

Toward the Interoperability of Language Resources

July 13-15, in conjunction with the 2007 LSA Summer Institute



Issue Statement

Gary F. Simons
SIL International

<http://www.sil.org/~simonsg/>

The greatest barrier to interoperation of language resources is the predominant culture within linguistics of individual researchers using idiosyncratic means to produce published results. In considering digital support for linguistics, it is easy to focus on stand-alone tools that will help linguists to keep doing their work of language documentation and description. But this is only the tip of the iceberg. We also need to be concerned about ensuring the long-term preservation of the digital resources developed by linguists. And we need to be concerned about fostering the use of those resources not just by linguists but by all 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 to support 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 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 above prescription focuses on achieving new heights of information sharing, but still has a bias toward linguists working individually. Pushing the envelop of linguistic practice even



further, we should be investigating how interoperation in the context of Web 2.0—the emerging platform for global collaboration—will allow our field to achieve new heights of productivity as virtual teams comprising seasoned scholars, linguists-in-training, native speakers of the language, and enthusiastic amateurs begin working in community to document and describe the languages of the world.

