

تہجیب

سیدنی لاندو

فرہنگ ہا، فن و ہنر فرہنگ نویسی

پیکرہ در فرہنگ نویسی

ترجمہ

فرہاد قربان زادہ، سعید رفسنجانی نژاد



فرهنگ‌ها
فن و هنر فرهنگ‌نویسی
ویراست دوم

دفتر اول
پیکره در فرهنگ‌نویسی

سرشناسه	: لاندو، سیدنی آی، ۱۹۳۳ - م.
عنوان و نام پدیدآور	: فرهنگ‌ها: فن و هنر فرهنگ‌نویسی / سیدنی آی، لاندو؛ فرهاد قربان‌زاده، سعید رفسنجانی نژاد. تهران: کتاب بهار، ۱۳۹۶.
مشخصات نشر	: ج.
مشخصات ظاهری	: ج.
فروست	: کتاب بهار؛ ۱۵. فرهنگ‌نویسی؛ ۲.
شابک	: 978-600-96833-4-5
وضعیت فهرست‌نویسی	: فیبا
یادداشت	: عنوان اصلی: <i>Dictionaries: The Art and Craft of Lexicography</i> , 2 nd ed, 2001
یادداشت	: واژه‌نامه.
یادداشت	: کتاب‌نامه: ص. [۱۰۵]-۱۰۶.
یادداشت	: نمایه.
مندرجات	: ج. ۱. پیکره در فرهنگ‌نویسی
موضوع	: واژه‌نامه‌نویسی
موضوع	: Lexicography
موضوع	: دایرةالمعارف‌ها و واژه‌نامه‌ها - - تاریخ و نقد
موضوع	: Encyclopedias and dictionaries - History and criticism
شناسه افزوده	: قربان‌زاده، فرهاد، ۱۳۶۳ - ، مترجم
شناسه افزوده	: رفسنجانی نژاد، سعید، ۱۳۶۱ - ، مترجم
رده‌بندی کنگره	: ۱۳۹۶ ف۴/ل۳۲۷p
رده‌بندی دیویی	: ۴۱۳/۰۲۸
شماره کتابشناسی ملی	: ۴۷۲۲۸۲۸

سیدنی آی. لاندو

فرهنگ‌ها
فن و هنر فرهنگ‌نویسی
ویراست دوم

دفتر اول
پیکره در فرهنگ‌نویسی

ترجمه
فرهاد قربان‌زاده
سعید رفسنجانی‌نژاد

تجرب

تهران ۱۳۹۶

فرهنگ‌نویسی ۲
کتاب بهار ۱۵

این کتاب ترجمه‌ای است از
DICTIONARIES
The Art and Craft of Lexicography
Second Edition
Chapter 6: The corpus in lexicography
Sidney I. Landau
Cambridge University Press
2001

کتاب

سیدنی آی. لاندو
فرهنگ‌ها: فن و هنر فرهنگ‌نویسی
ویراست دوم
دفتر اول: پیکره در فرهنگ‌نویسی
ترجمه
فرهاد قربان‌زاده، سعید رفسنجانی‌نژاد
ویرایش زبانی: محمد صاحبی
چاپ نخست: ۱۳۹۶
شمارگان: ۳۰۰
چاپ و صحافی: پردیس دانش
همه حقوق چاپ و نشر این کتاب محفوظ و مخصوص کتاب بهار است.

مراکز فروش: کتاب‌فروشی توس (۶۶۴۶۱۰۰۷)، کتابسرای افتخارزاده (۶۶۹۵۵۸۹۵)،
نشر چاپار (۸۸۸۹۹۶۸۰)
کتاب بهار (۰۹۱۲۱۲۲۶۰۰۸)
Ketabebaharpub94@gmail.com

فهرست

۷	پیشگفتار مترجمان.....
۹	پیشگفتار مؤلف.....
۱۱	تاریخچه پیشرفت پیکره‌های زبانی.....
۱۵	مخالفت چامسکیایی با تحلیل‌های کمی.....
۱۹	پیدایش زبان‌شناسی پیکره‌ای.....
۲۸	تغییرات فنی.....
۳۱	پیکره‌های نسل دوم.....
۴۰	پیکره‌های ملی و تخصصی.....
۴۵	کاربرد پیکره در فرهنگ‌نویسی نوین.....
۵۸	انتخاب شاهد در فرهنگ‌نویسی انگلیسی به‌عنوان زبان دوم.....
۶۲	همایند.....
۷۵	کاربردهای دیگر پیکره.....
۷۸	محدودیت‌های پیکره.....
۸۱	گردآوری پیکره.....
۸۲	گردآوری متن.....
۸۴	ساختار پیشنهادی انواع متن‌های پیکره.....
۹۰	جامعیت.....
۹۳	ابزارهای واژگانی.....
۱۰۰	کاربرد پیکره در آینده.....
۱۰۵	منابع.....
۱۰۷	واژه‌نامه فارسی به انگلیسی.....
۱۰۹	واژه‌نامه انگلیسی به فارسی.....
۱۱۱	نمایه.....
۱۱۳	نام کسان، کتاب‌ها، پیکره‌ها، نهادها و جاها.....

پیشگفتار مترجمان

کتاب پیش رو ترجمه فصل «پیکره در فرهنگ‌نویسی» از کتاب فرهنگ‌ها؛ فن و هنر فرهنگ‌نویسی^۱ (ویراست دوم) اثر شناخته‌شده سیدنی آی. لاندو^۲ است. چاپ نخست کتاب در سال ۱۹۸۴ و ویراست دوم آن با بازبینی و بازنویسی کلی و افزوده‌های بسیار در سال ۲۰۰۱ منتشر شد. کتاب اصلی از این فصل‌ها تشکیل شده‌است: ۱. فرهنگ چیست؟؛ ۲. تاریخچه کوتاه فرهنگ‌نویسی انگلیسی؛ ۳. عناصر پایه‌ای فرهنگ‌ها و دیگر کتاب‌های مرجع؛ ۴. تعریف‌نگاری؛ ۵. کاربرد؛ ۶. پیکره در فرهنگ‌نویسی (کتاب پیش رو)؛ ۷. تألیف فرهنگ؛ ۸. مسائل حقوقی و اخلاقی در فرهنگ‌نویسی. در این اثر به بسیاری از پرسش‌های فرهنگ‌نویس در زمینه پیکره زبانی پاسخ داده می‌شود.

