

ارسال: ۱۴۰۰/۸/۱۱

پذیرش: ۱۴۰۰/۱۰/۱

10.22034/nf.2026.313192.1120

انواع واحد‌های چندواژه‌ای در زبان فارسی

سعید رفیعی خضری* (پژوهشگر فرهنگستان زبان و ادب فارسی، تهران، ایران)

چکیده: واحدهای تک‌واژه‌ای (واژه‌ها) واحدهایی زبانی هستند که نقش نظم‌دهی و مقوله‌بندی تجارب بشری را در قالب واحدهای نمادین و نام‌گذاری بر مفاهیم تازه بر عهده دارند. در زبان‌های طبیعی، افزون بر واحدهای تک‌واژه‌ای، سازه‌های زبانی‌ای وجود دارند که اگرچه به لحاظ ساخت‌واژی با واحدهای تک‌واژه‌ای مطابقت ندارند، دقیقاً همان نقش را ایفا می‌کنند. این سازه‌ها عبارت‌های چندواژه‌ای هستند، واحدهایی واژگانی که از توالی دو یا بیشتر از دو واژه ساخته می‌شوند و به لحاظ واژگانی، نحوی، معنایی، کاربردشناختی و آماری، نسبت به مؤلفه‌های پایه، در عناصر سازنده خود بانشان هستند. این واحدها بخش قابل‌توجهی از واژگان زبان را به خود اختصاص می‌دهند و، با توجه به ترکیب آزاد واژه‌ها در آنها، ویژگی‌های صوری و نقشی منحصر به فردی پیدا می‌کنند و شامل انواع متنوعی می‌شوند. میزان کاربرد این واحدها و اهمیت آن‌ها در زبان به هیچ وجه کمتر از واحدهای تک‌واژه‌ای نیست، به همین دلیل شناخت و توصیف آنها، نه تنها در حوزه زبان‌شناسی، که در حوزه‌های بسیار دیگری، از جمله گروه‌شناسی، واژه‌شناسی، واژگان‌نگاری، زبان‌شناسی رایانه‌ای، زبان‌روان‌شناسی، زبان‌جامعه‌شناسی و پردازش زبان طبیعی، توجه پژوهشگران را به خود جلب کرده است. در این مقاله، پس از مقدمه، به اهمیت واحدهای چندواژه‌ای، لزوم شناخت و توصیف آن‌ها پرداخته شده، سپس تعاریفی از واحدهای چندواژه‌ای و معیارهایی برای استخراج آن‌ها به دست داده شده و سرانجام، بر پایه اصول و قواعد مطرح‌شده، انواع واحدهای چندواژه‌ای معرفی و طبقه‌بندی شده است. این معرفی و طبقه‌بندی براساس داده‌ها، در چهارچوب زبان امروز و بر اساس مشاهدات هم‌زمانی، صورت گرفته است.

* s_rafieykezry@yahoo.com

کلیدواژه‌ها: واحدهای تک‌واژه‌ای، واحدهای چندواژه‌ای، چندواژگی، واحد واژگانی، انواع واحدهای چندواژه‌ای.

۱. مقدمه

در زبان‌ها «واحدهای واژگانی»^۱ از طریق «قواعد صرفی-نحوی»^۲ معانی را در قالب گروه‌ها و جمله‌ها «رمزگذاری»^۳ می‌کنند. پاسخ این پرسش که «واژه چیست؟» در وهله اول ساده به نظر می‌رسد، اما بسیار پیچیده است. آیا «دست» و «دست‌ها» دو واحد مجزاً هستند یا دوگونه از یک واحد؟ در سنت زبان‌شناسی و «واژگان‌نگاری»^۴ اینها دو «صورت‌واژه»^۵ صرفی از یک واژه (واحد تک‌واژه‌ای)^۶ هستند و تفاوت نحوی و معنایی هر دو، طبق فرایند جمع بستن اسم در زبان فارسی، قابل تشخیص است. حال وضعیت دو واحد زبانی «دست بالا» و «دست پایین» چگونه است؟ برای فارسی‌زبانی که با واحدهای زبانی «دست»، «بالا» و «پایین» به صورت مجزاً آشناست، اما هرگز با دو واحد «دست بالا» و «دست پایین» روبه‌رو نشده، حدس زدن معانی آن‌ها، یعنی «حداکثر» و «حداقل» بسیار سخت است. برای اینکه بتوان معانی این واحدها را بازیابی کرد، باید به شکلی در «واژگان ذهنی»^۷ وجود داشته باشند. واحدهایی نظیر «دست بالا» و «دست پایین»، که مشخصه‌های ویژه‌ای دارند و این مشخصه‌ها از واحدهای سازنده آن‌ها قابل تشخیص نیست، واحدهای چندواژه‌ای^۸ نامیده می‌شوند (رفیعی خضری، ۱۳۹۴، ص ۹۶).

مطالعه و بررسی واحدهای چندواژه‌ای در سال‌های اخیر دامنه گسترده‌ای یافته است، اما سابقه آن به اوایل دوره ساختارگرایی یا حتی پیش‌ساختارگرایی برمی‌گردد (Masini, 2019). وقتی به اصطلاحات گوناگونی نگاه می‌کنیم که زبان‌شناسان برای نامیدن این پدیده وضع کرده‌اند، درمی‌یابیم با پدیده‌ای نامتجانس، چندوجهی و متنوع روبه‌رو هستیم. این تنوع اصطلاحات نشان‌دهنده فقدان سنت و ارتباطی یکپارچه میان حوزه‌های پژوهشی، عدم قطعیتی بنیادی در خصوص جایگاه این واحدها در توانش زبانی^۹ و پیچیدگی‌های مفهومی آن‌هاست (ibid). پژوهشگران در حوزه‌های گروه‌شناسی،^{۱۰} واژه‌شناسی،^{۱۱} واژگان‌نگاری، اصطلاح‌شناسی،^{۱۲}

- | | | |
|--------------------------------|--------------------------|---------------------|
| 1. lexical unit | 2. morphosyntactic rules | 3. encoding |
| 4. lexeme | 5. lexicography | 6. word form |
| 7. mono/ individual word units | 8. mental lexicon | 9. multi word units |
| 10. linguistic competence | 11. phraseology | 12. Lexicology |
| 13. terminology | | |

زبان‌شناسی پیکره‌ای،^{۱۴} زبان‌روان‌شناسی،^{۱۵} زبان‌جامعه‌شناسی،^{۱۶} زبان‌آموزی^{۱۷} و پردازش زبان طبیعی^{۱۸} به تحلیل و مقوله‌بندی این واحدها پرداخته‌اند (ibid).

حتی شرحی اجمالی و گذرا از آنچه صورت گرفته، در قالب پیشینه پژوهش، کاری بس دشوار است. در اینجا مجبوریم کاملاً ناخواسته از برخی درگذریم و تنها به برخی از آنها ارجاع دهیم که بیشتر در راستای هدفمان قرار می‌گیرند.

برای معرفی انواع واحدهای چندواژه‌ای و طبقه‌بندی آن‌ها در بخش چهارم، نخست در بخش دوم، با استفاده از منابعی چند، به اهمیت و لزوم شناخت و توصیف آن‌ها خواهیم پرداخت. در بخش سوم، ضمن ارائه تعاریفی برای این واحدهای زبانی، معیارهایی برای استخراج آن‌ها به دست خواهیم داد. در بخش پنجم، نتیجه‌گیری، نیز به نکات کلی و مهم مطرح شده در بخش‌های قبل اشاره خواهیم کرد.

