

1. What are the most pressing needs among possible cyberinfrastructure and/or digital tools for linguistics?

The biggest problem with computer aided linguistics today is that using computer tools is no where near as easy as using a pen and pencil. Computer tools do not have anything like the flexibility of such a simple and elegant system. For this reason, many linguists will not even try to use them.

The next biggest problem is that computer systems do not have any system of cataloguing comparable to that of a simple set of templates and a good filing system. True, computers have sophisticated RDBMS's, but it is nearly impossible to put linguistics data into a serious database that works on multiple computer platforms, has even the slightest bit of an intuitive interface, and is free for students to use.

Anyone who can read English and has linguistics training can quickly learn to use an English catalogue of any language . To use a database you need the correct operating system and the right kind of computer. Programming a database is surprisingly hard, and in any case, the process of making a relational database work is not anything like the way a non-computer program might imagine. Once a database is set up, it is equally hard to change anything. Since every language is different, each researcher will want different data structures, so a generic structure won't help. Alternatives such as shoebox can help a lot, but they also have a failing which leads to the third problem.

Linguists and programmers have been responding to the first two issues for years, but they have done so in a piecemeal fashion. In an ideal situation, linguists want a computer program that will record sound files, save the raw data, let them clean the data up, label and segment the data, tag the sentences, words, morphemes, glosses, and segments, update information in a CVS style environment, search through the whole mess intuitively, and print standard orthography, phonetics, morphology, gloss, translations and citations right into any word processing document I'll be using for publication.

The situation is better than it was even a few years ago; many of these tasks can now be done on computer. Record a file using Audacity and a good microphone/preamp, clean the data with Audacity, transcribe using Transcriber 1.5.2, label and segment data in PRAAT, record the syntax/morphology in shoebox - it all helps. But this method is missing pieces of the puzzle. It lacks easy backup of historical work, lacks an easy way to extract data into publishable forms, severely lacks in intuitive organization let alone behavior, and isn't even multi-platform anymore. We're grateful for these tools, but we believe it is possible make tools that get wha we need so that data on an endangered can be entered more easily, updated more safely, and shared more effectively.