در صفحه‌های آغازین کتاب اصلی آمده‌است:

ویرایش دوم اثر دوران‌ساز سیدنی آی. لاندو توصیف جامع و کاملاً به‌روزی از چگونگی پژوهش درباره فرهنگ‌ها و فرهنگ‌نویسی را ارائه می‌دهد. در این اثر توجه ویژه‌ای به فناوری رایانه‌ای شده‌است که به شیوه‌های گوناگونی فرهنگ‌نویسی نوین را تغییر داده‌است. فصلی کاملاً جدید به این اثر افزوده شده‌است و همه فصل‌ها روزآمد و دوباره سامان‌دهی شده‌اند تا بتوانند تغییرات به‌وجودآمده در فرهنگ‌نویسی را نشان دهند. لاندو در فرهنگ‌نویسی دانش کاربردی اثرگذاری دارد. او همه ویژگی‌های فرهنگ‌ها را بررسی کرده و شرح داده‌است. در این شرح و بررسی، مثال‌های فراوانی از جدیدترین فرهنگ‌های آمریکایی و بریتانیایی ذکر و تاریخچه‌ای از فرهنگ‌نویسی انگلیسی نیز گنجانده شده‌است. این کتاب کاملاً براساس دانش فرهنگ‌نویسی معاصر نوشته شده‌است. سبک نگارش آن خواندنی است و در آن از به‌کارگیری بیان تخصصی غیرضروری و هر نوع اصطلاح تخصصی پرهیز شده‌است. این اثر برای خوانندگانی که تنها به دانستن درباره فرهنگ‌ها علاقه‌مندند (و دانش خاصی هم درباره این موضوع ندارند) و همچنین فرهنگ‌نویسان حرفه‌ای جذاب است.

در ادامه این مطلب آمده‌است:

سیدنی آی. لاندو سرویراستار پیشین شعبه آمریکای شمالی انتشارات دانشگاه کیمبریج است و از سال ۱۹۶۱ به فرهنگ‌نویسی می‌پردازد. او با بسیاری از مجله‌های تخصصی این حوزه همکاری داشته‌است و ویراستار فرهنگ بین‌المللی پزشکی و زیست‌شناسی^۳

1. *Dictionaries; The Art and Craft of Lexicography*

2. Sidney I. Landau

3. *International Dictionary of Medicine and Biology*

(۱۹۸۶) در سه جلد و فرهنگ انگلیسی آمریکایی کیمبرج^۱ (۲۰۰۰) است.

※

به درخواست ناشر، فصل‌های هشتگانه کتاب را چند همکار فرهنگ‌نویس در دست ترجمه دارند. پیش‌بینی می‌شود ترجمه فارسی کتاب در چهار مجلد کم‌حجم منتشر شود و چاپ دوم آن، شامل تمامی متن کتاب، در یک مجلد انتشار یابد. فصل «پیکره در فرهنگ‌نویسی» پیش از دیگر فصل‌های کتاب آماده شده‌است. از سوی دیگر، پیکره زبانی مبحثی است که می‌تواند مستقلاً نیز مطالعه شود و ارتباط تنگاتنگی با دیگر فصل‌های کتاب ندارد.

مقاله‌ها و منابع فارسی در زمینه پیکره‌های زبانی بسیار اندک‌شمارند و تعداد بسیار کمی از همین منابع اختصاصاً درباره پیکره زبانی در فرهنگ‌نویسی هستند. خوشبختانه، انجمن زبان‌شناسی ایران در سال‌های ۱۳۹۴ و ۱۳۹۵ دو همایش با موضوع زبان‌شناسی پیکره‌ای برگزار کرد که مجموعه مقاله‌های منتشرشده آن همایش‌ها گامی در پیشبرد زبان‌شناسی پیکره‌ای در ایران بوده‌است.

اکنون چند مؤسسه و گروه فرهنگ‌نویسی در ایران مشغول تألیف فرهنگ هستند. تاکنون در فرهنگ‌های فارسی، آن‌گونه که باید، به پیکره توجه نشده و همین مسئله کاستی‌های بسیاری را در فرهنگ‌های فارسی پدید آورده‌است. برای ویرایش فرهنگ‌های شناخته‌شده پیشین، مانند لغت‌نامه دهخدا و فرهنگ فارسی معین، نیز به‌ناگزیر باید از پیکره‌های زبانی کلان بهره برد تا نتیجه کار علمی‌تر شود.

مترجمان گاه نکته‌های تازه‌ای را در پانوشته افزوده‌اند که از شکاف شانزده‌ساله چاپ کتاب اصلی و ترجمه فارسی آن می‌کاهد.

در پایان کتاب دو واژه‌نامه کوتاه انگلیسی به فارسی و فارسی به انگلیسی، شامل اصطلاحات تخصصی زبان‌شناسی پیکره‌ای و فرهنگ‌نویسی افزوده شده‌است.

آقایان الوند بهاری و دکتر محمد صاحبی این کتاب را پیش از چاپ خوانده‌اند که از ایشان سپاسگزاریم.

فرهاد قربان‌زاده - سعید رفسنجانی‌نژاد

پیشگفتار مؤلف*

امروزه رایانه‌ها فرهنگ‌نویسی را نیز، به همراه دیگر چیزها، متحول کرده‌اند. بر اثر این تحولات ناگزیر به ویرایش دوم این اثر بودم.

بخت با من یار بود که از سال ۱۹۹۴ تا ۲۰۰۰ یک‌سره در تألیف فرهنگ جدیدی برای زبان‌آموزان خارجی، به نام فرهنگ انگلیسی آمریکایی کیمبریج، شرکت داشتم. در طی تألیف این اثر، با تغییرات بسیار زیادی در فرهنگ‌نویسی آشنا شدم که از فناوری رایانه سرچشمه می‌گرفت. هنگام تألیف کتاب پیشین خود، فرهنگ چندجلدی پزشکی به نام فرهنگ بین‌المللی پزشکی و زیست‌شناسی (John Wiley 1986)، آشنایی کمی با رایانه و فرهنگ‌نویسی داشتم. هرچند، آن تجربه در مقایسه با تجربه‌های اخیرم بسیار ناچیز بود. هنگامی که در میانه ویرایش فرهنگ پزشکی بودم و ویرایش نخست این اثر نوشته شده بود، ارجاعات زیادی به فرهنگ‌نویسی تخصصی، به‌ویژه پزشکی، وجود داشت. در این ویراست، برخی شاهدهای مربوط به تعریف‌های تخصصی را کنار گذاشتم، ولی روی هم‌رفته، به صورتی گسترده به فرهنگ‌نویسی تخصصی پرداخته‌ام.