۲. اهمیت واحدهای چندواژه‌ای، لزوم شناخت و توصیف آن‌ها

شناخت و توصیف زبان طبیعی، تحلیل، تقطیع و پردازش آن، بدون شناخت و توصیف واحدهای چندواژه‌ای، امکان‌پذیر نیست. در هر نظام زبانی‌ای که در آن تحلیل درستی از زبان بشری مد نظر باشد، توجه به واحدهای چندواژه‌ای ضرورت است.

داش اسمال: ... باز خدا پدر این ز بیامرزه... اگر گیر اوناش بیفتی که هرطور بخان می‌چپونن بعد از اینکه، چن دفعه سند به قول خودشون تجدید کردی اُهی تنزیل دادی اُنوق گروی رُ که بالا می‌کشن سهل دو قورت و نیمشون هم باقی است چه خواهی گفت! (رفیعی خضری، ۱۳۹۷، ص ۴۱)؛

این روزها با شیرین گرجی درافتاده‌ام... کاسه آبخوری خانم را من شکسته بودم... به گردن شیرین انداختم هم فحش شنید و هم مجبور شد تاوان بدهد. البته او هم برای من مایه می‌گیرد و می‌بینم که مدام با لیلیا که نقداً لولهنگش پیش خانم خیلی آب می‌گیرد تنگ گوشی با هم حرف می‌زنند. (همان، ص ۱۲۱).

همان گونه که در این مثال‌ها دیده می‌شود، واحدهای چندواژه‌ای واژگانی و نحوی (که با حروف سیاه مشخص شده‌اند)، همانند واحدهای تک‌واژه‌ای، در طیف وسیعی از صورت‌های نحوی و واژگانی (اسمی، قیدی، صفتی، فعلی...) ظاهر می‌شوند و «چندواژه‌ای‌های واژگانی»،^{۱۹} مانند «گیر کسی

14. corpus linguistics

15. psycholinguistics

16. Sociolinguistics

17. language acquisition

18. natural language processing

19. lexical multi-word units

افتادن، بالا کشیدن، دو قورت و نیم، دو قورت و نیم کسی باقی بودن یا برای کسی مایه گرفتن» و جز اینها، به لحاظ معنایی، معنایی غیر از جمع معنای مؤلفه‌های سازنده خود پیدا می‌کنند. اگرچه واحدهای چندواژه‌ای همواره به هم پیوسته هستند، در برخی از آنها، مانند «لولهنگ کسی آب گرفتن» در مثال بالا، امکان وارد کردن سازه‌هایی در میان سازه‌های اصلی وجود دارد. در خصوص اهمیت واحدهای چندواژه‌ای، لزوم شناخت و توصیف آنها می‌توان به نکات زیر اشاره کرد:

بسامد و شمار زیاد: برخی از پژوهشگران واحدهای چندواژه‌ای را، به لحاظ داشتن بسامد و شمار زیاد، با واحدهای تک‌واژه‌ای برابر می‌دانند (Baldwin and Kim, 2010, p 208, quoted from Jackendoff, 1997; Tschichold and Pawels, 2000). برخی دیگر بسامد و شمار واحدهای چندواژه‌ای را نسبت به واحدهای تک‌واژه‌ای حداقل ده به یک می‌دانند و معتقدند در همه فرهنگ‌ها، در زیر هر واحد تک‌واژه‌ای به عنوان مدخل، طبق قاعده، شماری از واحدهای چندواژه‌ای فهرست می‌شوند که در آنها واحدهای تک‌واژه‌ای هسته یا یکی از عناصر سازنده هستند؛ برای مثال، آب در «آب از آب تکان نخوردن»، «آب آوردن»، «آب افتادن»، «آب‌باریکه»، «آب برداشتن» (Melčuk, 1995, p 169). ملکووک معتقد است که مردم با واحدهای چندواژه‌ای صحبت می‌کنند، نه با واحدهای تک‌واژه‌ای (ibid).

نوع^{۲۰}: واحدهای چندواژه‌ای به لحاظ نوع هم دست کمی از واحدهای تک‌واژه‌ای ندارند و توزیع آنها همانند است. افزون بر این، صورت‌های جدیدی از واحدهای چندواژه‌ای نیز همپای توسعه زبان، به وجود می‌آیند، برای مثال: «به بام هوا سقف ساختن»، «انگشت به دهان کسی مالیدن»، «چکش کسی برق ندادن»، «چاشت یک بنگی نشدن»، «مثل شتر کف کرده» و... و بر شمار آنها می‌افزایند (Baldwin and Kim, 2010, p 208, quoted from Gates, 1988; Fazly and others, 2007).

ایجاد تمایزات ظریف معنایی و گسترش واژگانی: واحدهای چندواژه‌ای در ایجاد تمایزهای ظریف معنایی و گسترش واژگانی نقش مهمی دارند و بسیار بعید است زبانی بدون نوعی از این واحدها توسعه یابد. واحدهای چندواژه‌ای در افزایش زبان‌آوری و درک زبان کاربرد گسترده‌ای دارند و گونه کاربرد زبانی بر اثر استفاده از این واحدها بانیشان^{۲۱} می‌شود؛ برای مثال، کاربرد واحدهای چندواژه‌ای ممکن است زبان را کم‌وبیش «غیررسمی»^{۲۲} یا «محاوره‌ای»^{۲۳} کند، مانند «غزل خداحافظی را خواندن» در مقابل «مردن»؛ «مثل شتر کف کرده» در مقابل «وامانده»؛ «چاشت یک بنگی نشدن»

20. type

21. marked

22. informal

23. colloquial

در مقابل «کم و ناچیز بودن» و... یا به لحاظ «سبکی»^{۲۴} و «منطقه‌ای»^{۲۵} تفاوت‌های قابل ملاحظه‌ای ایجاد کند؛ مانند «چشم‌دوختن» در مقابل «زل زدن»؛ «چشم به راه» در مقابل «چشم‌انتظار»؛ «تیر و رد» در مقابل «مرغوب و نامرغوب»؛ «تیر و طایفه» در مقابل «کس و کار» و «فک و فامیل»... (Baldwin and Kim, 2010, p 208, quoted from Fillmore and others, 1988; Liberman and Sproat, 1992; Nunberg and others, 1994; Dirven, 2001).

پردازش یا تقطیع زبان طبیعی، ابهام‌زدایی نحوی، برچسب‌دهی معنایی (به‌ویژه در ترجمه):
پژوهش‌های گسترده‌ای در زمینه‌های زیر صورت گرفته است که حکایت از اهمیت واحدهای چندواژه‌ای، لزوم شناخت و توصیف آن‌ها دارند. پردازش یا تقطیع زبان طبیعی؛ به دست دادن انگاره‌هایی برای شناخت و توصیف واحدهای چندواژه‌ای با هدف روانی و استحکام در به‌کارگیری یا درک بهتر زبان طبیعی؛ درک تأثیرات شناخت واحدهای چندواژه‌ای در ابهام‌زدایی نحوی تا درک معنایی؛ شناخت نقش داده‌های مربوط به واحدهای چندواژه‌ای کاملاً واژگانی شده در ساده‌سازی تحلیل ساختار نحوی جمله‌هایی که در آن‌ها این واحدها به‌کاررفته‌اند؛ پی بردن به اهمیت تأثیراتی که شناخت دقیق واحدهای چندواژه‌ای در برچسب‌دهی معنایی دارد و درک تأثیرات دقیق نحوی و کاربردشناختی واحدهای چندواژه‌ای در ترجمه، به‌ویژه ترجمه ماشینی^{۲۶} (Baldwin and Kim, 2010, p 208, quoted from Mazaki and others, 1993; Gerber and Yang, 1997; Matsuo and others, 1997; Dirven, 2001; Baldwin and others, 2004; Piao and others, 2003).

