Standardization in Andean languages

Rodolfo Cerrón-Palomino
Universidad Nacional Mayor de San Marcos

This paper discusses current efforts to standardize the two major Andean languages: Quechua and Aymara. The author reflects on the difficulties in gaining consensus on standardizing the phonology, grammar and lexicon. He offers specific recommendations on how to deal with neologisms and proposes the creation of a pan-Andean entity to resolve the issue of standardization.

Preface

Today, I will discuss the state of the art in relation to language standardization as far as the so-called "major languages" of the Andes, Quechua and Aymara, are concerned. I will focus particularly on work done at the level of graphization, grammatication and lexication, in the senses defined by Ferguson (1968) and Haugen (1983). As will be seen, most of the decisions taken in such matters result from a purely descriptive treatment, largely "leaving the language alone," and therefore without a real concern for aspects such as standardization, codification and language development. As such, Andean languages are still treated as transitional mediums towards Hispanization and not as ends in themselves.

Such a purely descriptive and synchronic treatment of language standardization is due, in my opinion, to two well-known and extremely deeply-rooted biases: the emphasis placed on dialectal differences among local varieties, and the neglect of a long tradition of grammatical and lexicographic studies of the languages concerned which go back to the middle of the sixteenth century.

As for the first bias, there is the strong proclivity to register, sometimes in a detailed fashion, the differentiating features which set the local varieties apart (the dialectologist acting as a language codifier would like to see the results of his fieldwork reflected in his "standardization"). This atomizing vision of language doesn't allow for the discovery, beyond observable but mostly superficial differences, of
common and underlying forms and features which, no doubt, comprise the communicative competence of the speakers. According to the second bias, the descriptivist acting as codifier neglects the philological and grammatical traditions of the major Andean languages, overlooking them as if the languages had been recently discovered. The antinormativist dogma ("leaving the language alone") prevents the descriptivist from realizing that, in matters of codification, most of the problems which are presently discussed were actually pointed out and discussed in the past, and some of the solutions proposed then could perfectly well be reconsidered with great advantage today (Mannheim, 1984; López, 1988). I will make some observations in relation to the experience accumulated so far in matters of corpus planning.

**Phonological and grammatical aspects**

What I see at present—in spite of the disruptive activities of foreign entities (for example, the Summer Institute of Linguistics (SIL) and the Comisión de Alfabetización y Literatura en Aymara [Commission on Aymara Literacy and Literature]) as well as of local institutions (academies of Quechua)—is a general consensus in favor of the use of a unified alphabet not only within each country, but also at the international level, thanks to the joint efforts of institutions committed to that goal. It is to be expected that in spite of the fact that the respective alphabets were developed following more or less common criteria—phonological, sociological, pedagogical and practical—and without dismissing the fact that the sociolinguistic and dialectal realities are different within each country (as far as Quechua is concerned, the situation is more complex in Peru than in Ecuador, and in Ecuador more than in Bolivia), the alphabets are nevertheless not free from traces typical of the descriptivistic dogma.

Thus, for example, there still persist transcriptivist tendencies which seek to imitate pronunciation exactly, overlooking the fact that codification presupposes the development of alternative registers for the language, aside from or parallel to the oral system. Implicit in this practice is another more subtle fact—that, consciously or unconsciously, what is being proposed is a notational system for the foreigner or for the Spanish-speaker, both of whom are ignorant of the languages being codified, but who would benefit from a transcriptional system as an aid to pronunciation. However, we know that no alphabet—at least within the so-called literary languages—teaches us how to pronounce. What we must ask at this point is: For whom is the writing system being proposed?
Another descriptivist trace which still persists among the proposed alphabets (especially in the Bolivian case and, until quite recently, also in the Peruvian) is that they are introduced as if they were phonological inventories (with points and manners of articulation), and not real alphabets. Aspects such as spelling or even the names of the characters are totally neglected. All of this seems to indicate that, unconsciously, there is not the slightest intention of putting a writing system into practice for the languages concerned; again, a transitional solution is being adhered to. This also explains why the need to prepare a manual of orthography was never perceived, as if the process of graphization concludes with the postulation of a phonemic-graphemic inventory only. This is particularly astonishing, especially when we know that there have been several previous attempts at formulating rules of orthography and punctuation, such as those proposed by the Bolivian Quechulist Berrios (1904: iv). This indicates to what extent it is detrimental to neglect the grammatical tradition mentioned above.

