

Presley A. Ifukor  
Doctoral Student, Computational Linguistics Group  
Institute of Cognitive Science  
University of Osnabrueck, Germany  
Email: pifukor@uos.de

## **HOW DOCUMENTATION TOOLS CAN MAKE LANGUAGE RESOURCES MORE READILY AVAILABLE FOR LINGUISTIC RESEARCH**

I shall respond to the theme above by suggesting how documentation tools can enhance the scientific and computational applications of linguistic resources for research. Such tools should be developed to cater for as many languages as possible, including African tonal languages. I examined some of these issues at a conference in Vienna, Austria last December. Participating in the Digital Tools Summit for Linguistics (DTS-L) will be a good opportunity for translating those ideas to practical use.

As Hans d'Orville (1996: 484; *Tackling Information Poverty*) rightly remarks, Information and knowledge...have strategic consequences for the global power constellation... Comparative advantages are henceforth expressed in the ability of countries to collect, process, tie together, act upon, and disseminate information through communications, information processing technologies and complex information systems.

Proper documentation and dissemination of resources, therefore, have far-reaching positive implications for modern linguistic research everywhere in the world. Suffice it to say that efficient documentation tools and pragmatic dissemination practices are germane to “the organizational process for acquiring, organizing, and communicating both tacit and explicit knowledge so that others may use the knowledge to be more effective and productive” (P. Gray, 2000; *Knowledge Management Overview*, CRITO Working Paper). Many frontline linguists and language researchers, technologically-speaking, are at the receiving end. Hence, the need for simple but sophisticated digital tools.

Right from the outset, at the very planning and developmental stage, the envisioned documentation tools should be multilingual in scope and nature. For a start, the tools can be made available in the following languages: English, French, Spanish, German, Chinese, Japanese and Swahili. Language researchers in developed and developing countries are ideally proficient in one of these languages and are therefore able to make use of the multilingual tools.

Methodologically, an ethnocomputing philosophy should be adopted. Ethnocomputing is an eclectic computing framework that incorporates a cultural perspective in the problem solving methods, conceptual categories, structures, and models used to represent data or other computing practices (cf. M. Tedre *et al* , 2006; *Ethnocomputing: ICT in Social and Cultural Context*. Communications of the ACM **49**). Such an approach will make the emerging tools to be of use to the highest number of researchers possible and complement the multilingual character of the tools.

The documentation tools can meet the needs of liberal-oriented linguists if the tools are: simple and easy to learn (probably with the aid of tutorials and practical examples); free and open source (whenever payment is required, it should be subsidized for researchers in developing countries); portable and platform-neutral; highly Unicode-compatible; flexible and use extensible meta-language; and, if they can embed a simple scripting language for data manipulation and transformation. Lastly, an active global collaborative research structure should be put in place.

I hereby wish to request a graduate student participant's airfare and housing subsidy.