از آنجا که تجربه‌های جدیدترم در زمینه فرهنگ‌نویسی انگلیسی به‌عنوان زبان دوم بوده‌است، در این ویرایش، در مقایسه با ویراست نخست، مثال‌های زیادی از این حوزه انتخاب شده‌است. بریتانیا در انتشار فرهنگ برای زبان‌آموزان خارجی پیشگام بوده‌است و هنوز هم بزرگ‌ترین و پیشرفته‌ترین فرهنگ‌های این حوزه در بریتانیا تألیف می‌شود؛ از این رو، مجبور بودم با فرهنگ‌های مهم انگلیسی به‌عنوان زبان دوم که در بریتانیا تألیف شده‌اند کاملاً آشنا شوم. در اغلب کتاب‌های فرهنگ‌نویسی بیشتر به فرهنگ‌های آمریکایی پرداخته می‌شود، ولی من در نخستین ویرایش این اثر به فرهنگ‌نویسی بریتانیایی توجه بیشتری کردم. در این ویراست، به فرهنگ‌نویسی بریتانیایی، به‌ویژه به فرهنگ‌های انگلیسی به‌عنوان زبان دوم، بیشتر پرداخته‌ام، اما از فرهنگ‌های دیگر غافل نبوده‌ام. با وجود این، اعتراف می‌کنم که به فرهنگ‌های آمریکایی دانش عمیق‌تری دارم و نمی‌توانم ادعا کنم که در همه حوزه‌های فرهنگ‌نویسی به یک اندازه به فرهنگ‌های بریتانیایی پرداخته‌ام.

* در اینجا آن بخش‌هایی از پیشگفتار مؤلف آمده‌است که با فصل «پیکره در فرهنگ‌نویسی» ارتباط دارد (مترجمان).

از دهه ۱۹۸۵ تا ۱۹۹۵ دنیای فرهنگ‌نویسی دستخوش تغییرات بسیاری شد. این تغییرات نخست از توانایی بسیار بیشتر فایل‌نگهدارها^۱ و سپس هارددرایو^۲ در رایانه‌های رومیزی شخصی سرچشمه می‌گیرد. بیشترین و مهم‌ترین اثر این تغییرات را در توسعه مجموعه‌های الکترونیکی عظیمی می‌توان یافت که زبان طبیعی را نشان می‌دهند (که *corpora* «پیکره‌ها» نامیده می‌شود و مفرد آن *corpus* «پیکره» است و در لاتینی به معنی «پیکر» است). از این پیکره‌ها به گونه‌ای برای بررسی و تحلیل کاربرد زبان استفاده می‌شود که پیش‌تر امکان‌پذیر نبود. کاربرد پیکره تغییرات زیادی را در بررسی و تألیف فرهنگ‌ها پدید آورد. فصل ششم کتاب با عنوان «پیکره در فرهنگ‌نویسی»^۳ در این ویراست کاملاً جدید است و موارد زیر را شرح می‌دهد: تاریخ گسترش پیکره، چگونگی کاربرد پیکره در فرهنگ‌نویسی نوین و چگونگی گردآوری آن. همچنین درباره کاربرد پیکره در آینده نیز حدس‌هایی زده‌ام.

فناوری رایانه در چگونگی تألیف و ویرایش فرهنگ‌ها تغییرات بنیادینی پدید آورده است. فرهنگ‌ها به صورت برخط^۴ تألیف می‌شوند و ویراستاران مجبور نیستند با یکدیگر در یک مکان کار کنند. در ویرایش هرروزه فرهنگ‌ها، از پیکره به شیوه‌های گوناگونی استفاده می‌شود و نرم‌افزارهای رایانه‌ای پیچیده برخی کارهای فرهنگ‌نویسی، مانند بازبینی ارجاع درونی^۵، را آسان کرده‌اند. اکنون فرایند برنامه‌ریزی برای فرهنگ، نسبت به گذشته، تغییر بسیاری کرده است. امروزه، کارشناسان رایانه در این فرایند حضور دارند و یاد گرفتن سخت‌افزار رایانه امری ضروری است.

به دلیل تأثیرهای بنیادین پیکره در فرهنگ‌نویسی، نقش مخزن شاهد^۶ (مجموعه‌ای از شاهدها که از مجله‌ها، کتاب‌ها و مانند آن‌ها رونویسی یا حروف‌نگاری می‌شود)، که زمانی تنها منبع اصلی تعریف‌نگاری بودند، تغییر کرده است.

1. file servers

2. hard drive

۳. برابر با کل کتاب پیش رو (مترجمان).

4. on-line

5. cross-reference

6. citation file

پیکره در فرهنگ‌نویسی

در پژوهش‌های زبانی به مجموعه‌ای از متن‌های گردآوری‌شده برای تحلیل ویژگی‌های آن‌ها «پیکره» گفته می‌شود. امروزه منظور از پیکره در فرهنگ‌نویسی پیکره الکترونیکی است که معمولاً در آن شمار زیادی واژه از منابع گوناگون وجود دارد. پیش از عصر رایانه، تا مدت‌ها، هدف از گردآوری پیکره‌ها بررسی متن‌های گوناگونی مانند کتاب مقدس و آثار شکسپیر و تولید واژه‌نامه‌های آن‌ها بود. این فرایند نیازمند تلاش انسانی بسیاری بود، زیرا تک‌تک واژه‌ها برای مشخص کردن جایی که در آن به کار رفته‌است و برآورد بسامدهای آن‌ها با دست رونویسی می‌شد. با این حال، برخی پیکره‌های نخستین حجم زیادی داشتند.^۱

تاریخچه پیشرفت پیکره‌های زبانی

پیکره ادوارد ال. تورندایک^۲ نخستین پیکره نوین و بزرگ زبان انگلیسی است که برای پژوهش‌های واژگانی (و نه تحلیل‌های ادبی) تهیه شد. این پیکره ۴/۵ میلیون کلمه‌ای در سال ۱۹۲۱ و با نام واژه‌نامه معلم^۳ انتشار یافت و پس از آن در دهه‌های بعدی در چندین مرحله بر حجم آن افزوده شد و تعداد واژه‌های آن به هجده میلیون رسید. تورندایک و اروینگ لورج^۴ این پیکره را با نام واژه‌نامه سی هزار کلمه‌ای

1. concordance

2. B. T. S. Atkins, Beth Levin, and A. Zampoli, "Computational Approaches to the Lexicon: an Overview," in B. T. S. Atkins and A. Zampoli (eds.), *Computational Approaches to the Lexicon* (Oxford University Press, 1994), p. 21.

با استناد به پژوهشی بر روی زبان آلمانی که در اواخر سده نوزدهم انجام شد و شامل یازده میلیون واژه می‌شود.

3. Edward L. Thorndike

4. *Teacher's Word Book*

5. Irving Lorge

معلم^۱ منتشر کردند. این پیکره شامل چندین فهرست از واژه‌های گردآوری‌شده از مجله‌ها و متن‌های مربوط به نوجوانان و بسامد آن‌ها است. این کتاب به گونه‌ای طراحی شده‌است که مربیان و معلمان با کمک آن درمی‌یابند هریک از واژه‌ها چه اندازه رایج‌اند تا در سطح معینی از آن‌ها استفاده کنند.

