

التعريف

هو دراسة اللغة على أساس مجموعات كبيرة من (Corpus Linguistics) إن علم اللغة النصوصي أو العيني استخدامات اللغة "الواقعية" المخترنة في قواعد بيانات محوسبة (مجاميع) تم إنشاؤها من أجل البحث اللغوي. ويعرف (Nordquist, 2018) أيضا باسم الدراسات المستندة إلى قواعد البيانات اللغوية

المجاميع هي مجموعة من البيانات اللغوية يتم تجميعها كنصوص مكتوبة أو نسخ مكتوبة لكلام مسجل. الغرض الرئيسي للمجاميع هو التأكد من الفرضيات حول اللغة. على سبيل المثال، لتحديد الاختلاف في استخدام صوت، كلمة، أو تركيب نحوي معين. يتعامل علم اللغة النصوصي مع مبادئ وممارسة استخدام المجاميع في دراسة اللغة (Crystal, 1992: 85).

ينظر بعض علماء اللغة إلى علم اللغة النصوصي كأداة بحث أو منهجية، ويراها آخرين كنظام أو نظرية في حد ذاته. أن الإجابة على سؤال ما إذا كان علم اللغة (Kuebler and Zinsmeister, 2015)) وبخلص كيلر وزينسميستر النصوصي نظرية أو أداة هي ببساطة أنه يمكن أن يكونا كليهما. ويعتمد ذلك على كيفية تطبيق علم اللغة النصوصي

على الرغم من أن الطرق المستخدمة في علم اللغة النصوصي تم تبنيها لأول مرة في أوائل الستينيات، إلا أن مصطلح علم اللغة النصوصي لم يظهر حتى الثمانينيات

أمثلة وملاحظات

إن علم اللغة النصوصي هو عبارة منهجية تضم عددًا كبيرًا من الطرق المرتبطة التي يمكن استخدامها من قبل العلماء ذوي الميول النظرية المختلفة. من ناحية أخرى ، لا يمكن إنكار أنه عادة ما يتم ربط علم اللغة النصوصي أيضًا بوجهة نظر معينة حول اللغة. جوهر هذه النظرة هو أن قواعد اللغة تستند إلى الاستخدام، وأن التغييرات تحدث عندما يستخدم المتحدثون اللغة للتواصل مع بعضهم البعض. فالحجة هي أنك إذا كنت مهتمًا بكيفية عمل لغة معينة، مثل اللغة الإنجليزية، فسيكون من الجيد دراسة اللغة وهي قيد الاستخدام. إحدى الطرق الفعالة للقيام بذلك هي استخدام (Hans Lindquist, 2009) منهجية علم اللغة النصوصي

ازدهرت دراسات قواعد البيانات اللغوية منذ عام 1980، حيث أصبحت التقنيات والحجج الجديدة لصالح استخدام قواعد البيانات اللغوية أكثر وضوحًا. وتستمر هذه الطفرة حاليًا، كما أن كلتا مدرستي علم اللغة النصوصي لازالت تتطور. إن علم اللغة النصوصي لازال في حالة نضوج من الناحية المنهجية ونطاق اللغات التي يعالجها علماء علم (McEnery and Wilson, 2001) اللغة النصوصي يتزايد سنويًا

علم اللغة النصوصي في الفصل الدراسي

في سياق الفصول الدراسية ، تعد منهجية علم اللغة النصوصي ملائمة للطلاب من جميع المستويات لأنها تدرس اللغة من أسفل وتتطلب خبرة مكتسبة قليلة جدًا للبدء بها. حتى الطلاب الذين يدخلون إلى البحث اللغوي من دون أدوات نظرية يتعلمون بسرعة كبيرة لتقديم فرضياتهم على أساس ملاحظاتهم بدلا من المعرفة المتلقية، واختبارها بالأدلة (Tognini-Bonelli, 2001) المقدمة من قواعد البيانات اللغوية

وللاستفادة من موارد قواعد البيانات اللغوية، يحتاج المعلم إلى تكيف عادي مع الروتين المتضمن في استرجاع (Sinclair, 2004) المعلومات من قواعد البيانات اللغوية، والأهم من ذلك، التدريب والخبرة في كيفية تقييم هذه المعلومات (2004).