**Standardization in Quechua**

Now, I would like to outline some of the aspects which have not been thoroughly considered in the formulation of the alphabets of Quechua, and which are relevant to attaining an authentic codification.

a. In the treatment of the syllable final stops in Cuzco-Puno as well as in Bolivian Quechua, here we still see the concretist-transcriptionist practice, which doesn't incorporate theoretical advances made by variationist theory and its notion of pandialectal grammars (Bailey, 1975), nor does it take into account the postulates of the ethnography of communication and its notion of communicative competence. In the case of the Quechua language, there is a considerable amount of dialectal evidence which allows postulating abstract segments, overriding their more concrete manifestations, without necessarily postulating segments attributed to the protolanguage. The advantages of the proposed solution can be seen, for example, in the fact that it is no longer necessary to add new graphs such as sh, f for j. Thus, for example, words such as 'five', 'wing' or 'hut' are written as pichqa, rapra and ch'uklla instead of pishqa, rafra and ch'ujila, respectively. Note that by taking into account such a proposal, it is perfectly possible to unify the writing system of Southern Peruvian Quechua, as is being done presently, and as was also been implemented (although gradually) in Bolivia. This kind of solution affects not only the syllable final consonants as such, but also the standardization of certain suffixes which suffered from sporadic changes, such as the genitive, the inclusive, the obviative subordinator, etc.
b. Another aspect which should be contemplated is the treatment of cases of polymorphism. It is my conviction that, in such cases, one form should be postulated, preferably the most conservative one. Take, for instance, the durative -chka or the dynamic -yku. As is well-known, both suffixes present a large amount of variation: -ska- -sqa- -sya- -sha- -sa-, for the first case, and -yku- -yu- -y-, for the second. In cases such as these, it does not seem to me recommendable to postulate forms which, although perfectly valid in oral speech, clearly deviate from more conservative and easily "recoverable" forms even in the most innovative variety, depending on the style of the speech used. Otherwise, it will be simply impossible to standardize the writing system; writers will keep writing as they wish.

c. Yet another aspect which should be reconsidered is the eclecticism assumed by some Bolivian Quechuists in relation to the issue raised in the writing of three or five vowels (cf., for example, Albó 1987). According to the official Bolivian alphabet, one can either write with three or five vowels. This decision, far from contributing to the standardization of the language, actually evades the problem and creates chaos in the writing system. Thus, the very same scribe can write sunqu, sunqa, sonqa, sonqu ('heart') or wiqi, wiqe, weqi ('tear'), and so on. I contend that, in such cases, one has to be more decisive even if it means taking a vote among the members of the committee in charge of standardization. Incidentally, this type of decision has a long tradition in the history of Spanish, going back at least to the beginnings of the XVI century, as Nebrija ([1517] 1977) pointed out.

d. It is also time to reconsider certain letters which were eliminated with negative consequences for the pan-Quecha unification of the writing system. Again, this has been done by neglecting the philological-grammatical tradition of the Spanish language as well as that of Quechua. Thus, we have, for example, the elimination of h or the argument against the use of k and w. To argue that the letter h is silent in Spanish and thus cannot be employed in Quechua (or in Aymara) is a clear misunderstanding of the problem. One has to ask again: for whom is the writing system being developed? Due to preconceived opinions such as these, the alphabets of Ecuador and Bolivia incorporate j instead of h, and, in the first case, hu instead of w. The use of the letters k and w was questioned in Ecuador based on their supposedly recent English and/or German origin. I would like to point out that the letter k was used in the XVII century by the Spanish grammarian Gonzalo Korreelas. As for the w, it has been used in Quechua and Aymara at least since 1821. These preconceived opinions hinder any attempt towards a real unification of the Quechua writing system along the whole Andean area.
e. We should also take into account the urgent task of preparing manuals of orthography. For this purpose, it is obviously important to make a careful study of the written materials available. Most of them, however, are not detailed enough to serve as good illustrations of the variety of uses of punctuation marks, for example. It should be recalled that the literary masterpieces of the past do not serve as viable indicators for punctuation since the underlying principles differed. As I have said, however, there exist pioneering efforts within this area, such as that of Berrios.

