



پیکره‌های زبانی و ترجمه

شیوه‌ها، مفاهیم و کاربردها

مهسا پهلوان زاده فیینی

مهرداد واشقانی فرامانی



پیکره‌های زبانی و ترجمه

**Corpora and Translation:
Methods, Concepts and Applications**

Mehrdad Vasheghani Farahani
Mahsa Pahlevanzade Fini
© Logos Publications 2023



پیکره‌های زبانی و ترجمه

شیوه‌ها، مفاهیم و کاربردها

مهرداد واشقانی فراهانی

مهسا پهلوان زاده فینی



سرشناسه: واشقانی فراهانی، مهرداد، ۱۳۶۵-

عنوان و نام پدیدآور:

پیکره‌های زبانی و ترجمه: شیوه‌ها، مفاهیم و کاربردها/

مهرداد واشقانی فراهانی، مهسا پهلوان زاده فینی.

مشخصات نشر: تهران، لوگوس، ۱۴۰۲

مشخصات ظاهری: ۱۳۵ ص، ۱۴/۵ × ۲۱/۵ س.م.

شابک: ۹۷۸-۶۲۲-۵۶۱۳-۱۴-۰

وضعیت فهرست‌نویسی: فیبا

موضوع: زبان‌شناسی پیکره‌ای، نرم افزار اسکچ انجین، زبان‌شناسی پیکره‌ای -- داده‌پردازی

شناسه افزوده: پهلوان زاده فینی، مهسا، ۱۳۶۸-

رده‌بندی کنگره: ۲۱۲۸

رده‌بندی دیویی: ۴۱۰/۱۸۸

شماره کتاب‌شناسی ملی: ۹۳۷۴۱۴

پیکره‌های زبانی و ترجمه: شیوه‌ها، مفاهیم و کاربردها

مهرداد واشقانی فراهانی و مهسا پهلوان زاده فینی

ویراستار علمی و ادبی: شیلان شفیعی		طراح جلد: سید مهدی حسینی
چاپ اول:	شابک:	شمارگان:
۱۴۰۲	۹۷۸-۶۲۲-۵۶۱۳-۱۴-۰	۳۰۰



برای اطلاعات بیشتر کد فوق را اسکن کنید و به وبسایت نشر لوگوس بروید.

قیمت: ۱۵۰۰۰۰ تومان

تمامی حقوق این اثر محفوظ است. تکثیر یا تولید مجدد آن کلاً یا جزئاً، به هر شکل اعم از چاپ، فتوکپی، اسکن، صوت، تصویر یا انتشار الکترونیک بدون اجازه مکتوب از نشر لوگوس ممنوع است.

ایمیل: info@irlogos.com

تماس با مرکز پخش: ۰۹۰۲ ۱۵۴ ۰۰۴۲

فروشگاه برخط لوگوس: www.irlogos.com

فهرست

۱.....	مقدمه
۵.....	فصل اول. در آمدی بر زبان شناسی پیکره‌ای
۵.....	مقدمه
۸.....	ویژگی‌های زبان شناسی پیکره‌ای
۹.....	پیکره بنیاد یا مشتق از پیکره؟
۱۱.....	چه افرادی می‌توانند از پیکره‌های زبانی استفاده کنند؟
۱۲.....	پیکره‌های زبانی و ترجمه
۱۴.....	نمونه‌هایی از تحقیقات صورت گرفته در مطالعات ترجمه پیکره محور
۱۶.....	نتیجه‌گیری
۱۷.....	فصل دوم. انواع پیکره و ویژگی‌های آن
۱۷.....	مقدمه
۱۸.....	پیکره‌های مقایسه‌ای
۲۰.....	پیکره‌های عمومی (ناظر)
۲۱.....	پیکره‌های مرجع
۲۲.....	پیکره‌های برچسب خورده
۲۲.....	پیکره‌های تاریخی
۲۳.....	پیکره‌های زبان آموز
۲۴.....	پیکره‌های تخصصی
۲۵.....	پیکره‌های گفتاری
۲۶.....	پیکره‌های یک بار مصرف
۲۷.....	پیکره‌های چندوجهی
۲۷.....	پیکره‌های هم تراز
۲۸.....	نتیجه‌گیری
۲۹.....	فصل سوم. ملاحظات در طراحی و ساخت پیکره‌های زبانی
۲۹.....	مقدمه

۲۹.....	ویژگی های پیکره.....
۳۱.....	زبان واقعی.....
۳۱.....	جمع آوری داده.....
۳۲.....	حجم و اندازه.....
۳۳.....	زمان تولید.....
۳۴.....	ویژگی بازنما بودن.....
۳۵.....	توازن.....
۳۷.....	قابل پردازش بودن.....
۳۸.....	مستندسازی اطلاعات.....
۳۸.....	حالت گفتاری یا نوشتاری.....
۳۸.....	مسائل حقوقی در طراحی پیکره.....
۳۹.....	محدودیت های زبان شناسی پیکره ای.....
۴۱.....	نتیجه گیری.....
۴۳.....	فصل چهارم. شیوه های رایج در داده کاوی پیکره.....
۴۳.....	مقدمه.....
۴۳.....	فهرست بسامد واژگان در پیکره.....
۴۵.....	خطوط واژه نشان در پیکره.....
۴۶.....	واژگان کلیدی در متن.....
۴۷.....	نسبت نوع و نمونه در پیکره.....
۴۸.....	باهم آیی ها.....
۴۹.....	ان-نگاشت ها و خوشه های واژگانی.....
۵۰.....	بررسی گرایش به استفاده از واژه یا عبارت در پیکره.....
۵۱.....	شرح دستوری واژگان در متن.....
۵۲.....	نتیجه گیری.....
۵۳.....	فصل پنجم. طراحی و ساخت پیکره های هم تراز در ترجمه.....
۵۳.....	مقدمه.....