التحليلات الكمية والنوعية

تعتبر التقنيات الكمية ضرورية للدراسات القائمة على قواعد البيانات اللغوية. فعلى سبيل المثال، إذا كنت ترغب المرء الإنجليزية، فسيحتاج إلى معرفة عدد المرات التي 'large' و 'big' في مقارنة استخدامات اللغة لأنماط كلمتي تحدث فيها كل كلمة في قاعدة البيانات، وكم عدد الكلمات المختلفة التي يتزامن حدوثها مع كلا هاتين الصفتين (المتراقات) ، ومدى انتشار كل من هذه التراكيب، وهذه كلها قياسات كمية

إن جزءًا هامًا من المنهج القائم على مجموعة قواعد البيانات اللغوية يتخطى الأنماط الكمية لاقتراح تفسيرات“ وظيفية تشرح أسباب وجود هذه الأنماط. ونتيجة لذلك، فإن قدرًا كبيرًا من الجهد في الدراسات المستندة إلى قواعد (Biber et al., 2004) البيانات اللغوية مكرس لشرح الأنماط الكمية وتجسيدها

يتم استخدام الأساليب الكمية والنوعية لعلم اللغة النصوي مع بعضها مجتمعة على نطاق واسع. كما أنه من خصائص علم اللغة النصوي أن يبدأ بالنتائج الكمية، ويتجه نحو التحليلات والنتائج النوعية. ولكن قد يكون لهذا الإجراء عناصر دورية. ومن المستحسن إخضاع النتائج الكمية للتدقيق النوعي ومحاولة تفسير لماذا يحدث تردد نمط معين، مثلًا. ومع ذلك، قد يكون التحليل النوعي (الاستفادة من قدرة الباحث على تفسير عينات من اللغة وهي في سياقها) وسيلة لتصنيف الأمثلة في قاعدة بيانات معينة حسب معانيها؛ وقد يكون هذا التحليل النوعي هو المدخل إلى تحليل كمي (Leech et al. 2012) إضافي يعتمد على المعنى

الخلاصة

إن علم اللغة النصوي نهج لدراسة اللغة التي يبتعد بشكل جذري عن حدس الباحث (الذي غالبًا ما يكون غير موثوق) حول ما هو صحيح نحويًا أو لا، أو ما هو شائع أو غير شائع ، وما إلى ذلك ، من خلال النظر بدلاً من ذلك إلى ما يحدث فعلاً عند استخدام الناس للغة. يتم تسجيل هذا الاستخدام اللغوي في مجاميع (قواعد بيانات لغوية)، والتي هي في الأساس مجموعات كبيرة من النصوص، حيث يفهم من كلمة 'نص' إما النص المكتوب أو قطعة خطاب منطوقة تكون عادة في بيئة عفوية (على سبيل المثال جلسة محادثة). وعادة ما تحتوي المجاميع على كل من اللغة المنطوقة والمكتوبة. قد تحتوي مجاميع اليوم على عدة ملايين من الكلمات، أو حتى أكثر من مليار كلمة. وهي تسمح لنا بمقارنة استخدام اللغة في الأنواع و اللهجات الخاصة المختلفة (منطوق مقابل مكتوب، أو الأدب الخيالي مقابل الكتابة الأكاديمية، الخ). هناك أيضا مجاميع تاريخية تحتوي على مواد لغوية من فترات مختلفة. وهناك مجاميع موازية تحتوي على نصوص بأكثر من لغة واحدة. وبالطبع، من الواضح أن تفضيلا اختيار مجموعة (قاعدة بيانات لغوية) واحدة على أخرى سيعتمد على سؤال البحث

إن طريقة المجاميع اللغوية إذن، إذا كان هناك بالفعل طريقة واحدة، هي استخدام قاعدة بيانات لغوية للإجابة على سؤال لغوي. تعتبر أبحاث قواعد البيانات اللغوية مناسبة بشكل مثالي للإجابة عن الأسئلة حول التردد والتوزيع في الأنواع واللهجات الخاصة المختلفة. في الواقع ، هناك العديد من الطرق المحددة التي يتم تطبيقها من قبل علماء علم

اللغة النصوصي، بدءاً من مجرد عد العناصر المعجمية التي تحدث في مجموعة البيانات اللغوية إلى ترميز الجمل المستخرجة من قواعد البيانات اللغوية لمجموعة من المعلمات التي تهم الباحث. لتجسيد الخيار الأخير وجعله أكثر واقعية قليلاً ، لنفترض أن الباحث يريد التحقيق في استخدام المبني للمجهول في اللغة الإنجليزية ، عندئذٍ يمكنه أولاً أن يسترجع من مجموعة قواعد بيانات اللغة الإنجليزية كل الجمل التي تحتوي على التصريف الثالث للفعل (وهو أمر ممكن لأن معظم قواعد البيانات اللغوية اليوم تبينها بوضع علامة عليها تبين نوع الكلمات) والتخلص من جميع الأمثلة التي ليس لها صلة. قد ينظر بعد ذلك في كل جملة ذات صلة على حدة، ويشير يدويًا، في ورقة اكسل على سبيل وما إذا كان ، 'by' المثال، إلى ما إذا كان الفاعل حي أو جماد، وما إذا كانت هناك عبارة جار ومجرور بحرف الجر الفاعل يشير إلى نفسه أو إلى كيانًا مختلفًا كفاعل للجملة السابقة (على افتراض أنه يمكنه استرجاع جزء من سياق الجمل)، وهكذا دواليك.

