



دارای درجه علمی-پژوهشی

مجموعه حالات روغن همایس ملی
زبان‌شناسی و سوره‌ای

به کوشش: آزاده میرزائی





بہ کوشش
آزادہ میرزائی



سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران

همایش ملی زبان‌شناسی بیکره‌ای (دومین: ۱۳۹۵: تهران)
مجموعه مقالات دومین همایش ملی زبان‌شناسی بیکره‌ای / به کوشش: آزاده میرزانی؛
ایرگزارکننده: انجمن زبان‌شناسی ایران.
تهران: نشر نویسه پارس، ۱۳۹۵.
۳۱۴ ص:، مصور، جدول، نمودار.

۹-۳۴-۷۰۳۰-۶۰۰-۹۷۸

فیبیا

زبان‌شناسی بیکره‌ای -- کنگره‌ها
Corpus (Linguistics) -- Congresses
زبان‌شناسی -- ایران -- کنگره‌ها
Linguistics -- Iran -- Congresses

میرزانی، آزاده، ۱۳۵۸ - ، گردآورنده.
انجمن زبان‌شناسی ایران.

۱۳۹۵ / ۲۵۸ / ۲۱۲۸

۴۱۰/۱۸۸

۴۴۴۷۴۵۵

سرشناسه

عنوان و نام پدیدآور

مشخصات نشر

مشخصات ظاهری

شابک

وضعیت فهرست نویسی

موضوع

موضوع

موضوع

موضوع

شناسه افزوده

شناسه افزوده

رده بندی کنگره

رده بندی دیویی

شماره کتابشناسی ملی

زبان‌شناسی پیکره‌ای

مجموعه مقالات دومین همایش ملی زبان‌شناسی پیکره‌ای

ویراستار چکیده‌های انگلیسی: دکتر معصومه استاجی

طراح جلد، گرافیک، صفحه‌آرایی و ناظر فنی چاپ: محمد محرابی

www.mehraabi.com / 091251693545

ناشر: نشر نویسه پارسی

دفتر انتشارات: ۰۲۱-۷۷۰۵۳۲۴۶

فروشگاه: ۰۲۱-۶۶۹۵۷۱۳۲

سامانه پیام کوتاه: ۳۰۰۰۴۵۵۴۵۵۴۱۴۲

وبگاه: www.neveeseh.com

نوبت چاپ: اول، ۱۳۹۵

شمارگان: ۳۰۰ نسخه

شابک: ۹۷۸-۶۰۰-۷۰۳۰-۳۴-۹

چاپ و صحافی: روز

قیمت: ۳۵۰۰۰ تومان

کلیه حقوق محفوظ و متعلق به «نشر نویسه پارسی» است.
تکثیر و انتشار این اثر با قسمتی از آن به هر شیوه، بدون مجوز قبلی و کتبی
ممنوع و مورد پیگیری قانونی قرار خواهد گرفت.



دارای درجه علمی - پژوهشی نمایه شده در پایگاه استنادی علوم جهان اسلام (ISC)

دارای درجه علمی - پژوهشی، بر اساس بند ۱ ماده ۳ آیین نامه نحوه برگزاری و ساماندهی
همایش‌های علمی ابلاغیه شماره ۳/۱۶۸۲۱۶/۱۳۹۳/۹/۱۱ مورخ ۱۳۹۳/۹/۱۱ وزارت علوم، تحقیقات و فناوری.



مجموعه مقالات دومین همایش ملی زبان‌شناسی بیکرهای

برگزارکننده

انجمن زبان‌شناسی ایران

با همکاری

پژوهشگاه علوم انسانی و مطالعات فرهنگی

پایگاه استنادی علوم جهان اسلام (ISC)

نشر نویسه پارسی

سازمان همایش

دکتر آزاده میرزائی دبیر علمی

فرامرز مژدگانلو دبیر اجرایی

کمیته علمی

دکتر محمود بی‌جن‌خان عضو هیأت علمی دانشگاه تهران

دکتر محمد بحرانی عضو هیأت علمی دانشگاه صنعتی شریف

دکتر پروانه خسروی‌زاده عضو هیأت علمی دانشگاه صنعتی شریف

دکتر ویدا شقاقی عضو هیأت علمی دانشگاه علامه طباطبائی

دکتر مصطفی عاصی عضو هیأت علمی پژوهشگاه علوم انسانی و

مطالعات فرهنگی

دکتر مسعود قیومی عضو هیأت علمی پژوهشگاه علوم انسانی و

مطالعات فرهنگی

دکتر شهرام مدرس خیابانی عضو هیأت علمی دانشگاه آزاد اسلامی واحد

کرج

دکتر آزاده میرزائی عضو هیأت علمی دانشگاه علامه طباطبائی

هیئت داوران

دکتر آریتا افراشی عضو هیأت علمی پژوهشگاه علوم انسانی و

مطالعات فرهنگی

دکتر محمد بحرانی	عضو هیأت علمی دانشگاه صنعتی شریف
دکتر فاطمه بهرامی	عضو هیأت علمی دانشگاه شهید بهشتی
دکتر پروانه خسروی زاده	عضو هیأت علمی دانشگاه صنعتی شریف
دکتر ویدا شقاقی	عضو هیأت علمی دانشگاه علامه طباطبائی
دکتر شهرام مدرس خیابانی	عضو هیأت علمی دانشگاه آزاد اسلامی
	واحد کرج
دکتر آزاده میرزائی	عضو هیأت علمی دانشگاه علامه طباطبائی