چنان‌که در این مرور کوتاه دیده می‌شود، شناخت و توصیف واحدهای چندواژه‌ای، به دلایلی بسیار، اهمیت ویژه‌ای دارد، از جمله بسامد و شمار زیاد؛ داشتن صورت‌های صرفی و نحوی گوناگون همانند واحدهای تک‌واژه‌ای؛ نقش مهم در پردازش و تقطیع زبان طبیعی؛ آموزش زبان و توسعه‌ی فراگیری زبان دوم؛ روانی و استحکام در به‌کارگیری زبان یا درک بهتر آن؛ ابهام‌زدایی نحوی و درک معنایی و نشان دادن تمایزات سبکی و بانسان ساختن گونه‌های کاربردی زبان؛ و ترجمه، به‌ویژه ترجمه ماشینی. در بخش بعد، نخست به چند تعریف پرسامد از واحدهای چندواژه‌ای اشاره می‌کنیم، تعاریفی که مورد توافق اکثر زبان‌شناسان فعال در عرصه شناخت و توصیف این واحدهاست، و سپس معیارهایی چند برای استخراج واحدهای چندواژه‌ای به دست می‌دهیم.

24. stylistically

25. regionally

26. machine translation(MT)

۳. تعریف واحدهای چندواژه‌ای و معیارهایی برای استخراج آن‌ها

چندواژگی پدیده‌ی واحدی نیست و بر مبنای اینکه بر کدام یک از جنبه‌های آن تأکید شود یا اصلاً چه سازه یا چه گروه‌واژه‌هایی^(۱) چندواژه تلقی شوند تعریف آن تفاوت می‌کند (Villavicencio and Kordoni, 2012). واحدهای چندواژه‌ای ترکیب‌های متداول و تکراری واژه‌ها در زبان روزمره‌اند. در این تعریف، بر ویژگی‌های بسامد و ترکیبی بودن تأکید شده است (Firth, 1957). بسیاری از پژوهشگران در عرصه مطالعات چندواژه‌ای‌ها واحدهای چندواژه‌ای را واحدهایی می‌دانند که ویژگی «تجزیه‌پذیری»^(۲) داشته باشند و به لحاظ واژگانی، نحوی، معنایی، کاربردشناختی و آماری «بانشان» باشند (Baldwin, 2004; Kim, 2005; Baldwin and Kim, 2010; Chakraborty, 2010).

در این تعریف «تجزیه‌پذیری» ناظر بر قابلیت تجزیه شدن به واژه‌های مستقل است؛ برای مثال: «حاجی مقوّا» (کنایه از آدم نحیف و لاغر) قابلیت تجزیه شدن به واژه‌های «حاجی» و «مقوّا» را دارد و به این اعتبار چندواژه‌ای تلقی می‌شود یا «ازکاردرآمده» (ماهر و باتجربه) قابلیت تجزیه شدن به واژه‌های «از»، «کار»، «در» و «آمده» را دارد و «بانشانی» ناظر بر «انحراف»^(۳) از ویژگی‌های پایه در اجزای سازنده است، انحرافی که ممکن است تنها در یکی از سطوح فوق اتفاق بیفتد یا در همه سطوح؛ برای مثال: «ازمابهران» به لحاظ نحوی بانشان است، زیرا در ماهیت صفتی است که از ترکیب غیرعادی حرف اضافه «از» و ضمیر «ما» و صفت تغییرمقوله‌یافته و جمع‌بسته‌شده «بهرتر» ساخته شده است. به لحاظ معنایی بانشان است، زیرا معنای کل از معنای تک‌تک اجزا قابل پیش‌بینی نیست، یعنی معنای ۱. موجودات نامرئی، مانند جن و پری، ۲. آنان که از جهت وضع زندگی و رفاه از دیگران برترند، ۳. آنان که، به علت داشتن ثروت، مقام یا خویشاوندی با صاحبان مقام و مانند آن‌ها، مورد توجه خاص قرار می‌گیرند (انوری، ۱۳۸۱، ذیل «ازمابهران»). به لحاظ کاربردشناختی نیز بانشان است، زیرا در بافت‌ها یا موقعیت‌های خاص به کار می‌رود.

در برخی پژوهش‌ها، در بخش ویژگی‌های زبانی واحدهای چندواژه‌ای، نسبتاً مفصل به ویژگی «بانشانی» پرداخته شده است. در این پژوهش‌ها «ترکیب‌پذیری»^(۴) را مفهومی نزدیک به بانشانی در نظر می‌گیرند و آن را میزان پیش‌بینی ویژگی‌های کلی واحدهای چندواژه‌ای بر مبنای ویژگی اجزای آن‌ها می‌دانند. هرچه این ویژگی‌ها بیشتر قابل پیش‌بینی باشند، میزان ترکیب‌پذیری بیشتر و میزان بانشانی کمتر است و برعکس (Baldwin and Kim, 2010, p 269-273)؛ برای مثال، میزان ترکیب‌پذیری در «آب روی آتش ریختن» بیشتر از میزان آن در «آب توی / در دل کسی تکان نخوردن» است، یعنی معنای «آب روی آتش ریختن» از معنای اجزای سازنده آن بیشتر قابل دست‌یابی است تا معنای «آب توی / در

دل کسی تکان نخوردن»، در نتیجه «آب روی آتش ریختن» نسبت به «آب توی / در دل کسی تکان نخوردن» بیشتر ترکیب‌پذیر و کمتر بانشان است.

اگرچه ترکیب‌پذیری در اکثر پژوهش‌ها تنها به سطح معنایی اختصاص یافته، این ویژگی به تمام سطوح، از واژگانی و نحوی گرفته تا آماری، قابل تعمیم است. بانشانی واژگانی در هر زبانی زمانی اتفاق می‌افتد که یکی از سازه‌های واحد چندواژه‌ای یا هیچ کدام از سازه‌های آن بخشی از واژگان آن زبان نباشد (ibid)؛ برای مثال، «انشاء‌الله» در زبان فارسی به لحاظ واژگانی بانشان است. بانشانی واژگانی به قطع منجر به بانشانی نحوی و معنایی نیز می‌شود. بانشانی آماری هم با بسامد زیاد و «تثبیت‌شدگی»^{۳۰} همراه است (ibid)؛ برای مثال، «راز و نیاز» واحدی چندواژه‌ای است با بسامد زیاد و صورتی تثبیت‌شده که صورت واژگانی دیگری مانند «نیاز و راز» پیدا نمی‌کند.

برخی از پژوهشگران واحدهای چندواژه‌ای را واحدهایی می‌دانند که از توالی واژه‌ها ساخته می‌شوند و در برخی سطوح تحلیل زبانی به عنوان واحدی منفرد عمل می‌کنند. آن‌ها بر ویژگی‌های خاص،^{۳۱} مانند کاهش شفافیت معنایی-نحوی، کاهش ترکیب‌پذیری معنایی، کاهش انعطاف نحوی، ناقض قواعد نحوی عام بودن، میزان زیاد واژگانی‌شدگی و میزان زیاد «قراردادی بودن»^{۳۲} تأکید می‌کنند (Villavicencio and Kordoni, 2012, quoted from Clazolary and others, 1988). برخی نیز تنها بر جنبه قراردادی بودن واحدهای چندواژه‌ای تأکید می‌کنند. در اینجا قراردادی بودن ناظر بر نوع یا بافت خاص بیان چیزهاست (Manning and Shütze, 1999)؛ برای مثال، «صبح به‌خیر» واحدی قراردادی برای سلام و احوالپرسی است. این نوع کاربرد واحد چندواژه‌ای را به لحاظ کاربردشناختی بانشان می‌کند.