پیکره اولیه دیگری که حجم درخور توجهی دارد پیکره ارنست هورن^۲ به نام واژگان نوشتاری پایه: ده‌هزار واژه پرکاربرد نوشتاری^۳ است. این تئکنگاری، که در سال ۱۹۲۶ به چاپ رسید، بیشتر بر اساس نامه‌های شخصی و چاپ‌شده از منابع و حوزه‌های بسیار گوناگون گردآوری شده‌است. این پژوهش^۴ بسیار پیچیده است. برخی ابعاد این تلاش را بر مبنای شرح شیوه کاری زیر می‌توان حدس زد:

واژه‌ها به این صورت دسته‌بندی شدند: در ابتدا واژه‌ها بر روی برگه‌هایی ۸/۵×۱۳ ضبط و از سمت چپ برگه‌ها الفبایی شدند و به همراه حروف آغازین واژه‌های کلیدی، مانند *ab*، *acc*، *add*، *aff* و *ag* ثبت شدند. تخمین کمابیش دقیق فضایی خالی که باید میان واژه‌های راهنما قرار داده می‌شد با بهره‌گیری از تجربه دسته‌بندی‌های پیشین امکان‌پذیر گردید. از این‌رو، واژه‌ها به ترتیب الفبایی نوشته شدند. مجموعه‌ای از برگه‌های الفبایی شده و دارای فاصله که شامل ده‌هزار واژه می‌شد بسیار رضایت‌بخش بود.^۵

از آنجاکه این بررسی ۵،۱۳۶،۸۱۶ واژه به کاررفته در متون را شامل می‌شد، بیش از پانصد بسته از این نوع باید با دقت آماده می‌شد.

برخی پیکره‌های اولیه برای بررسی بسامد آواهای گفتار طراحی شده بودند. گادفری دیوئی^۶، حامی املای ساده‌شده، در سال ۱۹۲۳ بسامد نسبی آواهای گفتاری انگلیسی^۷ را منتشر کرد (در عنوان آن به‌عمد *e* ناملفوظ از آخر *relative* حذف شده‌است). او آواهای گفتاری صد‌هزار واژه را از منابع گسترده گوناگون بررسی کرد. همان‌گونه که خود شرح می‌دهد، پیش از او پژوهش‌های مشابهی انجام شده بود، از جمله پژوهش

1. *The Teacher's Word Book of 30,000 Word*

2. Ernest Horn

3. *A Basic Writing Vocabulary: 10,000 Words Most Commonly Used in Writing*

4. Ernest Horn, *A Basic Writing Vocabulary: 10,000 Words Most Commonly Used in Writing* (Iowa City: College of Education, University of Iowa, 1926), p. 22.

5. Godfrey Dewey

6. *Relativ Frequency of English Speech Sounds*

ویلیام دوئیلت ویتنی^۱ در سال ۱۸۷۴ که بعدها دستورنویس مبتکری شد و به ویرایش فرهنگ سده^۲ پرداخت.^۳

پیکره‌های پیشالکترونیکی، که بیشترین تأثیر را بر فرهنگ‌نویسی گذاشتند، برای آموزش انگلیسی به زبان‌آموزان خارجی استفاده شدند. این حوزه آموزشی هم‌اکنون با عنوان آموزش زبان انگلیسی^۴، و به بیان دقیق‌تر انگلیسی به‌عنوان زبان دوم^۵ و انگلیسی به‌عنوان زبان خارجی^۶ شناخته می‌شود. به معنای دقیق کلمه، انگلیسی به‌عنوان زبان دوم شامل بررسی زبان انگلیسی در نقاطی از جهان است که در آنجا مردم به زبان انگلیسی صحبت می‌کنند، اما انگلیسی زبان بومی آن‌ها نیست. درحالی‌که موضوع انگلیسی به‌عنوان زبان خارجی بررسی زبان انگلیسی مردمی است که در کشورهای غیرانگلیسی‌زبان زندگی می‌کنند. امروزه فرهنگ‌های زبان‌آموزان خارجی به هر دو گروه اختصاص دارد، اما این فرهنگ‌ها در گذشته فرهنگ‌های انگلیسی به‌عنوان زبان خارجی بودند و تألیف آن‌ها در ژاپن آغاز شد. در این زمینه، اثر پیکره‌ای مایکل وست^۷، هارولد ای. پالمر^۸ و ای. اس. هورنبی^۹، در دهه ۱۹۳۰ تبدیل به فرهنگی دوران‌ساز شد و در آن دوره بر فرهنگ‌های انگلیسی به‌عنوان زبان دوم اثری ماندگار گذاشت. «در مؤسسه پژوهشی آموزش زبان انگلیسی توکیو»^{۱۰}، که در آن کارهای پژوهشی هورنبی از میانه دهه ۱۹۳۰ محوریت داشت، پژوهشی معروف به حرکت «کنترل واژگان» عمیق‌ترین و ماندگارترین تأثیر را بر تاریخ اولیه فرهنگ‌های زبان‌آموز گذاشت.^{۱۱} هارولد پالمر چهره‌ای پیشرو در زمینه محدود کردن واژگان برای زبان‌آموزان خارجی بود. او در ابتدای دهه ۱۹۳۰ درباره گزینش واژگان گزارشی ارائه کرد.^{۱۲} در اکتبر ۱۹۳۴ با ابتکار

1. William Dwight Whitney

2. *The Century Dictionary*

3. Godfrey Dewey, *Relativ Frequency of English Speech Sounds* (Cambridge, Mass.: Harvard University Press, 1923). Reference is made to Whitney's essay on p. 7.

4. *English Language Teaching: ELT*

5. *English as a Second Language: ESL*

6. *English as a Foreign Language: EFL*

7. Michael West

8. Harold E. Palmer

9. A. S. Hornby

10. Tokyo Institute for Research in English Teaching

11. A. P. Cowie, "A. S. Hornby, 1898-1998: A Centenary Tribute," *International Journal of Lexicography* 11:4 (December 1998), 254-55.

12. *Ibid.*, p. 255.

