




Metaphor Identification in Persian: Annotation, Data Analysis, and Reliability Assessment for Compiling a Metaphor Corpus for Persian

Mohammad Saeid Miri ¹  Ph.D. Candidate in Linguistics, Allameh Tabataba'i University, Tehran, Iran

Abstract

This is becoming increasingly necessary for machines to be able to comprehend figurative language as artificial intelligence and natural language processing continue to advance. Metaphors are one of the figurative language forms that are difficult for machines to grasp. Developing metaphor corpora, which will enable machines to be trained using them, is the initial stage in the process of enhancing metaphor comprehension. The Metaphor Identification Procedure Vrije Universiteit, known as MIPVU, is a method that can be used to annotate metaphors. In the present study, MIPVU is evaluated in order to compile a Persian metaphor corpus. A collection of scholarly papers and news articles was gathered and annotated. The reliability of the procedure was subsequently evaluated using the Kappa coefficient and Cochran's Q. According to the results of the investigation, MIPVU can annotate Persian metaphors precisely and reliably ($\kappa=0.964$). Consequently, this procedure offers a reliable method for compiling a metaphor corpus.

Keywords: Metaphor, Metaphor Corpus, Reliability, Semantic Annotation.

1. ms_miri@outlook.com (Corresponding Author)

How to cite: Miri, M. S. (2024). Metaphor Identification in Persian: Annotation, Data Analysis, and Reliability Assessment for Compiling a Metaphor Corpus for Persian. *Language and Linguistics*, 19(38), 181- 208. doi: [10.30465/LSI.2024.47498.1730](https://doi.org/10.30465/LSI.2024.47498.1730)

1. Introduction

Recent advances in Natural Language Processing (NLP) have made it necessary for computers to understand figurative language, especially metaphor. However, metaphor is problematic for computers because it compares two seemingly unrelated domains (like THEORY and BUILDING in “It's not a strong theory”). Compiling metaphor-annotated corpora is one of the first steps in improving computer understanding of metaphor. Prior to this, a metaphor identification method is required, with clear and explicit protocols to mark words as (non-)metaphorical.

Metaphor Identification Procedure Vrije Universiteit (henceforth MIPVU, Steen et al., 2010) is one way of annotating linguistic metaphors. It is widely used in corpus and computational approaches to metaphor and in building metaphor corpora (e.g., for English, Chinese, and Slovenian). As of yet, there is no metaphor corpus available for Persian. This paper evaluates MIPVU for compiling a Persian metaphor corpus. The question is, to what extent can MIPVU precisely identify metaphors in Persian, or, in other words, how reliable is it?

To answer the question, a collection of news and academic texts was collected, and three analysts tagged the data using the Persian version of MIPVU (Miri, 2022). After that, the procedure's reliability was calculated using Cohen's and Fleiss' Kappa and Cochran's Q. Analysts also found some cases that did not correspond to MIPVU protocols.

2. Literature Review

Philosophers and rhetoricians have traditionally defined metaphor as a linguistic device, a deviation from normal language use, used to make an impression on the reader or listener. The prevailing view persisted until the late 20th century, especially before the introduction of Conceptual Metaphor Theory (henceforth CMT, Lakoff & Johnson, 1980). In contrast with the traditional view, metaphor in CMT is not merely a matter of language but represents “a cross-domain mapping in the conceptual system” (Lakoff, 1993, p. 2). Since then, numerous cognitive and/or corpus linguists have applied CMT to linguistic data to identify metaphors in discourse. The first attempt was the Metaphor Identification Procedure (MIP, Pragglejaz Group, 2007), which soon became the most commonly used method. It provided the basis for more elaborate versions, such as MIPVU, for identifying different types of metaphor, such as deliberate metaphor (Steen, 2011), and metaphor in various modalities, such as visual metaphor (Šorm & Steen, 2018). MIP is based on the contrast between the basic and contextual meanings. Steen et

al. (2010) stepped further and revised MIP, proposing an extended version called Metaphor Identification Procedure Vrije Universiteit (henceforth MIPVU). Several metaphor corpora have been developed based on MIPVU (for English, Krennmayr & Steen, 2017; for Chinese, Lu & Wang, 2017; and for Slovenian, Antloga, 2020). Despite this, there have only been a few studies on metaphor identification in Persian (Assi et al., 2022; Miri, 2023), and no corpus has yet been published.

3. Method

A 7,758-token¹ corpus of news (from Hamshahri Corpus, Aleahmad et al., 2009) and academic (from Pazhooheshname Corpus, Alayiaboozar et al., 2021) texts was collected and pre-processed. Three Persian native speakers annotated the corpus: the author, a Master's student in Linguistics, and an undergraduate student in Spanish. To evaluate their skill in identifying metaphors, the analysts took a 66-token test before annotating corpus data. Each text was annotated twice by two analysts (the author did this for both news and academic texts).

Then, the reliability was assessed using Cochran's Q, Cohen's Kappa, and Fleiss' Kappa in two phases: first, for the 66-token test, and second, for the corpus data. Cochran's Q could only be calculated for the first phase since there were fewer than three analysts per text. Reliability was measured in two rounds: before discussion (annotating without knowing each other's labels), and after discussion (revising some disagreed cases after a discussion session). Finally, margins of errors were estimated in five stages: (i) words with 'w' tag, (ii) compound verbs, (iii) polywords, (iv) borderline cases, and (v) tokens with metaphor-related tags.

The Persian Advanced Learner's Dictionary (Assi, 2019), a corpus-based contemporary dictionary of Persian, was used to identify the (more) basic and contextual meaning. In instances where the meanings were not found in the primary dictionary or where there was not a clear distinction between senses, we used two additional sources: Sokhan Dictionary (Anvari, 2002) and FarsNet (Shamsfard et al., 2010).

4. Results

The analysts mostly reached a consensus while annotating the corpus data. The agreement percentage during the first phase was 92.2%, and the agreement percentage during the second phase was 94.7% before discussion (BD) and 98.3% after discussion (AD). Fleiss' Kappa is 0.842 for the first phase, and Cohen's Kappa is 0.864 (BD) and 0.959 (AD) for the second

phase (in total). Table 1 shows the results. There are two reasons for the difference in lexical units after discussion: deciding about some borderline cases after discussion and inadvertent mistakes in demarcation of lexical units (this happened in the data analysis stage).

Table 1
Reliability Assessment Results in Each Phase.

Phase	Data	No. of lexical units	Agreement percentage			κ	Cochran's Q (df=2)	
			Not MRW	MRW	Total			
1	66-token test	51	76.5	15.7	92.2	0.842	2.00	
2	News	BD	3507	74.5	20.6	95.1	0.863	-
		AD	3506	74.9	23.3	98.2	0.952	-
	Academic	BD	3091	65.3	28.8	94.1	0.863	-
		AD	3107	68.1	30.4	98.5	0.965	-
	Total	BD	6598	70.2	24.5	94.7	0.864	-
		AD	6613	71.7	26.6	98.3	0.959	-

* BD: Before Discussion, AD: After Discussion, MRW: Metaphor-Related Word

A margin of error calculation was performed in five steps, as previously mentioned: (i) words with 'w' tag (0.3%), (ii) compound verbs (0.06%), (iii) polywords (0.15%), (iv) borderline cases (0.045%), and (v) tokens with metaphor-related tags (0.25% 'met' but incorrectly tagged 'not met'; 0.16% 'not met' but incorrectly tagged 'met'; and 0.015 'implicit met' but incorrectly tagged 'not met'). Table 2 shows the reliability result after reducing the margin of error.

Table 2
Reliability Assessment Result (Final).

Data	No. of lexical units	Agreement percentage			κ
		Not MRW	MRW	Total	
Corpus data	6619	71.8	26.8	98.6	0.964

5. Discussion

Even though analysts agreed with each other in most cases, some disagreements remained. There is disagreement among analysts regarding the basic meaning of Persian prepositions. For example, analysts disagreed on the basic meaning of به (/be/ ‘to, for’) and برای (/baɾɒje/ ‘for’). Another example was حال (/hal/ ‘[psychological] state, present [time]’) in the context در همین حال (/daɾ hamin hal/ ‘meanwhile’). Although the temporal sense is the basic meaning, it is unclear whether the contextual meaning is related to state or time.

6. Conclusion

Overall, this study evaluated MIPVU for building a Persian metaphor corpus and highlighted areas for possible improvement. By using MIPVU, linguistic data was annotated, and reliability was assessed. According to the results, it is a reliable method for compiling a metaphor corpus for Persian. MIPVU performed well and replicated the results of other reliability tests in other languages (see Nacey et al., 2019). This procedure reduces the possibility of disagreements between analysts due to its explicit and straightforward protocols. Of course, MIPVU is not flawless and may need minor modifications, depending on the language. Nevertheless, it is still in the middle of its development and will improve as it deals with more data from other languages.

Acknowledgments

This study would not have been completed without the assistance and guidance of Dr. Azadeh Mirzaei, who guided me throughout the research process, and Chadi Raheb and Hossein Hazrati, who annotated the corpus.

Notes

¹ The number of tokens is not the same as the number of lexical units (for demarcation of lexical units, see Steen et al., 2010).