برخی هم واحدهای چندواژه‌ای را واحدهایی با بیش از یک واژه و نیازمند تفسیر می‌دانند که لازمه این تفسیر شناخت تک‌تک واژه‌های سازنده و روابط میان آن‌هاست. اینها همچنین واحدهای چندواژه‌ای را توالی‌های تکرارپذیر می‌دانند و بر «فهرست‌پذیر بودن»^{۳۳} این واحدها انگشت می‌گذارند (Fillmore and others, 1988; Fillmore, 2003; Katamba, Francis and John Stanham, 2006).

پژوهشگرانی هستند که از مشکل بودن فراگیری و کاربرد واحدهای چندواژه‌ای برای سخن‌گویان غیربومی می‌گویند، حتی اگر تسلط کافی بر واژگان عام و قواعد نحوی داشته باشند، و معتقدند که کاربرد واحدهای چندواژه‌ای به کلام روانی و طبیعی بودن می‌بخشد؛ از ترجمه‌ناپذیری

30. institutionalization

31. idiosyncrasy

32. Conventionality

33. listable

واژه به واژه واحدهای چندواژه‌ای می‌گویند و از نامتقارن بودن آن‌ها در میان زبان‌های گوناگون (Villavicencio and Kordoni, 2012)؛ برای مثال، «باکمال تعجب» در زبان فارسی در مقابل «surprisingly» در زبان انگلیسی یا «باتمام وجود» در زبان فارسی در مقابل «heart and soul» در زبان انگلیسی یا «ازپافتاده» در زبان فارسی در مقابل «flat out» در زبان انگلیسی... .

پژوهشگران ویژگی‌های دیگری نیز برای واحدهای چندواژه‌ای برمی‌شمارند که شامل تمام واحدهای چندواژه‌ای نمی‌شود. این ویژگی‌ها عبارت‌اند از قابلیت بازگویی با یک واژه،^{۳۴} قابلیت که از طریق آن واحد چندواژه‌ای با یک واژه بازگو می‌شود، برای مثال: «غزل خداحافظی را خواندن» با «مردن» یا «به چاک زدن» با «گریختن»؛ قابلیت ضرب‌المثلی شدن،^{۳۵} قابلیت که از طریق آن واحد چندواژه‌ای موقعیتی به لحاظ اجتماعی مورد علاقه یا تکرار شونده را توصیف می‌کند یا به‌طور ضمنی شرح می‌دهد این علاقه به سبب شباهت یا ارتباط آن موقعیت با صحنه‌ای است که روابط و چیزهای عینی و ساده را دربرمی‌گیرد، برای مثال: «سگ صاحبش را شناختن» که به صورت کنایی برای توصیف موقعیت‌هایی به کار می‌رود که نشان‌دهنده آشفتگی و شلوغی یا هرج و مرج است یا «سوراخ دعا را گم کردن» که برای توصیف موقعیت‌هایی به کار می‌رود که نشان‌دهنده ناشیگری کردن یا عوضی رفتن راه رسیدن به مقصود است؛ و ویژگی‌های زبرزنجیری یا هموندی^{۳۶} (الگوهای تکیه...)، زمانی که تعیین‌کننده واحد چندواژه‌ای هستند و منجر به معنای خاصی می‌شوند، برای مثال، ترکیبات «اسم+ صفت‌نسی»، زمانی که ناظر بر معنای خاصی باشند، الگوی تکیه‌ای متفاوتی دارند و واحدی چندواژه‌ای را رقم می‌زنند، مانند «گل سرخی» زمانی که معنای آن ناظر بر شخص خاصی باشد (Baldwin and Kim, 2010, p 273-275).

تعاریفی که از برخی منابع برای واحدهای چندواژه‌ای به دست داده شد کم‌وبیش در همه منابع دیده می‌شوند و در واقع ناظر بر ویژگی‌های این گونه واحدها هستند. معیارهای شناخت و استخراج واحدهای چندواژه‌ای بر مبنای همین ویژگی‌ها و در قالب همین تعاریف پا می‌گیرند. افزون بر معیارهای تجزیه‌پذیری، بانسانی، ترکیب (نا)پذیری، تثبیت‌شدگی، قراردادی بودن، قابلیت بازگویی با یک واژه، قابلیت ضرب‌المثلی شدن، قابلیت تغییر گونه کاربردی، فهرست‌پذیری، عناصر زبرزنجیری، می‌توان به معیارهایی نظیر «ثبات نحوی-واژگانی»،^{۳۷} «میزان آشنایی»^{۳۸}، «انعطاف‌پذیری»^{۳۹} و «صورت نوشتاری»^{۴۰} نیز اشاره کرد (Moon, p 8). شایان ذکر است که همه این معیارها پیوستاری

34. single-word paraphrability

35. Proverbiality

36. prosody

37. lexico-syntactic fixedness

38. familiarity

39. flexibility

40. orthographic form

هستند و به یک میزان در همه واحدهای چندواژه‌ای دیده نمی‌شوند؛ برای مثال، ترکیب‌ناپذیر کامل، مانند «اسفنج سبز شدن از کله کسی»؛ نیمه‌ترکیب‌پذیر، مانند «استخوان نرم کردن»؛ و ترکیب‌پذیر، مانند «اشک کسی را در آوردن». اکنون، با استفاده از دو معیار تجزیه‌پذیری و باننشانی (واژگانی، نحوی، معنایی، کاربردشناختی، آماری) چندواژه‌ای بودن چند واحد زبانی را در جدول ۱ به نمایش می‌گذاریم.

جدول ۱ به‌کارگیری چند معیار برای شناخت و استخراج واحدهای چندواژه‌ای

باننشانی					تجزیه‌پذیری	
واژگانی	نحوی	معنایی	کاربردشناختی	آماری		
-	+	+	+	+	+	بی‌برو (و) برگرد
-	+	+	+	+	+	یک‌بخته
+	-	+	+	+	+	خاله‌خان‌باجی
-	-	+	+	+	+	روی داریه ریختن

در جدول ۱، برای مثال، «بی‌برو (و) برگرد» واحدی چندواژه‌ای است، زیرا به واحدهای «بی»، «برو»، «(و)»، «برگرد» تجزیه‌پذیر است؛ به لحاظ واژگانی بی‌نشان است، زیرا اجزای آن در واژگان زبان فارسی وجود دارند؛ به لحاظ نحوی بانشان است، زیرا هیچ یک از اجزای آن قید نیستند، اما کل واحد قید است و در ساخت آن از الگوهای متداول ساخت قید پیروی نشده است؛ به لحاظ معنایی بانشان است، زیرا معنای «بی‌برو (و) برگرد» (بی‌شک، حتماً) حاصل معنای تک‌تک اجزای سازنده آن نیست؛ به لحاظ کاربردشناختی بانشان است، زیرا در بافت‌های خاص به کار می‌رود؛ و به لحاظ آماری نیز بانشان است، زیرا از بسامد بالا و تثبیت‌شدگی برخوردار است (رفیعی خضری، ۱۳۹۴، ص ۱۰۵ و ۱۰۶).

تا اینجا به اهمیت واحدهای چندواژه‌ای و به برخی تعاریف و معیارهایی برای شناخت و استخراج آن‌ها اشاره کردیم. در بخش بعد به معرفی و طبقه‌بندی انواع واحدهای چندواژه‌ای می‌پردازیم.