کمیته اجرایی

زهره ابراهیم بانکی، ریحانه منظوری، محمدحسن ترابی، سلیمه زمانی،
سهیلا ایزدی، مریم محمدی

فهرست مطالب

۹	پیشگفتار
۱۱	بررسی لیست کلمات عمومی جدید (ان جی اس ال) در پیکره طراحی شده برای کتاب های پراسپکت: (کتاب های آموزش زبان انگلیسی پایه متوسطه اول در ایران) مجید الهی شیروان - پیمان نصرآبادی - رضا پرهیزگار
۲۵	پیکره موجودیت های نامدار کارانس، تلاشی برای توسعه ی نیمه خودکار پیکره زبان محمد حسن بشری موحد - سید مرتضی خالقی - بهروز مینایی بیدگلی
۴۱	پیکره گفتار محاوره ای زبان فارسی امروز محمود بی جن خان
۵۱	آماده سازی دادگان گفتار محاوره ای فارسی جهت مدل سازی ناشیوایی ها در بازشناسی گفتار محاوره ای محمد حسن آبادی - محمد بحرانی - حسین صامتی
۶۹	تحلیل مدارهای هویت در سند استراتژی امنیت ملی ۲۰۱۵ بریتانیا امیر رضائی پناه - سمیه شوکتی مقرب
۹۳	پیکره زبان آموز: مبانی، روش شناسی، الگوی طراحی و تولید سعید صفری
۱۲۵	بررسی اصطلاحات مرکب در پیکره وابستگی زبان فارسی مریم فعال همدانچی - سعیده قدردوست نخچی
۱۴۷	بررسی ساختارهای قید: پژوهشی پیکره بنیاد سعیده قدردوست نخچی - ندا پورمرتضی خامنه - پری ناز دادرس - سلیمه زمانی

-
- ۱۶۷ تهیه خودکار زیرپیکره از پیکره زبانی
مسعود قیومی
- ۱۸۵ جایگاه بندهای قیدی تصدیقی در زبان فارسی
پریسا کیانی
- ۲۱۳ پیکره تیترا خبرگزاری های فارسی زبان
آزاده میرزائی - پگاه صفری
- ۲۲۵ مفهوم سازی استعاره «نفرت» در زبان فارسی: رویکردی
پیکره ای - شناختی
امیرسعید مولودی - غلامحسین کریمی دوستان - محمود بی جن خان
- ۲۵۳ نخستین پیکره جامع زبان فارسی: از آغاز تا امروز
مصطفی عاصی
- ۲۶۵ گامی به سوی پیکره های در سطح گفتمان زبان فارسی
آزاده میرزائی
- ۲۶۷ **The Roadmap of Iranian Corpus of Academic
Written English (ICAWWE)**
Ali Beikian- Akbar Hessabi- Mehran Borzoufard
- ۲۹۷ چکیده انگلیسی مقالات

دومین همایش ملی زبان‌شناسی پیکره‌ای در ۲۰ آبان ۱۳۹۵ در پژوهشگاه علوم انسانی و مطالعات فرهنگی به همت انجمن زبان‌شناسی ایران و با همکاری نشر نویسه پارسی و پایگاه استنادی علوم جهان اسلام برگزار شد. برای این همایش ۱۸ مقاله از دانشگاه‌ها و مراکز تحقیقاتی مختلف به دبیرخانه همایش ارسال شد که تقریباً تمامی محورهای همایش را پوشش می‌دادند. پس از داوری ۱۵ مقاله برای ارائه سخنرانی و پوستر برگزیده شدند. زبان‌شناسی پیکره‌ای حوزه نوپایی است که با ورود رایانه و اینترنت با شتاب مسیر پیشرفت را طی می‌کند. سازوکارهای مطمئن و کارا سبب شده است تا این حوزه به سرعت جای خود را در میان حوزه‌های دیگر زبان‌شناسی باز کند. امروزه استفاده از پیکره‌های زبانی نقش مهمی در تحقیقات زبان‌شناسی و دیگر حوزه‌های پژوهشی دارد که با زبان سروکار دارند. همین سبب شده است تا از یک سو تهیه انواع پیکره‌ها از اقبال زیادی برخوردار باشند و از سوی دیگر استخراج اطلاعات از آنها با دو منظور پژوهش‌های زبان‌شناختی و پردازش زبان طبیعی توسط ماشین، علاقمندان زیادی را به خود جلب کند. پس از برگزاری موفق نخستین همایش ملی زبان‌شناسی پیکره‌ای و ایجاد فرصتی مناسب برای هم‌اندیشی متخصصان و علاقمندان این حوزه، اکنون در دومین همایش ملی زبان‌شناسی پیکره‌ای شاهد ارائه مقاله‌های ارزشمندی هستیم که به صورت تخصصی ابعاد مختلف این حوزه را واکاوی می‌کنند و همپای دانش روز جهان با رویکردی داده‌محور به موضوع زبان فارسی می‌پردازند. نگاهی گذرا به مجموعه مقالات دو همایش زبان‌شناسی پیکره‌ای و توجه به سطح مقالات ارائه‌شده و تنوع آنها نشان می‌دهد که زبان‌شناسی پیکره‌ای در ایران در همه ابعاد و حوزه‌هایش به صورت شایسته‌ای در حال گسترش است.