۵۴.....	رژیم جمع‌آوری داده برای طراحی پیکره‌های هم‌تراز.
۵۵.....	ارزیابی پیکره هم‌تراز.
۵۵.....	هم‌تراز کردن متون در زبان مبدأ و مقصد.
۵۷.....	سوگیری و جهت ترجمه در پیکره‌های هم‌تراز.
۶۰.....	ساخت یک پیکره هم‌تراز در نرم‌افزار اسکچ انجین.
۶۸.....	نتیجه‌گیری.
۶۹.....	فصل ششم. آمار در زبان‌شناسی پیکره‌ای.
۶۹.....	مقدمه.
۷۰.....	مفاهیم مقدماتی در آمار.
۷۰.....	داده چیست؟
۷۰.....	انواع داده.
۷۲.....	مفهوم متغیر.
۷۲.....	متغیر مستقل.
۷۳.....	متغیر وابسته.
۷۳.....	متغیر تعدیل‌کننده.
۷۳.....	متغیر کنترل.
۷۴.....	مفهوم بسامد.
۷۴.....	آزمون‌های پارامتریک و ناپارامتریک.
۷۵.....	انواع آزمون‌های پارامتریک.
۷۷.....	ضریب همبستگی پیرسون.
۷۷.....	آزمون‌های ناپارامتریک.
۷۷.....	آزمون کای دو یا کای اسکوئر یا خی دو.
۷۸.....	آزمون یومان-ویتنی.
۷۸.....	آزمون کروسکال-والیس.
۷۸.....	نتیجه‌گیری.
۷۹.....	فصل هفتم. زبان‌شناسی پیکره‌ای و ترجمه در عمل.

۷۹	 مقدمه
۷۹	 محاسبه نوع و نمونه در پیکره هم تراز
۸۲	 معادل یابی و بررسی معادل در پیکره هم تراز
۸۶	 بررسی واژه «مسئولیت» و معادل های آن در پیکره هم تراز
۹۹	 ترجمه باهم آیی ها در زبان انگلیسی و فارسی
۱۰۲	 نتیجه گیری
۱۰۳	 فصل هشتم. معرفی پیکره های تک زبانه و هم تراز در زبان فارسی
۱۰۳	 مقدمه
۱۰۴	 پیکره های زبان فارسی
۱۱۳	 نتیجه گیری
۱۱۵	 فصل نهم: نتیجه گیری
۱۱۷	 فهرست منابع
۱۱۷	 منابع زبان فارسی
۱۱۸	 منابع انگلیسی
۱۲۵	 نمایه

مقدمه

مطالعات ترجمه رشته‌ای نوپاست که به بررسی جنبه‌های مختلف ترجمه مکتوب و شفاهی می‌پردازد. با توجه به حرکت پرجهش تولید علم، همین‌طور پیشرفت‌های صورت‌گرفته این رشته در ایران از جمله برگزاری دوره‌های دکتری و گسترش روزافزون این رشته تحصیلی در دانشگاه‌های مختلف کشور در کنار برگزاری کنفرانس‌های پژوهشی-تحقیقاتی، نیاز به منابع پژوهشی و آموزشی نوشته‌شده به زبان فارسی توسط محققین این رشته بسیار احساس می‌شود. از جمله زمینه‌های تحقیقاتی بکر در حوزه مطالعات ترجمه در ایران، زبان‌شناسی پیکره‌ای و ترجمه یا به عبارتی مطالعات ترجمه پیکره‌محور^۱ است. با توجه به نقش پیکره‌ها در مطالعات زبانی-به‌طور عام- و مطالعات ترجمه-به‌طور خاص- کتاب حاضر با رویکردی آموزشی تلاشی برای معرفی، آموزش و بهره‌مندی علاقه‌مندان به زبان‌شناسی پیکره‌ای در مطالعات ترجمه است. با این حال، سایر علاقه‌مندان به دیگر گرایش‌های رشته‌های زبان انگلیسی، زبان‌شناسی همگانی، زبان‌شناسی کاربردی، آموزش زبان فارسی به غیرفارسی‌زبانان و زبان و ادبیات فارسی نیز می‌توانند از محتوای این کتاب بهره ببرند. همین‌طور، مطالعه کتاب حاضر برای افرادی که به زبان‌هایی غیر از زبان انگلیسی

تسلط دارند، پیشنهاد می‌شود.

کتاب پیش رو از فصل‌های مختلفی تشکیل شده است و خواننده با خواندن کامل کتاب به درکی نسبتاً جامع از زبان‌شناسی پیکره‌ای و مطالعات ترجمه پیکره‌محور خواهد رسید. در حقیقت، کتاب به گونه‌ای طراحی و تدوین شده است که بدون نیاز به مدرس نیز قابلیت کاربردی خود را حفظ کند و خواننده با خواندن دقیق کتاب بتواند در زمینه زبان‌شناسی پیکره‌ای و مطالعات ترجمه پیکره‌محور تحقیق و پژوهش کند.

از جمله الزامات کار در زبان‌شناسی پیکره‌ای بهره‌گیری از نرم‌افزارهای تخصصی جهت تحلیل متون است. به همین منظور، یکی از نرم‌افزارهای تحلیل متون به نام «اسکچ انجین»^۱ که به باور نویسندگان حاضر کامل‌ترین و بهترین نرم‌افزار زبان‌شناسی پیکره‌ای در دنیا محسوب می‌شود، در کتاب پیش‌رو معرفی می‌شود. با این حال، نرم‌افزارهای دیگری نیز وجود دارند که خواننده کتاب می‌تواند از آن‌ها استفاده کند. بنابراین تمام مثال‌ها و نمونه‌های این کتاب بر اساس این نرم‌افزار گردآوری و تدوین شده است. همچنین با هدف تسلط خواننده کتاب به نرم‌افزار، شیوه کار با آن به تفصیل تشریح شده است. ذکر این نکته مهم است که اگر چه این نرم‌افزار دارای قابلیت‌های بسیاری است اما از آنجایی که تأکید کتاب بر روی زبان‌شناسی پیکره‌ای و ترجمه است، عمده قسمت‌های آموزش نرم‌افزار و مثال‌های به کاررفته محدود به تهیه پیکره‌ای هم‌تراز^۲ و تحلیل آن‌هاست. همچنین برای آموزش کار با نرم‌افزار و ارائه مثال‌های گوناگون از نمونه‌های استخراج شده از پیکره‌ای هتمرازشده توسط خود نویسندگان استفاده شده است؛ با این حال نمونه‌هایی از دیگر پیکره‌ها نیز در کتاب یافت می‌شوند که عمدتاً استخراج شده از پیکره ملی بریتانیا^۳ است.