۴. انواع واحدهای چندواژه‌ای و طبقه‌بندی آن‌ها

واحدهای چندواژه‌ای پدیده واحدی نیستند، اما، همان‌گونه که دیدیم، ویژگی‌های مشترکی دارند و بر مبنای همین ویژگی‌ها، افزون‌بر به دست دادن معیارهایی برای شناخت، معرفی و استخراج، می‌توان انواعشان را مشخص و آن‌ها را طبقه‌بندی کرد. بدیهی است که این امر خود در جهت

شناخت و توصیف بهتر این واحدهای زبانی مؤثر است.

بر همین اساس، انواع واحدهای چندواژه‌ای زیر در زبان فارسی تشخیص پذیر هستند، واحدهایی که کم‌وبیش در زبان‌های دیگر نیز وجود دارند و در منابع گوناگون نیز از آن‌ها نام برده شده است (برای مثال: Baldwin and Kim, 2010; Moon, 1988; Melčuk, 1995; Grant, and Bauer, 2004; Fellbaum, 2007; Hendrickx, Iris, Arália Mendes and Sandra Antunes, 2010): ترکیبات،^{۴۱} هماینها،^{۴۲} جفت‌واژه‌های همپایه،^{۴۳} اصطلاح‌واژه‌ها، کلیشه‌های متداول یا قراردادی،^{۴۴} زبانزدها،^{۴۵} گفته‌ها یا حکم،^{۴۶} تشبیهات،^{۴۷} اصطلاحات،^{۴۸} نام‌ها،^{۴۹} عبارات‌های خارجی،^{۵۰} کنش‌های گفتاری غیر مستقیم،^{۵۱} نقل قول‌ها^{۵۲} و ضرب‌المثل‌ها^{۵۳}.

در زیر نمونه‌هایی چند برای انواع واحدهای چندواژه‌ای به دست داده شده است. برخی نمونه‌ها از باطنی (۱۳۸۱ و ۱۳۹۲)، ذوالفقاری (۱۳۸۸)، صدری‌افشار (۱۳۸۸) و فرهنگ واژه‌های مصوّب (۱۳۹۱) برگرفته شده‌اند:

ترکیبات: جهنّم‌دره، آشفته‌بازار، امیرآخور، بناگوش، پسر بچه؛

هماینها^(۴۱): چراغ قرمز، آب دهان، برق شیطنت، برگ چغندر؛

اصطلاح‌واژه‌ها: به آخر خط رسیدن، تلب شدن، از دنیا بی خبر، سر کسی شیره مالیدن؛

جفت‌واژه‌های همپایه^(۴۲): آزمون و خطا، کم‌وبیش، پس‌وپیش، حساب‌و‌کتاب، وقت‌وبی‌وقت؛

عبارات‌های خارجی: خالی‌الذهن، خارق‌العاده، علی‌الخصوص، فوق‌الذکر، ماشاءالله (معین،

۱۳۷۱، ج ۴، بخش دوم)؛

نام‌ها: عالی‌قاپو، برج طغرل، مسجدشاه، سبزه‌میدان، ولی‌عصر؛

تشبیهات: مثل خر تو گل موندن، مثل سگ پاسوخته، مثل سنگ‌پای قزوین، مثل بازار شام، مثل

فیل و فنجون؛

صورت‌های صرفی: از کوره در رفتن، از کوره دررفته، تب لازم،^(۴۳) تب لازمی،^(۴۴) تنگ‌دل کسی^(۴۵)؛

اصطلاحات: توزیع بریده (ریاضی)، خمیره چرب (فناوری غذا)، رزم‌بانگ (نظامی)، کاشینه

درون‌استخوانی (دندان‌پزشکی)، ژرف‌ساخت (زبان‌شناسی)؛

41. compounds

42. collocations

43. binomials

44. routine formulae

45. catchphrases

46. sayings

47. similes

48. terms

49. names

50. foreign expressions

51. indirect speech acts

52. Quotations

53. proverbs

کلیشه‌های متداول و قراردادی: نفرین‌ها (الهی جزّ جگر بزنی)، دعاها (خدا رحمتش کند)، صلوات‌ها (حق پدر صلوات فرست را بیامزد)، سوگندها (این تن بمیره)، چاووشی‌ها (ندای شوق‌امام غریب می‌آید / ز خاک کربلا بوی سیب می‌آید)، قربان‌صدقه‌ها (قربونش برم که خال داره / حکم سپهسالارداره)، وردها (گوش شیطان کر)، گزاره‌های آغازی و پایانی داستانی (یکی بود یکی نبود)، کلیشه‌های متداول (تا چشم به هم بزنی)، بازی‌سروده‌ها (عموزنجیرباف)، رجزها (به کوری چشم بدخواه)، بازی‌های زبانی (قاضی به غزاف رفت و قضا دید و قضا کرد و غذا خورد)، تهدیدها (بلایی سرت بیارم که آن سرش ناپیدا)، خواب‌گزارای‌ها (مگر در خواب بینی پنبه‌دانه)؛

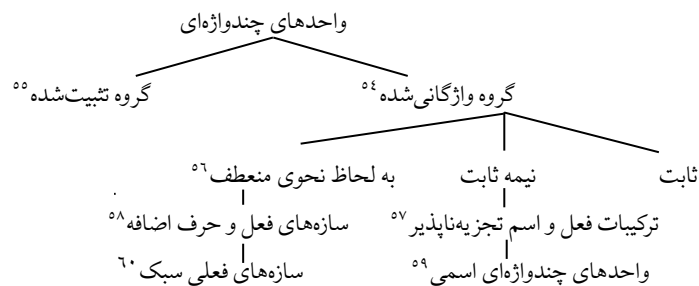
زبانزدها: آره تو بمیری، برو بابا، اشتباه نشه، برو به جهنم، حرف نداره، آره و آلا؛
گفته‌ها یا حکم: خدا گر ز حکمت ببندد دری / ز رحمت گشاید در دیگری، آب دریا را اگر نتوان کشید / هم به قدر تشنگی باید چشید، زبان سرخ می‌دهد سر سبز بر باد، شکسته استخوان داند بهای مومیایی را؛

ضرب‌المثل‌ها: با گرگ دنبه می‌خورد با چوپان گریه می‌کند، خر بیار معرکه بار کن، سرت تو لاک خودت باشه، غوره نشده مویز شده، هر که بامش بیش برفش بیشتر؛
کنش‌های گفتاری غیر مستقیم: در بازه (هوا سرده)، نمک نداره (نمکدان را بده)، زنگ می‌زنند (برود را باز کن)، دارم گوش می‌کنم (صحبت نکن)؛

نقل‌قول‌ها: ممکن است شامل هر نقل‌قولی بشود، اما بیشتر ناظر بر نقل‌قول‌هایی است که یا مخاطبان با آنها آشنا هستند یا از افراد شناخته‌شده‌اند. این نوع نقل‌قول‌ها، به لحاظ بسامد و تثبیت‌شدگی، به پای کلیشه‌های متداول و قراردادی نمی‌رسند و به همین دلیل واژگانی نمی‌شوند و فهرست‌ناپذیر باقی می‌مانند؛ برای مثال، «به قدر کافی زنده و هر چه کمتر فضل‌فروشانه»^(۷) (میرعابدینی، ۱۳۷۸، ص ۱۷).

با توجه به انواع فوق، می‌توانیم در رده اول واحدهای چندواژه‌ای را به دو طبقه تقسیم کنیم: «طبقه واژگانی‌شده»، که در آن واحدهای چندواژه‌ای دارای ویژگی‌های ترکیب‌پذیری، بانشانی، ثبات و تثبیت‌شدگی هستند و فهرست‌پذیر و طبقه «واژگانی‌نشده»، که در آن واحدهای چندواژه‌ای فاقد برخی از این ویژگی‌ها، مانند بانشانی آماری و ثبات، هستند و فهرست‌ناپذیر.