مقالات این دوره از همایش پوشش مناسبی است از محورهای همایش که موضوعات زیر را دربرمی‌گیرد: مطالعات زبان‌شناختی پیکره-بنیاد، استفاده از پیکره‌ها در تحلیل گفتمان، کاربرد پیکره‌ها در آموزش زبان، پردازش زبان طبیعی بر اساس هوشمندسازی ماشین، ساخت پیکره‌های برچسب‌خورده در سطح گفتمان، تولید دادگان گفتاری، پیکره گفتار محاوره‌ای و پیکره‌های زبان آموز.

لازم است از بزرگانی که در شکل‌گیری این همایش تلاش کرده‌اند، تشکر کنم؛ نخست از رئیس انجمن زبان‌شناسی ایران، سرکار خانم دکتر بلقیس روشن، از مسئولان محترم پژوهشگاه علوم انسانی و مطالعات فرهنگی، رئیس پژوهشکده زبان‌شناسی پژوهشگاه جناب آقای دکتر مصطفی عاصی، دبیر اجرایی جناب آقای فرامرز مؤدکانلو و همچنین شورای اجرایی همایش آقایان محمدحسن ترابی، عباس رجب‌سلمانی و خانم‌ها زهرا ابراهیم‌بانکی، سلیمه زمانی، شاهد زندی‌راد، آذین شهریاری‌فرد، مریم محمدی، طاهره همتی، ذلیخا عظیم‌دخت که بزرگوارانه در اجرای همایش بسیار کوشیدند، از جناب آقای امیر احمدی مدیر محترم نشر نویسه پاریسی برای همراهی‌هایشان در اطلاع‌رسانی‌های همایش و همچنین چاپ مجموعه مقالات.

در پایان شایسته است از کمیته علمی و هیأت داوران همایش صمیمانه تشکر کنم، همچنین از استادان، دانشجویان و پژوهشگران گرانقدر که با ارسال مقاله خود به غنای علمی این همایش افزودند و دستاوردهای پژوهشی خود را با علاقمندان این حوزه در میان گذاشتند صمیمانه سپاسگزاری می‌کنم.

آزاده میرزائی

آبان ۱۳۹۵

بررسی لیست کلمات عمومی جدید^۱ (ان جی اس ال) در پیکره طراحی شده برای کتاب‌های پراسپکت کتاب‌های آموزش زبان انگلیسی پایه متوسطه اول در ایران)

مجید الهی شیروان^۲

پیمان نصرآبادی^۳

رضا پرهیزگار^۴

چکیده

کتاب‌های جدید آموزش زبان انگلیسی برای دوره اول متوسطه در ایران حدود سه سال است که با عنوان پراسپکت چاپ شده و مورد استفاده قرار می‌گیرد. علی‌رغم نیاز به بررسی این مجموعه از کتاب‌های درسی از منظر فراگیری کلمات، هنوز هیچ تحقیق در این مورد انجام نشده است. یک معیار برای بررسی این مجموعه منطبق بودن کلمات آن با لیست‌های گوناگون کلمات رایج زبان انگلیسی مبتنی بر پیکره‌ها می‌باشد. یکی از پرسامدترین لیست کلمات استخراج شده از پیکره‌های بزرگ زبان انگلیسی، لیست کلمات عمومی جدید (ان جی اس ال)^۱ می‌باشد. هدف این تحقیق بررسی میزان تطابق مجموعه پراسپکت با ان جی اس ال است. برای انجام این کار، پیکره سه کتاب از مجموعه پراسپکت بر اساس دستورالعمل‌های روشمند طراحی شد. سپس بسامد ان جی اس ال براساس نوع و نمونه کلمات در پیکره طراحی شده با استفاده از نرم‌افزار انت کانک به عنوان یک ابزار رایگان تجزیه و تحلیل پیکره و همچنین نرم‌افزار رنج، مورد بررسی قرار گرفته است. نتایج بررسی

^۱ NGS L, New General Service List

^۲ دکتر آموزش زبان انگلیسی، دانشگاه بجنورد، Elahimajid64@gmail.com

^۳ کارشناس ارشد آموزش زبان انگلیسی، دانشگاه بجنورد

^۴ کارشناس ارشد آموزش زبان انگلیسی، دانشگاه حکیم سبزواری

نشان داد که مجموعه پراسپکت با بسامد لغات آن‌جی‌اس‌ال سازگار نیست. در پایان مثال‌های بیشتر و پیشنهاد برای گسترش مطالب تکمیلی مورد بحث قرار گرفت.

