاندارد نشان داد که در سطح برچسب‌های ساخت‌واژی صفت اندازه «زیاد» در سطوح «بن مضارع، اسم مرکب، صفت مرکب و مرکب» دارای نمره استاندارد بالاتر از ۲۸۴ است که این نمرات تفاوت معناداری را بین «زیاد» با صورت‌های «زیادتر، فراوان‌تر، زیادتری، فراوان‌تری» ایجاد کرده است. در سطح برچسب‌های نحوی



صورت تصریفی «فراوان‌تر» در سطوح برچسب‌های «خبری، امری، پایه، پیرو» با نمره‌های استانداردی بالاتر از ۲ تفاوت معناداری را با سایر صورت‌ها ایجاد کرده است. درنهایت، در سطح برچسب‌های معنایی، صفت عالی «فراوان‌ترین» در سطوح برچسب‌های «انتزاعی، عینی، اسم عام، اسم گزاره‌ای» به ترتیب با نمرات استاندارد ۰.۳۹، ۰.۶۶، ۰.۳۹ و صفت تفضیلی «فراوان‌تر» در سطح برچسب «اسم مکان» با نمره استاندارد ۰.۳۹ تفاوت معناداری را ایجاد کرده‌اند. علاوه بر این، در سطح برچسب‌های معنایی صفت اندازه «زیاد» در سطوح برچسب‌های «ایستا، کنشی و لحظه‌ای» با نمرات استاندارد ۰.۶۱، ۰.۱۶، ۰.۲۲ باعث ایجاد تفاوت معناداری نسبت به سایر صورت‌های تصریفی شده است.

۷. پی‌نوشت‌ها

1. size adjectives
2. behavioral profile analysis
3. Divjak & Gries
4. near-synonymy
5. Deese
6. Charles & Miller
7. Justeson & Kats
8. Lee
9. Biber& et al
10. Dixon
11. Jones & et al
12. Otani
13. Moser
14. Kuznetsova
15. id-tags
16. Atkins
17. Hanks
18. lemmas
19. annotation
20. co-occurrence
21. hierarchical agglomerative cluster analysis
22. activity
23. stative
24. achievement
25. accomplishment
26. Lakoff & Johnson

27. Turner
28. Antccconc
29. concordance
30. Kerjcie & Morgan
31. manual annotation

۸ منابع

- انوری، ح.، و احمدی گیوی، ح. (۱۳۸۷). *دستور زبان فارسی ۲*. ویرایش سوم. تهران: فاطمی.
- عاصی، م.، رضاقلی فامیان، ع.، و آقاجانی، د. (۱۳۸۵). *بهسوی طراحی شبکه و ازگانی صفات زبان فارسی. زبان و زبان‌شناسی*، ۴، ۱۳۶-۱۲۵.
- عبدالکریمی، س. (۱۳۹۳). بررسی هم‌معنایی گزاره‌ای، توصیفی و نسبی میان فعل‌های ساده فارسی و متناظرهای معنایی مرکب آن‌ها. *علم زبان*، ۲، ۱۴۳-۱۵۸.
- عبدالکریمی، س. (۱۳۹۴). بررسی انواع هم‌معنایی میان فعل‌های ساده فارسی و متناظر صوری - معنایی مرکب آن‌ها. *جستارهای زبانی*، ۷، ۲۰۱-۲۲۷.
- کاویانی فردزاده، ا. (۱۳۹۶). بررسی چندمعنایی چهار پیشوند غیرفعلی نفی در زبان فارسی براساس چارچوب چندمعنایی اصولی و رویکرد رفتاری. *پایان‌نامه کارشناسی ارشد. دانشگاه شیراز*.
- ماهوتیان، ش. (۱۳۷۸). *دستور زبان فارسی از دیدگاه رده‌شناسی*. ترجمه م. سمائی. تهران: نشر مرکز.
- Lee, S. (1994). *Untersuchungen zur Valenz des Adjektivs in der deutschen Gegenwartssprache*. Berlin: Lang.

References:

- Abdolkarimi, S. (2014). A review on propositional, descriptive and partial synonymy between Persian simple verbs and their corresponding compound verbs. *Language Science*, 2, 143-158. [In Persian].
- ----- (2016). Study of different kinds of synonymy between Persian non-compound verbs and their formal-semantic equivalent compound verbs. *Language*

Related Research, 6, 201-207. [In Persian].