کتاب از فصل‌های مختلفی تشکیل شده است؛ در فصل اول، مقدمه و تاریخچه‌ای کوتاه از زبان‌شناسی پیکره‌ای ارائه می‌شود. هدف از این فصل آشنا نمودن خوانندگان با تاریخچه، مفهوم، انواع رهیافت‌ها در مطالعه پیکره‌ای و

کاربردهای گوناگون زبان‌شناسی پیکره‌ای از جمله در مطالعات ترجمه است. همین‌طور، ویژگی‌های پژوهش‌های حوزه زبان‌شناسی پیکره‌ای نیز در فصل اول معرفی می‌شوند. با خواندن این فصل خواننده با مفاهیمی از جمله «پیکره»، «انواع پیکره» و تاریخچه زبان‌شناسی پیکره‌ای آشنا می‌شود. همین‌طور، در این فصل به تعدادی از پژوهش‌های شاخص صورت‌گرفته در زمینه مطالعات ترجمه پیکره‌محور اشاره خواهد شد تا خواننده بتواند از پژوهش‌های نوین در این حوزه مطلع گردد.

در فصل دوم، انواع و اقسام پیکره‌های زبانی و نحوه استفاده از آن‌ها در متن‌پژوهی و مطالعات زبانی معرفی می‌شوند. هدف از نگارش این فصل، آشنانمودن خواننده با انواع مختلف پیکره‌های زبانی، ویژگی‌ها و نحوه به‌کارگیری آن‌هاست. فصل سوم به نکاتی اشاره دارد که در زمان ساخت و تهیه پیکره‌های زبانی باید به آن‌ها توجه کرد. با خواندن این فصل انتظار می‌رود تا خواننده بتواند مقدمات و الزامات تهیه پیکره‌های زبانی را فراگیرد. فصل چهارم به شیوه‌های رایج در داده‌کاوی می‌پردازد. در این فصل تکنیک‌های مقدماتی در حوزه داده‌کاوی و متن‌پژوهی با استفاده از نرم‌افزار تخصصی زبان‌شناسی پیکره‌ای معرفی می‌شوند. فصل پنجم به تفصیل درباره طراحی و ساخت پیکره‌های هم‌تراز مختص مطالعات ترجمه می‌پردازد. انتظار می‌رود با خواندن این فصل خواننده بتواند مقدمات تهیه پیکره‌های هم‌تراز شده با هدف تحقیق در مطالعات ترجمه را فراگیرد. همین‌طور، نحوه کار با نرم‌افزار اسکچ انجین و تهیه پیکره هم‌تراز در این نرم‌افزار در این فصل توضیح داده می‌شود. فصل ششم کتاب حاضر به مسئله آمار در زبان‌شناسی پیکره‌ای می‌پردازد. در این فصل مفاهیم مقدماتی در آمار از جمله داده و انواع آن، متغیر، بسامد به همراه آزمون‌های پارامتریک و ناپارامتریک توضیح داده خواهد شد. فصل هفتم کتاب پیش‌رو مربوط به زبان‌شناسی پیکره‌ای و ترجمه در عمل است. در این فصل به تفصیل درباره مباحث مختلفی از جمله معادل‌یابی و بررسی معادل در پیکره‌های هم‌تراز اشاره خواهد شد. فصل هشتم نیز به معرفی مهم‌ترین و معروف‌ترین پیکره‌های تک‌زبانه و هم‌تراز شده در زبان فارسی می‌پردازد. هدف از نگارش این فصل معرفی پیکره‌های به‌روز تک‌زبانه یا هم‌تراز شده در زبان فارسی به خواننده است. فصل

نهم که فصل پایانی کتاب است به جمع‌بندی فصول پیشین و نتیجه‌گیری کلی اختصاص دارد.

در پایان از خانم دکتر شیلان شفیع‌ی استادیار مطالعات ترجمه در گروه زبان انگلیسی دانشگاه شهید چمران اهواز، بابت ویراستاری علمی و ادبی این کتاب تشکر و قدردانی می‌شود. امید است کتاب حاضر مورد استقبال خوانندگان قرار بگیرد و به بسط حوزه مطالعات ترجمه پیکره‌محور در زبان فارسی کمک کند. در پایان، هر گونه سؤال، پیشنهاد یا انتقاد در رابطه با محتوای کتاب را می‌توانید به آدرس الکترونیکی نویسنده^۱ ارسال کنید.

مهرداد واشقانی فراهانی

مهسا پهلوان‌زاده فینی

فصل اول

در آمادی بر زبان شناسی پیکره‌ای

مقدمه

امروزه بیشترین حجم داده‌های زبانی با هدف تحقیق و پژوهش با استفاده از ابزارهای الکترونیکی جمع‌آوری و تحلیل می‌شود (Ji, Oakes, Defeng & Hareide, 2017). از جمله شیوه‌های روبه‌پیشرفت در مطالعات زبان، زبان شناسی پیکره‌ای است. به بیان ساده، زبان شناسی پیکره‌ای^۱ شیوه و رهیافتی توصیفی^۲ (Ädel, 2020) در زبان شناسی است که در آن با کمک نرم‌افزارهای خاص کامپیوتری متون گفتاری^۳ یا نوشتاری^۴ را می‌توان در حجم وسیع گردآوری و تحلیل کرد (Zanettin, 2012). از منظر ریشه‌شناسی، واژه پیکره در زبان لاتین به معنای بدن انسان است. در واقع، پیکره^۵ به معنای جمع‌آوری نظام‌مند^۶ و هدف‌مند^۷ داده‌های زبانی با هدف بررسی آن‌ها - با کمک نرم‌افزارهای کامپیوتری - و یافتن پاسخ‌هایی روشن و قابل دفاع است.