در نمودار زیر (۴.۱)، واحدهای چندواژه‌ای، به‌طور خاص، بر پایه ویژگی‌های معنایی و نحوی طبقه‌بندی شده‌اند (Sag and others 2003; Baldwin and Kim, 2010, p 268, follow Bauer, 2019).



نمودار ۴.۱ طبقه‌بندی واحدهای چندواژه‌ای

در این نمودار، واحدهای چندواژه‌ای در رده اول بر مبنای دو ویژگی واژگانی‌شدگی و تثبیت‌شدگی به دو طبقه و سپس چندواژه‌ای‌های واژگانی‌شده خود، بر مبنای ویژگی ثابت، به سه طبقه «ثابت»، «نیمه‌ثابت» و «به لحاظ نحوی منعطف» تقسیم شده‌اند. در طبقه «ثابت» ها واحدهایی قرار می‌گیرند که هیچ گونه تغییر صرفی - نحوی پیدا نمی‌کنند و هیچ عنصری در ساختار آن‌ها وارد نمی‌شود؛ برای مثال، هرکی هرکی، جاتنگ کن، پشت قبالة کسی انداختن. در طبقه «نیمه‌ثابت» ها واحدهایی قرار می‌گیرند که به لحاظ ترتیب و ترکیب واژه‌ها بسیار مقیدند، اما تغییرات اندکی می‌پذیرند؛ برای مثال، پشت کسی را خالی کردن (پشت کسی را حسابی خالی کردن)، همین آس در کاسه بودن (همین آس و همین کاسه بودن). در طبقه «به لحاظ نحوی منعطف» ها واحدهایی قرار می‌گیرند که دستخوش تغییرات نحوی می‌شوند؛ برای مثال، در/توی پوست خود نگنجیدن، پشت پا به / بر چیزی زدن، دست از سر کسی برداشتن (از سر کسی دست برداشتن)، از چیزی دست شستن (دست از چیزی شستن). قابلیت انعطاف، با توجه به انواع واحدها، به نحو معنی‌داری تغییر می‌کند.

به نظر نگارنده، در نمودار فوق «گروه واژگانی‌شده» می‌تواند «تثبیت‌شده» باشد یا نباشد، همان‌که با عنوان «بانسانی آماری» از آن یاد می‌کنند. از سوی دیگر، همان‌گونه که در بخش ۳ اشاره شد، همه این ویژگی‌ها پیوستاری‌اند و انواع واحدهای چندواژه‌ای می‌توانند به درجاتی نشان‌دهنده هر یک از این ویژگی‌ها باشند. به سخن دیگر، ثابت، نیمه‌ثابت و به لحاظ نحوی منعطف از جنس ویژگی واژگانی‌شده نیستند و نمی‌توانند مبنای طبقه‌بندی واحدهای چندواژه‌ای قرارگیرند؛ برای مثال، اگر

54. lexicalized phrase

55. institutionalized phrase

56. syntactically flexible

57. non-decomposable verb and noun combinations

58. verb-particle constructions

59. nominal MWE

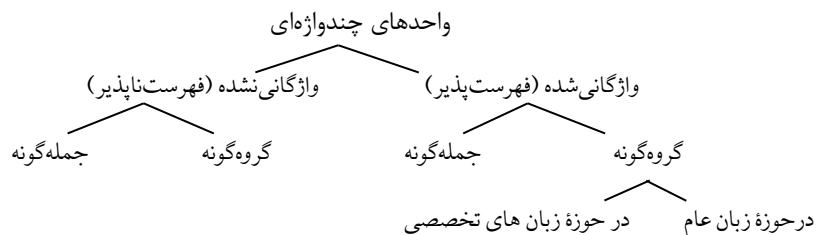
60. Light-verb constructions

ما اصطلاح‌واژه را، که واحدی چندواژه‌ای است، در نظر بگیریم، ممکن است ثابت، نیمه‌ثابت یا به لحاظ نحوی منعطف باشد؛ برای مثال، زاغ‌سیاه کسی را چوب زدن (ثابت)؛ انگشت به دندان گذاردن / گرفتن، انگشت حیرت به دندان گذاردن / گرفتن (نیمه‌ثابت)؛ دست به سروگوش کسی کشیدن / به سروگوش کسی دست کشیدن (به‌لحاظ نحوی منعطف) یا هریک از واحدهای چندواژه‌ای دیگر، مانند «ترکیب‌های فعل و اسم تجزیه‌پذیر»، که در نمودار بدان‌ها اشاره شده است، ممکن است اصطلاح‌واژه باشند یا نباشند.

در ردهٔ اول، می‌توانیم واحدهای چندواژه‌ای را به دو طبقه تقسیم کنیم: طبقهٔ «واژگانی‌شده»، که در آن واحدهای چندواژه‌ای ویژگی‌های ترکیب‌پذیری، بانشانی، ثبات و تثبیت‌شدگی دارند و فهرست‌پذیرند و طبقهٔ «واژگانی‌نشده»، که در آن واحدهای چندواژه‌ای فاقد برخی از این ویژگی‌ها هستند، ویژگی‌هایی مانند بانشانی آماری، ثبات... و فهرست‌ناپذیری.

در ردهٔ دوم، می‌توانیم هر دو طبقهٔ واژگانی‌شده و واژگانی‌نشده را به دو زیرطبقهٔ «گروه‌گونه» و «جمله‌گونه» تقسیم کنیم. در طبقهٔ گروه‌گونه چندواژه‌ای‌هایی قرار می‌گیرند که همانند گروه هستند و در طبقهٔ جمله‌گونه چندواژه‌ای‌هایی همانند جمله. جمله‌گونه‌ها تغییرات نحوی، مانند مجهول‌سازی^{۶۱} و موصولی‌سازی^{۶۲}، را نمی‌پذیرند و تنها در سطح واژگانی ممکن است دستخوش تغییر شوند. این واحدها صورت‌های صرفی هم ندارند.

در ردهٔ سوم، «گروه‌گونه‌های واژگانی‌شده» را می‌توان به دو طبقهٔ «حوزهٔ زبان عام» و «حوزهٔ زبان‌های تخصصی» تقسیم کرد. طبیعی است که در چنین انگاره‌ای بسیاری از نوع‌ها ممکن است با یکدیگر هم‌پوشی داشته یا به هر دو شکل گروه‌گونه و جمله‌گونه وجود داشته باشند. نمای کلی نمودار در شکل ۴.۲ به دست داده شده است.



نمودار ۴.۲ طبقه‌بندی واحدهای چندواژه‌ای

در این نمودار و در طبقه واژگانی شده و گروه‌گونه و در حوزه زبان عام، ترکیبات، هماینها، اصطلاح‌واژه‌ها، جفت‌واژه‌های همپایه، عبارت‌های خارجی غیراصطلاح، نام‌ها، تشبیهات و صورت‌های صرفی و در حوزه زبان تخصصی اصطلاحات قرار می‌گیرند. در طبقه واژگانی شده و جمله‌گونه واحدهای چندواژه‌ای کلیشه‌های متداول و قراردادی، زبانزدها، گفته‌ها یا حکم و ضرب‌المثل‌ها قرار می‌گیرند.