شناسایی استعاره در زبان فارسی: برچسب‌گذاری، تحلیل داده، و ارزیابی پایایی برای ساخت پیکره استعاره فارسی

دانشجوی دکترای زبان‌شناسی، دانشکده ادبیات فارسی و زبان‌های
خارج، دانشگاه علامه طباطبائی، تهران، ایران

محمدسعید میری^۱

چکیده

با پیشرفت چشم‌گیری که در هوش مصنوعی و پردازش زبان طبیعی رخ داده است، نیاز به درک بیان‌های غیر تحت اللفظی توسط ماشین بیش از گذشته حس می‌شود. استعاره یکی از این بیان‌هاست که ماشین همواره در درک آن به مشکل برمی‌خورد. برای بهبود این فرایند باید دست به ساخت پیکره استعاره زد تا با آن بتوان استعاره را به ماشین آموخت. روال شناسایی استعاره دانشگاه آزاد آمستردام (MIPVU) یکی از روش‌های شناسایی و برچسب‌گذاری استعاره است. موضوع پژوهش حاضر ارزیابی این روال برای ساخت پیکره استعاره فارسی است. برای این منظور، پیکره‌ای از متون خبری و دانشگاهی گردآوری و بر اساس روال گفته‌شده توسط کارشناسان برچسب‌گذاری شد. سپس پایایی پژوهش با معیارهای آماری ضریب کاپا و Q کوکران حساب شد. نتایج حاصل از بررسی پیکره نشان می‌دهد که روال شناسایی استعاره توانسته است با پایایی مناسب ($K=0.964$) استعاره‌های فارسی را شناسایی و برچسب‌گذاری کند و در نتیجه، روش مناسبی برای ساخت پیکره استعاره است.

کلیدواژه: استعاره، پیکره استعاره، پایایی، برچسب‌گذاری معنایی.

۱- مقدمه

در عصر حاضر، گسترش فناوری‌های مبتنی بر هوش مصنوعی و پرکاربرد شدن آن‌ها در زندگی روزمره انسان، این نیاز را ایجاد کرده که ماشین بتواند پیچیدگی‌های معنایی زبان انسان را درک کند. بنابراین، امروزه ارتقای توانایی ماشین برای درک معانی پیچیده یکی از موضوعات مهم در زبان‌شناسی رایانشی و پردازش زبان طبیعی است. برای رسیدن به چنین هدفی ابتدا باید مجموعه‌ای از داده‌های زبانی که برچسب‌های معنایی خورده‌اند، تهیه کرد.

استعاره یکی از انواع پیچیدگی معنایی است که در آن یک مفهوم به کمک مفهومی دیگر و بر پایه مفهوم‌سازی غیرمستقیم و نوعی مشابهت بیان می‌شود (استین^۱، ۲۰۰۷). برای این که رایانه بتواند استعاره را درک کند و تشخیص دهد، در گام اول باید با داده‌های زبانی که برچسب‌های مربوط به استعاری بودن دارند، آموزش داد. از این رو، ساخت پیکره استعاره در ارتقای درک پیچیدگی‌های معنایی توسط رایانه مؤثر است. برای ساخت چنین پیکره‌ای باید روشی برای شناسایی استعاره داشت که به اندازه کافی صریح و کم‌ابهام باشد و توافق میان کارشناسان برچسب‌گذار به میزان قابل قبولی باشد. یکی از روش‌های برچسب‌گذاری استعاره، استفاده از روال شناسایی استعاره دانشگاه آمستردام^۲ (MIPVU)، از این پس «روال شناسایی استعاره»، استین و همکاران، ۲۰۱۰) است. این روال علاوه بر این که استعاره را مشخصاً تعریف کرده و گام‌های مشخصی برای شناسایی استعاره تعریف کرده، در پژوهش‌های پیکره‌بنیاد نیز به کار رفته است و چندین پیکره استعاره نیز بر اساس این روش ساخته شده است. در حال حاضر سه پیکره استعاره بر اساس همین روال برای زبان‌های انگلیسی (کرنمایر^۳ و استین، ۲۰۱۷)، چینی (لو^۴ و وانگ^۵، ۲۰۱۷) و اسلونیایی (آنتلواگا^۶، ۲۰۲۰) تهیه و منتشر شده است. پژوهش‌هایی درباره ساخت پیکره استعاره مبتنی بر همین روال برای زبان‌های دیگر منتشر شده، اما هنوز پیکره‌های آن‌ها در دسترس قرار نگرفته است. گرچه تشخیص و برچسب‌گذاری استعاره، چنان که گفته شد، اهمیت زیادی در هوش مصنوعی، پردازش زبان طبیعی، و زبان‌شناسی رایانشی و پیکره‌ای دارد، تا کنون پیکره‌ای برای استعاره‌های زبان فارسی منتشر نشده است.

این روال دستورالعمل‌های صریحی برای شناسایی و برچسب‌گذاری استعاره دارد، اما ویژگی‌های خاص زبان‌ها باعث شده که تقریباً برای هر زبان نسخه‌ای مجزا داشته باشد. در

1 G. J. Steen

2 Metaphor Identification Procedure Vrije Universiteit (MIPVU)

3 T. Krennmayr

4 X. Lu

5 B. P.-Y Wang

6 Š. Antloga

بعضی زبان‌ها شیوه تقطیع واحدهای واژگانی (به عنوان واحد بررسی استعاره) متفاوت از نسخه اولیه روال است، یا منابع آن زبان (مثلاً فرهنگ‌های لغت یا وردنت^۱) ویژگی‌ها یا محدودیت‌های خاصی دارند که در نسخه اولیه روال پیش‌بینی نشده است. بنابراین، ضروری است که روال متناسب با زبان بهینه شود. برای نمونه، فعل‌های مرکب فارسی ویژگی‌هایی دارند که تقطیع واحدهای واژگانی‌شان برای شناسایی استعاره نمی‌تواند بر اساس نسخه اولیه روال انجام شود. این فعل‌ها بیش از یک موردواژه‌آند و در پیکره‌ها معمولاً اجزای‌شان را از هم جدا می‌کنند. حال اگر قرار باشد به یک فعل مرکب برچسب استعاری اختصاص بدهیم، آن را باید به جزء فعلی داد یا غیرفعلی؟ آیا باید فعل مرکبی همچون هماهنگ کردن را یک موردواژه و یک واحد واژگانی حساب کرد یا دو موردواژه (هماهنگ و کردن) و یک واحد واژگانی؟ اگر بین اجزای فعل مرکب فاصله بیفتد چه؟ نسخه اولیه روال پاسخ روشنی برای این موارد ندارد و حتی دستورالعمل‌های مربوط به فعل‌های گروهی^۲ نیز نمی‌تواند همه پرسش‌ها را پاسخ دهد. بنابراین، در ساخت پیکره استعاره فارسی نیز باید به ویژگی‌های خاص زبان فارسی توجه کرد و در مواردی که روال پاسخ مشخصی ندارد، باید تصمیم تازه‌ای گرفت.

هدف پژوهش حاضر آزمودن روال معرفی شده برای ساخت پیکره استعاره فارسی است. پرسش اصلی پژوهش این است که روال شناسایی استعاره مورد بحث تا چه اندازه در تشخیص استعاره‌های زبانی فارسی دقیق و بی‌ابهام است. به دیگر سخن، میزان پایایی^۳ روال در تشخیص استعاره‌های فارسی چقدر است. برای پاسخ به این پرسش، مجموعه‌ای از متون خبری و دانشگاهی جمع‌آوری و توسط سه کارشناس برچسب‌گذاری شد. برچسب‌گذاری پیکره بر اساس نسخه فارسی این روال (میری، ۱۴۰۲) انجام شده است. سپس میزان توافق بین کارشناسان با معیارهای آماری محاسبه پایایی به دست آمد. همچنین، کارشناسان حین برچسب‌گذاری استعاره‌ها به مواردی برخوردند که نه با نسخه اولیه روال سازگاری داشتند، و نه با نسخه فارسی. این موارد در مقاله بررسی و برای بعضی از این چالش‌ها پیشنهادهایی ارائه خواهد شد.

۲ پیشینه پژوهش

استعاره را فیلسوفان و زبان‌شناسان به اشکال گوناگون تعریف کرده‌اند. یکی از قدیمی‌ترین تعاریف‌ها متعلق به ارسطوست. او استعاره را اطلاق ویژگی‌ای به یک نام می‌دانست که به چیزی دیگر تعلق دارد (ارسطو، ۳۳۰ ق. م. / ۱۹۹۵: ۱۴۵۷b). استعاره از نظر او چیزی زینتی، جدا از

1 WordNet

2 token

3 phrasal verbs

4 reliability

کاربرد معمول زبان، و ابهام‌زا بود. این تعریف کلی در کنار ویژگی‌هایی که ارسطو ذکر کرده است، تا اواخر قرن بیستم دیدگاه غالب محققان این حوزه بود؛ این که استعاره اساساً زبانی است و نوعی انحراف^۱ از کاربرد معمول یا کشاکشی^۲ معنایی است (اورتنی^۳، ۱۹۹۳)، می‌توان آن را از کلام جدا کرد، و هدفش اثر گذاشتن بر روی خواننده یا شنونده است. اما طرح نظریه استعاره مفهومی^۴ (لیکاف^۵ و جانسون^۶، ۱۹۸۰) و پس از آن انتشار مقاله نظریه استعاره معاصر^۷ (لیکاف، ۱۹۹۳) در اواخر قرن بیستم، این دیدگاه را تا حد زیادی تغییر داد. به نظر زبان‌شناسان شناختی استعاره نه در نظام زبان، که در نظام مفهومی^۸ ما ریشه دارد و بر ادراک، تفکر، و عمل ما اثر می‌گذارد. افزون بر این، استعاره نه فقط در «الفاظ بدیع شاعرانه»، بلکه «در زبان عادی روزمره نیز بسیار به کار می‌رود» (لیکاف، ۱۹۹۳: ۲۰۳).

