Matthias Urban  
Center for Advanced Studies “Words, Bones, Genes, Tools”  
University of Tübingen, Germany

**Abstract**

The Central Andes of Peru and Bolivia belong to those areas of the world whose linguistic landscape has been shaped decisively by language contact. Scholars have long recognized the intricate similarities between the Quechuan and Aymaran languages, and the consensus view is to attribute these to convergence arising from long periods of stable bilingualism in prehistory. As more recent research shows, also the languages of Northern Peru were shaped by language contact before their extinction. As a result, contact effects need to be taken into account as a general factor in studies of the linguistic history and prehistory of the Central Andes since without them crucial aspects of the lexical and grammatical structure of Central Andean languages cannot be properly understood. Thus, language contact constitutes a privileged and overarching point of access to understanding the linguistic prehistory of the region as a whole.

*Keywords*: Central Andean languages, contact linguistics, anthropological linguistics, ethnicity.

1. **Introduction**

As understood in the present article, the Central Andes comprise the coast and highlands of all of Peru and the Andean part of Bolivia, in which the basin of Lake Titicaca is located (cf. Stanish 2001: 41). It is a region characterized and defined by a long-standing and continuous cultural trajectory towards societal complexity that set in as early as the third millennium BC. The ultimate representative was the famous Inca Empire, which exerted control over the entire Central Andes as defined here and beyond. Periods of relative cultural homogeneity and at least in some cases also political unification, such as the Inca hegemony, are called Horizons by Andean archaeologists. For other points of time, however, one can distinguish roughly three different subareas of the Central Andes with different cultural expressions.

---

1 This work was supported by funding from the German Research Foundation (DFG), grant number UR 310/1-1.
One of these was centered in what is today Northern Peru, where a succession of complex, maritively adapted societies were based in the fertile river valleys intersecting the narrow strip of arid desert on the Peruvian coast. Another subarea of the Central Andes was located in the highlands of Central and Southern Peru, where subsistence was based heavily on pastoralism in addition to agriculture. From this area the Inca would ultimately emerge. A third, southern hotspot of cultural development in the Central Andes can be discerned on the altiplano around Lake Titicaca.

Linguistically, at the eve of Spanish conquest the Central Andes presented a mosaic of languages coexisting in geographical and social space, with some salient differences in the structure of the linguistic landscape between regions which correspond roughly to the different archaeologically recognizable subareas distinguished above (cf. fig. 1).

While the distribution of languages in Southern Peru and Bolivia was mosaic-like (in reality, in fact, likely more mosaic-like than the map in figure 1 suggests) due to the social organization of Andean societies, and while bi- or multilingualism was extensive (cf. Mannheim 1991: 33), in terms of language families the linguistic diversity was and is relatively modest. As far as southern Peru is concerned, we know of languages belonging to the shallow Quechuan and Aymaran families as well as of an unaffiliated language, Puquina. All three were also spoken in Bolivia, in addition to the small Uru-Chipaya language family. When one includes the eastern Andean slopes, in Northern Peru, in contrast, we know of as many of fourteen languages distinguishable at least by name: Tallán, Sechura, Mochica, and Quingnam on the coast, Culli in the highlands, and Hibo, Cholón, and the virtually undocumented languages Xoroca, Chirino, Bagua, Patagón, Sácata, and Copallín on the eastern slopes near Jaén. In addition, still today three different varieties of Quechua are spoken in Northern Peru, even though at least two are severely endangered. Very little can be said about the wider genealogical affiliations of the non-Quechuan languages of Northern Peru. Hibo and Cholón are generally considered to form a small language family of modest geographical extension (though cf. the doubts expressed by Torero 2002). The others, however, are best considered linguistic isolates, i.e. languages that cannot be shown to be related to any other language (see Torero 1993 for some suggestions for wider affiliations of Patagón, Chirino, Xirino and Copallín, and Sácata, based on very little data). Thus, also on the genealogical level, Northern Peru was originally more diverse linguistically than the Central and Southern highlands and the altiplano of Bolivia. What is more, toponymy betrays the presence of further languages in Chachapoyas and the highlands of Cajamarca (Torero 1989). These languages may have been related to Cholón (Urban to appear). Yet, as we will see now, there is one overarching commonality between the languages once spoken in the different areas of the Central Andes: the importance of language contact in shaping lexicon and, as far as is discernible, structure. For reasons of space, I here provide only a very general summary of relevant phenomena at a high level of abstraction.