- AleAhmad, A., Amiri, H., Darrudi, E., Rahgozar, M., & Oroumchian, F. (2009). Hamshahri: A standard Persian text collection. *Knowledge Based Systems*, 22 (5), 382-387.
- Anthony, L. (2011). AntConc (Version 3.2.2.1) [Computer Software]. Waseda University, Tokyo, Japan. Retrieved from. http://www.antlab.sci.waseda.ac.jp/antconc_index.html.
- Anvari, H. & AhmadiGhivi, H. (2008). *Persian Grammar* 2. 3rd Edition. Tehran: Fatemi Publications. [In Persian].
- Assi, M., Gholifamian, A. R., & Aghajani, D. (2006). Towards building a word-net for Persian adjectives. *Language and Linguistics Journal*, 4, 125-136. [In Persian].
- Atkins, B. T. S. (1987). Semantic ID-tags: Corpus Evidence for Dictionary Senses. In *Proceedings of the Third Annual Conference of the UW Centre for the New Oxford English Dictionary*. 17–36.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Brown, K., & Miller, J. (2013). *The Cambridge dictionary of linguistics*. New York: Cambridge University Press.
- Charles, W. G., & Miller, G. A. (1989). Contexts of antonymous adjectives. *Applied Psycholinguistics*, 3, 357–375.
- Croft, W. (2012). *Verbs: aspect and causal structure*. Oxford university press.
- Deese, J. (1964). The associative structure of some common English adjectives. *Journal of Verbal Learning and Verbal Behavior*, 5, 347-357.
- Divjak, D. (2010). *Structuring the lexicon: A clustered model for near-synonymy*. Berlin/New York: Walter de Gruyter GmbH & Co. KG.
- Divjak, D., & Gries, Th. (2006). Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory*, 2 (1), 23-60.
- Dixon, R. W. (2004). *Adjective Classes in Typological Perspective*. In Dixon,

- R.M.W., & Aikhenvald, A (eds.) Adjective Classes: A Cross linguistic Typology.
- Golshaie, R. (2016). A corpus study on identification and semantic classification of light verb constructions in Persian: the case of the light verb xordan 'to eat/ collide'. *Language Sciences*, 57, 21-33.
 - Gries, Th. S. (2010). Behavioral profiles: a fine-grained and quantitative approach in corpus-based lexical semantics. *The Mental Lexicon*, 5, 323-346.
 - Gries, Th. S., & Divjak, D. (2009). Behavioral profiles: a corpus-based approach to cognitive semantic analysis. In *New directions in Cognitive Linguistics*. Evans, V., & Pourcel, S (eds). 57–75.
 - Gries, Th. S., & Otani, N. (2010). Behavioral profiles: A corpus-based perspective on synonymy and antonymy. *ICAME Journal*, 34, 121- 150.
 - Hanks, P. (1996). Contextual dependency and lexical sets. *International Journal of Corpus Linguistics*, 1 (1), 75–98.
 - Jones, S., Paradis, C., Murphy, M. L., & Willners, C. (2007). Googling for opposites: a web-based study of antonym canonicity. *Corpora*, 2, 129-154.
 - Justeson J. S., & Kats, S. M. (1991). Co-occurrences of antonymous adjectives and their contexts. *Computational Linguistics*, 1 (1), 1–19.
 - Kaviyani Fardzade, E. (2017). *An investigation of the polysemy of four negative non-verbal prefixes in Persian language according to Principled Polysemy Framework and Behavioral Profiles Approach*. Master thesis. Shiraz University. Shiraz. [In Persian].
 - Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological measurement*, 30, 607-610.
 - Kuznetsova, J. (2015). *Linguistics profiles*. Berlin/Boston: Walter de Gruyter. GmbH.
 - Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago: University of Chicago Press.
 - Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about*

- the mind.* Chicago: University of Chicago Press.
- ----- (1993). *the Contemporary Theory of Metaphor*. In A. Ortony (Ed). *Metaphor and Thought*. Cambridge: Cambridge University Press. 202-251.
 - Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago and London: The University of Chicago Press.
 - Lee, S. (1994). Studies on the valence of the adjective in contemporary German. Berlin: Lang. [In German].
 - Mahootian, Sh. (1999). *The Persian grammar from typological perspective*. Translated by: Mehdi Samaei. Tehran: Nashr-e Markaz. [In Persian].
 - McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Theory, method and practice*. Cambridge: Cambridge University Press.
 - Moser, Ch. (2014). *Synonymy, plesionymy and sameness of meaning: A corpus-based behavioral profiles analysis of adjectives of size*. Master Thesis. University of Wien.