f. The extreme fluctuation of the laryngeals (aspirated and glottalized consonants) within so-called "Inca-Quechua," where the same word registers these consonants or not depending on the dialectal area, should make us consider whether it is worthwhile to represent them in the writing system for these varieties which in fact share a considerable common and uniform vocabulary. An alternative favoring a less differentiated solution to the problem of the above mentioned consonants was pursued in the XVI century. The best proof that such a solution works are the thousands of pages of literary works in Quechua (and Aymara) produced within a 65-year span (1584-1649). This practice—directed towards the re-unification of Quechua—was hindered by the creole and mestizo Quechuists, mainly from Cuzco, who endorsed the mistaken idea that orthography must reflect the actual pronunciation of the language. The consequences of that campaign were detrimental to the so-called General Quechua developed by the former Quechuists. I think that a solution in the spirit of the colonial grammarians should be reconsidered, if the intention is to orthographically unify the language. But this presupposes, obviously, overcoming prejudices like the preference for a "minimal pair" orthography (according to which, if there is a pair of lexemes which differ by one distinctive element, and even if one of these elements occurs only in a few words, a new letter is proposed to accommodate that difference). This minimal pair orthography disregards the fact that one doesn't write lists of words only, and that the context prevents any ambiguous interpretation. It is appropriate to recognize here the Ecuadorian solution to the problem of the aspirated consonants: the writing system simply ignores them.

**Standardization in Aymara**

With respect to this language, I should mention that there has been no systematic attempt at codifying it so far. Instead, the usages of traditional grammarians and scribes prevail, as well as an adherence to usages introduced by the followers of the descriptivist tendency, especially at the phonological and morphological levels. As for the latter, it is a curious fact that a taxonomic grammar such as that of Hardman,
Vásquez, and Yapita (1988) was adopted as if it were a standardized grammar. This is clearly a gross error because the linguistic facts presented in it correspond mostly to the speech of a few idiolects and, even worse, there is no attempt at standardizing the language. It is not surprising that the authors didn't seem worried at all about the many alternants, even though the basic forms could have been easily derived. On the contrary, a sort of exotic vision is predominant throughout the whole book, where levelled forms are presented with indexes going from 1 to 5, ignoring the fact that there are dialects or registers within the same speech where such apparently levelled forms are clearly distinguished. For example, the first and second person verbal suffixes in La Paz Aymara are neutralized in a unique form, -ta, whereas in the Huancané (Perú) and Northern La Paz varieties they are distinguished as -tha and -ta, respectively. Standardization in Aymara is only in its beginning stages. I will point out some of the problems which should be taken into account regarding standardization of this language.

a. What is urgently needed is a codification of the language, a normative grammar, because the existing manuals, be they traditional or modern, are intended only as guides for learning Aymara as a second language. As such, they were not written in order to standardize the language at all. In such cases, as it is easy to realize, there is no concern for the evaluation of competing forms, for example.

b. At the graphization level, in spite of assertions made by the proponents of the official alphabet (approved in 1983) claiming that it is phonemic, it clearly turns out to be phonetic once it is put into practice. Of course it is "phonemic" as far as the segmental inventory is concerned; but when it is employed in actual writing, one cannot avoid representing words and forms variably, with different shapes, as a result of the operation of morphosyntactically-conditioned phonological processes. Since most of the syntactically-conditioned vowel truncation in Aymara is predictable, there is no reason to "transcribe" the alternations; they are easily recoverable. In trying to reflect the pronunciation in the writing system, what one sees is a subliminal concern for the foreigner or for the non-Aymara-speaking population. Thus, the written system is conceived as an aid to the non-Aymara speaker. Obviously, in the face of a sentence such as uka ch'iyara anuxa qutaru hali 'that black dog runs towards the lake,' only those who don't know the language could possibly read it out letter for letter, instead of uka ch'iyar anux qutar hali, which is the actual pronunciation. Once more, the question arises: for whom is the alphabet intended?

c. As I have said, the narrow phonemic description (or better, transcription) of the language, as a result of the lack of a real standardization, leads to
the proliferation of pseudo-homophones (such as, for example, the five -ta suffixes); and to an unnecessary number of long vowels. These can be easily avoided provided we focus on more conservative dialects where no lengthening compensation arises. Such is the case of the first person future marker. Whereas in La Paz it is realized as vowel lengthening, in the northern dialects its realization is -nha, and so on (thus, sara 'I will go' instead of the conservative form sara-nha).

d. As for the practical and sociological criteria cited by the proponents of the official alphabets of Peru and Bolivia, it would not be unfair to say that the decision favoring the use of dieresis for marking vowel lengthening and the use of x to represent the postvelar consonant were completely unfortunate. Not only is it easy to omit the dieresis (as already happens in Spanish, where even in linguists omit it in a word such as lingüística) but the use of x makes the language overly exotic, especially in view of the fact that a similar consonant exists in Spanish, which is written as j. But once the h was supplanted by j to represent the glottal aspirated consonant, it became necessary to select another letter for the postvelar. As a result, similar words in Quechua and Aymara are written differently (thus, for example, hucha 'fault' in Quechua and jucha in Aymara).