-
1. Corpus Linguistics
 2. Descriptive
 3. Spoken
 4. Written
 5. Corpus
 6. Systematic
 7. Purposeful

در تعریفی مشابه می‌توان پیکره را به عنوان مجموعه‌ای بزرگ و هدف‌مند از داده‌های طبیعی زبان گفتار و نوشتار قلمداد کرد که با استفاده از نرم‌افزار زبان‌شناسی پیکره‌ای قابلیت بررسی و جست‌وجو دارد (Cheng, 2012).

اولین بار نلسون فرانسیس بود که واژه پیکره را برای فرایند جمع‌آوری متون به شیوه‌ای نظام‌مند استفاده کرد (Teubert & Čermakova, 2007). با این حال، اولین تلاش‌ها برای طراحی و ساخت پیکره در دهه ۶۰ میلادی صورت گرفت که پیکره براون^۱ در دانشگاه براون^۲ آمریکا تهیه شد. این پیکره با حجمی بالغ بر ۱ میلیون واژه حاوی مطالبی گردآوری شده از متون کتاب‌خانه‌ای (حدوداً ۵۰۰ متن) خود دانشگاه براون بود. متون به کار رفته در این پیکره همگی نثر و هر متن حاوی حدوداً دوهزار کلمه بود. منابع گوناگونی در طراحی پیکره براون استفاده شدند که می‌توان به مطبوعات، کتاب، داستان، رمان و ... اشاره کرد (Sekhar Dash & Arulmozi, 2018). پس از طراحی و تهیه پیکره براون، پیکره‌های دیگری نیز به مرور زمان به وجود آمد که می‌توان به پیکره لانکستر-اوسلو-برگن^۳ اشاره کرد که در دهه ۷۰ میلادی طراحی و ساخته شد. این پیکره که توسط همکاری مشترک دانشگاه‌های لانکستر-اوسلو (انگلستان-نروژ) طراحی شده بود حدوداً ۱ میلیون واژه را دربرمی‌گرفت و هدف طراحان و سازندگان این پیکره ارائه پیکره‌ای همسو با پیکره براون و همین‌طور بررسی زبان انگلیسی بریتانیایی^۴ بود (Leech, 2007).

با وجود طراحی پیکره‌هایی که در بالا به آن‌ها اشاره شد، زبان‌شناسی پیکره‌ای تا اوایل دهه ۸۰ میلادی در حاشیه قرار گرفت و مورد علاقه زبان‌شناسان نبود. علت این امر در دسترس نبودن نرم‌افزارهای کامپیوتری لازم برای همگان و سختی‌های جمع‌آوری داده‌های زبانی بود که کار طراحی و ساخت پیکره‌های زبانی را بسیار سخت، زمان‌بر و پرهزینه می‌کرد. علاوه بر آن تا دهه ۸۰ بیشترین تمرکز زبان‌شناسان بر روی زبان‌شناسی صورت‌نگرا^۵ و بررسی توانش زبانی^۱ بود که بیشتر بر جنبه‌های

1. Brown Corpus

2. Brown University

3. Lancaster-Oslo-Bergen

4. British English

5. Formalist Linguistics

ساختاری و صوری زبان تمرکز داشت؛ حال آنکه اساس زبان‌شناسی پیکره‌ای بیشتر بر مبنای زبان‌شناسی نقش‌گرا^۲ و بررسی کنش‌زبانی^۳ استوار است (مک انری و هاردی، ۲۰۱۲). در همین راستا، همان‌طور که بریر (۲۰۱۱) می‌نویسد، چامسکی زبان‌شناس برجسته آمریکایی، معتقد به بررسی توانش‌زبانی به جای کنش‌زبانی است؛ به این معنا که هر نوع پیکره طبیعی زبان دارای نوعی انحراف است چرا که پیکره‌زبانی اساساً به بررسی توانش‌زبانی نمی‌پردازد، بلکه هدف از آن بررسی الگوهای زبان حقیقی یا همان کنش‌زبانی است.

با این حال، در اواخر دهه ۸۰ میلادی بود که اولین فرهنگ لغت زبان انگلیسی با استفاده از پیکره‌های زبانی تهیه و تدوین شد. در حقیقت، فرهنگ لغات زبان انگلیسی کوپیلد^۴ اولین نسل از فرهنگ‌های انگلیسی‌زبانی بود که در آن برای ارائه نمونه و مثال‌های حقیقی زبان از پیکره بهره می‌بردند. این پیکره توسط همکاری مشترک دانشگاه بیرمنگهام و انتشارات کالینز تحت مدیریت پروفیسور سینکلر^۵ طراحی و تدوین شد. پس از این پیشرفت بزرگ در حوزه زبان‌شناسی پیکره‌ای و با گسترش روزافزون فناوری و علوم کامپیوتر در دهه ۹۰ میلادی پیکره‌هایی بسیار بزرگ‌تری در مقایسه با پیکره براون و پیکره کوپیلد طراحی و ساخته شد (میرزایی، ۱۳۹۶). شاید بتوان گفت که مهم‌ترین پیکره‌ای را در دهه ۹۰ میلادی در زبان انگلیسی طراحی و ساخته شد، پیکره ملی بریتانیا است. این پیکره که از آن به عنوان «جنگل بی‌ان‌سی»^۶ نیز یاد می‌شود در دهه ۹۰ میلادی با همکاری مشترک دو دانشگاه آکسفورد^۷ و لانکستر^۸ انگلستان با همراهی کتابخانه ملی انگلستان به گونه‌ای گردآوری، طراحی و ساخته شد که با بیش از ۱۰۰ میلیون واژه طیف متنوعی از ژانرهای مختلف را در گفتمان نوشتاری و گفتاری پوشش می‌داد.

