

Linguist-Driven Tools

The following list is excerpted from the draft document, "Linguist-Driven Tools: A Manifesto," given by Arienne M. Dwyer as a plenary lecture at the annual meeting of the Deutsche Gesellschaft für Sprachwissenschaft, 24 February 2006, Bielefeld, Germany.

Comments and additions encouraged: anthlinguist@ku.edu

Current functionality gaps

- Metadata input devices and tools; metadata conversion software
- Output tools
- Querying and analysis tools (e.g. semistructured queries and statistical tools)
- Analytical tools linking texts, media, and lexica
- Versioning tools

Strategic directions for the field

- Use regional workshops and funding mechanisms to encourage cooperative work between language engineers and ordinary working linguists;
- Explore tools development from a USER FUNCTIONALITY perspective, taking into account that users cannot be experts in all stages of linguistic data processing (discovering/collecting, digitizing, annotating, querying, migrating, archiving, and disseminating);
- Tool design must address the non-English-speaking, community-based segment of users;
- The alleged distinction between descriptive and theoretical linguistics is artificial; annotated descriptive data **are** theoretical, and we need to promote the idea that theory need description and vice-versa;
- Empirical, corpus-based research is coming in from the eccentric fringe of linguistics to occupy center stage,
- We should train our students to **expect** that tools and technologies will change rapidly;
- They should also expect the majority of these tools to share basic sound design features (e.g. open standards, open source, extensible, cooperatively developed, etc.);
- Tools must be shown to be useful before they are adapted by linguists;
- It is BOTH linguists' and tools developers' responsibility to jointly plan out tools whose basic functions (at least) are easy to use;
- It is our responsibility as linguists to clearly state our tools needs (i.e. Functionalities), and to not avoid new tools just because they seem too technically challenging;
- Pretty good practice can be good enough for the time being, but we should require it of ourselves and our students;
- Linguists need to work together with experienced service providers (e.g. libraries) to begin to develop archives.