مایکل وست، که در هند انگلیسی تدریس کرده بود، بنیاد کارنگی^۱ در نیویورک سیتی همایشی برگزار کرد.^۲ این همایش به موضوع گزینش واژگان در آموزش انگلیسی می‌پرداخت. در میان افراد دعوت‌شده، پالمر و ادوارد تورندایک در مقام مشاور حضور داشتند. پس از آن، پالمر، وست و لورنس فوسیت^۳ با کمک هورنبی گزارش موقتی درباره گزینش واژگان^۴ را تهیه کردند که در سال ۱۹۳۶ منتشر شد و از پژوهش‌های اثرگذار بود. گزارش موقتی از پژوهش‌های واژگانی تورندایک بسیار موشکافانه‌تر بود، زیرا در آن به معنی و هویت دستوری (مانند تمایز میان «fly» «حشره» و معنی‌های گوناگون فعل fly) نیز توجه شده بود. سپس، مایکل وست با کمک بسامد معنایی‌ای که اروینگ لورج تهیه کرده بود، از بخشی از گزارش مقدماتی برای بازبینی و گسترش فهرست کاربردهای عمومی واژه‌های انگلیسی^۵ استفاده کرد.^۶ در این اثر، با توجه به تفکیک معنی‌ها در فرهنگ انگلیسی آکسفورد^۷، معنی‌های واژه‌ها تفکیک و بسامد هر کدام مشخص شده‌است. از آنجاکه تعداد واژه‌های این اثر زیاد نیست، کاربرد محدودی در فرهنگ‌نویسی داشته‌است، اما در زمینه کارهای لازم برای ارتقای پژوهش‌های پیکره‌ای نمونه‌ای درخور توجه است. مایکل وست و جیمز اندیکوت^۸ در سال ۱۹۳۵ نخستین فرهنگ انگلیسی به‌عنوان زبان خارجی را با نام فرهنگ انگلیسی روش نوین^۹ منتشر کردند. در این اثر «با استفاده از ۱۴۹۰ واژه» معنی ۲۴,۰۰۰ عنصر زبانی تعریف شده‌است.^{۱۰}

1. Carnegie Corporation

2. Graeme Kennedy, "Preferred Ways of Putting Things with Implications for Language Teaching," in Jan Svartvik (ed.), *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82* (Berlin: Mouton de Gruyter, 1992), pp. 335-73.

3. Lawrence Faucett

4. *The Interim Report on Vocabulary Selection*

5. *A General Service List of English Words*

6. Michael West (ed.), *A General Service List of English Words* (London: Longmans, Green, 1953).

"The General Service List" was originally published as Part V of the *Interim Report* in 1936. The 1953 version utilizes Irving Lorge's *The Semantic Count of the 570 Commonest English Words* (New York: Institute of Psychological Research, Teachers College, Columbia University, 1949).

7. *OED*

8. James Endicott

9. *The New Method English Dictionary*

10. Kennedy, "Preferred Ways of Putting Things," p. 339.

اثر چارلز سی. فریس^۱ که از آثار ابتکاری و پیشرو در حوزه پژوهش‌های پیکره‌ای بود نیز موجب پیشرفت‌هایی در این حوزه شد. در کتاب وی با عنوان دستور انگلیسی آمریکایی^۲ (۱۹۴۰) متن چندین هزار نامه بررسی شده‌است. پس از این کتاب، او برای توصیف دستور زبان در کتاب ساخت زبان انگلیسی: درآمدی بر ساخت جمله‌های انگلیسی^۳ (۱۹۵۲) گفت‌وگوهای تلفنی را تحلیل و مکتوب کرد. خواننده‌ای که گمان می‌کند همه این پژوهش‌های گسترده جذاب، اما غیرکاربردی‌اند کاملاً در اشتباه است. بررسی‌های فرای و نیز پژوهش‌های گریم کندی^۴ نشان داد ۹۵ درصد بیشتر متن‌های نوشتاری را ۴۰۰۰ تا ۵۰۰۰ واژه گوناگون (یا نوع^۵) و ۸۵ درصد آن‌ها را ۱۰۰۰ «نوع» تشکیل می‌دهد. همچنین، حدود ۶۰ درصد گفت‌وگوها از ۵۰ واژه نقش‌نمای بسیار پربسامد است و حدود ۴۰ درصد ۱۰۰۰ واژه پربسامد انگلیسی اسم هستند. پیش از آنکه بررسی‌های پیکره‌بنیاد اهمیت تمرکز بر واژه‌های پربسامد را نشان دهد، کتاب‌های درسی و دستور زبان‌های آموزش انگلیسی بیشتر به واژه‌های کم‌کاربرد توجه می‌کردند. پژوهش‌های جدید، هرچند به‌طور محدود، وضعیت واقعی کاربرد زبان را روشن‌تر می‌کنند.

مخالفت چامسکیایی با تحلیل‌های کمی

برخی زبان‌شناسان نفوذ فراگیر نظریه‌های نوام چامسکی^۶ را علت مخالفت رایج با تحلیل‌های کمی زبان‌شناختی، به‌ویژه در ایالات متحده، در دهه‌های ۱۹۵۰ تا ۱۹۸۰، می‌دانند. کتاب ساخت‌های نحوی^۷ چامسکی و کتاب‌های بعدی وی مبتنی بر این فرض بود که چون کاربردهای بالقوه زبان (که او کنش^۸ می‌نامد) بی‌پایان است، توانایی (توانش^۹) سخنگویان بومی برای تولید چنین تنوع بی‌پایانی را تنها با درک شمی قاعده‌های دستوری تحلیل کرد. این قاعده‌ها شالوده تولید چنین پاره‌گفتارهایی را می‌سازند. در کتاب ساخت‌های نحوی برخی از این قاعده‌ها یا «گستارها» تصریح

1. Charles C. Fries

2. *American English Grammar*

3. *The structure of English: an Introduction to Construction of English Sentences*

4. Graeme Kennedy

5. *type*

6. Noam Chomsky

7. *Syntactic Structures*

8. *performance*

9. *competence*

شده‌است که اگر درست فهمیده شوند، می‌توانند صورت خاصی از هر پاره‌گفتار بالقوه را تبیین کنند. این کتاب موفقیتی درخشان بود (و هست)، نظریه‌ای خردمندانه و ماهرانه که به دنبال تبیین همه پاره‌گفتارهای بالقوه بود. این نظریه بر بیشتر گروه‌های دانشگاهی که به بررسی زبان می‌پرداختند تأثیر فراوانی گذاشت. نظریه چامسکی نفوذ بسیاری داشت و هر پژوهشگر زبان‌شناسی که با اصول بنیادین دستور گشتاری هم‌رأی نبود آینده خود را مبهم می‌دید.

