

تسليط الضوء على تنوع اللسانيات وتطورها

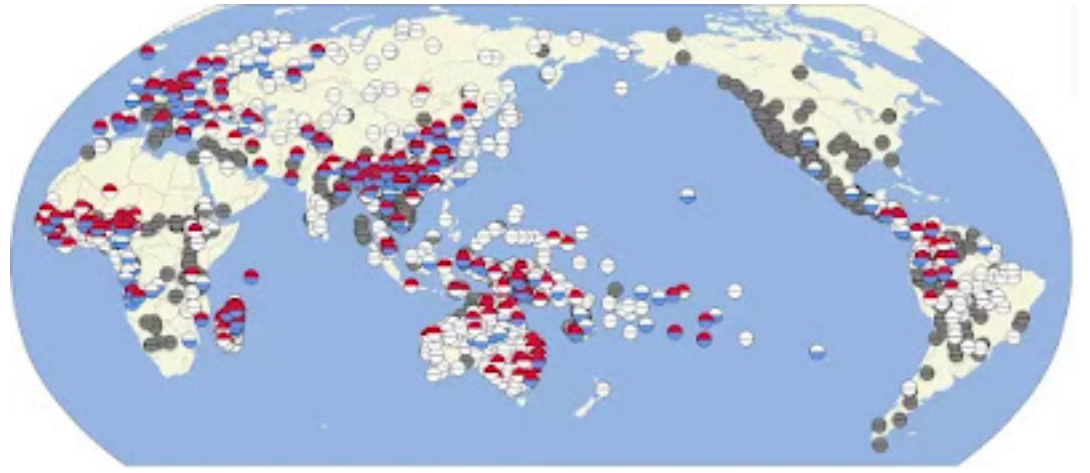




يحتوي بنك معجم المفردات Lexibank

على بيانات معجمية موحدة للمعايير لأكثر من 2000 لغة وسوف تمهد الطريق لرؤى جديدة في تطور التنوع اللغوي.

هل صحيح أن العديد من اللغات في العالم تستخدم كلمات مشابهة لكلمة "ماما" و "بابا" ل "أمي" و "أبي"؟ إذا كانت اللغة تستخدم نفس الكلمة بالضبط لكل من "الذراع" و "اليد"، فهل تستخدم أيضًا نفس الكلمة بالضبط لكل من "الساق" و "القدم"؟ كيف تتمكن اللغات من استخدام عدد قليل نسبيًا من الكلمات للتعبير عن العديد من المفاهيم؟ أنشأ فريق متعدد التخصصات من اللسانيين وعلماء الحوسبة وعلماء النفس قاعدة بيانات عامة كبيرة يمكن استخدامها لدراسة هذه الأسئلة والعديد من الأسئلة الأخرى بمساعدة الأساليب الحاسوبية.



الشكل: تستخدم العديد من اللغات في العالم كلمات مثل "ماما" أو "بابا" للإشارة إلى "أمي" و "أبي". اللغات التي فيها كلمة "ماما" ملونة باللون الأحمر، واللغات التي فيها كلمة "بابا" ملونة بالأزرق.

معايير جديدة وبرامج جديدة

بنك المفردات المعجمية Lexibank يوفر بيانات في شكل قوائم كلمات موحدة لأكثر من 2000 لغة مختلفة. "العمل على البنك المعجمي Lexibank تزامن مع الدفع باتجاه نموذج (استمارات) بيانات أكثر اتساقًا في قواعد البيانات اللسانية. وبالتالي يمكن أن يخدم بنك المفردات المعجمية Lexibank كمثال واسع النطاق على فوائد توحيد المعايير ومحفزًا لمزيد من توحيد المعايير،" وفقًا لتقرير روبرت فوركل Forkel Robert، الذي أشرف على الجانب الحوسبي من جمع البيانات. "فقد قررنا إنشاء معاييرنا الخاصة، والتي أسميناها نماذج البيانات لكل اللغات، والتي تستخدم حاليًا بنجاح في العديد من المشاريع التي يشارك فيها قسمنا".