Lexication

In relation to the lexical codification of the Andean languages, after the monumental work of González Holguín ([1608] 1952) for Quechua, and Bertonio ([1612] 1984) for Aymara, we have not seen a lexical codification as exhaustive or with such careful semantic control. Since then, several vocabularies and lexicons written for the benefit of the Spanish-speaking reader have appeared, all of which try to accommodate the Andean semantic base within the lexemic molds of the Spanish language. Thus, standardization of this aspect is a task yet to be begun. It is shocking to realize that the need to compile dictionaries in Quechua and Aymara with a thorough semantic control of the entries in the same language has only been felt during the last two decades. Here also, unwillingly or not, the implicit posture of the lexicographers was in favor of assimilationism. That is why I welcome the efforts of Ecuadorian colleagues for their Shimiyuc Panca (Ministerio de Educación y Cultura, 1982), which, in spite of its provisional character, constitutes the very first attempt at providing us with a monolingual Quichua lexicon which is not simply ancillary to Spanish. It was in this spirit that we also conceived our Vocabulario razonado, now in press (cf. Ballón Aguirre, Cerrón-Palomino and Chambi).
It should be clear that most, if not all, of the modern dictionaries cannot be taken as models of lexical standardization (not even those which have come out lately for Aymara). They also suffer from the old bias of the structural-descriptivist trend, and, of course, from the phonological analysis imposed on the language. Some modern dictionaries are lexicons prepared exclusively with a linguistic-dialectological criterion, more concerned with the registration of forms than with the semantic definitions of the entries. The error lies, however, in the fact that those dictionaries are taken as a frame of reference for correct spelling. It does not take much time to realize that the notion of norm is completely absent. Not only do the same authors lack coherence in their own orthographic notation, but, paradoxically, deviating forms are taken as archetypes while the genuine ones appear as mere variants. Thus Cusihuamán (1976), for example, gives mihuy 'to eat', puhu 'spring', wahay 'to call', etc. as the "basic" forms for mikhuy, pukyu and waqway, respectively. In addition, there is no concern for unifying the vocabulary, since the prevailing tendency is to stress the dialectal situation, showing the differences rather than pointing out the similarities. Again, it is the interest of the linguist rather than that of the codifier which prevails. In this respect, it is comforting to note again the efforts made by the Ecuadorian colleagues, who are trying hard to lexically unify the dialects (turning competing lexical items into synonyms). I too have tried to formulate a Common Southern Quechua Basic Dictionary (Cerrón-Palomino, 1990), which unfortunately has had no possibility of being published so far. Also in this spirit, and as a first approach, I have prepared, with the aid of two other colleagues, a vocabulary related to agricultural activities, trying to define the entries on the basis of definitions provided in Quechua by informants (Ballón Aguirre, Cerrón-Palomino & Chambi, in press).

With respect to efforts thus far on lexical elaboration, there is a general consensus favoring a self-reliant or nativistic solution in coining new terms for new concepts. There is, in that sense, a rather nationalistic attitude, which challenges any kind of indiscriminate borrowings. Such an option appears explicitly in, for example, Montaluisa (1980) and Zúñiga (1987). The latter summarizes the conclusions arrived at during the First Workshop on Quechua and Aymara Writing held in 1983.

The nativistic solution has been severely criticized by those who adhere to assimilationism, the SIL members among them. Weber (1987), an SIL linguist, strongly criticizes the conclusions arrived at in the above mentioned workshop, as being too puristic. Weber's is, actually, a defense of free borrowing. As can be seen through proposals made so far in lexical elaboration, whether in Ecuador or in Peru, the nativistic solution clearly differs from that of traditional purism. In no way is it intended
to purge the native lexicon through a witch-hunt, which is clearly absurd. On the other hand, one should not forget the internal colonial situation which characterizes the Andean societies and their languages. It is not the same to be a purist in a context of oppression where there is large scale or massive borrowing in a vertical fashion, from the dominant language to the dominated one, as among any of the European languages, where, grossly speaking, borrowings go horizontally and symmetrically from one language to another. As Ninyoles (1975) would say, in such a situation, not being a purist is the same as being disloyal to one's own language. Therefore, an option in favor of purism doesn't exclude the borrowing of words provided, on the one hand, that the loans already form an integral part of the lexicon (= spontaneous borrowing); and, on the other hand, that when faced with new concepts, the loans are selectively chosen, after the nativistic resources have been exhausted (= programmed borrowings).

