

Multilingual Content Development for eLearning in Africa

e-Learning has a very important role to play in education and capacity-building in Africa, as the continent works towards meeting the millennium goal of “education for all”. However, this requires a substantial investment in building the requisite infrastructure as well as developing the human resource capacity required for development of relevant content as well as service delivery. By adopting a continental-wide strategy for the provision of e-learning programs, it is possible to take advantage of economies of scale where the large initial investment is subsequently offset by lower incremental costs as more students are accommodated into more diversified learning programs.

Given Africa’s diversity with respect to linguistic environments, successful continental-wide e-learning initiatives have to address the important issue of multilinguality. Since any successful learning experience must be designed and centred on the needs of the learner, the critical importance of providing educational content, digital or otherwise, in the learner’s language cannot be overstated. By developing tools and techniques which not only support the creation and delivery of high quality digital content for e-learning, but also exploit language technology resources to provide this content in linguistically diverse learning environments at no additional cost, WebALT’s core technology proffers a sound technological solution to meet this challenge. WebALT’s language technology can be adopted and extended to support content development and delivery in African languages, enabling the adoption of continental-wide e-learning programs at all levels of education and in different subject areas.

The WebALT Project

Web Advanced Learning Technologies¹, WebALT is a European Digital Content on the Global Network Consortium bringing together the universities of Helsinki, Cologne, Eindhoven, Barcelona, and the commercial partner Maths4More. The objective of the WebALT project is to develop the software infrastructure and language-independent content for an XML database of mathematical problems to be used in undergraduate university courses in mathematics. The project combines computer algebra tools, web learning techniques and language technology (LT) to produce live language-independent didactical material for advanced learning of mathematics. The technological solutions developed by the partners are versatile and robust, making it possible to apply them directly for middle and high school mathematics instructions, and with specific adaptation, can even be applied to different learning areas.

WebALT’s main components include:

- i. A Problem Tree XML Database where problems are annotated with subproblems making it possible to provide an adaptive learning environment where students receive automated private instruction.
- ii. Course content definitions (XML metadata) which facilitate shared creation of educational materials, with an additional feature which simplifies the task of locating relevant exercise material from the database.
- iii. Multilingual content delivery which is facilitated by the definition of formal language-independent vocabularies which enable the mark-up of mathematical

content and language generation techniques to automatically produce natural language variants in different languages e.g. English, Spanish, French etc.

WebALT's Multilingual Technology

To produce a verbalization of mathematics in a variety of linguistic styles and languages, WebALT relies on natural language (NL) generation from an OpenMath² encoding. OpenMath is a standard aimed at supporting a semantically rich interchange of mathematics among varied computational software tools such as computer algebra systems, theorem provers, and tools for visualizing or editing mathematical text (Caprotti and Carlisle, 1999). It provides an unambiguous language-independent representation of the exercise problem. Furthermore, by extending OpenMath with problem-related vocabulary and OpenMath attributes for NL grammar hints (Caprotti et al., 2000), flexibility in the sentence flavour and grammatical constructs of the resulting natural language variants is achieved.

WebALT's multilingual generator has been implemented using the Grammatical Framework³ (GF), a multilingual parsing and generation framework. The application grammars define a mathematical concept lexicon as well as language-specific generation rules which define how a given OpenMath symbol (or any combination of such symbols) is verbalized into natural language, based on specified NLG hints. Currently, the WebALT generator has a good multilingual capability covering English, Finnish, Swedish, Spanish, French and Italian with work underway for Dutch, Catalan and Portuguese. The following example shows the multilingual rendering of an OpenMath-encoded math problem:

Math *linalg1:outerproduct(linalg2:vector(1,2,3),linalg2:vector(4,5,6))*

English Calculate the outer product of the vector [1, 2, 3] and the vector [4, 5, 6].

French Calculer le produit extérieur du vecteur [1, 2, 3] et du vecteur [4, 5, 6].

Spanish Calcula el producto exterior del vector [1, 2, 3] y el vector [4, 5, 6].

Finnish Laske vektorin [1, 2, 3] ja vektorin [4, 5, 6] ulkotulo.

Italian Calcola il prodotto esterno del vettore [1, 2, 3] e del vettore [4, 5, 6].

Swedish Beräkna yttre produkten av vektorn [1, 2, 3] och vektorn [4, 5, 6].