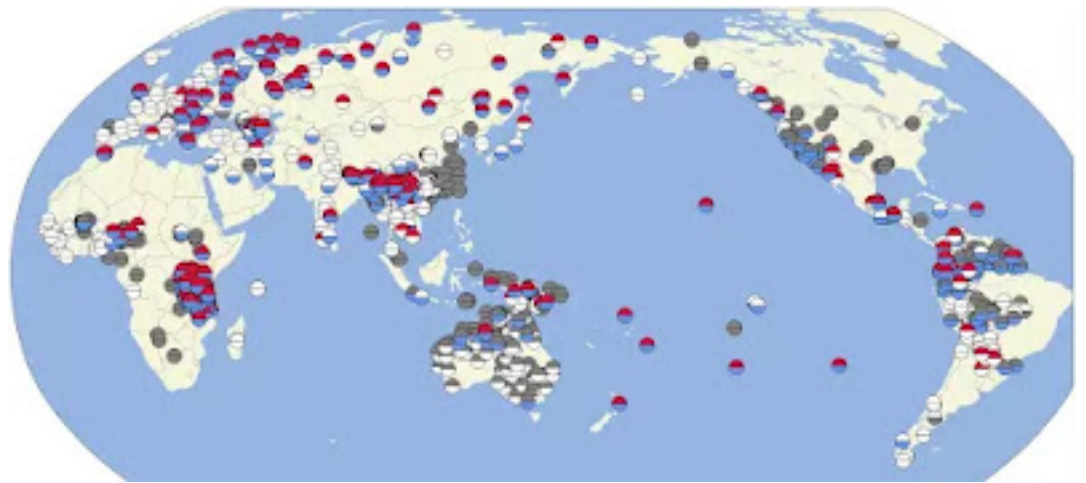
المعايير الجديدة التي اقترحها الفريق مصحوبة بوسائل برمجية جديدة تسهل بشكل كبير سير عمل workflows علماء اللسانيات. يقول يوهان ماتيس ليست List Mattis-Johann، الذي أشرف على الجزء العملي من تنسيق البيانات: "لقد صممت مهام سير عمل جديدة بمساعدة الكمبيوتر تمكننا من جعل مجموعات البيانات اللغوية الحالية قابلة للمقارنة." "باستخدام مهام سير العمل هذه، قمنا بزيادة كفاءة توحيد معايير البيانات وتنسيقها بشكل دراماتيكي."

التعرف على أنماط تطور اللغة

بالإضافة إلى جمع البيانات المعيارية للغات ومشاركتها، صمم الباحثون أيضًا تقنيات حوسبية جديدة للإجابة على أسئلة بخصوص تطور التنوع اللغوي. بيّن الباحثون كيف يمكن استخدام هذه الأساليب وذلك بعمل حوسبة لمدى اختلاق اللغات أو توافقها بالنسبة لستين (60) ميزة مختلفة.

بفضل تمثيل بيانات اللغات الموحدة المعايير، أصبح من السهل حاليًا التحقق من عدد اللغات التي تستخدم كلمات مثل "ماما" و "بابا" لـ "أمي" و "أبي"، كما ذكر جوان ليست. ويضيف سيمون ج. غرينهيل J. Simon, Greenhill، أحد مؤسسي مشروع بنك المفردات المعجمية Lexibank: "اتضح أن هذا النمط يمكن أن يوجد بالفعل في العديد من لغات العالم وفي مناطق مختلفة جدًا." "نظرًا لأن جميع اللغات التي فيها هذا النمط لا ترتبط ارتباطًا وثيقًا ببعضها البعض، فإنها تعكس تطورًا متوازيًا ومستقلًا تمامًا، كما اقترح عالم اللسانيات الروسي الكبير رومان ياكوبسون في عام 1968".

توسيع البيانات وتطوير أساليب جديدة



الشكل: تستخدم العديد

من اللغات التي تستخدم نفس الكلمة لـ "الذراع" و "واليد" (الملونة باللون الأحمر) تستخدم أيضًا نفس الكلمة لـ "الساق" و "القدم" (ملونة باللون الأزرق).

مجموعة البيانات الجديدة وسمات اللغة المحوسبة أو توماتيكيًا ستساهم في تكوين رؤى جديدة عن الأسئلة المفتوحة في التنوع اللساني وتطور اللغة. يقول ليست: "لا أحد يعتقد أن التحليل يجب أن يتوقف عند الأمثلة التي سقناها في ورقتنا." "بل بالعكس، نأمل أن تحفز هذه الأمثلة علماء اللسانيات وعلماء النفس وعلماء التطور للبناء عليها وذلك توسيع البيانات وتطوير أساليب جديدة،" يضيف فوركل.

حتى في دراستهم الحالية، قدم الباحثون نتائج تستدعي إجراء دراسات مستقبلية. " عند البحث عن أي اللغات تستخدم نفس الكلمة لـ "الذراع" و" اليد"، وجدنا أن هذه اللغات عادةً ما تستخدم نفس الكلمة أيضًا لـ "الساق" و"والقدم"، كما ذكر ليست. "على الرغم من أن هذا قد يبدو بمثابة صدفة سخيفة، إلا أنه يبين أن معجم مفردات اللغات البشرية غالبًا ما يكون منظمًا أكثر بكثير مما قد يفترضه المرء عند دراسة لغة واحدة بعزلة عن غيرها من اللغات".