در این نمودار و در طبقه واژگانی نشده و گروه‌گونه واحدهای چندواژه‌ای ترکیبات، هماینها و گروه‌های نحوی و در طبقه واژگانی نشده و جمله‌گونه واحدهای چندواژه‌ای کنش‌های گفتاری غیرمستقیم، نقل‌قول‌ها و بندها (دستوری و کلامی) قرار می‌گیرند.

۵. نتیجه‌گیری

واحدهای واژگانی، از طریق قواعد صرفی-نحوی، معانی را در قالب گروه‌ها و جمله‌ها رمزگذاری می‌کنند. این واحدهای واژگانی ممکن است تک‌واژه‌ای یا چندواژه‌ای باشند. در زبان، افزون بر واحدهای تک‌واژه‌ای، مجموعه گسترده‌ای از واحدهای چندواژه‌ای وجود دارد که بخش قابل توجهی از واژگان زبان را به خود اختصاص می‌دهد. واحدهای چندواژه‌ای واحدهایی هستند که از توالی دو یا بیشتر از دو واژه ساخته می‌شوند و در طیف وسیعی از صورت‌های نحوی و واژگانی ظاهر می‌شوند. این واحدها به لحاظ واژگانی، نحوی، معنایی، کاربردشناختی و آماری نسبت به مؤلفه‌های پایه در عناصر سازنده خود بان نشان هستند و معنایی غیر از جمع معنای مؤلفه‌های سازنده خود پیدا می‌کنند.

از دلایل مهمی که لزوم شناخت این واحدها و انواع آن‌ها را اجتناب‌ناپذیر می‌سازند عبارت‌اند از بسامد و شمار زیاد، داشتن نقش مهم در پردازش و تقطیع زبان طبیعی و در آموزش زبان و توسعه در فراگیری زبان دوم و در ترجمه، به‌ویژه ترجمه ماشینی، روانی و استحکام در به‌کارگیری زبان یا درک بهتر آن، ابهام‌زدایی نحوی و درک معنایی و نشان دادن تمایزهای سبکی و بان نشان ساختن گونه‌های کاربردی زبان.

واحدهای چندواژه‌ای مجموعه‌ای از واحدهای زبانی نامتجانس هستند، اما ویژگی‌های مشترکی دارند، مانند تجزیه‌پذیری، بان‌شانی، ترکیب (نا)پذیری، قراردادی بودن، قابلیت بازگویی با تک‌واژه، قابلیت ضرب‌المثلی شدن، قابلیت تغییر دادن گونه کاربردی زبان، فهرست‌پذیری، صورت نوشتاری و جز اینها که می‌توان بر مبنای این ویژگی‌ها انواع آن‌ها را مشخص کرد. ترکیبات، هماینها، جفت‌واژه‌های همپایه، اصطلاح‌واژه‌ها، کلیشه‌های متداول یا قراردادی،

زبانزدها، گفته‌ها یا حکم، تشبیهات، اصطلاحات، نام‌ها، عبارات‌های خارجی، کنش‌های گفتاری غیرمستقیم، نقل‌قول‌ها و ضرب‌المثل‌ها انواع واحدهای چندواژه‌ای در زبان فارسی هستند که در دو طبقه واژگانی شده (فهرست‌پذیر) و واژگانی نشده (فهرست‌ناپذیر) و به دو شکل گروه‌گونه و جمله‌گونه ظاهر می‌شوند. بدیهی است که امکان هم‌پوشی بسیاری از این نوع‌ها با یکدیگر وجود دارد.

پی‌نوشت‌ها

(۱) پیش‌تر، در دو مقاله «ماهیت گروه‌واژه‌ها: مطالعه‌ای در ساختار گروه‌واژه‌های فارسی» در هشتمین همایش زبان‌شناسی ایران و «آموزش گروه‌واژه‌ها در یادگیری زبان دوم: بحثی نظری» در دومین همایش ملی آموزش زبان فارسی و زبان‌شناسی، معادل «گروه‌واژه» را برای اصطلاح idiom به کار بردم، اما بعدها، بر اثر مطالعات بیشتر، به این نتیجه رسیدم که idiom خود ممکن است یک گروه‌واژه (lexical phrase یا phrasal lexeme) باشد؛ لذا بعد از آن معادل «اصطلاح‌واژه» را، که مقلوب «واژه اصطلاحی» است، برای اصطلاح idiom مناسب‌تر یافتم.

(۲) تفاوت همایندها با ترکیبات در این است که در همایندها معنی یکی از سازه‌ها، بر اساس آنچه با آن همایند می‌شود، تغییر می‌کند؛ برای مثال، واژه چراغ در باهم‌آیی با واژه قرمز علامتی در راهنمایی‌وراندگی می‌شود.

(۳) در مقاله‌های پیشین، برای اصطلاح binomial لفظ دوجزئی را به کار بردم، اما با مطالعه و آشنایی بیشتر با این واحدهای چندواژه‌ای، لفظ جفت‌واژه هم‌پایه را مناسب‌تر یافتم. در فارسی به این واحدها ترکیب عطفی گفته می‌شود که لفظ مناسبی نیست، زیرا این لفظ ناظر بر فرایند ساخت این واحد زبانی است نه نام آن.

(۴) «تب متناوب که دائماً عارض شخص شود، خاصه در بیماری سل». (نجفی، ۱۳۷۸، ص ۲۹۵)

(۵) «(به طعنه و تحقیر) به رنگ زرد». (نجفی، ۱۳۷۸، ص ۲۹۵)

(۶) «قید، نزدیک کسی و تقریباً چسبیده به او». (نجفی، ۱۳۷۸، ص ۳۲۹)

(۷) بازیل پ. نیکیتین (۱۹۶۰-۱۸۸۵)، خاورشناس لهستانی تبار روسی.

منابع

انوری، حسن (۱۳۸۱)، فرهنگ بزرگ سخن، ۸ ج، تهران: سخن.
باطنی، محمدرضا و دستیاران (۱۳۸۸)، فرهنگ همایندها پویا (انگلیسی-فارسی)، تهران: فرهنگ معاصر.
_____ با دستیاری زهرا احمدی‌نیا (۱۳۹۲)، فرهنگ معاصر پویا (اصطلاحات و عبارات رایج: فارسی-انگلیسی)، تهران: فرهنگ معاصر.

ذوالفقاری، حسن (۱۳۸۸)، فرهنگ بزرگ ضرب‌المثل‌های فارسی، ۲ ج، تهران: معین.
رفیعی خضری، سعید (۱۳۹۱)، «ماهیت گروه‌واژه‌ها: مطالعه‌ای در ساختار گروه‌واژه‌های فارسی»، مجموعه مقالات

هشتمین همایش زبان‌شناسی ایران، به کوشش محمد دبیرمقدم، ج ۱، تهران: دانشگاه علامه طباطبایی، صص ۳۲۶-۳۱۵.

_____ (۱۳۹۲)، «آموزش گروه‌واژه‌ها در یادگیری زبان دوم: بحثی نظری»، مجموعه مقالات دومین همایش ملی آموزش زبان فارسی و زبان‌شناسی، شیراز، ۲۵ و ۲۶ اردیبهشت، صص ۱۲۳-۱۳۴.

_____ (۱۳۹۴)، «رده‌شناسی واحدهای چندواژه‌ای در زبان فارسی»، مجموعه مقالات برتر سومین همایش ملی زبان‌شناسی و آموزش زبان فارسی، به کوشش سید محمد حسینی و شهلا شریفی، تهران: نویسه پارسی، صص ۹۵-۱۱۵.

_____ (۱۳۹۷)، فرهنگ اصطلاحات زبانی (در ادبیات داستانی و نمایشی)، ج ۱ (۱۳۰۰-۱۲۵۰ ش)، تهران: فرهنگستان زبان و ادب فارسی.

صدری‌افشار، غلامحسین (۱۳۸۸)، فرهنگ زبانزدهای فارسی، تهران: مازیار.