از دهه ۱۹۸۰ میلادی تا کنون، نظریه استعاره مفهومی مهم‌ترین چارچوب نظری برای مطالعه استعاره است و پژوهش‌های داده‌بنیاد و پیکره‌بنیاد بسیاری در این فضای نظری به انجام رسیده است. با وجود نقاط قوت این نظریه برای مطالعه استعاره، اجرایی ساختن آن برای شناسایی استعاره در گفتمان با پرسش‌های روش‌شناختی مهمی روبه‌روست: این که واحد شناسایی استعاره در این بررسی‌ها چیست، استعاره را چگونه می‌توان از دیگر موارد مشابه (همچون مجاز) تمییز داد، در موارد مرزی چه تصمیمی باید گرفت، و این که پایایی این بررسی‌ها چگونه است (استین، ۲۰۰۷، ۲۰۱۱؛ کرنمایر و استین، ۲۰۱۷). به نظر کووچش^۹ (۲۰۱۶)، رایج‌ترین و جدی‌ترین انتقاد از این نظریه، انتقادهای روش‌شناختی است: این که چگونه استعاره را می‌توان در گفتمان تشخیص داد و با داده‌های واقعی آن را آزمون. نکته دیگری که باید به آن توجه کرد این است که در اغلب پژوهش‌های پیکره‌بنیاد مبتنی بر استعاره مفهومی، یا یک برچسب‌گذار همه داده را برچسب زده است، یا اگر برچسب‌گذارها بیش از یک نفر بوده است، در پایان برچسب‌های یکدیگر را بررسی نکرده‌اند. به عبارتی دیگر، پایایی بسیاری از این پژوهش‌ها مشخص نیست و معمولاً گزارشی از آن وجود ندارد.

تلاش برای دستیابی به روشی کاربردی برای شناسایی استعاره در پیکره از دهه ۱۹۹۰ میلادی آغاز شده است و همچنان ادامه دارد. می‌توان دهه ۱۹۹۰ و دهه اول قرن بیست و یکم

1 deviance

2 tension

3 A. Ortony

4 Conceptual Metaphor Theory (CMT)

5 G. Lakoff

6 M. Johnson

7 The Contemporary Theory of Metaphor

8 conceptual system

9 Z. Kövecses

را دورهٔ تتبعات نظری پژوهشگران این حوزه دانست. در این دوران عمدهٔ کار پژوهشگران یافتن بهترین چارچوب نظری برای شناسایی استعاره بوده است. بخشی از این تلاش‌ها در کتاب پژوهیدن و به‌کارگیری استعاره^۱ (کمرون^۲ و لو^۳، ۱۹۹۹) آمده است. نقطهٔ عطف این حوزه، یا به عبارتی نخستین تتبع عملی، انتشار مقالهٔ گروه پراگلیز^۴ (۲۰۰۷) و معرفی روال شناسایی استعاره^۵ (از این پس «روال پراگلیز») بود که مبنای بسیاری از پژوهش‌های شناسایی استعاره قرار گرفت. مبنای تشخیص استعاره در این روال، وجود تفاوت محسوس میان معنی پایه‌ای (معنی عینی‌تر، انسان‌محور، بدنی‌تر، و قدیمی‌تر) و معنی بافتی است.

گرچه روال پراگلیز صریح است و توانست توفیقاتی به دست آورد، اما اشکالات آن موجب شد که مورد بازنگری قرار گیرد. مهم‌ترین دلایل بازنگری وقوع اختلاف نظر میان کارشناسان، پایایی متوسط، و شامل نشدن مصادیق دیگری از استعاره (روایی^۶ نامناسب) بود. سرانجام استین و همکاران (۲۰۱۰) نسخه‌ای دیگر از روال پراگلیز را معرفی کردند که ایرادهای نسخهٔ پیشین را نداشت: روال شناسایی استعارهٔ دانشگاه آزاد آمستردام. این روال تفاوت‌های زیادی با روال پراگلیز دارد. این تفاوت‌ها را می‌توان در چهار دسته جای داد:

(۱) در حدود و ثغور استعاره‌ها؛ به این معنی که آیا تشبیه و نوعی از قیاس را می‌توان

بررسی کرد یا خیر،

(۲) در سطح بررسی؛ یعنی این که آیا باید ساخت مفهومی را در نظر گرفت یا خیر،

(۳) در استفاده از معیار قدمت برای تعیین معنی پایه‌ای(تر)،

(۴) در دسته‌بندی استعاره‌ها.

روال شناسایی استعاره دستورالعمل‌های جزئی‌تری برای تعیین واحدهای واژگانی، تعیین معنی بافتی و پایه‌ای(تر)، تمایز میان معنی پایه‌ای(تر) و بافتی، تضاد و شباهت میان این معانی دارد. در شکل ۱ مراحل شناسایی استعاره با این روش به طور خلاصه آمده است.

1 Researching and Applying Metaphor

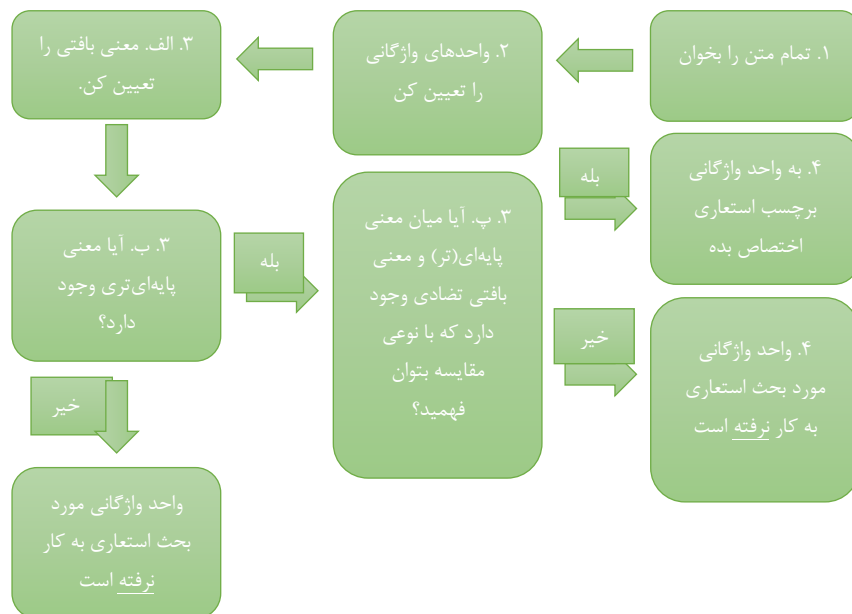
2 L. J. Cameron

3 G. Low

4 Praggeljaz Group

5 Metaphor Identification Procedure (MIP)

6 validity



شکل ۱. روال شناسایی استعاره (بر اساس دورست^۱ و همکاران، ۲۰۱۳: ۷۹، با کمی تغییر)

روال شناسایی استعاره بعدها به تهیه پیکره استعاره دانشگاه آزاد آمستردام^۲ (کرنامیر و استین، ۲۰۱۷) انجامید. همچنین، پژوهش‌های دیگری نیز بر اساس همین روال منتشر شده است (برای نمونه، کال^۳، ۲۰۱۲؛ دورست^۴، ۲۰۱۱، کرنامیر، ۲۰۱۱). در بخش قبل صحبت از این شد که روال شناسایی استعاره نسخه‌های مختلفی دارد. بخش عمده این نسخه‌ها در نیسی^۵ و همکاران (۲۰۱۹) معرفی شده است. بادریژلوا^۶ و همکاران (۲۰۱۳)، لو و وانگ (۲۰۱۷) و آنتلوگا (۲۰۲۰) نیز، به ترتیب، بر روی پیکره استعاره زبان روسی، چینی، و اسلونیایی کار کرده‌اند. در میان آثار فارسی نیز پژوهش‌های معدودی بر اساس این روال دیده می‌شود. عاصی و همکاران (۱۴۰۱) از این روال برای بررسی استعاره‌های فارسی در گفتمان آموزش استفاده کرده‌اند. میری (۱۴۰۲) نیز نسخه‌ای از روال ارائه داده که با ویژگی‌های زبانی فارسی سازگار است. هر دو مقاله در کلیات، که همان سازگار ساختن روال شناسایی استعاره با ویژگی‌های

1 A. G. Dorst
 2 VU Amsterdam Metaphor Corpus
 3 A. A. Kaal
 4 A. G. Dorst
 5 S. Nacey
 6 Y. Badryzlova

زبان فارسی است، مشابهند، اما در جزئیات برجسب‌گذاری، دسته‌برچسب^۱ معرفی شده، فرهنگ لغت مورد استفاده، و پیکره پژوهش با هم تفاوت دارند. افزون بر این، مقاله عاصی و همکاران (۱۴۰۱) به بررسی استعاره‌های زبانی در گفتمان آموزش می‌پردازد، حال آن که میری (۱۴۰۲) ملاک‌های شناسایی استعاره و خود روال را بررسی می‌کند.

۳- چارچوب نظری

استین (۲۰۰۷، ۲۰۱۶) در توضیح رویکرد خود بر این نکته تأکید می‌کند که روال شناسایی استعاره به بررسی استعاره در زبان می‌پردازد. او برای این که منظور خود را از بررسی استعاره در زبان روشن کند، به سه تفکیک اشاره دارد: تفکیک میان مطالعه (۱) نشانه زبانی و رفتار زبانی، (۲) نظام زبان و کاربرد زبان، و (۳) کاربرد زبان استعاری و بیان حاوی نگاشت میان حوزه‌ای.

تفکیک اول بین ساخت‌های واژگان-دستوری و نقش‌های نشانه، و فرایندهای روان‌شناختی یا اجتماعی حین استفاده برخط زبان، که همان رفتار است. تفکیک دوم میان این دو مفهوم برقرار است: کارکرد استعاری پدیده‌ای در بافتی مشخص حاصل تنش میان معنی پایه‌ای و بافتی (کاربرد)، و بخشی از شبکه اصلی ساخت‌های زبان (نظام). سومین تفکیک میان دیگر صورت‌های زبانی استعاری است که مستقیماً بیان می‌شوند (مثلاً تشبیه)، و بیان‌های حاوی نگاشت میان حوزه‌ای در اندیشه است که معنایی غیرمستقیم همراه خود دارند.