-
1. Competence
 2. Functionalist
 3. Performance
 4. COBUILD Dictionary
 5. Sinclair
 6. BNC Jungle
 7. Oxford
 8. Lancaster

نمونه دیگر از تأثیر پیشرفت‌های صورت گرفته در علوم کامپیوتر در تهیه پیکره‌های زبانی، پیکره انگلیسی آمریکایی معاصر^۱ است که در دهه ۹۰ میلادی کار تهیه آن آغاز شد و تا سال ۲۰۱۵ ادامه داشت. این پیکره بزرگ بیش از ۴۵۰ میلیون واژه دارد و توسط مارک دیویس، استاد مطرح زبان‌شناسی پیکره‌ای در دانشگاه بریگم یانگ^۲ طراحی و ساخته شده است. این پیکره بیش از ۱۶۰ هزار متن را در خود جای داده است و انتخاب متون آن به گونه‌ای بوده که در بازه زمانی بین سال‌های ۱۹۹۰ تا ۲۰۱۵ میلادی، ۲۰ میلیون واژه از هر سال استخراج و به پیکره اضافه شده است (Cheng, 2012).

پس از آن تاریخ و با گسترش و پیشرفت روزافزون علوم کامپیوتر و فناوری‌های نوین، زبان‌شناسی پیکره‌ای پیشرفت قابل ملاحظه‌ای کرده است (Nevalainen, 2013)؛ به نحوی که امروزه کاربرد زبان‌شناسی پیکره‌ای به حیطه‌هایی خارج از علوم زبان نیز گسترش پیدا کرده و محققان در رشته‌هایی از جمله تاریخ، آموزش، جامعه‌شناسی، حقوق، علوم سیاسی و علوم ارتباطات از امکانات و رهیافت‌های موجود در زبان‌شناسی پیکره‌ای در تحقیقات خود استفاده می‌کنند. همین‌طور، امروزه حجم داده‌های جمع‌آوری شده برای تهیه پیکره‌های زبانی به قدری بزرگ شده است که پیکره‌هایی با حجم و اندازه ۱۰۰ میلیون واژه عادی و متوسط در نظر گرفته می‌شوند (Timmis, 2015).

ویژگی‌های زبان‌شناسی پیکره‌ای

زبان‌شناسی پیکره‌ای ذاتاً بر روی کنش زبانی (استفاده عینی که ما از زبان می‌کنیم) متمرکز است و بر روی توانش زبانی (دانش ذاتی فرد از زبان) تمرکز نمی‌کند (Mautner, 2009). با این وجود، همه مجموعه‌های گردآوری شده از داده‌های زبانی را نمی‌توان الزاماً پیکره زبانی نامید. در واقع، پیکره به عنوان مجموعه‌ای از داده‌های زبانی که به صورت هدفمند جمع‌آوری شده باشد، دارای ویژگی‌های است که آن را از دیگر شیوه‌های جمع‌آوری داده‌های زبانی مثل آرشیو^۳ متمایز می‌کند (Hunston, 2002).

1. Corpus of Contemporary American English

2. Brigham Young University

3. Archive

(2002). آن‌چه که زبان‌شناسی پیکره‌ای را از دیگر شیوه‌های جمع‌آوری داده‌های زبانی متمایز می‌کند، به غیر از جمع‌آوری داده‌های زبانی نوع تحلیلی است که باید بر روی داده‌ها صورت بگیرد؛ به این معنا که تحلیل داده‌ها در زبان‌شناسی پیکره‌ای صرفاً بر اساس استنتاج و تحلیل داده‌های گردآوری شده است، نه خارج از آن (واعظیان، ۲۰۱۴). به طور کلی بایبر و دیگران (۱۹۹۸) موارد زیر را از جمله مهم‌ترین ویژگی‌های زبان‌شناسی پیکره‌ای برمی‌شمردند. ذکر این نکته مهم است که عدم رعایت هر یک از موارد زیر، می‌تواند کل تحقیق پیکره‌ای محقق را با چالش روبه‌رو کند. بنابراین، موارد زیر باید در طراحی، ساخت و تحلیل پیکره توسط محقق کاملاً رعایت شوند.

- زبان‌شناسی پیکره‌ای عملی تجربی است؛ به این معنا که از الگوهای واقعی زبان در بافت طبیعی‌شان^۱ بهره می‌برد.
- زبان‌شناسی پیکره‌ای از مجموعه‌ای بزرگ و هدفمند از متون طبیعی تشکیل شده است.
- زبان‌شناسی پیکره‌ای از کامپیوتر و فناوری‌های تحلیل متن بسیار بهره می‌برد.
- زبان‌شناسی پیکره‌ای از رهیافت‌های تحلیل کمی و کیفی بهره می‌برد.

پیکره‌بنیاد یا مشتق از پیکره؟

به طور کلی تحقیقات در زبان‌شناسی پیکره‌ای به دو نوع عمده تقسیم می‌شوند: تحقیقات پیکره‌بنیاد^۲ و تحقیقات مشتق از پیکره^۳. نخستین بار تونینی-بونلی^۴ در سال ۲۰۰۱ این تفاوت را مطرح کرد. به طور کلی تحقیقات پیکره‌بنیاد در چند مورد با تحقیقات مشتق از پیکره تفاوت دارند (Teubert, 2005). اگر چه تفاوت‌هایی بین این دو رهیافت در زبان‌شناسی پیکره‌ای وجود دارد، بر اساس گفته‌های مک انری، ژیا و تونو (2006، p. 11) «تفاوت آشکار بین دو رهیافت پیکره‌محور و مشتق از پیکره در مطالعات زبانی در حقیقت فازی است». در واقع، نمی‌توان به راحتی هر تحقیق در زمینه زبان‌شناسی پیکره‌ای را در یکی از دو رهیافت موجود قرار داد و در اکثر موارد،

1. Natural Context
 2. Corpus-based
 3. Corpus-driven
 4. Tognini-Bonelli

تحقیق پیکره‌ای می‌تواند هم‌زمان دارای ویژگی‌هایی از هر دو رهیافت باشد.