پیروان چامسکی مخالف هر نوع رویکرد کمی در پژوهش‌های زبانی بودند و گاهی نیز این رویکردها را استهزا می‌کردند، زیرا این رویکردها نمی‌توانستند انواع بی‌پایان کاربردهای بالقوه زبانی را توضیح دهند. آن‌ها می‌گفتند فقط سخنگوی بومی می‌تواند چنین کاری کند. گریم کندی می‌گوید نفوذ چامسکی رویکردهای آماری به زبان را سرکوب کرد و در نتیجه از نفوذ پژوهش‌های پیکره‌بنیاد بر روی واژگان «از دهه ۱۹۵۰ تا احوای آن ... [در] دهه ۱۹۸۰ کاسته شد»^۱ پژوهشگر دیگری که درباره این دوره از چشم‌انداز مسلط دهه ۱۹۹۰ بحث کرده‌است می‌نویسد: «چامسکی زبان‌شناسی را شاخه‌ای از روان‌شناسی شناختی می‌داند ... که می‌توان آن را بر پایه داده‌های شمی و جمله‌های جداگانه استوار کرد. داده‌های پیکره‌ای گویا نیستند و بررسی زبان با استفاده از پیکره اساساً جذاب نیست»^۲ او می‌افزاید: «دیدگاه چامسکیایی در ۲۵ سال گذشته تغییر چشمگیری نکرده‌است»^۳ جفری لیچ^۴، یکی از پیشگامان پژوهش‌های پیکره‌ای، درباره گسست در این حوزه از پژوهش‌های زبان‌شناسی می‌نویسد: «ممکن است این گسست به‌طور مشخص از اواخر دهه ۱۹۵۰ آغاز شده باشد. چامسکی بود که زبان‌شناسی پیکره‌ای نسل اول را به حاشیه راند. دیدگاه وی درباره کامل نبودن پیکره و کامل بودن شم زبانی در نسل بعدی زبان‌شناسان نظری تبدیل به عقیده متعارف شد»^۵ لیچ از مقاله چامسکی در دانشگاه تگزاس^۶ در سال ۱۹۵۸ نقل می‌کند که هر پیکره

1. *Ibid.*, p. 340.

2. Michael Stubbs, "British Tradition in Text Analysis: From Firth to Sinclair," in Mona Baker, Gill Francis, and Elena Tognini-Bonelli (eds.), *Text and Technology: In Honour of John Sinclair* (Philadelphia: John Benjamins, 1993), p. 3.

3. *Ibid.*

4. Geoffrey Leech

5. Geoffrey Leech, "The State of the Art in Corpus Linguistics," in Karin Aijmer and Bengt Altenberg (eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik* (London: Longman, 1991), p. 8.

6. University of Texas

طبیعی بسیار تحریف و دستکاری می‌شود. در این پیکره‌ها برخی عبارات‌ها به علت بدیهی بودن، برخی به علت غلط بودن، و نیز برخی دیگر به علت مؤدبانه نبودن حذف شده‌اند. به همین علت پیکره چیزی بیش از فهرست نیست.^۱

به نوشته لیچ، اصطلاح زبان‌شناسی پیکره‌ای^۲ در دهه ۱۹۵۰ وجود نداشته‌است. زبان‌شناسانی که خود را ساخت‌گرا می‌نامیدند (برگرفته از زبان‌شناسی ساخت‌گرا^۳) به مدت بیست سال و تا زمان انقلاب رایانه‌ای به کنار رانده شدند. به گفته لیچ، اصطلاح زبان‌شناسی پیکره‌ای نخستین بار در دهه ۱۹۸۰ به کار رفت. او آن را نوعی روش پژوهشی می‌داند و نه شاخه‌ای از زبان‌شناسی.^۴ من این روش را پژوهشی زبانی می‌دانم که در آن متن‌ها در پرونده‌های رایانه‌ای گردآوری می‌شود و با به‌کارگیری روش‌های آماری می‌توان این متن‌ها را تحلیل کرد. امروزه زبان‌شناسی پیکره‌ای گاهی شاخه‌ای از زبان‌شناسی رایانشی^۵ به حساب می‌آید و ممکن است پژوهشی بر روی زبان طبیعی توصیف شود که از راه‌کارهای علم رایانه استفاده می‌کند. همان‌گونه که خواهیم دید، این مطلب که پیکره‌های زبانی از اواخر دهه ۱۹۵۰ تا ۱۹۸۰ به کلی از پیشرفت بازایستادند کاملاً درست نیست، اما در آن زمان اوضاع در دانشگاه‌های آمریکایی به گونه‌ای بود که با این روش خردمندانه پژوهش مخالفت می‌شد و پژوهش‌های پیکره‌بنیاد به ندرت در ایالات متحده یا بریتانیا انجام شد.

از همه رویدادهای خاص این دوره، هیچ‌یک به اندازه نقد معروف چامسکی تغییر جو نظری در برخورد با پژوهش زبانی تجربه‌بنیاد را به‌طور کامل ارزیابی نمی‌کند. این نقد چامسکی بر کتاب بی. اف. اسکینر^۶ با عنوان رفتار کلامی^۷ در سال ۱۹۵۷، دقیقاً همان سال انتشار کتاب ساخت‌های نحوی چامسکی، منتشر شد. این نقد در معتبرترین مجله زبان‌شناسی، *Language*، نهاد رسمی انجمن زبانی آمریکا^۸، انتشار یافت. حتی در مجله‌ای که به چاپ نقدهای مفصل معروف بود، نقد سی‌وسه‌صفحه‌ای چامسکی از

1. *Ibid.*

2. *Corpus Linguistics*

3. *Structural linguistics*

4. Geoffrey Leech, "Corpora and Theories of Linguistic Performance," in Svartvik (ed.), *Directions in Corpus Linguistics*, pp. 105-22.

5. *computational linguistics*

6. B. F. Skinner

7. *Verbal Behavior*

8. Linguistic Society of America

طولانی‌ترین نقدها بود.^۱ اسکینر پیش‌ازاین در جایگاه روان‌شناس رفتارگرای پیشرو در دوران خود کاملاً شناخته‌شده بود؛ تجربه‌های او با کبوترها نشان داده بود که با «شرطی‌سازی عامل»^۲ رفتار این حیوانات را می‌توان کنترل کرد. او در کتاب دانش و رفتار انسانی^۳ به دنبال نشان دادن این مطلب بود که چگونه اصول پایه‌ای رفتار عامل درباره انسان نیز صدق می‌کند. اسکینر از مدت‌ها پیش‌تر به زبان علاقه داشت و افراد بسیاری در انتظار کتاب او درباره این موضوع بودند. این کتاب آشکارا بر آن بود که گزارشی اساسی با موضوعی مهم به قلم یکی از روان‌شناسان تجربه‌گرای پیشرو آن دوره باشد. در این زمان اسکینر حدوداً پنجاه‌وپنج‌ساله و در اوج شهرت بود، اما چامسکی به تازگی سی‌سالگی را پشت سر گذاشته بود.