أما بالنسبة للتحليل ، إذا أخذت الخيار الثاني على سبيل المثال ، يمكن بعد ذلك معالجة البيانات (التي تشكل أساسًا مجموعة كبيرة من القيم للمتغيرات) إحصائياً للحصول على صورة أوضح لكيفية استخدام المبني للمجهول ومتى يتم استخدامه. يجب الأخذ في الاعتبار أنه قد يكون هناك طرق أفضل للإجابة على هذا السؤال بمساعدة مجموعة قواعد البيانات. من الناحية المثالية، يجب على المرء أيضًا تضمين جمل المبني للمعلوم في البحث، لمعرفة الظروف التي يكون فيها احتمال تعزيز استخدام جملة المبني للمجهول أو لا (مقارنة بالمبني للمعلوم). كما أن هناك أشكال أخرى للتحليل. فبدلاً من ذلك، قد لا يهتم الباحث بالنتائج العددية الدقيقة ويفضل بدلاً من ذلك تحليلاً "نوعياً" متعمقاً لأمثلة مختارة مسترجعة من مجموعة قواعد البيانات

غالبًا ما يكون تفسير النتائج التي تم الحصول عليها من خلال البحث في المجاميع اللغوية هو الجزء الأصعب، ولكن لا ينبغي أن يكون كذلك. إن إعطاء الأرقام فقط من أجل إعطاء أرقام ليس مثيراً للاهتمام. ولهذا السبب، إذا كان لدى الباحث سؤال واضح من البداية ، فإن طريقة البحث الدقيقة، والتحليل الكمي أو النوعي وتفسير النتائج يجب أن يقدموا أنفسهم بشكل طبيعي. على سبيل المثال ، إذا رغب المرء في التحقق من الادعاء الذي غالباً ما يُسمع أن المبني للمجهول يعتبر نموذجي للكتابة الأكاديمية فقط، وأنه يجب تجنبه في الأماكن الأخرى، فيمكنه إخضاع ذلك للاختبار في مجموعة قواعد بيانات لغوية. قد يكون التفسير الصحيح للنتائج الذي توصل إليها هو أنه يكشف أن المبني للمجهول لا يقتصر على السجلات الرفيعة، وأنه في الواقع طبيعي تمامًا في الكلام غير الرسمي (على سبيل المثال I'm stuck : The room's been done).

المراجع

Crystal, David. 1992. *An Encyclopedic Dictionary of Language and Languages*. Oxford.

Douglas Biber, Susan Conrad, and Randi Reppen, *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge University Press, 2004.

Elena Tognini-Bonelli, *Corpus Linguistics at Work*. John Benjamins, 2001.

Geoffrey Leech, Marianne Hundt, Christian Mair, and Nicholas Smith, *Change in Contemporary English: A Grammatical Study*. Cambridge University Press, 2012.

Hans Lindquist, Corpus Linguistics and the Description of English. Edinburgh University Press, 2009.

John McHardy Sinclair, How to Use Corpora in Language Teaching, John Benjamins, 2004.

Kuebler and Zinsmeister (2015). Corpus Linguistics and Linguistically Annotated Corpora.

Nordquist, Richard. "corpus linguistics." ThoughtCo, Dec. 7, 2018, [thoughtco.com/what-is-corpus-linguistics-1689936](https://www.thoughtco.com/what-is-corpus-linguistics-1689936).

Tony McEnery and Andrew Wilson. Corpus Linguistics, Edinburgh University Press, 2001.

شارك هذا الموضوع

- انقر للمشاركة على فيسبوك (فتح في نافذة جديدة)
- (فتح في نافذة جديدة) اضغط للمشاركة على Google+
- اضغط للمشاركة على تويتر (فتح في نافذة جديدة)
- (فتح في نافذة جديدة) اضغط لتشارك على LinkedIn
- اضغط لإرسال هذا الموضوع لصديق بواسطة البريد الإلكتروني (فتح في نافذة جديدة)
- (فتح في نافذة جديدة) اضغط للمشاركة على Tumblr
- اضغط للطباعة (فتح في نافذة جديدة)