کلیدواژه‌ها: لیست کلمات عمومی جدید، پیکره پراسپکت، انت‌کانک^۱، رنج^۲

۱. مقدمه

پیکره به عنوان ابزاری جدید برای آموزش زبان در چند دهه اخیر نقش پررنگ‌تری را در زبان‌شناسی کاربردی ایفا کرده است. پیکره را می‌توان طبق گفته بایبر، کنارد و رپن^۳ مجموعه‌ای از متون معتبر که به طور طبیعی جمع‌آوری و به صورت الکترونیکی ذخیره شده‌اند و قابل دسترسی هستند تعریف کرد. اولین پیکره توسط براون^۴ در سال ۱۹۶۱ ایجاد شد. این پیکره شامل یک میلیون کلمه بود که به مرور زمان و پیشرفت تکنولوژی پایه‌ای برای ایجاد پیکره‌های دیگر شد، مانند کوکا^۵ که شامل ۴۵۰ میلیون کلمه است. استابز^۶ (۱۹۹۶) معتقد است که برای رسیدن به افق‌های جدید آموزش زبان باید از زبان‌شناسی پیکره‌ای کمک گرفت. زبان‌شناسی پیکره‌ای طبق تعریف روش‌شناسی زبان‌شناسی است و زبان را از طریق پیکره بررسی می‌کند. این موضوع زبان‌شناسی کاربردی اهمیت ویژه‌ای دارد.

زبان‌شناسی پیکره‌ای از نظر کاربرد در دو دسته طبقه‌بندی می‌شود. اولین طبقه‌بندی آن کاربرد مستقیم زبان‌شناسی پیکره‌ای توسط معلمان و دانش‌آموزان است که می‌توانند از آن در کلاس به منظور یادگیری بهتر زبان استفاده کنند. جانز^۷ این استفاده را دی‌دی‌ال^۱ می‌نامد. دومین طبقه‌بندی آن

^۱ AntConc

^۲ Range

^۳ Biber, Conrad, Reppen,

^۴ Brown

^۵ COCA (Corpus of Contemporary American English)

^۶ Stubbs

^۷ Johns

کاربرد غیرمستقیم زبان‌شناسی پیکره‌ای توسط محققان و نویسندگان است که می‌توانند آثار یا تحقیقاتی مبتنی بر پیکره‌ها خلق کنند. توجه این مقاله بر دومین طبقه‌بندی معطوف شده، که درصدد است میزان تطابق واژگان کتاب تازه منتشرشده مقطع اول متوسطه (پراسپکت) را با واژگان ان‌جی‌اس‌ال (که در بخش بعدی آنرا بیشتر توضیح می‌دهیم) بررسی کند. همچنین در ادامه به معرفی نرم‌افزارهای استفاده‌شده برای مقایسه کتاب پراسپکت و ان‌جی‌اس‌ال و شیوه جمع‌آوری اطلاعات می‌پردازیم و در پایان به راه‌حلهایی برای رفع نواقص کتاب اشاره می‌کنیم.

۲. پیشینه تحقیق

همانطور که گفته شد زبان‌شناسی پیکره‌ای زبان‌شناسی روش‌شناسی است که در آن زبان را می‌توان به واسطه پیکره بررسی کرد. پیکره‌های زیادی در طول زمان ایجادشده که فواید بیشماری هم به همراه داشته‌اند. یکی از این فواید ظهور جی‌اس‌ال^۱ است که در سال ۱۹۵۳ توسط وست^۲ ایجادشد. جی‌اس‌ال شامل ۲۰۰۰ کلمه است که طبق گفته کاکسهد^۴ ۸۰ درصد متون علمی را که او مطالعه کرده است پوشش می‌دهد. جی‌اس‌ال با وجود ارزشمندی‌اش، به دلیل استفاده از پیکره‌ای محدود برای استخراج واژه، کاملاً قابل اعتماد نیست. مشکل محدودبودن این لیست با ظهور ان‌جی‌اس‌ال^۵ برطرف شد. ان‌جی‌اس‌ال را براون و کالیگان^۶ طراحی کردند که ۹۰ درصد متن‌های جامع زبان انگلیسی را پوشش می‌دهد و از پیکره‌ای شامل ۲۷۳ میلیون کلمه تشکیل شده است. با تمام این تفاسیر محاسبه تمام کلمات یک

^۱ DDL (Data Driven Learning)

^۲ GSL (General Service List)

^۳ West

^۴ Coxhead

^۵ New General Service List

^۶ Browne and Culligan

کتاب و بسامد آنها بسیار مشکل و زمان‌بر است ولی راه‌حلی مناسب برای این کار وجود دارد. با نرم‌افزارهایی می‌توان این کار را انجام داد و پیکره‌ای از کتابهای درسی یا هر کتابی به دست آورد. رومر^۱ (۲۰۰۶) این چنین گفته است: "کتابهای درسی خودشان می‌توانند تبدیل به پیکره‌هایی و با زبان انگلیسی رایج مقایسه شوند".