به طور کلی، تحقیقات پیکره‌محور از بالا به پایین^۱ است؛ به این معنا که محقق با استفاده از نظریه‌های موجود و با کمک پیکره‌های زبانی به رد، پذیرش، بازبینی یا بسط یک نظریه یا فرضیه خاص می‌پردازد (Lan as cited in Sin-wai, 2015). به عبارت دیگر، در شیوه پیکره‌محور، محقق از قبل فرضیه‌ای را بر اساس نظریه‌های موجود مطرح کرده و بر همان اساس سؤال یا سؤالات تحقیق‌اش را ارائه می‌کند. سپس با بررسی داده‌های زبانی موجود در پیکره سؤال یا سؤالات را پاسخ داده و نظریه مطرح شده در تحقیق را رد یا تایید می‌کند. همچنین، در شیوه پیکره‌محور، محقق از قبل می‌داند یا مشخص می‌کند که به دنبال کدام عناصر زبانی در پیکره است؛ به همین دلیل این رویکرد را رویکردی استنباطی می‌نامند (Bibel, Engberg, Ruano & Sosoni, 2019). به عنوان مثال، محقق از قبل مشخص می‌کند که در تحقیق‌اش به دنبال نقش‌نامه‌های گفتمان^۲ یا خوشه‌های واژگانی^۳ است، در واقع، محقق پژوهش را به بررسی عناصر زبانی خاص محدود کرده است. تفاوت سوم بین دو رهیافت پیکره‌محور و مشتق از پیکره در آن است که در اولی معمولاً محقق با پیکره‌های برجسب‌گذاری شده^۴ سروکار دارد. در واقع، محقق از پیکره برجسب‌گذاری شده به این دلیل استفاده می‌کند که می‌داند از قبل دقیقاً به دنبال بررسی چه عناصری در زبان است؛ بنابراین با استفاده از پیکره‌های برجسب‌گذاری شده محقق به دنبال واژگان، عناصر یا ساختارهای خاصی است. به همین دلیل از این نوع رهیافت در زبان‌شناسی پیکره‌ای بیشتر با عنوان «شیوه یا روش‌شناسی»^۵ یاد می‌کنند.

در رهیافت مشتق از پیکره، بر خلاف رهیافت پیکره‌محور، محقق از قبل هیچ فرضیه و نظریه‌ای ندارد؛ اساس این رهیافت این است که محقق بدون هیچ نوع پیش‌فرض ذهنی به بررسی پیکره می‌پردازد. در این رهیافت که اصطلاحاً از پایین به بالاست، پاسخی که محقق در پایان بررسی به دست می‌آورد، به عنوان نظریه یا فرضیه

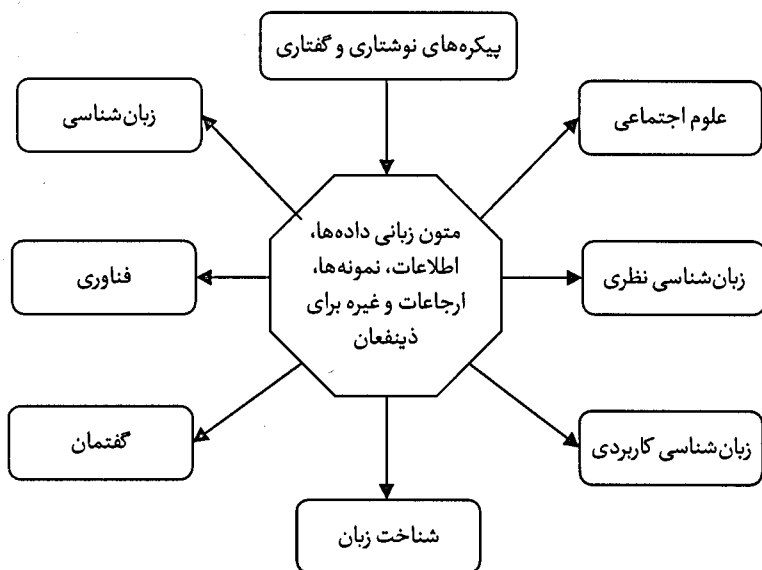
1. Up to down
2. Discourse Markers
3. Lexical Bundles
4. Tagged
5. Methodology

به دست آمده از خود تحقیق ارائه و مطرح می‌شود. همین‌طور، در این رهیافت محقق معمولاً از پیکره‌های برجسته‌گذاری نشده^۱ (خام) استفاده می‌کند. علت این کار آن است که محقق از قبل مشخص نکرده که به دنبال چه عناصر زبانی است و بدون تصمیم از قبل گرفته شده به بررسی پیکره می‌پردازد. بنابراین، این رهیافت را رویکردی اکتشافی یا نوعی فرضیه زبانی^۲ می‌نامند (همان، ۲۰۱۹).

