

Issue Statement

Gary F. Simons
SIL International
<http://www.sil.org/~simonsg/>

In considering digital tools for linguistics, it is easy to focus on tools that will help linguists to do their work of language documentation and description. But this is only the tip of the iceberg. We need to be equally concerned about ensuring the long-term preservation of the digital resources developed by linguists and about fostering the use of those resources by the citizens of cyberspace who are increasingly coming to include the people whose languages are being documented and described. Thus the cyberinfrastructure for linguistics needs tools for four sets of key players:

1. Linguists need tools to help them create digital language documentation and description in archival forms that are suitable for long-term preservation.
2. Archives need tools that help linguists deposit the results of their work and that make the results available in appropriate ways over the web to the public and to interoperating services.
3. Users of language resources (including linguists and language community members as well as the public at large) need web-based tools (i.e. interoperating services) that make it easy for them to find, access, and manipulate language resources no matter where in the world they happen to be housed in an archive.
4. Service providers need tools (which are interoperation standards and development tools based on those standards) that allow them to build the end-user services that interoperate globally across the holdings of all known language archives.

The first of these categories tends to get the most focus in discussions of digital tools for linguistics. While there seems to be broad support for the notion that we should develop a global cyberinfrastructure of domain-specific services interoperating across all known language archives, many regard that as a distant dream. A common argument is that the global infrastructure is a moot point until the individual language resources exist in best practice archival forms that are interoperable; therefore our highest priority should be to develop tools for the linguist that will automatically generate the results of their work in such archival forms. The problem with this strategy is that it does not provide any motivation for linguists to adopt such tools and to deposit their results in archives. People do not change their ways of working without compelling reasons.

Thus I believe that building the infrastructure for global interoperation is the higher priority, and this means focusing on the latter three categories above. We need to create such compelling web-based services that they become focal points for language and linguistics in cyberspace. Then linguists and language archives alike would be queuing up to learn how they could adapt their ways of working to turn the resources they create and curate into resources that are plugged into the cyberinfrastructure.