

Abstract Submission: SOAS Conference December 2007  
New techniques and opportunities for documenting and describing languages

Aymara on the Internet: A Step toward Interoperability and User Access  
Sue Legg, Howard Beck and Elizabeth Lowe with M.J. Hardman  
The University of Florida, Gainesville Florida

Linguists are by nature and by disciplinary necessity often very independent scholars whose work is intended to capture the distinctiveness of languages and cultures, as Gary Simons noted in his plenary address to the TLTR (Toward the Interoperability of Language Resources) workshop at Stanford University in July, 2007, the theme of which was the need for interoperability of tools and common data identification standards. Simons noted that the sheer volume of work to be done to collect, document, and analyze dying indigenous languages is daunting. Moreover, language repositories are moving to a digital form, and increasingly they will be built in a collaborative environment in which professional linguists and the public will have some form of direct access to contribute, analyze and use the data. In many respects, the experience of the Aymara on the Internet project has shared this goal while struggling with the day-to-day perspectives of a multidisciplinary team on how to reach it. This paper addresses the vision of the team and how that vision evolved.

The Aymara on the Internet Project builds upon Professor M.J. Hardman's field research in Peru on the endangered Jaqi languages (Aymara, Jaqaru and Kawki) in the 1950s as well as the classroom teaching materials she developed through a series of grants dating from the 1970s. A new grant from the U.S. Department of Education (2004-2007) funded the current phase of the project. The demand for online materials in these languages is growing exponentially as Indigenous movements gain strength in the Andean region. The original vision was to convert these materials into a high quality and flexible web-based delivery format for language learning. The vision expanded, however, when the possibilities of an ontology-based database backend for the instructional materials was introduced to the project planning team. The realization that the students as well as other linguists would have easy access to complex linguistic structures was encouraging, and the project took on a new dimension.

Thus, the project evolved from a PowerPoint online presentation much like other language learning software applications to a methodology that offers not only language learning, but also an XML metadata driven archive of the materials, a web services architecture for sharing a generic database tool for linguistic data collection and analysis, and a new way of thinking about how data and tools to manipulate data will become shareable in a diverse community of linguists, educators and the general public. While we have developed an ontology management system (OMS) comprised of many software systems (authoring tools, visualization tools, physical storage managers, eLearning), it is not necessary to think of the ontology as physically residing within a specific database. Rather the ontology is a knowledge network distributed worldwide, using XML as an exchange format, with different parts of the ontology managed by different people and organizations.


PROJECT PHRASE PATTERNS CATEGORIES PHRASES WORDS MORPHEMES **EDIT**

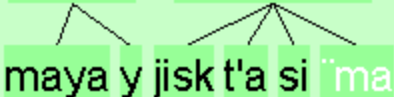
← I Dialog (Phrases)

Label  
Introduction  
Instructions  
Notes  
Patterns  
J. Winus tiyas , mama.  
W. Tata mayay jiskt'asir  
J. Kunach mam?  
W. Kurijiruran kunas sut  
J. Piruti Apas satawa. J  
W. Jisa juparuwa , alujt  
J. A .. alujamintux k"ä is  
W. Jayat tata?  
J. Janiw jayäkiti... k"ä is  
W. A... jisa.  
J. Ukawa. Nayax ukäk s  
W. Ya tata.  
J. Janit jutirikta?  
W. Janiwa , jich"akipuni  
J. Kawkinkiritas mama?  
W. Istarusunirusankiritw  
J. Nayax akankiritwa ...  
W. Walkiw , uk"amax q'

Enter Phrase: W. Tata mayay jiskt'asima [jiskt'a'si`ma] .

English Gloss: W. Sir, may I ask you a question?  
English Morph Annotation:  
Spanish Gloss: W. Señor, disculpe una pregunta.  
Spanish Morph Annotation:

Record Listen Load Sound File  Notes Flags

W. tata mayay jiskt'asima .  
  
 maya y jisk t'a si `ma

`ma

1>2p future

Base Morpheme: `ma  root  suffix vv

nominal  verbal stem derivation  verbal person derivation  verbal inflectional  thematic  sentence  independent