در این مقاله، از کتاب‌های پراسپکت پیکره‌ای ایجاد کردیم و کلمات آن را با کلمات ان‌جی‌اس‌ال، که براساس پیکره‌ای بزرگ و معتبر از زبان انگلیسی است، مقایسه و بررسی کنیم.

۳. بیان مسئله

بنا به گفته سینکلر^۲، پایه‌گذار استفاده از پیکره در آموزش زبان، کلمه‌ها خارج از متن معنی خاصی را نمی‌رسانند و معانی فقط در صورتی تأثیرگذار خواهند بود که کلمه مورد نظر میان زنجیره‌ای از کلمات دیگر به کار رود. بر همین اساس محققان برای زبان انگلیسی پیکره‌های متفاوتی را طراحی کرده‌اند. پیکره مجموعه بسیار بزرگ و طبقه‌بندی‌شده از جملاتی است که به طور طبیعی در هر زبان به کار می‌روند. این مجموعه غالباً بر روی رایانه‌ها ذخیره شده و با استفاده از نرم‌افزارهای خاص می‌توان کلمه مورد نظر را در میان انبوه جملات جستجو کرد. یکی از پیکره‌های بزرگ کوکا نام دارد که دسترسی به آن به صورت آنلاین می‌باشد. تصویر ۱، نتیجه جستجوی کلمه *Drive* را در پیکره کوکا نشان می‌دهد.

^۱Romer

^۲Sinclair

list questions. In this case, the advocate should formulate questions mentally during the *drive* to the facility. Then, at the ER, one can list additional questions, and show that teachers care--and, in so doing, feed students' *drive* to succeed. The authors offer steps for adding or reinstating subjective standards, relying solely on objectivity seems to forsake a part of humans' *drive* to succeed: the empathetic side driven by emotion, will happen to other forms and areas we value if we do? Will games *drive* out literature? Will multimodal literacies *drive* out traditional print forms? Will games *drive* out literature? Will multimodal literacies *drive* out traditional print forms? Of course we must take care of the textually complex characteristics: # Then she says, Your daddy and me used to *drive* this old road on his bike. # You did? I say. I Mia's past as a member of a loving, musically gifted family taking a *drive* to visit friends. From the opening pages, Mia navigates between their development should # 1. utilize well-defined language of effective classroom learning and teaching to *drive* the professional development expertise (see Figure 9), and a video of his comic sprites. # Gerard's *drive* to show evidence of his skills and expertise in creating comics, and his desire

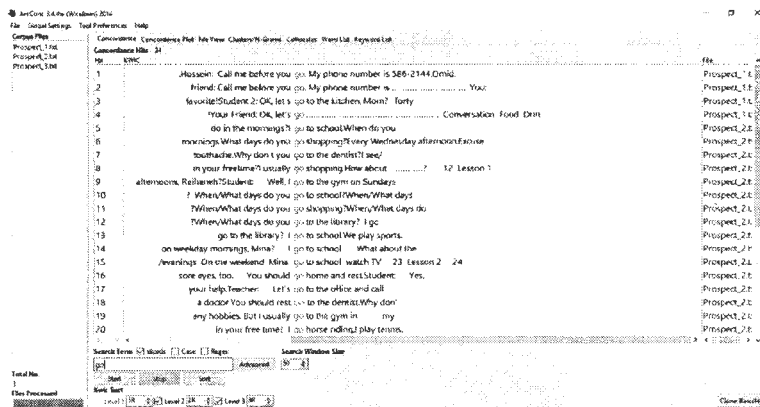
تصویر ۱. نتایج جستجوی کلمه Drive در پیکره کوکا

همانطور که در تصویر ۱ مشخص است، تعریف و بررسی کلمه *Drive* در بافت‌های مختلف اطلاعات بیشتری را در مقایسه با زمانی که این کلمه به تنهایی و جدا از بافت به کار رفته باشد در اختیار قرار می‌دهد؛ بنابراین طبیعی است که معنی تحت‌اللفظی آن (رانندگی کردن) برای آموزش این کلمه روش مناسبی نیست.

پیکره کتاب‌های پراسپکت را ما بر همین اساس ساخته‌ایم. این پیکره شامل تمامی کلمات و جملاتی است که در مجموعه سه جلدی کتاب‌های پراسپکت استفاده شده‌اند. این پیکره با انجام ویرایش‌های فراوان روی فایل پی‌دی‌اف این کتاب‌ها و تبدیل آن‌ها به فایل با پسوند docx ساخته شده است. با استفاده از نرم‌افزارهای خاص می‌توان از این پیکره استفاده کرد و اطلاعات ارزشمندی را استخراج نمود. به عنوان مثال می‌توان کلمه‌ای را جستجو کرد و دید این کلمه در کدامیک از قسمت‌ها و در کدام بافت‌ها در کتاب به کار رفته است.