چه افرادی می‌توانند از پیکره‌های زبانی استفاده کنند؟

هر کسی که به دنبال بررسی و تحقیق الگوهای حقیقی زبان در حجمی بزرگ باشد، می‌تواند از یک یا چند نوع پیکره گوناگون زبانی استفاده کند. امروزه محققین با توجه به پیشرفت‌های صورت گرفته در زمینه علوم کامپیوتر و نرم‌افزارهای داده‌کاوی به راحتی می‌توانند انواع و اقسام متون را در حجم‌های متفاوت بررسی کنند. بنابراین، محققین در زمینه‌های گوناگونی از جمله تحلیل و بررسی متون، مقایسه متون در یک یا چند زبان، تحلیل متون تاریخی، تحلیل زبان گفتار و نوشتار، تحلیل انتقادی گفتمان^۳، ماشین‌های ترجمه و زبان‌شناسی رایانشی^۴، تدریس زبان دوم به زبان‌آموزان، تهیه و تدوین کتب آموزشی^۵، نگارش فرهنگ‌های تک‌زبانه و دوزبانه (ترجمه)، واژه‌نگاری^۶ و ادبیات می‌توانند از پیکره‌های زبانی استفاده کنند. با توجه به اهداف مختلف تحقیقاتی، محققین از یک یا چند نوع از پیکره‌های متفاوت بهره می‌جویند که در فصل آینده به تفصیل درباره آن سخن خواهیم گفت. شکل زیر به کاربردهای گوناگون زبان‌شناسی پیکره‌ای و ارتباط این حوزه با دیگر حوزه‌های تحقیقاتی اشاره دارد:

-
1. Untagged
 2. Linguistic Theory
 3. Critical Discourse Analysis
 4. Computational Linguistics
 5. Material Development
 6. Lexicography



شکل ۱. کاربردهای زبان‌شناسی پیکره‌ای (برگرفته از Dash & Arulmozi, 2018)

پیکره‌های زبانی و ترجمه

در طی دو دهه گذشته، کاربرد زبان‌شناسی پیکره‌ای هم در ترجمه عملی و هم در نظریه‌پردازی این رشته نقش بسیار مهمی ایفا کرده است (Zanettin, 2013). در دنیای نظریه‌پردازی مطالعات ترجمه، اولین بار مونا بیکر^۱ در دهه ۹۰ میلادی از زبان‌شناسی پیکره‌ای در تحلیل متون ترجمه‌شده استفاده کرد (Saladanha, 2011). به همین دلیل از مونا بیکر به عنوان مبتکر حوزه مطالعات ترجمه پیکره‌محور نام برده می‌شود. به نظر بیکر، نقش زبان‌شناسی پیکره‌ای در مطالعات ترجمه از آن منظر مهم است که می‌توان با کمک پیکره‌های زبانی به ماهیت متن ترجمه‌شده به عنوان عملی ارتباطی پی برد (Baker, 1993). با هدف انجام تحقیق در همین زمینه، بیکر با طراحی و ساخت پیکره‌ای هم‌تراز^۲ در سطح جمله به بررسی تفاوت‌های موجود در

۱. استاد ممتاز رشته مطالعات ترجمه در دانشگاه منچستر انگلستان

متن ترجمه‌شده و زبان اصلی پرداخت. در واقع، او با تهیه پیکره‌ای انگلیسی-عربی که در سطح جمله هم‌تراز شده بود، نظریه معروف خود با نام جهانی‌های ترجمه^۱ را برای اولین بار مطرح کرد. از نظر بیکر جهانی‌های ترجمه در چهار سطح قابل رویت هستند:

- تصریح ترجمه^۲ به فرایندی اطلاق می‌شود که در آن مترجم اطلاعات تلویحی در متن مبدأ را به صورت آشکارتری در متن مقصد بیان می‌کند. هدف از این کار می‌تواند حذف یا کم‌رنگ کردن فاصله‌های فرهنگی-زبانی در زبان مبدأ و مقصد باشد. به عنوان مثال، افزودن واژگان در ترجمه، ابهام‌زدایی از ضمائر و ساخت‌های نحوی از نمونه‌های این فرایند است.
- ساده‌سازی ترجمه^۳ که می‌تواند در سطح واژگانی، نحوی و معنایی روی دهد؛ به این معنا که به طور متوسط طول جملات در زبان مقصد کوتاه‌تر از زبان مبدأ شود. همین‌طور بر اساس این اصل، تراکم واژگانی در متون ترجمه‌شده نسبت به متن اصلی کمتر است که می‌تواند در نتیجه حذف واژگان تکراری در ترجمه باشد.
- عادی‌سازی ترجمه^۴ به معنای برجسته‌کردن ساخت‌های زبان مبدأ در ترجمه و تطبیق آنها با ساخت‌های زبان مبدأ با ساخت‌های زبان مقصد است. به عبارت دیگر، عناصر و ساخت‌های رایج در زبان مقصد به جای عناصر و ساخت‌های زبان مبدأ استفاده می‌شوند. نتیجه این فرایند، به کار بردن ساخت‌های نحوی و واژگانی متداول‌تر در زبان مقصد است.
- هم‌گرایی^۵ به این معناست که متون ترجمه‌شده دارای ویژگی‌هایی هستند که آن‌ها را از متون غیرترجمه‌ای متمایز می‌کند. به عبارت دیگر، متون ترجمه‌شده از نقطه نظر تراکم واژگانی، نوع ساخت‌های به کار رفته نحوی و غیره به یکدیگر شباهت‌هایی دارند.

1. Translation Universals

2. Explication

3. Simplification

4. Normalization

5. Levelling Out

نمونه‌هایی از تحقیقات صورت گرفته در مطالعات ترجمه پیکره‌محور

با توجه به نقش رو به گسترش پیکره‌های موازی در ترجمه، در چند سال گذشته تحقیقات بسیاری در این حوزه صورت گرفته است (برای نمونه مراجعه کنید به شعبانی، ۲۰۰۸؛ موسوی میانگاه، ۲۰۱۰؛ موسوی میانگاه و محمدی، ۲۰۱۱؛ موسوی رضوی، ۲۰۱۱؛ حسابی، ۲۰۱۲؛ واعظیان و اسفندیاری، ۲۰۱۷؛ زند رحیمی و شریفی مقدم، ۲۰۱۷؛ واشقانی فراهانی و وحید دستجردی، ۲۰۱۹). در این تحقیقات پیکره‌های زبانی هم‌تراز شده نقش اصلی را در پژوهش ترجمه داشته‌اند. در ادامه به تعدادی از این موارد در زبان فارسی خواهیم پرداخت.