این نقد ویرانگر تنها به دنبال اثبات غیرقابل‌اعتماد بودن بسیاری از فرض‌های پایه‌ای اسکینر و در نتیجه نامعتبر بودن نتیجه‌گیری‌های او نبود (هرچند در این کار نیز موفق بود)، بلکه در پی آن بود که آن‌چنان کل کتاب اسکینر را به باد انتقاد بگیرد که هیچ مطلبی را از قلم نیندازد. در این مقاله، محدودیت‌های نقد زیر پا گذاشته شده‌است و نویسنده، حتی در این دوره آغاز راه، نشان می‌دهد که در بررسی پژوهش‌های زبانی به روشی واحد سخت پایبند است و بررسی پیکره‌بنیاد پژوهش‌های زبانی را ابزاری موجه نمی‌داند. در صفحه‌های پایانی نقد، چامسکی می‌نویسد «روشن است که افزون بر درج عناصر واژگانی در قالب‌های دستوری، موضوعات دیگری نیز در ساختار جمله دخیل هستند. هر رویکردی به زبان که این فرایندهای عمیق‌تر را نادیده بگیرد امکان ندارد بتواند در توضیح رفتار واقعی زبان موفقیت‌های بیشتری به‌دست آورد».^۴ او همچنین نوشت: «به نظر می‌رسد که ما جمله را عنصر زبانی جدیدی در نظر می‌گیریم، نه به این علت که جمله چند عنصر مشابه را با روش‌هایی ساده [تأکید از من است] به یکدیگر پیوند می‌دهد، بلکه به این علت که جمله با دستور زبانی تولید شده‌است که هرکس به‌نوعی و در قالب خاصی آن را درونی کرده‌است».^۵ چرا باید جمله‌ها به‌سادگی تشخیص داده شوند؟ چرا توانیم عنصری جدید را به‌گونه‌ای بسیار پیچیده به‌صورت

1. Noam Chomsky, Review of *Verbal Behavior*, by B. F. Skinner, *Language* 35:1 (1959), 26-58.

2. operant conditioning

3. *Science and Human Behavior*

4. *Ibid.*, p. 54.

اسکینر عمیقاً به تبیین دیگری از رفتار معتقد بود. او فکر می‌کرد رفتار خودانگیخته به اندازه ماده‌ای که خودبه‌خود به‌وجود آمده معتبر است.

5. *Ibid.*, p. 56.

جمله تشخیص دهیم؟ به نظر نمی‌رسد تنها بر مبنای این باور که فرایند تشخیص تناظر جمله‌ها آسان نیست، بتوان این نتیجه‌گیری را توجیه کرد که هرکس باید دستوری درونی شده داشته باشد تا به کمک آن جمله تولید کند. موفقیت فرهنگ‌نویسی پیکره‌بنیاد ما را کاملاً به این نتیجه سوق می‌دهد که تشخیص هر عنصر زبانی جدید، به علت اینکه آن عنصر از بعضی جهات، اغلب نه به صورت مستقیم یا ساده، شبیه عنصر آشنای دیگری است دقیقاً همان چیزی است که در یادگیری و کاربرد زبان اتفاق می‌افتد. به همین علت، زبان‌شناسان در سال‌های اخیر به کاربرد استعاره در کاربرد زبان روزمره بسیار توجه کرده‌اند.

نقد چامسکی برای همیشه به آن دسته از پژوهش‌های زبانی که رویکرد آن‌ها در راستای مسیر رفتار کلامی باشد پایان داد و به روشنی ادعا کرد برای بررسی فراگیری و کاربرد زبان فقط یک روش پذیرفته است.

پیدایش زبان‌شناسی پیکره‌ای

پیکره استانده امروزی ویراسته انگلیسی امریکایی^۱، نخستین پژوهش مهم رایانه‌بنیاد^۲، که طراحی جامعی داشت، بیشتر با نام «پیکره براون»^۳ شناخته می‌شود. دلیو. نلسون فرانسیس^۴ و هنری کوچرا^۵ در سال‌های ۱۹۶۳ و ۱۹۶۴ این پیکره را در دانشگاه براون^۶ گردآوری کردند. این پیکره شامل پانصد پاره متن دوهزارکلمه‌ای در چندین مقوله (یا ژانر)^۷ گوناگون نوشتاری است. این پیکره روی هم رفته کمی بیش از یک میلیون واژه (مورد)^۸ و حدود پنجاه هزار واژه متفاوت (نوع) دارد. پیکره براون را تعداد اندک شماری پژوهشگر مصمم و آینده‌نگر، با وجود تردیدها و بی‌توجهی‌های زیاد، گرد آوردند. به باور مؤلفان فرهنگ‌های انگلیسی و پژوهشگران زبان‌شناس تهیه این پیکره رویدادی نوآورانه بود. برخی پژوهشگران گفته‌اند که مجموعه‌ای از چند متن برگزیده می‌تواند شناختی از کاربرد زبان را فراهم آورد که با شیوه‌های دیگر نمی‌توان به آن شناخت

1. *Standard Corpus of Present-Day Edited American English*

2. computer-based

3. Brown Corpus

4. W. Nelson Francis

5. Henry Kučera

6. Brown University

7. *genre*

8. *token*

دست یافت. پیکره، برخلاف شم زبانی، سرنخ‌هایی از بافت را در اختیار ما قرار می‌دهد.

پیکره براون راه را برای پژوهش‌های مشابه انگلیسی بریتانیایی و نیز پیکره‌های بزرگ‌تر در دهه‌های بعد هموار کرد. امروزه پیکره‌ای یک میلیون کلمه‌ای کوچک به حساب می‌آید، اما در دهه ۱۹۶۰ چنین پیکره‌ای بزرگ محسوب می‌شد یا آن قدر بزرگ دانسته می‌شد که جامع در نظر گرفته شود. امروزه هیچ کس ادعا نمی‌کند که پیکره‌ای با این حجم می‌تواند شامل مجموعه‌ی جامعی از انگلیسی آمریکایی امروزی باشد، زیرا ممکن است در پیکره‌های یک میلیون کلمه‌ای، بسیاری از واژه‌ها و عبارت‌های رایج به ندرت یافت شوند یا حتی گاهی اصلاً یافت نشوند. همان‌گونه که ریچارد ویزکی^۱ ذکر کرده‌است، از واژه‌های رایجی که تنها دو بار یافت شده‌اند می‌توان به *abysmal* «افتضاح»، *checkup* «معاینه»، *landslide* «چشمگیر» و *rap* «سرزنش» اشاره کرد. واژه‌های *ballistics* «بالستیک، پرتاب‌شناسی»، *gnaw* «گاز زدن»، *invert* «وارونه کردن» و *radiate* «تابیدن» تنها یک بار در پیکره آمده‌اند.^۲ اگرچه مؤسسه‌های فرهنگ‌نویسی آمریکایی بلافاصله از پیکره براون مطلع شدند و نسخه‌های بسیاری از آن را تهیه کردند، این پیکره به سرعت بر شیوه فرهنگ‌نویسی تأثیر نگذاشت. در مقابل، این پیکره الگویی برای گردآوری پیکره‌های زبانی قابل ادغام شد. همچنین گردآورندگان این پیکره با سخاوت بسیار نسخه‌ای از آن را در قالب‌های گوناگون به قیمت تمام‌شده در اختیار دیگران قرار دادند. این امر باعث شد دیگران پیکره براون را سرمشق خود قرار دهند. جفری لیچ، بر مبنای الگوی این پیکره، در سال ۱۹۷۰ در دانشگاه لنکستر^۳ پژوهش بر روی انگلیسی بریتانیایی را بر عهده گرفت و در سال ۱۹۷۸ با همکاری دانشگاه اسلو^۴ با پشتیبانی مرکز رایانه‌ای نروژ^۵ در برگن^۶ آن را تکمیل کرد؛ این پیکره به پیکره ال.اُ.بی.^۷ معروف است. تمامی متن‌های استفاده‌شده در پیکره براون در سال ۱۹۶۱

1. Richard Venezky

2. Richard L. Venezky, "Computer Application in Lexicography," in Raven I. McDavid, Jr., and Audrey R. Duckert (eds.), *Lexicography in English* (New York: Academy of Sciences, 1973), p. 290.