روال شناسایی استعاره به ترتیب نشانه‌بنیاد، کاربرد بنیاد، و متمرکز بر زبان است (استین، ۲۰۱۶). تعریف استین از استعاره را می‌توان این‌گونه خلاصه کرد: استعاره نگاشتی میان حوزه‌ای در ساحت زبان است. دلیل تأکید استین بر زبان برای تقابل با/اندیشه یا نگاشت میان حوزه‌ای در ساحت اندیشه نیست، بلکه او می‌خواهد با تمرکز بر زبان، دیگر مصادیق زبان استعاری را نیز که لزوماً حاوی نگاشت میان حوزه‌ای در ساحت اندیشه نیستند، از جمله بعضی استعاره‌های عمدی^۲ (استین، ۲۰۱۱)، در این تعریف بگنجانند.

همان‌گونه که پیداست، استین دغدغه‌ای روش‌شناسانه در برخورد با استعاره دارد. برای او مهم است که استعاره به گونه‌ای تعریف شود که تا حد امکان تمام مصادیق مرسوم را شامل و از دیگر بیان‌ها (مثلاً مجاز) جدا شود. نشانی از این دغدغه را می‌توان در ساخت و پرداخت چارچوب نظری روال شناسایی استعاره دید. او دیدی تجربی به پژوهش در حوزه استعاره دارد و چارچوب نظری‌اش را مرحله به مرحله مانند مراحل پژوهش تجربی پیش می‌برد: مفهوم‌سازی،

1 tagset

2 deliberate metaphors

عملیاتی‌سازی، جمع‌آوری داده، تحلیل داده، و تفسیر یافته‌ها (استین، ۲۰۰۷؛ استین و همکاران، ۲۰۱۰). ابتدا تعریفی برای استعاره ارائه می‌کند که با آن بتوان پژوهش را آغاز کرد، بعد با در نظر گرفتن ملاحظات اجرایی، آن را مقید می‌کند، سپس به جمع‌آوری داده می‌پردازد و معیارها را در برخورد با داده پالوده‌تر می‌کند، بعد داده را تحلیل می‌کند، و در نهایت نتایج را تفسیر می‌کند.

در مرحله مفهوم‌سازی، مبنای تشخیص استعاره را بر معنی غیرمستقیم و شباهت می‌گذارد. منظور از معنی غیرمستقیم همان نگاشت از یک حوزه مفهومی به حوزه‌ای دیگر است؛ با این شرط که این غیرمستقیم بودن در کاربرد زبانی ساخت‌های مفهومی است و نیازی به آمدن یا نیامدن نشانه‌ای زبانی نیست. برای درک بهتر *قابل دفاع* را در «ادعای شما قابل دفاع نیست» در نظر بگیرید. معنی یا، به تعبیری دقیق‌تر، مفهوم‌سازی غیرمستقیم یعنی حوزه مفهومی بحث با حوزه جنگ فهمیده می‌شود و به جای این که از بحث کردن صحبت شود، به طرز غیرمستقیم درباره جنگیدن صحبت شده است. پایه این نگاشت میان حوزه‌ای شباهت میان این دو حوزه است. استین با معرفی معیار غیرمستقیم بودن در کاربرد زبانی ساخت مفهومی، توانست هم انواع مشهور مفهوم‌سازی غیرمستقیم با بیان غیرمستقیم (استعاره غیرمستقیم) را، و هم انواع مفهوم‌سازی غیرمستقیم با بیان مستقیم (استعاره مستقیم، مثلاً تشبیه یا قیاس) را توجیه کند. این باریک‌بینی در روال پراگماتیک به این صورت وجود ندارد و تنها در روال شناسایی استعاره مبنا قرار گرفته است. در روال پراگماتیک غیرمستقیم بودن فقط در سطح زبانی تعریف شده است. به همین دلیل است که در این روال نمی‌توان تشبیه، تمثیل، یا قیاس را شناسایی کرد.

در مرحله عملیاتی‌سازی، استین و همکاران (۲۰۱۰) به این نکته مهم اشاره می‌کنند که واحد شناسایی استعاره باید تعریف و ویژگی‌های آن دقیقاً مشخص شود. آن‌ها واحد واژگانی^۱ را مبنا قرار می‌دهند و برای شناسایی آن دستورالعمل مفصلی نوشته‌اند. در مرحله جمع‌آوری داده، به این نتیجه می‌رسند که مرجعی برای تشخیص معنی پایه‌ای (تر) و بافتی باید معرفی شود: فرهنگ لغت. در این مرحله، معیارهای صریحی برای تشخیص معانی و سازوکارهای تصمیم‌گیری در موارد مشکوک یا مرزی پیشنهاد شده است. در مرحله تحلیل داده، درباره روش تحلیل و ارزیابی پایایی صحبت می‌شود. روال شناسایی استعاره باید به گونه‌ای عمل کند که اختلاف‌ها حداقل باشد و در حد مطلوبی پایا باشد.

در نهایت، در مرحله تفسیر استعاره، بحث این است که آیا استعاره را باید در سطح زبانی تفسیر کرد یا در سطح ساخت مفهومی. در این روال به دو دلیل حرفی از جزئیات ساخت

¹ lexical unit

مفهومی استعاره‌ها به میان نمی‌آید و تنها تحلیل زبانی استعاره بررسی می‌شود: یک، به دلیل ظهور اختلاف نظرهای حل‌ناشدنی میان کارشناسان (در تشخیص حوزه‌ها یا فضاها) و دو، به دلیل نبود نظریه‌ای فراگیر در باب استعاره در زبان‌شناسی شناختی. در واقع در روال شناسایی استعاره نداشت میان حوزه‌ای را پذیرفته‌اند، اما تنها در سطح کاربرد زبانی.

توضیحات بالا را می‌توان این‌گونه جمع‌بندی کرد که روال شناسایی استعاره بر پایه‌ای نظری استوار است که نظام نشانه‌ای زبان را در سطح کاربرد بررسی می‌کند، استعاره را نوعی مفهوم‌سازی غیرمستقیم (مبتنی بر نگاشت میان حوزه‌ای) می‌بیند، واحد شناسایی‌اش واحد واژگانی است، معانی را با رجوع به فرهنگ لغت از هم تمییز می‌دهد، ارزیابی پایایی را به عنوان معیاری برای محک خود در نظر دارد، و سطح تحلیلش زبانی است.

۴- روش پژوهش

برای آزمودن کارایی شناسایی استعاره نیاز است که روال مورد نظر از صافی داده‌های واقعی زبان و آزمون‌های پایایی بگذرد. برای این منظور، پیکره‌ای شامل متون خبری و دانشگاهی فارسی تهیه شد. متون خبری از نسخه‌ی دوم پیکره‌ی همشهری (آل احمد^۱ و همکاران، ۲۰۰۹) و متون دانشگاهی از پیکره‌ی پژوهش‌نامه (علایی ابوذر و همکاران، ۱۴۰۰) انتخاب شد؛ ۳،۷۶۳ موردواژه برای زیرپیکره‌ی خبری، و ۳،۹۹۵ موردواژه برای زیرپیکره‌ی دانشگاهی، جمعاً ۷،۷۵۸ موردواژه (پس از بهنجارسازی). متون به صورت تصادفی گزینش شد. برای نمونه، از بین متون همشهری، بین اعداد شناسه‌ی هر متن، عددی به تصادف انتخاب شد. از آن‌جا که در پیکره‌ی پژوهش‌نامه متون از یکدیگر تفکیک نشده‌اند، جمله‌هایی ۱۰۰-واژه‌ای به تصادف جدا شد. داده‌های هر دو پیکره پیش از تحلیل استعاره بهنجار (نرمال) شدند. در مرحله‌ی بهنجارسازی این اقدامات صورت گرفت: اصلاح و یکسان‌سازی برچسب‌گذاری مقوله‌ی دستوری، یکدست کردن موردواژه‌ها (اصلاح فواصل)، اصلاح نویسه‌های غیرمعیار (مثلاً <ي> و <ك>).

۴.۱ فرهنگ‌های لغت

در فرایند شناسایی استعاره تشخیص معنی پایه‌ای و تمایز آن با معنی بافتی به کمک فرهنگ لغت انجام می‌گیرد؛ آن هم فرهنگ لغتی که پیکره‌بنیاد، عام، و معاصر باشد. اغلب فرهنگ لغت‌های فارسی چنین ویژگی‌هایی را توأمان ندارند؛ یا پیکره‌بنیاد نیستند، یا تخصصی‌اند، یا به معنی تاریخی واژه‌ها توجه دارند. در این پژوهش از فرهنگ زبان‌آموز پیشرفته فارسی (عاصی، ۱۳۹۸) به عنوان فرهنگ لغت اصلی (دارای هر سه ویژگی یادشده) استفاده شد. در مواردی

که فرهنگ لغت اصلی پاسخ مناسبی نداشت یا میان کارشناسان اختلافی بود، از فرهنگ بزرگ سخن (انوری، ۱۳۸۱) و *فارسنت* (شمس‌فرد^۱ و همکاران، ۲۰۱۰) به عنوان فرهنگ لغت‌های کمکی بهره بردیم.

