

Can a robot help save an endangered language?

Maximiliano DURAN

ELLIAD , Université de Franche Comté
Besançon, France
duran_maximiliano@yahoo.com

Abstract

A robot using artificial intelligence, a comprehensive set of linguistic resources and pedagogical functionalities may help to preserve Quechua. It can help in M.T of school texts and general culture documentation into Quechua. Written documentation, is essential to keep this language alive. I have been working on such a robot, for several years. I named it Yachaj/expert. The first stage of this project has the following functions: Automatic conjugation, lexical queries of Quechua-FR-SP; elementary spelling checking; and transliteration (alpha version) of texts written in the official spelling of Cuzco, Ecuador or Bolivia to that of Ayacucho and vice-versa.

Keywords: robot, quechua, Yachaj, automatic conjugation, transliteration

Runasimipi

Ama runasimi wañunampaj, allin qispichisqa, allin yachachisqa, llapan rikchaq runasimi cheqap-yachaykunawan, kikin-ruraqqa yanapakuwanchikmanmi runasimi unachaypi. Chaymi ñuqa, kay ñawpaq qanchis watakunapi runasimita huk kikin-ruraqta “yachachichkani”. Paymi yanapakullanman runasimipi, tukuy niraj yachaykunata, yachay-wasikunapi yachachiyta. Paytaqmi, yanapawananchik Fransespi, Castellanopi qellqakunata runasimiman tikrayta, chayna achkallaña runasimipi qellqasqa taqekuna kanampaq. Cheqap-yachaymanta qellqakuna, willakuy-yachaymanta, llimpi-taki-yachaykuna qellqakuna achkallaña runasimipi qellqasqa taqekuna rikurinampaq. Cuzco qelljqa qellajasjata ayakuchu qelqaman tikraypipas yanapawasunchik, kutiriyintapas.

1. A lack of compulsory quechua schooling

The rapid erosion of the Quechua language in just two generations (Fig. 1), in Peruvian territory, confirms its endangered language status. This relative decrease in population objectively shows the danger of Quechua becoming a dead language in the next two generations. History shows that Quechua has been the victim of many injustices of all kinds that have brought it to this tragic state: historical, administrative, social, linguistic and even psychological.

In Peru, there has been no State policy of compulsory schooling in Quechua in these regions. One consequence of this is the very large and traumatic negative social impact on this population. Since it is not obligatory to speak, and eventually to write correctly in QU, parents are inhibited from transmitting and deepening the knowledge of this language in their children, which inexorably implies the loss of their character of mother-paternal tongue. The language becomes marginalized. The lexical, morpho-syntactic level of the discourse is gradually impoverished. Socially, the child, and subsequently the adult Quechua-speaker, receives from his peers a discriminatory treatment and negative assessment from the monolingual Spanish-speaking speakers, instead of his bilingual knowledge (Quechua-Spanish) being positively weighted. Another consequence is that the incentive for the natural development of the language in the different fields of culture in the society is lost. Cultural production in QU becomes almost a clandestine activity and that is why very few people create literature or songs in this language nowadays; and to aggravate the tragedy, the few who still do, use a lexicon with many “loans” of the language of contact favoring the “quechuallano” or the “medio-lengua” (half quechua-half Spanish).

1.1 Quechua speaking population in Peru

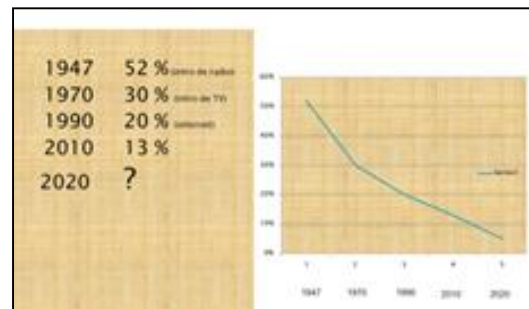


Figure 1: Fall of Quechua speaking population.

1.2 Other converging factors

The other factors that makes worse this erosion are:

- Incoherent spelling
- Scarcity of general cultural texts written in QU
- Social discrimination and racism against quechua speakers

2. Diversification of quechua spelling

In Peru, several spellings have been established by governmental decrees for the QU: one for the Cuzco region, others for Ayacucho and Ancash, among others. In Ecuador, Bolivia and Qrgentina (Santiago del Estero) have also decreed other orthographies. Thus, for the same QU word like language /simi there are the following spellings:

284shimi, simi, šimi, čimi. It is as if the various transcripts of

3.2 Morphology

Yachaj “knows” how to apply thousands of rules of QU morphology, and he knows how to apply them. It knows exhaustively all the nominal suffixes `SUF_N`, adjective suffixes `Suf_A`, and verb suffixes `inter` and `post positional IPS` and `PPS`, and knows how to make the grammatically valid combinations between two or more suffixes. It knows all the semantic values of about 240 language suffixes. In the following figure we show the extracts of the inflections (655) of a noun like `wasi/ house` and the verbal forms (7743) that Yachaj obtains by applying conjugation rules, of derivation with 1-3 suffixes.