یکی از نرم‌افزارهایی که امکان جستجو در پیکره را فراهم می‌آورد، انت کانک نام دارد. این نرم‌افزار توسط انتونی^۱ طراحی شده و یکی از بهترین نرم‌افزارها برای استفاده از پیکره و جستجوی کلمات در آنها می‌باشد. در تصویر ۲ ما با استفاده از این نرم‌افزار کلمه Go را در پیکره کتاب‌های پراسپکت جستجو کرده‌ایم:

^۱ Anthony



تصویر ۲. نرم‌افزار انت‌کانک: نتایج جستجوی کلمه Go در پیکره کتابهای پراسپکت

همانطور که مشاهده می‌شود این نرم‌افزار تمامی بافت‌هایی که کلمه Go در کتاب‌های پراسپکت به کار رفته است را نشان داده است. به این طریق می‌توان تمامی استفاده‌های این کلمه را به روش‌های مختلف آموزش داد.

نرم‌افزار دیگری که در استفاده از پیکره به کار می‌رود نرم‌افزار رنج می‌باشد. این نرم‌افزار توسط نیشن^۱ طراحی شده است و با استفاده از آن می‌توان بسامد و تعداد دفعاتی که هر کلمه در یک پیکره به کار رفته را مشاهده کرد.

استفاده از این نرم‌افزار به این شکل است که پیکره ساخته‌شده با پسوند TXT را به عنوان فایل ورودی به این نرم‌افزار می‌دهیم و سپس فایل خروجی را که آن هم یک فایل TXT است و شامل تک‌تک کلمات و بسامد هر کلمه در هر کتاب است را دریافت می‌کنیم. تصویر ۳، نمونه‌ای از این جستجو است.

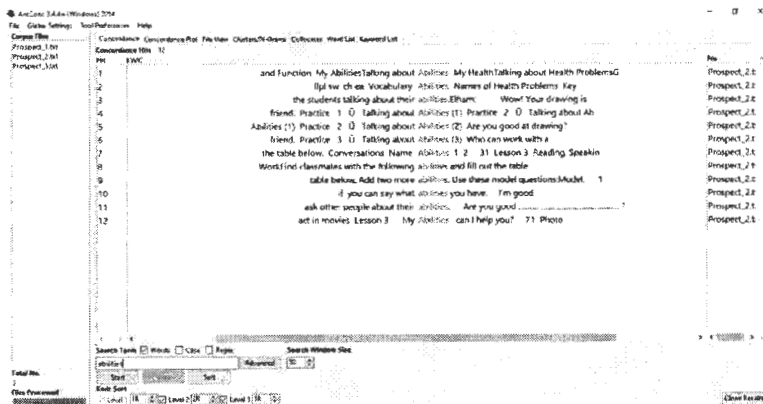
^۱ Nation

Types Found In Base List One

TYPE	RANGE	FREQ	F1	F2	F3
A	3	3754	1086	2006	662
ABILITIES	1	12	0	12	0
ABILITY	1	1	0	1	0
ABOUT	3	162	34	75	53
ACCOUNT	1	2	0	0	2
ACTIVITIES	2	18	0	16	2
ACTUALLY	2	5	0	2	3
ADD	1	1	0	1	0
ADDRESS	2	23	21	0	2
AFTER	1	1	0	0	1
AGAIN	3	5	3	1	1

تصویر ۳. لیست تعدادی از کلمات به کار رفته در کتابهای پراسپکت به همراه
بسامد کلمات استخراج شده با استفاده از نرم افزار رنج

با استفاده از این نتایج به عنوان مثال می بینیم که کلمه *Abilities* در کتابهای پراسپکت ۱۲ بار تکرار شده است. قابل ذکر است که این نتایج به طور کامل با نرم افزار قبلی یعنی انت کانک همخوانی دارد. یعنی اگر کلمه یاد شده را در نرم افزار انت کانک جستجو کنیم، ۱۲ بار تکرار این کلمه را می توانیم به همراه بافت های آن مشاهده نماییم.



تصویر ۴. نتایج جستجو کلمه *Abilities* در پیکره پراسپکت با استفاده از نرم افزار
انت کانک

شمارش تعداد جملات بالا نشان می‌دهد که کلمه *Abilities* دوازده بار در کتاب پراسپکت دو به کار رفته است. همانطور که در بالا گفته شد لیست ان‌جی‌اس‌ال، لیستی بسیار حائز اهمیت از کلماتی است که بیشترین بسامد را در زبان انگلیسی دارند. این لیست کلمات را بر اساس بسامد آن‌ها طبقه‌بندی کرده است. بنابراین بهتر است کتاب‌های طراحی‌شده برای آموزش زبان بر طبق این لیست باشند به این طریق که بسامد کلماتی که در این کتاب‌ها به کار می‌روند با بسامد همان کلمات در لیست ان‌جی‌اس‌ال همخوانی داشته باشند. به عنوان مثال اگر کلمه‌ای در لیست ان‌جی‌اس‌ال در تعداد دفعات تکرار (بسامد) رتبه اول را به خود اختصاص داده است، به این معنی است که آن کلمه در کتاب‌های آموزش زبان نیز باید بیشترین بسامد را داشته باشد و همچنین کلمات دیگر هم در بسامد و رتبه با این لیست همخوانی داشته باشند. در این صورت آن کتاب موفقیت خیلی بیشتری در آموزش زبان را خواهد داشت و کمک بیشتری به دانش‌آموزان خواهد کرد.