۴.۲ کارشناسان

سه کارشناس برجسب‌گذاری داده‌های پیکره را انجام دادند: (۱) دانشجوی کارشناسی ارشد زبان‌شناسی همگانی از دانشگاه گیلان، ۳۳ ساله، زن، (۲) دانشجوی کارشناسی ارشد زبان‌شناسی همگانی از دانشگاه علامه طباطبائی، ۲۷ ساله، مرد (نگارنده)، و (۳) دانشجوی کارشناسی زبان و ادبیات اسپانیایی از دانشگاه علامه طباطبائی، ۲۳ ساله، مرد. این کارشناسان پیش از برجسب‌گذاری آموزش‌های لازم را فرا گرفتند و پس از آموزش برای ارزیابی دقت آن‌ها، به آزمونی ۶۶-واژه‌ای پاسخ دادند. بین آزمون ۶۶-واژه‌ای و برجسب‌گذاری داده‌های پیکره یک هفته فاصله زمانی بود. کارشناس ۱ و ۲ متون دانشگاهی، و کارشناس ۳ و ۲ متون خبری را برجسب‌گذاری کردند. در واقع کل داده‌های پیکره دو بار جداگانه برجسب خورده است. این کار در دو مرحله صورت گرفت: پیش از بحث، و پس از بحث. در مرحله اول (پیش از بحث)، واحدهای واژگانی تعیین‌شده را طبق روال برجسب زدند (بدون اطلاع از برجسب‌های یکدیگر). در مرحله دوم (پس از بحث)، پس از اتمام برجسب‌گذاری کارشناسان با رجوع چندباره به دستورالعمل‌ها و فرهنگ لغت سعی کردند اختلاف‌ها را تا حد ممکن کم کنند و به تصمیم‌های یکسانی برسند.

کارشناسان بر اساس نسخه فارسی روال (میری، ۱۴۰۲) و دسته‌برجسب معرفی‌شده در آن برجسب‌گذاری کردند. گزیده‌ای از دسته‌برجسب معرفی‌شده در جدول ۱ آمده است. برجسب‌های واحدهای واژگانی (lexunit) شش نوع‌اند: واحد واژگانی ساده (w)، واژه چندجزئی^۲ (p)، فعل مرکب (cv)، نام خاص (n)، و صرف نظر شده (i). برجسب‌های استعاره هفت نوع‌اند: غیراستعاری (not met)، استعاره غیرمستقیم (indirect met)، استعاره مستقیم (direct met)، استعاره ضمنی (implicit met)، نشان استعاره (MFlag)، مورد مرزی (WIDLII^۳)، و صرف نظر تحلیل استعاره (DFMA^۴). در مواردی که بین استعاری/غیراستعاری بودن واحد واژگانی تردید هست، برجسب WIDLII («در صورت شک

1 M. Shamsfard

2 polyword

3 When In Doubt, Leave It In (WIDLII)

4 Discarded For Metaphor Analysis (DFMA)

نگهش دار» به واحد تعلق می‌گیرد. اگر اطلاعات بافتی به قدری ناقص باشد که نتوان معنی بافتی واحد واژگانی را شناسایی کرد، برچسب DFMA به آن تعلق می‌گیرد.

جدول ۱. دسته‌برچسب پیکره استعاره (بر اساس میری، ۱۴۰۲: ۲۸۳، با کمی تغییر)

توضیح	برچسب
واحد واژگانی: ساده (W)، چندجزئی (P)، فعل مرکب (CV)، نام خاص (N)، و صرف نظر شده (i)	lexunit
غیراستعاری	not met
استعاره غیرمستقیم	indirect met
استعاره مستقیم	direct met
استعاره ضمنی	implicit met
نشان استعاره	MFlag
موارد مرزی	WIDLII
صرف نظر شده از تحلیل استعاره	DFMA

۴.۳ پایایی

در ابتدای این بخش سخن از این رفت که برای آزمودن کارایی باید روال را با داده‌های واقعی زبان مواجه کرد تا عیارش را بتوان سنجید. با این حال، صرف رجوع به داده‌ها ملاک اتم نیست، بلکه باید روال را از صافی‌های مختلفی گذراند و آن را پالود. یکی از این صافی‌ها، بررسی پایایی است؛ به این معنی که در صورت تکرار یک آزمایش، میزان شباهت نتایج اندازه‌گیری تا چه اندازه خواهد بود (اوریت^۱ و اسکروندال^۲، ۲۰۱۰). در این پژوهش پایایی شامل بررسی توافق میان کارشناسان است که با دو معیار Q کوکران و کاپا (کاپای کوئن و فلایس) سنجیده می‌شود.

متغیرهای آزمون Q کوکران دو-حالتی‌اند؛ یعنی فقط می‌توانند دو مقدار بگیرند (مثلاً ۰ یا ۱). فرض صفر آزمون این است که تیمارها^۳ (در این پژوهش یعنی برچسب‌های هر کارشناس) هرکدام به یک اندازه مؤثرند. اگر مقدار Q از نظر آماری معنادار باشد، به این معنی است که اختلاف نظر به اندازه‌ای است که پایایی را از حد مطلوب دور می‌کند. این که چه چیزی «از نظر آماری معنادار» است، با پی-مقدار^۴ مشخص می‌شود. برای نمونه، اگر پی-مقدار کوچک‌تر از ۰.۰۵ باشد، از نظر آماری معنادار است (وجود شواهدی برای رد فرض صفر)، و اگر بزرگ‌تر

1 B. S. Everitt
2 A. Skron dal
3 treatments
4 p-value

از ۰.۰۵ باشد، از نظر آماری معنادار نیست (کمبود یا نبود شواهدی برای رد فرض صفر). برای انجام آزمون باید بیش از دو بررسی بر روی داده‌ها انجام گرفته باشد (یعنی درجه آزادی بیش از دو باشد). ضریب کاپا عددی بین ۱- و ۱+ است که برای بررسی میزان توافق میان کارشناسان استفاده می‌شود و دو نوع است: کاپای کوئن^۱ (برای دو کارشناس) و کاپای فلایس^۲ (بیش از دو کارشناس). نظر غالب این است که اگر مقدار کاپا از ۰.۶ بیشتر باشد، میزان توافق در حد متوسط و قابل قبول است. برای این که پایایی بهتری داشته باشیم، این عدد باید ۰.۷ یا بیشتر باشد.

پایایی داده‌ها در سه مرحله بررسی شد: پایایی آزمون ۶۶-واژه‌ای، پایایی پیش از بحث، و پایایی پس از بحث. در مراحل پیش و پس از بحث از Q کوکران استفاده نشد؛ به این دلیل که تعداد برچسب‌گذاران کمتر از سه بود (برای هر واحد واژگانی). برای محاسبه پایایی از نرم‌افزار SPSS نسخه ۲۶ (آی‌بی‌ام ۲، ۲۰۱۹) استفاده شد. برای محاسبه پایایی همچون دیگر نسخه‌های روال، در این پژوهش نیز به موارد استعاری (استعاره غیرمستقیم، مستقیم، و ضمنی) عدد ۱، و به موارد غیراستعاری عدد صفر اختصاص یافت. از آن‌جا که موارد مرزی، نشان استعاره، و موارد صرف نظر شده نمی‌توانند مقدار صفرویکی داشته باشند، در بررسی پایایی این داده‌ها کنار گذاشته شدند (اصلاح این موارد در بخش کاهش حاشیه خطا صورت می‌گیرد). همچنین از همه واحدهای واژگانی‌ای که برچسب نخوردند، اعم از نشانه‌های سجاوندی، اجزای واحدهای چندجزئی، و واحدهای دارای برچسب i صرف نظر شد.

۴.۴ کاهش حاشیه خطا

یکی از نشانه‌های صریح بودن روال شناسایی استعاره این است که حاشیه خطای آن بسیار کم باشد. در پژوهش حاضر داده‌های پس از بحث در ۵ مرحله بررسی شدند: (۱) برچسب‌گذاری واحدهای واژگانی با برچسب w، (۲) برچسب‌گذاری فعل‌های مرکب، (۳) برچسب‌گذاری واحدهای چندجزئی، (۴) بررسی موارد مرزی (WIDLII و DFMA)، و (۵) بررسی واحدهای واژگانی استعاری. در مرحله اول، برچسب واحدهای واژگانی (از نوع w) بررسی شد تا اگر موردی به اشتباه برچسب خورده، اصلاح شود. در مرحله دوم و سوم، واحدهای واژگانی با رجوع به فرهنگ لغت‌ها بار دیگر بررسی شدند. از آن‌جا که ممکن است بسیاری از واژه‌های مرکب نوساخته باشند، در مرحله چهارم موارد مرزی بررسی شدند. در مرحله پنجم، واحدهای واژگانی

1 Cohen's Kappa

2 Fleiss' Kappa

3 IBM Corp.

استعاری برای یافتن اشتباهات سهوی و تخطی از روال بررسی شدند. جزئیات هر مرحله در بخش بعدی آمده است.

۵ تحلیل داده‌ها

در بخش پیشین چگونگی انجام این پژوهش شرح داده شد. این بخش به تفصیل به بررسی داده‌های پیکره و نحوه برچسب‌گذاری اختصاص دارد. در بخش اول (۵.۱)، گزارشی اجمالی از نتیجه تحلیل داده می‌آید و در بخش دوم (۵.۲) به ترتیب موارد اتفاق و اختلاف نظر کارشناسان بررسی خواهد شد. بخش سوم (۵.۳) درباره بررسی موارد مرزی است. در بخش چهارم (۵.۴) درباره آزمون‌های پایایی صحبت خواهد شد. در بخش آخر (۵.۵) سازوکارهای کاهش حاشیه خطا و نتایج نهایی آمده است.