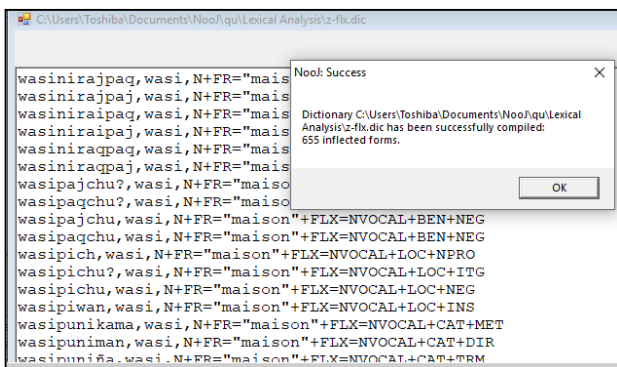


Figure 4: Inflectional forms of `wasi` (1-3 suf.) (655)

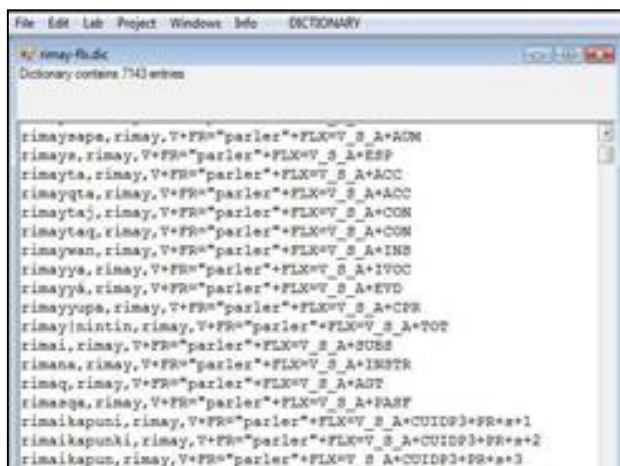


Figure 5: Inflected forms of verb `rimay` (1-3 suf.) (7743)

4. Spelling checker

The robot contains a user interface to get help in spelling correction of texts written in official QU of Ayacucho. Work is under way on spell-checkers for the other dialectal versions

5. Automatic inter-dialectal transliterator

Yachaj is able to carry out the inter-dialectal transliteration between the Cuzco>Ayacucho versions and vice versa, in the laboratory. Following the same technology, the programming of automatic transliteration has been initiated to help pass from an official spelling, Cuzco, Ancash, Ecuador, Santiago del Estero or Boliviano to Ayacuchano and vice versa. This system can be used to solve this writing controversy of "Into which written quechua? and allow us to overcome that divisive barrier that each idiosyncrasy claims to make prevail the quechua and its writing of 'its' region is correct and the rest is not. If a Cuzco way of writing is not readable to the Ayacucho-Quechua-speaking, by pressing the "transliterar" button to the Ayacuchano, he will be able to get his text in official Ayacucho spelling. And vice versa. Here is an excerpt from the automatic transliteration of a poem by Jayme Araoz Chacón (2008) written in official Cuzco spelling into Ayacucho spelling.

Cuzco spell	Ayacucho sp.	Castilian
Rumi	Rumi	<i>Piedra</i>
Waklaw chimpaman	Waklaw chimpaman	<i>La piedra que al frente lancé</i>
Rumi chanqasqay	Rumi chamqasqay	
Waklaw chimpaman	Waklaw chimpaman	<i>La piedra que al frente arrojé</i>
Rumi wikh'usqay	Rumi wischusqay	
Maytaq kunan	Maytaq kunan	
Rikhurimunchu	Rikurimunchu	<i>Porqué ahora no aparece?</i>
Maytaq kunan	Maytaq kunan	
Kutimunñachu	Kutimunñachu	<i>Porqué hoy ya no regresa?</i>
Chhaynallataqsi	Chaynallataqsi	
Warma munasqay	Warma munasqay	
Chhaynallataqsi	Chaynallataqsi	<i>De modo parecido</i>
Yana wayllusqay	Yana wayllusqay	<i>La mujer que quiero,</i>
Maytaq kunan	Maytaq kunan	<i>Del mismo modo mi</i>
Kutimunñachu	Kutimunñachu	<i>amada ves cómo ya</i>
Maytaq kunan	Maytaq kunan	<i>no vuelve, ves cómo</i>
rikhurimunchu	rikurimunchu	<i>hoy ya no se aparece</i>

6. Conclusion

We have presented the first stage of a quechua expert robot. To build it, we are using an important number of linguistic resources obtained mainly with the aid of the linguistic platform NooJ (Silberztein 2003, 2015). We are using also some machine learning techniques. For the time being it is capable to perform a certain number of pedagogical functionalities which may be useful for elementary school learning in quechua. The first stage of this project have the following functions: Automatic conjugation, lexical queries of Quechua-FR-SP; elementary spelling checking; and an alpha version of a transliterator of texts written in official spelling of Cuzco, into Ayacucho spelling and vice-versa.

7. Bibliographical References

- Araoz, J. (2008). Llaqta Taki Harawi. *UNICEF. Danny's EIRL. Cuzco Peru*,
- Dubois, J. et Dubois-Charlier F. (1997). *Les verbes français*, Larousse, Paris.
- Duran, M. (2009), *Dictionnaire Quechua-Français-Quechua*, Editions HC. Paris.
- Duran, M (2013), *Formalizing Quechua Noun Inflexion*. Formalizing Natural Languages with NooJ. Edited by A. Donabédian, V. Khurshudian and M. Silberztein. Cambridge scholars. Newcastle upon Tyne,.
- Duran, M.(2013) *Formalizing Quechua verbs Inflexion*. Proceedings of the NooJ 2013 International Conference, Saarbrücken,.
- Silberztein M (2015). *La formalisation des langues*. ISTE Editions. London.
- Silberztein M(2003). *NooJ Manual*. <http://www.nooj4nlp.net> (220 pages, updated regularly).