به همین منظور با استفاده از پیکره پراسپکت و نرم‌افزار رنج، که بسامد کلمات را به نمایش می‌گذارد، بسامد تمامی کلمات به‌کاررفته در این کتاب‌ها را استخراج و با لیست ان‌جی‌اس‌ال مقایسه کرده‌ایم.

۴. تحلیل داده‌ها

در ادامه بسامد چند کلمه را در کتاب‌های پراسپکت و بسامد همان کلمه‌ها را در لیست ان‌جی‌اس‌ال بررسی خواهیم کرد.

	A	B	C	D	E	F	G	H	I	J
1 A		3	3754	662	2006	1086		NGSL		
2 THE		3	504	211	175	118	1	27872		6
3 YOU		3	383	125	144	114	2	60910		1
4 AND		3	379	128	138	113	3	12810		10
5 S		3	316	80	93	143	4	30789		3
6 TO		3	282	122	81	79	5			
7 IS		3	266	95	89	82	6	29272		5
8 YOUR		3	247	75	46	126	7			
9 LESSON		3	207	98	61	48	8			
10 ABOUT		3	162	53	75	34	9	71		1243
11 CAN		3	158	43	73	42	10	2967		34
12 I		3	148	69	66	13	11	2890		41
13 MY		3	147	35	27	85	12	6820		17
14 IN		3	146	43	70	83	13			
15 WHAT		3	136	50	52	34	14	21142		7
16 DO		3	128	56	61	11	15	2204		48
17 ARE		3	123	46	52	25	16	6802		20
18 STUDENT		3	115	6	70	89	17			
19 TEACHER		3	113	17	54	42	18	391		244
20 HE		3	107	80	19	8	19	199		514
21 IT		3	102	52	43	7	20	11017		11
22 ASK		3	88	29	37	22	21	13772		10
23 OF		3	87	40	35	12	22	675		137

تصویر ۵. لیست تعدادی از کلمات به کار رفته در کتابهای پراسپکت و لیست ان جی اس ال به همراه بسامد و رتبه کلمات در تعداد دفعات تکرار

در این مثال ستون A تعدادی از کلمات به کار رفته در کتابهای پراسپکت می باشد. ستون B نشان دهنده این است که کلمه مورد نظر در کدام کتابها به کار رفته است. به عنوان مثال عدد ۳ نشان دهنده این است که کلمه مورد نظر در هر سه کتاب وجود دارد. ستون C تعداد کل دفعات تکرار هر کلمه در هر سه کتاب یعنی رتبه هر کلمه را با توجه به بسامد آن نشان می دهد. ستون D بسامد کلمه مورد نظر را تنها در کتاب پراسپکت ۱ نشان می دهد. ستون E بسامد کلمه مورد نظر را تنها در کتاب پراسپکت ۲ نشان می دهد. ستون F بسامد کلمه مورد نظر را تنها در کتاب پراسپکت ۳ نشان می دهد. ستون G رتبه هر کلمه را با توجه به بسامد آن در کتابهای پراسپکت نشان می دهد. ستون H بسامد همان کلمه را در لیست ان جی اس ال نشان می دهد و ستون I رتبه همان کلمه در بسامد ولی در لیست ان جی اس ال را نشان می دهد.

به مثال زیر توجه کنید:

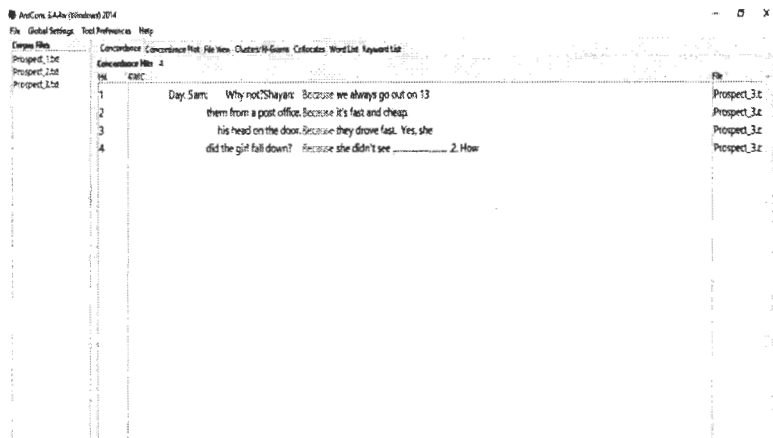
کلمه And در کتاب‌های پراسپکت ۳۷۹ بار تکرار شده است. بسامد این کلمه ۱۲۸ بار فقط در کتاب پراسپکت ۱ می‌باشد. همان کلمه در کتاب پراسپکت ۲، ۱۳۸ بار تکرار شده و همچنین بسامد آن در کتاب پراسپکت ۳، ۱۱۳ می‌باشد. با توجه به ستون G این کلمه رتبه سوم را در بسامد در کتاب‌های پراسپکت به خود اختصاص داده است و این در حالی است که رتبه آن کلمه در بسامد در لیست ان‌جی‌اس‌ال ۱۰ می‌باشد. این نشان می‌دهد که بسامد کلمه And در کتاب‌های پراسپکت تا حدودی با بسامد همان کلمه در لیست ان‌جی‌اس‌ال همخوانی دارد. حال به کلمه About توجه کنید.