I would like to discuss two aspects related to the process of coining words—of induced neologisms. One of them has to do with the limitations of the native option, and the other with the phonological and/or orthographical representation of foreign neologisms.

There are several alternatives which imply a self-reliant option: (a) derivation and composition, (b) semantic expansion, (c) rescue of words (either obsolete or from different dialects), (d) semantic calque, and (e) descriptive periphrasis. It is the latter which appears to be the least advisable, as proven by practice. In fact, coining by description results, formally, in either heavy phrases or even sentences; and semantically, in extremely descriptive and concrete expressions which are far from practical. This is especially incompatible with two of the most elementary properties that characterizes the lexicon of a standardized language: clarity and conciseness. In view of this, and once all the chances for an elegant and precise nativistic solution are exhausted, one might sooner resort to borrowing, especially if alternative loanwords spontaneously admitted already exist. Note how Bertonio ([1612] 1984: Prólogo), at the beginning of the XVII century, had already suggested a similar solution, when faced with the problem of the early Spanish borrowings in Aymara. In fact, he says: "Thus [the Indians] will understand better if we tell them cādelero 'candlestick,' or candrillo apanima 'bring the candlestick': rather than cādela saataaña apanima 'bring the instrument where the candle stands up'; for although the second is proper to the language, the other is best received and used ..." Note the extreme vagueness of the periphrastic expression 'instrument where the candle stands up,' as compared to 'candlestick.'
As for the second aspect—that related to the formal nativization of the neologisms of foreign origin—there doesn't seem to any general consensus yet. The positions adopted range from those who propose spelling (and pronunciation!) in terms of Spanish orthography to those who call for a full Quechuaization or aymarization, following the assimilatory mechanisms of the native languages. These extreme positions reflect, no doubt, opposing ideologies, and, each in turn guarantees a uniform solution—at least in theory—to the written representation problem of the loanwords, although it is easy to see which one would result in a higher cost in implementation. In the Peruvian case, they opted for full nativization of the borrowings. However, given the complex sociolinguistic situation of the country—with dialectal areas penetrated in different degrees by the Spanish language—I think that one should reconsider the categorical nature of the alternative chosen. Experience demonstrates that one single pattern of nativization for the whole language isn't workable. Although in the so-called "Indian blot," a full adaptation of the loanword to the canon of the recipient language seems justified (which is true for Aymara also), the same cannot be said for the remaining Quechua dialect areas. For these, a partial accommodation of the foreign words would be advisable, taking into account the full incorporation of foreign segments into the native phonological component of the dialects. Thus, for example, not even in the most remote areas of the Peruvian Central Highland, would it be possible to find speakers who would say *timplu* 'diploma=diploma', *kawittu* 'cabildo=council', *suyru* 'suegro=father-in-law', etc. instead of *dimplua*, *kawildu*, and *suydru*, respectively. It is clear then, that the solution to the problem lies in a non-unique pattern of writing loanwords, although it will impede an eventual unification of the orthography. Be that as it may, there does seem to be consensus that loanwords do not have to be written as in the donor language, as the assimilationists (including the academicists) would prefer. Here, a minimum of autonomy is required: each language should be written following its own orthographic system.

To conclude, it would be advisable, in view of the observations made, to assume a common task directed towards the standardization of the Andean languages in a more systematic and functional manner, avoiding *ad hoc* solutions arising during the preparation of pedagogical texts, as well as in the production of other types of materials, including translations. What is required is a permanent pan-Andean consultant entity in charge of standardization and working within a more centralized framework. This is fully justified because we are working within the same linguistic family (or within the same stock, if we include the Aru languages), and because there exists a general consensus in favor of a nationalistic option with regard to the process
of coining words; hence it should be obvious that the same solutions could be considered. In fact, an international workshop held at Santa Cruz de la Sierra, Bolivia, in October, 1989 proved that such a joint treatment of the problems could be extremely profitable (see Ministerio de Educación y Cultura [1990] for the final report).

1 Text of a lecture given at the Language in Education Division of the University of Pennsylvania (November 5, 1991). I thank Wolfgang Wölck, Utta von Gleich and Nancy Hornberger for their helpful comments and stylistic betterment of an earlier version of this paper.
References


Berríos, J.D. (1904). Elementos de gramática de la lengua keshua. La Paz: González y Medina Libreros Editores.