حسینی معصومی و غضنفری (۱۳۹۱) در تحقیقی به بررسی کاربرد پیکره‌های زبانی پویا^۱ در ترجمه اینترنتی متون فارسی به انگلیسی پرداختند. آن‌ها برای این کار متون فارسی را با استفاده از نرم‌افزار ترجمه گوگل به عنوان نرم‌افزار معیار بررسی کردند و در پایان به این نتیجه رسیدند که انواع خطاهای زبان‌شناختی و غیرزبان‌شناختی از جمله خطاهای فرهنگی بر ترجمه تولیدشده توسط ماشین‌های ترجمه اثرگذار است. در پژوهشی دیگر، فراهانی (۲۰۱۲) با استفاده از پیکره‌های موازی به بررسی کیفیت ترجمه قرآن کریم پرداخته است. در واقع، هدف او از انجام این تحقیق بررسی ترجمه انگلیسی اسامی الهی در قرآن کریم بوده و برای همین منظور، او ۳ ترجمه انگلیسی قرآن کریم را بررسی نموده و به این نتیجه رسیده که تراکم واژگانی و معنایی اسامی در زبان عربی و نحوه ترجمه آن‌ها مسئله اصلی در ترجمه بوده است. همین‌طور مشخص شد که مترجمان از دو راهکار هم‌معنایی و گسترش واژگانی به عنوان راه‌کارهای اصلی در ترجمه قرآن به زبان انگلیسی استفاده کرده‌اند.

در یکی از جدیدترین پژوهش‌های صورت گرفته در حوزه مطالعات ترجمه پیکره‌محور، واشقانی فراهانی و کاظمیان (۲۰۲۰) با استفاده از پیکره‌ای هم‌تراز شده به بررسی فرایند کم‌تخصیصی معنایی^۲ در ترجمه عناصر فراگفتمانی در گفت‌وگوهای TED پرداختند. آن‌ها با استفاده از پیکره‌ای هم‌تراز شده در زبان‌های انگلیسی و

فارسی به این نتیجه رسیدند که دو راهبرد کلی «ترجمه تحت‌اللفظی و کم‌تخصیصی معنایی» صورت گرفته است. نوع کم‌تخصیصی خود دارای ۳ نوع حذف، توسیع^۱ و تحدید^۲ است.

در همین راستا، واشقانی فراهانی و کاظمیان (۲۰۲۰) به بررسی ترجمه و نحوه شکل‌گیری تعامل بین نویسنده و خواننده در ترجمه عناصر فراگفتمانی پرداختند. با این هدف، آن‌ها پیکره‌ای هم‌تراز شده از ۳۰ گفت‌وگوی ترجمه‌شده در گفت‌وگوهای TED در موضوعات سیاسی را انتخاب کردند. با تهیه پیکره‌ای هم‌تراز شده و بررسی دقیق خطوط واژه‌نشان آن‌ها دریافتند که در ترجمه چهار نوع تغییر عمده رخ می‌دهد از جمله "explicit change" (im)، "information change" (dis)، " (in) visibility change" و "change" (de)emphasis.

در پژوهش پیکره‌محور دیگری، تقوی و هاشمی (۲۰۲۲) به بررسی جهانی‌های ترجمه با تمرکز بر تصریح و ساده‌سازی در ترجمه متون انگلیسی به فارسی پرداختند. آن‌ها، با استفاده از پیکره‌ای مقایسه‌ای^۳ در زبان فارسی و پیکره‌ای هم‌تراز شده در زبان‌های انگلیسی-فارسی، به این نتیجه رسیدند که در رابطه با ساده‌سازی متون ترجمه‌شده به زبان فارسی دارای تراکم واژگانی کمتری نسبت به پیکره زبان مقایسه‌ای زبان فارسی بودند؛ هرچند این تفاوت معنادار نبود. همین‌طور واژگان استفاده‌شده در زبان فارسی ترجمه‌شده به نسبت پیکره زبان فارسی مقایسه‌ای از تنوع بیشتری برخوردار بودند.

در یکی دیگر از تحقیقات صورت گرفته در زمینه مطالعات ترجمه پیکره‌محور، قانع و واشقانی فراهانی (۲۰۲۲) به بررسی نقش‌های گفتمانی و ترجمه نقش‌نمای گفتمانی "You Know" و ترجمه فارسی آن پرداختند. با این هدف، آن‌ها از دو پیکره بهره جستند: یکی پیکره تک‌زبان فارسی با نام «پیکره بیجن‌خان» که نقش پیکره مرجع را داشت و دیگری پیکره هم‌تراز شده انگلیسی-فارسی با نام «پیکره موازی میزان». با بررسی دو پیکره مذکور، آن‌ها به این نتیجه رسیدند که اگرچه نقش‌نمای گفتمانی

"You Know" در زبان انگلیسی دارای ۱۴ نقش در زبان انگلیسی است، اما در زبان فارسی این نقش نما ۱۷ نقش دارد. همین طور با بررسی خطوط واژه‌نشان در پیکره موازی آن‌ها به این نتیجه رسیدند که در ترجمه این نقش نمای گفتمانی از ۶ رهیافت مختلف ترجمه استفاده می‌شود.

نتیجه‌گیری

در فصل اول مطالبی درباره تاریخچه زبان‌شناسی پیکره‌ای، اولین پیکره‌های ساخته شده و همین طور مطالعات ترجمه پیکره‌محور ارائه شد. در کنار این مطالب، تعریف زبان‌شناسی پیکره‌ای و واژه پیکره بیان شد. همین طور، دو رهیافت اساسی زبان‌شناسی پیکره‌بنیاد و مشتق از پیکره به طور کامل معرفی و ویژگی‌های هر یک از آن‌ها تشریح و در کنار آن ویژگی‌های اصلی زبان‌شناسی پیکره‌ای شرح داده شد. سپس به کاربردهای زبان‌شناسی پیکره‌ای در علوم مختلف و مطالعات ترجمه پیکره‌محور و به تعدادی از مطالعات صورت گرفته در این زمینه نیز اشاره شد.