۵.۱ خلاصه نتایج

تحلیل داده‌ها یک بار برای آزمون ۶۶-واژه‌ای، یک بار برای متون خبری (پیش و پس از بحث)، و یک بار برای متون دانشگاه (پیش و پس از بحث) انجام گرفت. در متون خبری ۲۰.۶ (پیش از بحث) و ۲۳.۳ (پس از بحث) درصد از واحدهای واژگانی برچسب استعاری گرفته‌اند. در متون دانشگاهی این میزان به ۲۸.۸ (پیش از بحث) و ۳۰.۴ (پس از بحث) رسیده است. در مجموع (متون خبری و دانشگاهی)، پیش از بحث ۷۰.۲ درصد واحدهای واژگانی غیراستعاری و ۲۴.۵ درصد استعاری است. پس از بحث، ۷۱.۷ درصد از واحدهای واژگانی غیراستعاری و ۲۶.۶ درصد استعاری است. خلاصه‌ای از نتایج تحلیل داده در جدول ۲ آمده است.

جدول ۲. خلاصه نتایج تحلیل داده

درصد موارد (%)		تعداد واحدهای واژگانی	
استعاری	غ‌استعاری	واژگانی	داده‌ها و مراحل
۱۵.۷	۷۶.۵	۵۱	آزمون ۶۶-واژه‌ای
۲۰.۶	۷۴.۵	۳۵۰۷	پیش از بحث
۲۳.۳	۷۴.۹	۳۵۰۶	پس از بحث
۲۸.۸	۶۵.۳	۳۰۹۱	پیش از بحث
۳۰.۴	۶۸.۱	۳۱۰۷	پس از بحث
۲۴.۵	۷۰.۲	۶۵۹۸	پیش از بحث
۲۶.۶	۷۱.۷	۶۶۱۳	پس از بحث

۵.۲ موارد اتفاق و اختلاف نظر کارشناسان

در این بخش برای روشن شدن چگونگی برچسب‌گذاری استعاره چند نمونه از موارد اتفاق و اختلاف نظر میان کارشناسان بررسی خواهد شد. همانطور که گفته شد، در غالب موارد میان کارشناسان اتفاق نظر وجود داشت. با این حال، مواردی پیش آمد که کارشناسان با هم اختلاف نظر داشتند و حتی پس از بحث نیز اختلافات پابرجا ماند. در ادامه به بعضی موارد اشاره خواهد شد. لازم به ذکر است که توضیح تحلیل داده برای اولین مثال‌های اتفاق نظر و اختلاف نظر مفصل‌تر است. برای پرهیز از بیش‌گویی، موارد بعدی مختصر بیان شده‌اند.

واحد واژگانی *توان* در «هنگامی که آن توان از ما گرفته شود، تکلیفی بر ما نیست» یکی از مواردی بود که کارشناسان در برچسب‌گذاری آن اتفاق نظر داشتند و به آن برچسب غیراستعاری اختصاص دادند. در فرهنگ لغت اول این ذیل مدخل *توان* آمده است: «۱- (اسم) نیرو؛ قدرت؛ توانایی، ۲- (اسم) /ریاضی/ تعداد دفعات ضرب کردن یک عدد در خودش که به صورت عدد کوچکی بالای آن عدد نوشته می‌شود، ۳- (اسم) /فیزیک/ مقدار کار انجام‌شده در واحد زمان» (عاصی، ۱۳۹۸: ۲۰۱). معنی شماره ۱ عینی، انسان‌محور، و تقریباً خاص است (معنی پایه‌ای). معنی بافتی نیز با معنی پایه‌ای تفاوت ندارد و نیازی به بررسی تمایز و شباهت نیست (برای اطلاع از نحوه تصمیم‌گیری درباره تمایز و شباهت میان معانی می‌توان به استین و همکاران، ۲۰۱۰ رجوع کرد).

حرکت در «حرکت به سوی رستگاری» نمونه‌ای از استعاره غیرمستقیم است. معنی پایه‌ای مربوط به تغییر مکان در جهان فیزیکی است (توضیح اول و دوم فرهنگ زبان‌آموز). معنی بافتی حرکت، رفتار یا تلاش است (معنی چهارم فرهنگ زبان‌آموز و همچنین معنی ذیل دسته act در فارسی‌نت). تمایز کافی میان این دو معنی وجود دارد، زیرا که توضیح‌ها هر یک شماره‌ای جداگانه در فرهنگ لغت دارند. میان این دو معنی شباهت برقرار است. معنی پایه‌ای به جهان فیزیکی (عینی‌تر و دقیق‌تر) اشاره دارد و معنی بافتی عام‌تر و مبهم‌تر است. پس واحد واژگانی حرکت در این بافت برچسب استعاره غیرمستقیم (indirect met) می‌گیرد.

اکوسیستم در «ارزیابی اکوسیستم کلان‌داده در کشور» در فرهنگ زبان‌آموز مدخلی ندارد. برای این واژه در فرهنگ بزرگ سخن یک معنی بیشتر ذکر نشده است: «مجموعه گیاهان و جانورانی که در ناحیه‌ای خاص زندگی می‌کنند، همراه با روابط پیچیده بین آنها و محیط زیستشان» (انوری، ۱۳۸۱: ۵۱۹). معنی پایه‌ای همین است. *اکوسیستم* در این بافت به معنی وضعیت یک حرفه خاص است. بنابراین، می‌توان بین محیط زندگی گیاهان و جانوران (عینی‌تر) و وضعیت حرفه‌ای (انتزاعی‌تر) نوعی شباهت دید. کارشناسان به *اکوسیستم* برچسب استعاره غیرمستقیم داده‌اند.

میری | ۲۰۱

شایع‌ترین اختلاف میان کارشناسان در تعیین معنی پایه‌ای حروف اضافه است. برای نمونه، کارشناسان در تعیین معنی پایه‌ای به و برای اختلاف نظر داشتند. پژوهش‌های معدودی درباره معنی حروف اضافه فارسی انجام شده است (مثلاً گلفام و همکاران، ۱۳۸۸ و میرزائی، ۱۳۹۸ درباره معنی حرف اضافه /ز؛ مختاری و رضایی، ۱۳۹۲ درباره معنی حرف اضافه با؛ رضویان و خانزاده، ۱۳۹۳ درباره معنی حرف اضافه به؛ خانزاده، ۱۳۹۲ درباره معنی حروف اضافه روی، /ز، در، به، و تا). این پژوهش‌ها جدا از این که کم‌تعدادند، به دلیل این که بررسی‌شان موردی است و چارچوب‌های نظری متفاوتی دارند، هنوز نمی‌توانند مرجع مناسبی برای تعیین معنی پایه‌ای حروف اضافه فارسی باشند.

واحد واژگانی شامل در «تمامی مقدمات شامل وجود موارد زیر است» یکی دیگر از موارد اختلاف میان کارشناسان بود. در فرهنگ زبان‌آموز تنها یک معنی برای آن ذکر شده است: «دارای چیزی؛ حاوی؛ مشتمل بر» (عاصی، ۱۳۹۸: ۴۳۴). کارشناس ۱ معنی پایه‌ای را دارا بودن چیزی عینی و فیزیکی می‌دانست و اگر معنی بافتی به دارا بودن چیزی انتزاعی اشاره داشته باشد، برچسب او استعاره غیرمستقیم بود. اما استدلال کارشناس ۲ این بود که در هیچ‌یک از فرهنگ لغت‌ها اشاره‌ای به دارا بودن چیزی عینی و فیزیکی نشده است. به همین دلیل کارشناس ۲ به این واحد واژگانی برچسب غیراستعاری داده است. این اختلاف پس از بحث نیز حل نشد.

۵.۳ موارد مرزی

پیشتر گفته شد که در مواردی که کارشناس بر سر استعاری بودن یا نبودن واحد واژگانی به نتیجه‌ای نمی‌رسد، به آن برچسب WIDLII (مورد مرزی) اختصاص می‌دهد. کارشناسان مجموعاً در ۲۶ مورد (۹ مورد در متون خبری و ۱۷ مورد در متون دانشگاهی) تردید داشتند. این موارد در گزارش نتایج پایایی نیامده است و پس از کاهش حاشیه خطا به زیرپیکره اضافه خواهد شد.

تعیین برچسب حال در بافت «در همین حال» یکی از مواردی بود که کارشناسان تردید داشتند. معنی پایه‌ای حال زمان کنونی است، منتها مشخص نیست که حال در این بافت به معنی زمان است یا وضعیت. کارشناس ۲ علاوه بر این به خود معنی پایه‌ای حال نیز شک داشت. به نظر او معنی پایه‌ای می‌تواند وضعیت جسمی و حتی پایه‌ای‌تر از معنی زمانی باشد. تمایل کارشناسان در مجموع به برچسب غیراستعاری بود، اما به دلیل تردیدشان همزمان به آن برچسب مورد مرزی نیز اختصاص دادند. واژه‌های تخصصی و نوساخته متون دانشگاهی در

مواردی کارشناسان را به تردید انداختند. مثلاً رده در «یک رده در نرم‌افزار پروتژه» را کارشناسان به عنوان مورد مرزی برچسب زدند.

۵.۴ آزمون‌های پایایی

هدف از انجام آزمون‌های پایایی در این پژوهش سنجیدن میزان توافق میان کارشناسان در هر مرحله است. ضریب کاپا برای همهٔ مراحل تحلیل داده (آزمون ۶۶-واژه‌ای و پیکره) به دست آمده است، اما Q کوکران، از آنجا که حداقل به سه کارشناس احتیاج دارد، تنها برای مرحلهٔ آزمون ۶۶-واژه‌ای محاسبه شده است.