3. University of Lancaster

4. University of Oslo

5. Norwegian Computing Centre

6. Bergen

7. Lancaster-Oslo/Bergen Corpus (LOB Corpus)

منتشر شده بود. متن‌های گردآوری‌شده برای استفاده در پیکره لنگستر-اسلو/برگن نیز در همان سال انتشار یافته بودند؛ ژانرها نیز، تنها با اختلافات جزئی در تعداد نمونه‌های متعلق به هر کدام همانند یکدیگر انتخاب شده بودند. در ادامه، ژانرهای استفاده‌شده در پیکره براون به همراه تعداد نمونه‌های دوهزارکلمه‌ای در هر ژانر آمده‌است:^۱

۴۴	الف) روزنامه: گزارش
۲۷	ب) روزنامه: سرمقاله
۱۷	پ) روزنامه: نقد و بررسی
۱۷	ت) دین
۳۶	ث) مهارت‌ها و سرگرمی‌ها
۴۸	ج) دانستنی‌های عمومی
۷۵	چ) ادبیات، زندگی‌نامه و غیره
۳۰	ح) متفرقه
۸۰	خ) نوشته‌های علمی و فرهنگی
۲۹	د) داستان: عمومی
۲۴	ذ) داستان: معمایی و پلیسی
۶	ر) داستان: علمی
۲۹	ز) داستان: ماجراجویانه و وسترن
۲۹	ژ) داستان: رمانتیک و عاشقانه
۹	س) طنز

سندهای دولتی، گزارش‌ها دربارهٔ نهادها و صنایع، فهرست‌های توصیفی انجمن‌ها و خبرنامه‌های بنگاه‌های صنعتی همگی زیرمجموعهٔ مقولهٔ «متفرقه» هستند. هر مقولهٔ گسترده به زیربخش‌هایی تقسیم می‌شود تا مطمئن شویم به هیچ مقوله‌ای (مانند ورزش یا سیاست در زیربخش گزارش روزنامه) بیش‌ازاندازه پرداخته نشده یا مقوله‌ای نادیده گرفته نشده‌است.

پیکره لنگستر-اسلو/برگن امکان مقایسهٔ مستقیم کاربردهای معاصر انگلیسی بریتانیایی و آمریکایی را فراهم آورد؛ به همین علت، می‌توان با داده‌های آماری معتبر گفت که بسامد چه واژه‌هایی در گونه‌ای نسبت به گونهٔ دیگر بیشتر است. واژه‌های

1. sample

2. Henry Kučera and W. Nelson Francis, *Computational Analysis of present-Day American English* (providence: Brown University Press, 1967), p. xix.

membership «عضویت»، *college* «دانشکده» و *campus* «دانشگاه» در انگلیسی آمریکایی به مراتب پرکاربردتر از انگلیسی بریتانیایی است، درحالی‌که واژه‌های *introduced* «معرفی شده»، *introduction* «معارفه» و *pupils* «دانش‌آموزان» در انگلیسی بریتانیایی پرکاربردتر از انگلیسی آمریکایی است.^۱ افزون‌براین، در هر دو پیکره براون و لنگستر-اسلو/برگن می‌توان به بسامد واژه‌های خاص در ژانرهای خاص پی برد. ازاین‌رو، می‌توان انگلیسی بریتانیایی و آمریکایی را از دیدگاه ژانرهای یکسان نیز مقایسه کرد. برای نمونه، ممکن است متوجه شده باشید که واژه‌ای به انگلیسی آمریکایی، که در نوشته‌های سیاسی متداول است، در انگلیسی بریتانیایی اغلب در داستان به کار می‌رود. آشکار است که این نوع اطلاعات برای پژوهش‌های زبانی و به‌ویژه مقایسه دو گونه متفاوت ارزشمند است، اما اهمیت واقعی این دو پیکره در این موضوع نهفته است که برای بسیاری از پیکره‌های بزرگ‌تر که در نهایت از پیکره‌های پیش از خود الگوبرداری می‌کردند سنجه‌ای را پایه‌گذاری کردند. اگر دوراندیشی فرانسيس و کوچرا نبود، پیکره‌های بزرگ‌تری که امروزه مؤلفان فرهنگ‌ها از آن‌ها بهره می‌برند گردآوری نمی‌شدند یا پر از اشتباهات اجتناب‌پذیر بودند.

در بریتانیا، حتی پیش از اینکه پیکره براون به جریان بیفتد، رندالف کوئرک^۲ در یونیورسیتی کالج لندن^۳ کار بررسی کاربرد انگلیسی^۴ را آغاز کرده بود. سرنوشت این اثر نیز، مانند پیکره براون، این بود که نقطه عطفی در پیشرفت زبان‌شناسی پیکره‌ای شود، هرچند این اثر با رایانه گردآوری نشده بود و هنوز به‌صورت برگه‌دان‌های پرشماری با برگه‌هایی کاغذی به ابعاد حدوداً ۱۵×۱۰ سانتی‌متر در قفسه‌های پرونده نگهداری می‌شود. در سال ۱۹۵۹، کوئرک دوپست نمونه از انگلیسی بریتانیایی گرد آورد. هریک از این نمونه‌ها شامل پنج‌هزار کلمه می‌شدند که درمجموع پیکره‌ای یک‌میلیون کلمه‌ای را تشکیل می‌دادند. نیمی از این نمونه‌ها گفتاری و نیم دیگر نوشتاری بودند. از این پیکره برای پژوهش‌های دستوری و نحوه کاربردهای سخن‌گویان و نویسندگان تحصیل‌کرده بریتانیایی استفاده شد. در این میان، موضوع بسیار جالب میزان توجه به گفتار بود، که در آن هنگام در مقایسه با اطلاعات نوشتاری به آن بی‌توجهی

1. Knut Hofland and Stig Johansson, *Word Frequencies in British and American English* (Bergen: Norwegian Computing Centre for the Humanities, 1982).

2. Randolph Quirk

3. University College London

4. *Survey of English Usage*