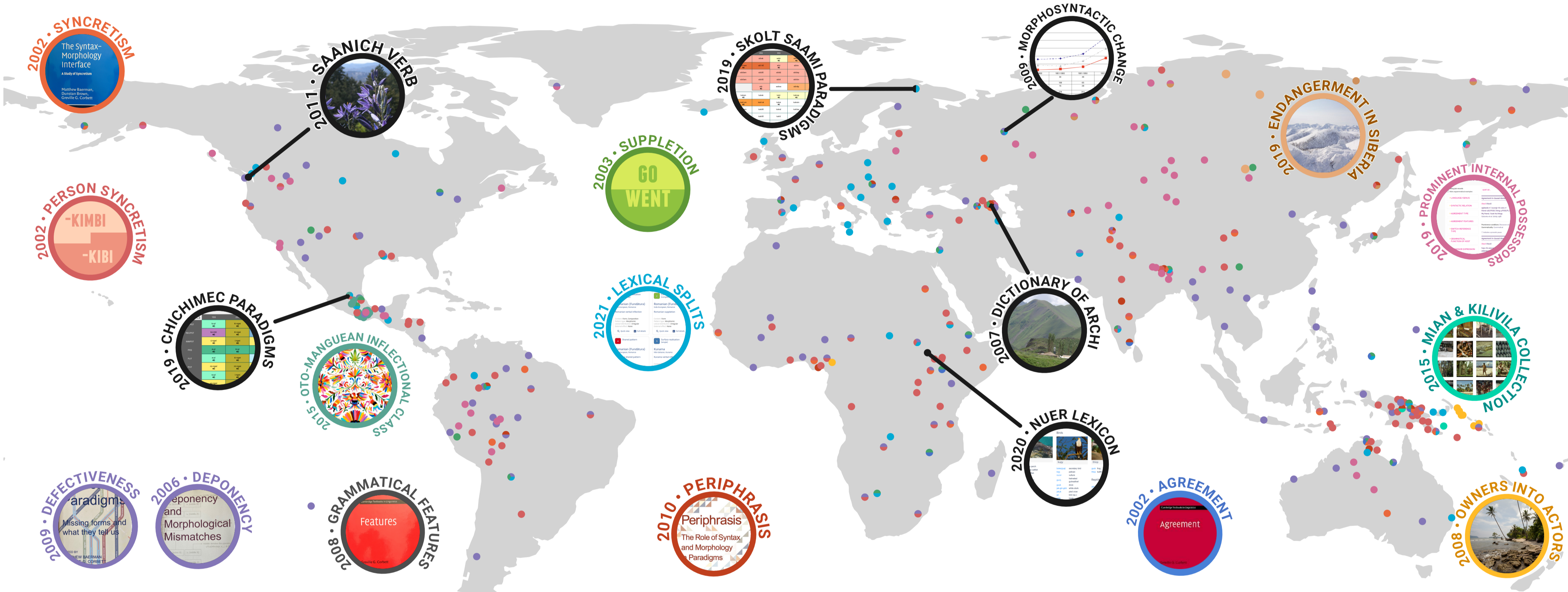


A collective work. Poster by Sacha Beniamine & Erich Round



Background

- ▶ In the 1990s most linguistic research was based on language data which were not publicly available.
- ▶ This made it impossible to verify theoretical claims.
- ▶ **SMG** committed instead to high-value databases that were freely accessible.

Methods

- ▶ Open data supporting replicability.
- ▶ Interactive visualizations of complex relationships.
- ▶ Dynamic engagement with language communities.
- ▶ DOIs and cite buttons promoting good citation practices.

Results

- ▶ 20 public databases created over 20 years.
- ▶ They document 325 languages around the world, across 100 families.
- ▶ Around 200 of these languages are vulnerable, endangered or extinct.
- ▶ Colleagues internationally report usage in teaching, research, and research design.
- ▶ Hundreds of citations.

Conclusions

- ▶ Early adoption of open data was one of the keys to SMG's success.
- ▶ Ensuring technical continuity is a challenge still awaiting a stable solution.
- ▶ We are always improving our open research practices, creating new standards and methodologies.

Many SMG members and collaborators have contributed to the databases: Matthew Baerman, Julia Baron, András Bány, Oliver Bond, Dunstan Brown, Lydia Byrne, Lynne Cahill, Marina Chumakina, Greville Corbett, Gerrit Jan Dimmendaal, Nick Evans, Roger Evans, Sebastian Fedden, Timothy Feist, Edward Garrett, R. J. Hayward, Andrew Hippisley, Nicholas Jenkins, Elena Kalinina, Anna Kibort, Alexander Krasovitsky, Alison Long, Marianne Mithun, Irina Monich, Irina Nikolaeva, Enrique Palancar, Bill Palmer, Harley Quilliam, Tatiana Reid, Sandy Ritchie, Erich Round, Helen Sims-Williams, Gregory Stump, Carole Tiberius, Claire Turner.