در آزمون ۶۶-واژه‌ای کارشناسان در ۹۲.۲ درصد موارد با هم اتفاق نظر داشتند و آن موارد اختلاف نیز بیشتر به دلیل عدم تسلط به روال بود. در مجموع داده‌های پیکره (خبری و دانشگاهی با هم) نیز کارشناسان در ۹۴.۷ (پیش از بحث) و ۹۸.۳ (پس از بحث) درصد موارد اتفاق نظر داشتند. شناسایی استعاره در متون خبری و دانشگاهی تقریباً بدون مشکل انجام شد و تنها در موارد اندکی کارشناسان با یکدیگر اختلاف داشتند. در مرحلهٔ برچسب‌گذاری متون خبری (پیش از بحث) از ۳,۵۰۷ واحد واژگانی ۳,۳۳۶ واحد (۹۵.۱ درصد) مورد توافق کارشناسان بود. این تعداد پس از بحث میان کارشناسان و رفع بعضی اختلاف‌ها به ۳,۴۴۴ (۹۸.۲ درصد) مورد از ۳,۵۰۶ واحد واژگانی رسید. در مرحلهٔ اول برچسب‌گذاری متون دانشگاهی (پیش از بحث) ۲,۹۰۸ واحد (۹۴.۱ درصد) از ۳,۰۹۱ واحد واژگانی مورد توافق کارشناسان بود. پس از بحث میان کارشناسان این تعداد به ۳,۰۶۱ (۹۸.۵ درصد) مورد از ۳,۱۰۷ واحد واژگانی رسید.

در آزمون ۶۶-واژه‌ای مقدار کاپای فلیس ۰.۸۴۲ به دست آمد. طبق تفسیر لندیس^۱ و کُک^۲ (۱۹۷۷) میزان توافق «تقریباً عالی» است. Q کوکران نیز ۲.۰۰ به دست آمده و پی-مقدار از ۰.۱ بزرگ‌تر شده است. یعنی شهادی برای رد فرض صفر وجود ندارد. در آزمون دوم، برای متون خبری کاپا ۰.۸۶۳ (پیش از بحث) و ۰.۹۵۲ (پس از بحث) گزارش شد (میزان توافق «تقریباً عالی»). در متون دانشگاهی نیز کاپا ۰.۸۶۳ (پیش از بحث) و ۰.۹۶۵ (پس از بحث) به دست آمد. در مجموع، پیش از بحث موارد اختلاف ۵.۳ درصد بوده است و میزان توافق نیز «قوی» ($\kappa=0.864$) است و از ۶,۶۱۳ واحد واژگانی پس از بحث، کارشناسان در ۹۸.۳ موارد اتفاق نظر داشته‌اند. نتایج نهایی آزمون‌های پایایی و جزئیات هر مرحله را می‌توان در جدول ۳

1 J. R. Landis

2 G. J. Koch

میری | ۲۰۳

مشاهده کرد. علت تغییر تعداد واحدهای واژگانی نسبت به تعداد اولیه دو چیز است: یک، تعیین تکلیف بعضی موارد مرزی و اضافه شدن به داده‌های مرحله آخر؛ و دو، اشتباهات سهوی در تعیین واحدهای واژگانی (این اتفاق در مرحله تحلیل داده افتاده است و نه کاهش حاشیه خطا).

جدول ۳. نتایج تحلیل داده و پایایی پژوهش در هر مرحله ($p > 0.10$, $df=2$)*.

آزمون	داده‌ها و مراحل	تعداد واحدهای واژگانی	درصد موارد (%)		درصد توافق		Q کوکران*	
			غ استعاری	استعاری	کلی	کاپا (κ)		
۱	آزمون ۶۶-واژه‌ای	۵۱	۷۶.۵	۱۵.۷	۹۲.۲	۰.۸۴۲	۲.۰۰	
۲	خبری	پیش از بحث	۳۵۰.۷	۷۴.۵	۲۰.۶	۹۵.۱	۰.۸۶۳	-
		پس از بحث	۳۵۰.۶	۷۴.۹	۲۳.۳	۹۸.۲	۰.۹۵۲	-
	دانشگاهی	پیش از بحث	۳۰۹.۱	۶۵.۳	۲۸.۸	۹۴.۱	۰.۸۶۳	-
		پس از بحث	۳۱۰.۷	۶۸.۱	۳۰.۴	۹۸.۵	۰.۹۶۵	-
	مجموع (دانشگاهی و خبری)	پیش از بحث	۶۵۹.۸	۷۰.۲	۲۴.۵	۹۴.۷	۰.۸۶۴	-
		پس از بحث	۶۶۱.۳	۷۱.۷	۲۶.۶	۹۸.۳	۰.۹۵۹	-

۵.۵ کاهش حاشیه خطا

کاهش حاشیه خطا در پنج مرحله انجام گرفت (همانند استین و همکاران، ۲۰۱۰). در مرحله نخست، واحدهای واژگانی دارای برچسب w بررسی شدند. از این میان ۲۰ مورد (۰.۳ درصد کل پیکره) به اشتباه برچسب خورده بودند. در مرحله دوم، حاشیه خطای فعل‌های مرکب، از میان ۳۱۳ فعل مرکب تنها ۴ مورد (۰.۰۶ درصد کل پیکره) به اشتباه برچسب خورده بود. در مرحله سوم، واژه‌های چندجزئی، ۱۰ مورد (۰.۱۵ درصد کل پیکره) به اشتباه برچسب خورده بودند که اصلاح شدند. در مرحله چهارم موارد مرزی بررسی شدند. از میان ۲۵ مورد مرزی، ۳ مورد اصلاح شدند (۰.۰۴۵ درصد کل پیکره). تغییرات در مرحله پنجم را می‌توان به سه دسته تقسیم کرد: اصلاح ۱۷ مورد (۰.۲۵ درصد کل پیکره) برچسب استعاری (با برچسب اشتباه غیراستعاری)، ۱۱ مورد (۰.۱۶ درصد) غیراستعاری (برچسب اشتباه استعاری غیرمستقیم)، و یک مورد (۰.۰۱۵ درصد) استعاری ضمنی (برچسب اشتباه غیراستعاری). نتیجه نهایی پایایی پس از کاهش حاشیه خطا در جدول ۴ آمده است.

جدول ۴. نتیجه نهایی پایایی پژوهش پس از کاهش حاشیه خطا

آزمون	داده‌ها	تعداد			درصد موارد (%)		توافق کاپا (K)
		واحد‌های واژگانی	غ‌استعاری	استعاری	کل		
۱	پیکره پژوهش	۶۶۱۹	۷۱.۸	۲۶.۸	۹۸.۶	۰.۹۶۴	

۶ نتیجه‌گیری

در مقاله حاضر به منظور آزمودن کارایی روال شناسایی استعاره دانشگاه آمستردام برای ساخت پیکره استعاره فارسی، متونی از گفتمان خبری و دانشگاهی جمع‌آوری و بر اساس نسخه فارسی این روال برچسب‌گذاری شد. در مرحله تحلیل داده، روال در مواجهه با داده‌های واقعی زبان فارسی قرار گرفت و چالش‌ها و ناسازگاری‌ها، اعم از کلی (تعیین معنی واژه‌های نوساخته) و مختص زبان فارسی (تعیین معنی پایه‌ای حروف اضافه فارسی) مشخص شد. پس از این در مرحله ارزیابی پایایی، میزان توافق بین کارشناسان با معیارهای کاپا و Q کوکران سنجیده شد و در پایان، از حاشیه خطا کاسته شد. با توجه به نتایج به‌دست‌آمده، می‌توان مدعی شد که روال شناسایی استعاره برای ساخت پیکره استعاره فارسی مناسب است. این روال با توجه به دستورات عمل‌های صریح و کم‌ابهامی که دارد، امکان بروز اختلاف نظرهای سلیقه‌ای میان کارشناسان را منتفی می‌کند. این روال با توجه به عملکرد خوبش در آزمون‌های پایایی، چه در این پژوهش و چه در باقی پژوهش‌ها در زبان‌های مختلف، و همچنین با توجه به انتشار چند پیکره بر همین بنیاد، روش مطمئنی برای برچسب‌گذاری و تحلیل استعاره است. البته این روال بی‌نقص نیست و همانطور که در بخش قبلی به آن اشاره شد، نیاز به تأمل بیشتر و تغییرات جزئی در دستورات عمل‌ها دارد.

انجام این پژوهش با محدودیت‌هایی نیز همراه بود. مهم‌ترین محدودیت به فرهنگ لغت‌های فعلی فارسی بازمی‌گردد. با وجود این که تلاش شد از فرهنگ‌های پیکره‌بنیاد و معاصر استفاده شود، اما این فرهنگ‌ها نیز نتوانستند به همه نیازهای پژوهش پاسخ بدهند. برای نمونه، مشخص نیست که چه میزان از مدخل‌های فرهنگ سخن به شیوه‌ای پیکره‌بنیاد تدوین شده‌اند. فرهنگ زبان‌آموز پیشرفته فارسی نیز با این که پیکره‌بنیاد و معاصر است، مدخل‌های بسیار کمی دارد. توضیحات مدخل‌ها در مواردی تعیین معنی پایه‌ای واحد واژگانی را با مشکل مواجه می‌کرد. مثلاً ذیل بعضی مدخل‌ها معنی زمانی و مکانی را یکجا و با یک توضیح ذکر می‌کند. چنین کاری تعیین معنی پایه‌ای را بسیار سخت می‌کند و در مواردی امکان تصمیم‌گیری را از کارشناسان می‌گیرد. پس از تحلیل داده، کارشناسان به این نتیجه رسیدند که اگر به جای

فرهنگ لغت‌ها از فارسانت به عنوان فرهنگ لغت اول استفاده می‌کردیم، احتمالاً فرایند برچسب‌گذاری سریع‌تر انجام می‌شد و در مواردی ابهام کم‌تر می‌شد (به دلیل تفکیک بهتر معانی). چنین مشی‌ای را در نسخه‌های متفاوت مبتنی بر روال شناسایی استعاره (مثلاً استعاره تصویری، شرم^۱ و استین، ۲۰۱۸؛ استعاره فیلمی، بُرت-میر^۲، ۲۰۱۹) نیز می‌بینیم. در این روال‌ها برای تعیین معانی از وردنت استفاده می‌شود.