English Gloss: 1>2p future

Spanish Gloss: 1>2p futuro

← ▶

Move Up

Move Down

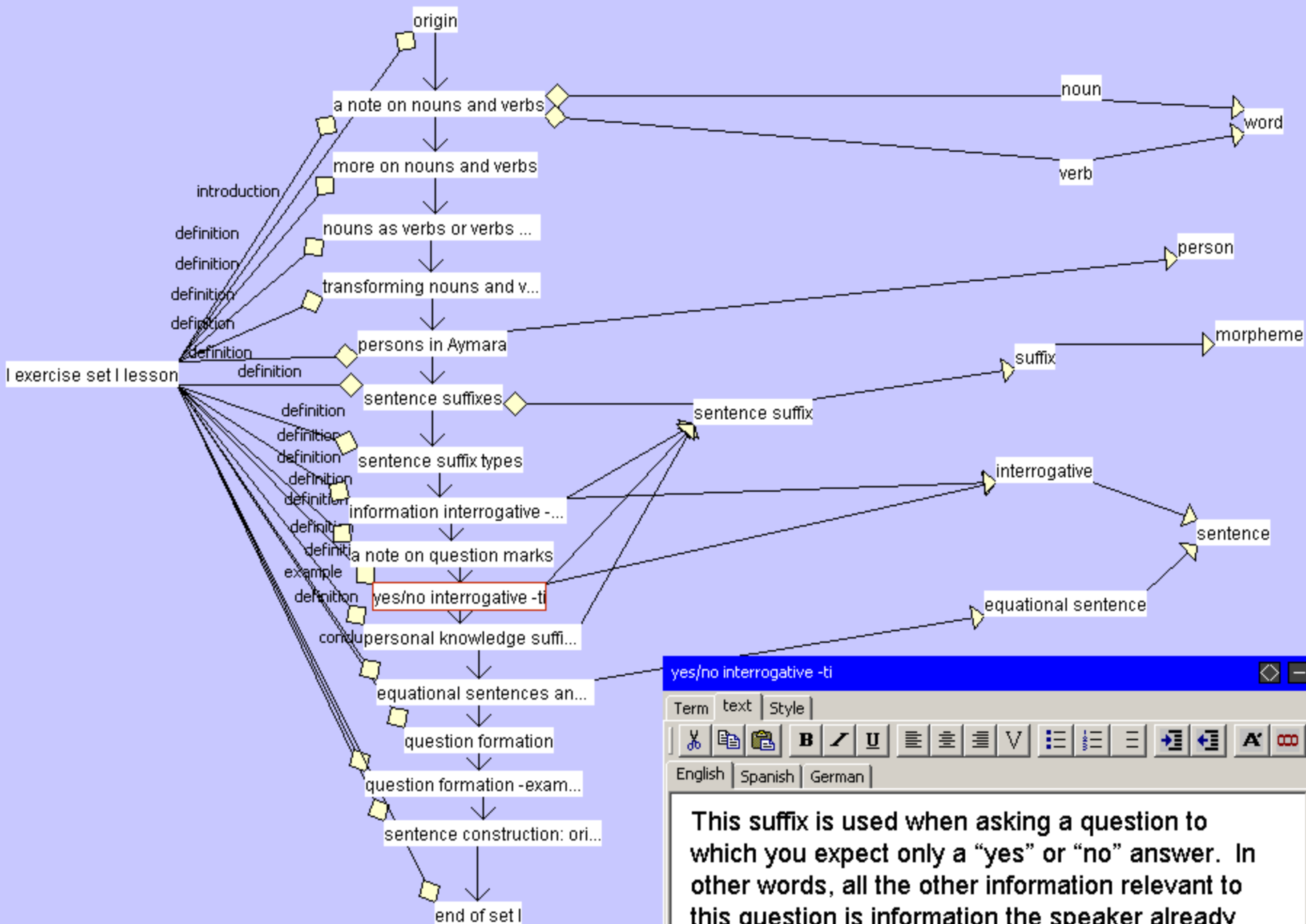
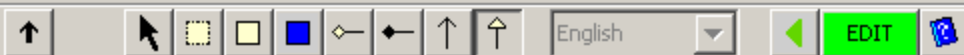
New

Delete

Save

Concordance Features

Feature	Value
source person	1
tense	future
target person	2



yes/no interrogative -ti

Term text Style



English Spanish German

This suffix is used when asking a question to which you expect only a “yes” or “no” answer. In other words, all the other information relevant to this question is information the speaker already knows.