فرهنگستان زبان و ادب فارسی (۱۳۹۱)، فرهنگ واژه‌های مصوب، دفتر نهم، تهران: نشر آثار فرهنگستان زبان و ادب فارسی.

معین، محمد (۱۳۷۱)، فرهنگ فارسی، ج ۶، تهران: امیرکبیر.

میرعابدینی، حسن (۱۳۸۹)، مروری بر تاریخ‌نگاری ادبیات معاصر (دفتر دوم). مجموعه ادب معاصر فارسی ۷، تهران: فرهنگستان زبان و ادب فارسی.

نجفی، ابوالحسن (۱۳۷۸)، فرهنگ فارسی عامیانه، ج ۲، تهران: نیلوفر.

Baldwin, Timothy (2004) 'Multiword expression,' CSSE. University of Melbourne, www.CSSE.unimelb.edu.au/tim/pubs/altss.

_____ and Su Nam.Kim (2010) 'Multiword expressions,' In Nitin Indurkha and Fred J.Damerau (eds), *Handbook of Natural language Processing*, Second Edition, USA: CRC Press, p 267-292.

Bauer, Laurie (2019), *Complex Lexical Units* (compounds and multi-word expressions), chapter 14, Volume 9 in the series *kowergens Divergens*, De Gruyter.

Chakraborty, Christine (2010), *Multiword Expressions*, India: Jadavpu University Press.

Fellbaum, Christine (2007), *Idioms and Collocations*, London: Continuum.

Fillmore, C (2003), 'An extremist approach to multi-word expressions,' In talk given at RCS, University of Pennsylvania.

Fillmore, C, P. Kay and M. C. O'Connor (1988), 'Regularity and idiomaticity in grammatical

- construction,' *Language* 64, p 501-538.
- Firth, John Rupert (1957) 'A synopsis of linguistic theory 1933-1955,' In *Studies in Linguistic Analysis*, edited by J. R. Firth 1-32, Oxford: Blackwell, p 1-32.
- Grant, L. and L. Bauer (2004) 'Criteria for re-defining idioms: Are we barking up the wrong tree?,' *Applied Linguistics*, 25, p 38-61.
- Green, Speire, Marie-Catherine de Marneffe and Christopher D. Manning. (2012), 'Parsing models for identifying multiword expressions,' *Computational Linguistics*, vol 39, no 1, p 1-33.
- Hendrickx, Iris, Arália Mendes and Sandra Antunes (2010), 'Proposal for multi-word expression annotation in running text,' *Proceeding of The Forth Linguistic Annotation Workshop, ACL 2010*, p 152-156.
- Katamba, Francis and John Stanham. (2006), *Morphology*, Second Edition, New York: Palgrave Macmillan.
- Kim, Su Nam (2005), *Statistical Modeling of Multiword Expressions*, Ph.D Completion Seminar, CSSE. University of Melbourne, Canada.
- Manning, Chris and Hinrich Schütze (1999), *Foundations of Statistical Natural Language Processing*, Cambridge: MIT Press.
- Masini, Francesca (1995), 'Multi-Word Expressions and Morphology,' published online, 2019, <http://cloi.org/10.1093/aerefore/9780199384655.013.611>
- Melčuk, Igor (1995), 'Phraseme in language and phraseology in linguistics' In M. Everaert and the others (eds), *Idioms: Structural and Psychological Perspectives*, Hillsdale, NJ: Lawrence Erlbaum Associates, p 167-233.
- Moon, Rosamond (1998), *Fixed Expressions and Idioms in English*, Oxford. New York: Clarendon Press.
- Sag, Ivan. A., Timothy Baldwin, Ann Copestak and Dan Flickinger (2002), 'Multiword expressions: A pain in the neck for NLP,' In *Proceedings of the 3rd International Conference on Intelligent Text Processing and Computational Linguistics (CIC Ling-2002)*, Mexico City, Mexic, p 1-15.
- Villavicencio, Aline and Valia Kordoni (2012), *Multiword Expressions in Theory and Practice*, Brazil: Federal University of Rio Grande.

Reference

- Anvari, Hasan (2002), *Farhang-e Bozorg-e Sokhan (the Great Dictionary of Sokhan)*, Tehran: Sokhan.
- Batani, Mohammad Rezā et al. (2009), *Farhang-e Hamāyandhā-y-e Pooya (englisi-fārsi) (Pooya Collocation Dictionary (English-Persian))*, Tehran: Farhang-e Mo'āser.
- _____ Assistant: Zahrā Ahmadiān (2013), *Farhang-e Mo'āser Pooya (englisi-fārsi) (Pooya Mo'āser Dictionary (English-persian))*, Tehran: Farhang-e Mo'āser.
- Mirābedini, Hasan (2010), *Moruri bar Tārkhnegāri-e Adabiyat-e Mo'āser (Daftar-e Dovvom) (A Review of Historygraphy of Contemporary Literature (second Book))*, Majmu'e-ye Adab-e Mo'āser-e Fārsi 7, Tehran: Farhangestān-e Zabān va Adab-e Fārsi.
- Mo'in, Mohammad (1992), *Farhang-e Fārsi (A Dictionary of Persian Words)*, Vol 6, Tehrān: Amir kabir.
- Najafi, Abolhasan (1999), *Farhang-e Fārsi-ye Amiyāne (A Dictionary of Persian Colloquialism)*, Tehrān: Niloufar.
- Rafiei khezri, Saeed (2012), "Māhiat-e Goroohvazhehā: Motāle'i dar Sākhtar-e Goroohvāzhehā-ye Fārsi (Study of Idioms: A Study about the Structure of Idioms in Persian)," Majmu'e Maghālāt-e Hashtomin Hamāyesh-e Zabānshenāsi, Ed. Mohammad Dabirmoghadam, Vol 1, Tehran: Daneshgāh-e Alāme-ye Tabātabā'i, p 315-326.
- _____ (2013), "Āmoozesh-e Goroohvazhehā dar Yadgiri-ye Zabān-e Dovvom (Teaching Idioms in Learning Second Language)," Majmu'e Maghālāt-e Dovvomin Hamāyesh-e Mellī-e Āmoozesh-e Zabān-e Fārsi va Zabānshenāsi, Shirāz, 25 & 26 Ordibehesht, p 123-134.
- _____ (2015), "Radeshenāsi-ye Vahedhā-ye Chanvazhe'i dar Zabān-e Fārsi (Typology of Multiword Units in Persian)," Majmu'e Maghālāt-e Bartar-e Sevvomin Hamāyesh-e Mellī-e Āmoozesh-e Zabān-e Fārsi, Ed. Seyed Mohammad Hosaini va Shahlā Sharifi, Tehrān: Nevise-ye Pārsi. p 95-115.
- _____ (2018), *Farhang-e Estelāhāt-e Zabāni, dar Adabiyāt-e Dāstāni va Namāyshi (A Dictionary of Idioms in Fictional and Dramatic Literature)*, Vol 1 (1871-1921), Tehrān: Farhangestān-e Zabān va Adab-e Fārsi.

Sadri Afshar, Gholāmhossein (2009), *Farhang-e Zabānzadha-ye Fārsi (A Dictionary of Persian Catchphrases)*, Tehran: Māziār.

Zolfaghāri, Hasan (2009), *Farhang-e Bozorg-e Zabolmasalhā-ye Fārsi (The Great Dictionary of Persian Proverbs)*, 2 Vols. Tehran: Mo'in.