	A	B	C	D	E	F	G	H	I	J
1	A	3	3754	662	2006	1086		NGSL		
2	THE	3	504	211	175	118	1	27872	6	
3	YOU	3	383	125	144	114	2	60910	1	
4	AND	3	379	128	138	113	3	12810	10	
5	S	3	316	80	93	143	4	30789	3	
6	TO	3	282	122	81	79	5			
7	IS	3	266	95	89	82	6	29272	5	
8	YOUR	3	247	75	46	126	7			
9	LESSON	3	207	98	61	48	8			
10	ABOUT	3	162	53	75	34	9	71	1243	
11	CAN	3	158	43	73	42	10	2967	34	
12	I	3	148	69	66	13	11	2890	41	
13	MY	3	147	35	27	85	12	6820	17	
14	IN	3	146	43	70	33	13			
15	WHAT	3	136	50	52	34	14	21142	7	
16	DO	3	128	56	61	11	15	2204	48	
17	ARE	3	123	46	52	25	16	6802	20	
18	STUDENT	3	115	6	70	39	17			
19	TEACHER	3	113	17	54	42	18	391	244	
20	HE	3	107	80	19	8	19	199	514	
21	IT	3	102	52	43	7	20	11017	11	
22	ASK	3	88	29	37	22	21	13772	10	
23	OF	3	87	40	35	12	22	675	137	

تصویر ۶. لیست تعدادی از کلمات به کاررفته در کتابهای پراسپکت و لیست

ان‌جی‌اس‌ال به همراه بسامد و رتبه کلمات در تعداد دفعات تکرار

با توجه به توضیحات بالا، این کلمه در بسامد رتبه نهم را در کتاب‌های پراسپکت به خود اختصاص داده است و این در حالی است که همین کلمه در لیست ان‌جی‌اس‌ال رتبه ۱۲۴۳ را داراست!
حال به کلمه **Because** توجه کنید.



تصویر ۷: نتایج جستجو کلمه **because** در پیکره کتاب‌های پراسپکت با استفاده از نرم افزار انت‌کانک

همان‌طور که مشاهده می‌شود این کلمه تنها ۴ بار و فقط در کتاب پراسپکت ۳ به کار رفته است و رتبه ۴۸۳ را در میان کلمات این سه کتاب داراست و این در حالی است که همین کلمه در لیست ان‌جی‌اس‌ال رتبه ۶۲ را به خود اختصاص داده که نشان دهنده نیاز به تکرار بیشتر کلمه **Because** در کتاب‌های پراسپکت می‌باشد.

وجود چنین ناهمخوانی‌هایی در بسامد کلمات (هنگام مقایسه کتاب‌های پراسپکت با لیست ان‌جی‌اس‌ال)، نشان‌دهنده نیاز به انجام تغییرات در بسامد کلمات در کتاب‌های پراسپکت برای کارایی بیشتر این کتاب‌ها و بهبود آموزش زبان می‌باشد.

۵. نتیجه‌گیری

با توجه به آنچه گفته شد، اینگونه به نظر می‌رسد که استفاده از پیکره در زبان و زبان‌آموزی افق‌های جدیدی را برای نه تنها مؤلفین کتاب‌های درسی بلکه معلمان و حتی دانش‌آموزان می‌گشاید. آموزش زبان با استفاده از پیکره در کشور ما تاکنون نقش بسیار کم‌رنگی را داشته اما با توجه به محتوای این مقاله اینگونه می‌توان دریافت که ترویج استفاده از پیکره و تشویق مؤلفین، معلمان و دانش‌آموزان به بهره‌گیری از خصوصیات ارزشمند پیکره، می‌تواند آموزش زبان را در کشور بهبود بخشیده و نتایج درخشانی را بر جای گذارد.

منابع

- Anthony, L. (2011). AntConc (Version 3.2. 2) [Computer Software]. Tokyo, Japan: Waseda University.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.
- Browne, C., Culligan, B., & Phillips, J. (2013). *The new general service list*. Cambridge: Cambridge University Press
- Coxhead, A. (2000). A new academic word list. *TESOL quarterly*, 34(2), 213-238.
- Johns, T. (2002). Data-driven learning: The perpetual challenge. *Language and Computers*, 42(1), 107-117.
- Johns, T.F. (1991). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T.F. Johns and P. King (Eds.), *Classroom concordancing* (pp. 27-45). Birmingham: ELR.

Nation, P. (2002). *Range and Frequency: Programs for Windows based PCs*. Wellington: Victoria University of Wellington

Römer, U. (2006). Where the computer meets language, literature, and pedagogy: Corpus analysis in English Studies. *How Globalization Affects the Teaching of English: Studying Culture through Texts*, 81-109.

Sinclair, John. *Corpus, concordance, collocation*. Oxford University Press, 1991.

Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Oxford: Blackwell.

West, M., & West, M. P. (Eds.). (1953). *ageneral service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. Addison-Wesley Longman Limited.