Education in Africa – The Language Question

In any discourse on education in Africa, the language question merits profound consideration, as there are many different language education models in use, and an equally diverse range of opinions on the subject. However, the general (predominant) experience has been that in most countries, children receive up to three years of Mother Tongue Education (MTE) (or other national language e.g. Swahili in Tanzania), followed by a switch to education in the former colonial language. Research in psycholinguistics and second language acquisition shows that such language models are inherently flawed since children are required to learn through a second language before they have

developed sufficient proficiency or competence in this language to make this possible. High teaching and training costs as well as lack of educational materials in African languages have been cited as some of the main reasons why many countries use colonial languages for educational instruction. A recent study⁴ in several African countries shows that contrary to popular wisdom, students who have longer MTE with teaching of a well-resourced second language as a subject, achieve a better proficiency in both their own language as well as the foreign language, compared to those learners who make an early switch to an international language. In addition, such students are more likely to perform better in mathematics and science subjects. The findings of the study clearly show that the most beneficial and effective education system in multilingual African nations is additive multilingual education which leads to high proficiency in African and international languages. Language Technology (LT) can be exploited to facilitate the realisation of such a multilingual educational environment, where African languages can be used increasingly for education – a prerequisite for development and learning which can only take place effectively through the languages that are familiar to the population.

WebALT's Resources for eLearning in Africa

Given the prevailing language model where European languages are widely used in education, WebALT's mathematical content can be readily deployed as is, in African universities and high schools that offer mathematical teaching in English, French or Spanish (with Portuguese forthcoming), boosting continental-wide e-learning initiatives in Mathematics. However, by enhancing this technology with African language capability, e-learning content can be delivered in various African languages as required, not just for mathematics but other key subject areas as well. This would make it possible to adopt e-learning at all levels of education, and would have a significant impact at the primary level where foreign language competence is not as strong as that of the mother tongue.

Extending WebALT's African Language Capability

GF forms the core of WebALT's multilingual capability, providing the language technology know-how for rendering mathematical problems into various European languages. Therefore extending WebALT's technology with African Language capability requires building GF resource grammars for the latter.

Development of GF resource grammars is a highly specialized task, requiring competence in the relatively new field of Language Technology, or a combination of expertise in Linguistics and Computer science and a good linguistic/computational knowledge of the target language. The biggest impediment to African Language Technology (ALT) is a lack of LT training and expertise in African languages, compounded by the lack of computational resources for most of these languages. Such resources include machine-readable dictionaries, computation-oriented descriptive grammars and corpora. Concerted efforts and investment must be made to raise the level of LT capability of African Languages by investing in both the required human expertise as well as the requisite technology.

By adopting Kwesi Kwaa Prah's view of 15 core languages⁵ for Africa that cover at least 85% of the African population and concentrating LT efforts on these, the daunting task of addressing Africa's linguistic diversity can be turned into a practical, achievable target. The viability of Prah's idea can be demonstrated by customizing the GF resource grammars for Swahili (currently under development) to be applicable for other languages in the shared Bantu language family, reducing the time and expertise that would otherwise be required to cover the entire family.

WebALT's core technology, enhanced with African Language capability provides a substantial breakthrough in the provision of culturally and linguistically sensitive content. This technology makes it possible to deploy continental-wide e-learning programs customized according to the linguistic requirements of the student, supporting effective learning and consequently, the overall development needs of Africa.

¹ www.webalt.net

² Caprotti, O., Buswell, S., Carlisle, D., Dewar, M., Gaëtano, M. and Kohlase, M. (ed.) 2004. *The OpenMath Standard V 2.0*. Report of the OpenMath Society. <www.openmath.org>

³ Ranta, A. 2004. Grammatical Framework, a Type-theoretical Grammar Formalism. *Journal of Functional Programming*, 14(2):145-189.

⁴ Optimizing Learning and Education in Africa – the Language Factor : A Stock-taking Research on Mother Tongue and Bilingual Education in Sub-Saharan Africa. <http://www.adeanet.org/meetings/en_Aug-locallanguages-2005.htm>

⁵ K.K. Prah (2002). *Language, Neo-colonialism and the African development challenge*. TRIcontinental, No. 150, Havana, Cuba. <http://www.casas.co.za/papers_language.htm >