شناسایی استعاره به شیوه‌ای که در این پژوهش بررسی شد، هنوز در میانه راه است. کارآمدی و عیار آن با تحلیل‌های پیش‌تر و از رهگذر برخورد با داده‌های بیش‌تر ارتقا می‌یابد. با توجه به این که روال شناسایی استعاره در پژوهش حاضر توانست عملکرد مطلوبی از خود نشان دهد و نتایج دیگر آزمون‌های پایایی در دیگر زبان‌ها را تکرار کند، علاوه بر این که می‌تواند روش مطمئنی برای ساخت پیکره باشد، از آن می‌توان در پژوهش‌های پیکره‌بنیاد استعاره نیز بهره برد. همچنین، پس از انتشار پیکره‌های مبتنی بر این روال، می‌توان با خودکارسازی فرایند مطالعه استعاره را به سمت تحلیل داده‌های بزرگ‌مقیاس^۳ یا کلان‌داده هدایت کرد.

سپاسگزاری

پژوهش حاضر بدون کمک و راهنمایی این افراد هرگز به انجام نمی‌رسید: دکتر آزاده میرزائی، که در مراحل مختلف پژوهش مرا راهنمایی کردند، و شادی راهب و حسین حضرتی، که داده‌های پیکره را برچسب‌گذاری کردند.

منابع

- انوری، حسن. (۱۳۸۱). فرهنگ بزرگ سخن. تهران: سخن.
- خانزاده، معصومه. (۱۳۹۲). بررسی چندمعنایی برخی حروف اضافه مکانی در زبان فارسی با رویکرد معنی‌شناسی شناختی [پایان‌نامه کارشناسی ارشد]. دانشگاه سمنان.
- رضویان، حسین، و خانزاده، م. (۱۳۹۳). «چندمعنایی حرف اضافه "به" در زبان فارسی با رویکرد معنی‌شناسی شناختی». *مطالعات زبان‌ها و گویش‌های غرب ایران*، ۲(۷)، ۵۷-۷۹.
- عاصی، سید مصطفی. (۱۳۹۸). فرهنگ زبان‌آموز پیشرفته فارسی. تهران: سمت.
- عاصی، سید مصطفی، بختیاری، فرزانه، گلفام، ارسلان، و نعمت زاده، شهین. (۱۴۰۱). «کاربرد شیوه نامه میپفو برای شناسایی استعاره‌های زبان فارسی با تکیه بر داده‌هایی از گفتمان آموزشی». *زبان پژوهی*، ۱۴(۴۲)، ۱۷۳-۲۰۱.

1 E. Šorm

2 L. Bort-Mir

3 large-scale

- علایی ابوذر، الهام، پاک‌نیت، نصرالله، حجت‌پناه، علی‌اصغر، زالی، مجتبی، و آقالویی آغمیونی، محمدهادی. (۱۴۰۰). «معرفی یک پیکره متنی تخصصی: پیکره پژوهش‌نامه». *پژوهش‌های زبان‌شناسی تطبیقی*، ۱۱(۲۲)، ۲۷۱-۲۸۹.
- گلفام، ارسلان، عاصی، سید مصطفی، آفاگل‌زاده، فردوس، و یوسفی راد، فاطمه. (۱۳۸۸). «بررسی حرف اضافه (از) در چارچوب معناشناسی شناختی و مقایسه آن با رویکرد سنتی». *زبان و زبان‌شناسی*، ۵(۱۰)، ۶۹-۸۰.
- مختاری، شهره، و رضایی، حدائق. (۱۳۹۲). «بررسی شناختی شبکه معنایی حرف اضافه با در زبان فارسی». *مجله زبان‌شناسی و گویش‌های خراسان دانشگاه فردوسی مشهد*، ۵(۲)، ۷۳-۹۴.
- میرزائی، آزاده. (۱۳۹۸). «کارکردهای معنایی و دستوری حرف اضافه «از» در زبان فارسی». *زبان‌شناسی*، ۱۵(۲۹)، ۲۱-۳۸.
- میری، محمدسعید. (۱۴۰۲). «ملاک‌های برجسب‌گذاری استعاره: گامی به سوی ساخت پیکره استعاره». *زبان پژوهی*، ۱۴(۴۷)، ۲۶۳-۲۹۳.

- Aleahmad, Abolfazl, Hadi Amiri, Ehsan Darrudi, and Farhad Oroumchian (2009). Hamshahri: A standard Persian text collection. *Knowledge-Based Systems*, 22, 382-387.
- Antloga, Špela (2020). "Metaphor Corpus KOMET 1.0." *Slovenian language resource repository CLARIN.SI*. <http://hdl.handle.net/11356/1293>
- Aristotle. (1995). *Poetics*. Trans. Stephen Halliwell. Harvard University Press. (Original work published 330 B.C.E)
- Badryzlova, Yulia, Natalia Shekhtman, Yekaterina Isaeva, and Ruslan Kerimov (2013). "Annotating a Russian corpus of conceptual metaphor: A bottom-up approach." In *Proceedings of the First Workshop on Metaphor in NLP* (pp. 77-86). Association for Computational Linguistics.
- Bort-Mir, Lorena (2019). *Developing, applying and testing FILMIP: The filmic metaphor identification procedure* [Ph.D. Thesis, Universitat Jaume I]. <http://rgdoi.net/10.13140/RG.2.2.18345.03688>
- Cameron, Lynne Joan & Graham Low (Eds.) (1999). *Researching and Applying Metaphor*. Cambridge University Press.
- Dorst, Aletta G. (2011). *Metaphor in Fiction: Language, Thought and Communication* [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam]. Uitgeverij BoxPress. <https://hdl.handle.net/1871/19629>
- Dorst, Aletta G., W. Gudrun Reijnierse, and Gemma Venhuizen (2013). "One small step for MIP towards automated metaphor identification?: Formulating general rules to determine basic meanings in large-scale approaches to metaphor". *Metaphor and the Social World*, 3(1), 77-99.
- Everitt, Brian & Anders Skronidal (2010). *The Cambridge dictionary of statistics* (4th ed). Cambridge, UK; New York: Cambridge University Press.

- IBM Corp. (2019). *IBM SPSS Statistics for Windows* (26.0) [Computer software]. IBM Corp.
- Kaal, Anna. (2012). *Metaphor in conversation* [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam]. BOXPress. <https://hdl.handle.net/1871/32987>
- Kövecses, Zoltán (2016). "Conceptual metaphor theory." In Elena Semino & Zsófia Demjén (Eds.), *The Routledge Handbook of Metaphor and Language*. Routledge.
- Krennmayr, Tina (2011). *Metaphor in newspapers* [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam]. LOT. <https://hdl.handle.net/1871/19774>
- Krennmayr, Tina & Gerard J. Steen (2017). "VU Amsterdam Metaphor Corpus." In Nancy Ide & James Pustejovsky (Eds.), *Handbook of Linguistic Annotation* (pp. 1053–1071). Springer Netherlands.
- Lakoff, George. (1993). "The contemporary theory of metaphor". In Andrew Ortony (Ed.), *Metaphor and Thought* (2nd ed., pp. 202–251). Cambridge: Cambridge University Press.
- Lakoff, George & Mark Johnson (1980). *Metaphors we live by*. University of Chicago Press.
- Landis, J. Richard & Gary J. Koch (1977). "The Measurement of Observer Agreement for Categorical Data." *Biometrics*, 33(1), 159.
- Lu, Xiaofei & Ben Pin-Yun Wang (2017). "Towards a metaphor-annotated corpus of Mandarin Chinese." *Language Resources and Evaluation*, 51(3), 663–694.
- Nacey, Susan, Aletta G. Dorst, Tina Krennmayr, and W. Gudrun Reijnierse (Eds.) (2019). *Metaphor Identification in Multiple Languages: MIPVU around the world* (Vol. 22). John Benjamins Publishing Company.
- Ortony, Andrew (1993). "Metaphor, language, and thought". In Andrew Ortony (Ed.), *Metaphor and Thought* (2nd ed., pp. 1–16). Cambridge University Press; Cambridge Core.
- Pragglejaz Group (2007). "MIP: A method for identifying metaphorically used words in discourse." *Metaphor and Symbol*, 22(1), 1–39.
- Shamsfard, Mehrnoush, Akbar Hesabi, Hakimeh Fadaei, Niloofar Mansoori, Ali Famian, and Somayeh Bagherbeigi, Elham Fekri, Maliheh Monshizadeh & Seyed Mostafa Assi (2010). "Semi automatic development of FarsNet: The Persian wordnet." *Proceedings of 5th Global WordNet Conference (GWA2010)*.
- Šorm, Ester, and Gerard J. Steen (2018). "Chapter 3. VISMIP: Towards a method for visual metaphor identification." In Gerard J. Steen (Ed.), *Visual metaphor: Structure and process* (pp. 47–88). John Benjamins Publishing Company.
- Steen, Gerard J. (2007). *Finding metaphor in grammar and usage: A methodological analysis of theory and research*. John Benjamins Pub. Co.
- _____ (2011). "The contemporary theory of metaphor—Now new and improved!" *Review of Cognitive Linguistics*, 9(1), 26–64.

- _____ (2016). "Identifying metaphors in language." In Elena Semino & Zsófia Demjén (Eds.), *The Routledge Handbook of Metaphor and Language*. Routledge.
- Steen, Gerard J., Aletta G. Dorst, J. Berenike Herrmann, Anna A. Kaal, and Tina Krennmayr (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. John Benjamins Publishing Company.