In this particular case, indeed, the inability to demonstrate such relationships may well result from the lack of adequate lexical and grammatical documentation for many of the languages. It seems safe to say though that if the languages were in reality related to others that relationship was not a particularly close one.
Figure 1. A reconstruction of the linguistic situation in the Central Andes at the point of European contact, from Urban (2018). Based on Cerrón-Palomino (2010), Urban (2019 b), and Torero (1986, 1990, 1993). The figure is for illustration purposes only; all linguistic boundaries shown are approximate and lowland languages further east are not shown.
2. Language contact and the Quechuan-Aymaran mesh

Scholars have for a long time noted the striking similarities between the Quechuan and Aymaran languages, and a significant amount of scholarly effort has been dedicated to understanding the genesis and nature of these similarities (recent contributions include, inter alia, Cerrón-Palomino 2003, Heggarty 2005, Adelaar 2010, 2012a, 2012b, 2017, Muysken 2012a, 2012b, van de Kerke und Muysken 2014, and Emlen 2017). The grammatical structures of both lineages show many similarities: at a rather coarse level of description, both Quechuan and Aymaran languages are agglutinating and exclusively suffixing, with extensive case marking on nouns and cross-referencing of subject and object and verbs. Similarities are also found in the vowel and consonant systems: a three-vowel system characterizes both the Quechuan and the Aymaran languages; the consonant inventory characteristically involves a uvular (lost in some Quechua varieties) and a retroflex affricate. But parallel structures can extend to such minute details that they are hardly explainable through chance factors, also because the particular language type represented by Quechuan and Aymaran is not too frequent in South America overall (Adelaar 2012: 607).3 What is more, the lineages share a significant amount of vocabulary, including basic vocabulary. The similarities are not shallow either: in fact, all of the aforementioned properties reconstruct to the respective proto-languages, with the qualification that many Quechuan case markers appear to have only been in the process of lexicalization at the proto-Quechua level (cf. Cerrón-Palomino 1987: 132–137) and that generally the Quechua morphological template gives the impression of being a more recent development than that of Aymaran (cf. Adelaar 2012a: 593, 2012b: 463). What to make of all this? The grammatical and in particular lexical similarities have suggested to some that at a deeper time level, both families are descendants of a still older common ancestor (Orr and Longacre 1968, cf. Campbell 1985). More recently, however, the consensus has emerged that the similarities are better explained through language contact. Although not framed as part of an answer to the question of the relationship between Quechuan and Aymaran, Emlen and Dellert’s (to appear) realization that some of the usually disyllabic proto-Quechua roots can be etymologized as consisting of smaller monomorphemic building blocks that must have been productively combinable at a very early stage of development of the Quechuan lineage is not easily reconciled with the idea of a common origin of Quechuan and Aymaran because Aymaran roots cannot be analyzed in this way. A genealogical relationship is always a costly hypothesis, but the alternative contact-based explanation of the similarities does not come cheap in this case either, for it means that speakers of the respective ancestor languages, proto-Quechua and proto-Aymara, must have been engaged in long-term language contact that was capable of leading to a wide-ranging convergence of the grammatical systems, or, given the larger transparency and regularity of the Quechuan grammatical apparatus, the reorganization of proto-Quechua’s grammatical structure according to an Aymaran model (cf. Adelaar 2012a: 593, Adelaar 2012b: 465). To complicate matters further, the Quechuan-Aymaran contact relationship harks not only back to the respective proto-languages, but further later layers of contact influences at more local levels overlay the so-called initial convergence (Adelaar 2012b). These arose as speakers of individual Quechuan and Aymaran varieties, after the split of the respective lineages and the dispersal of the language families from their inferred homeland in Central Peru to other parts of the central Andes, came into contact once more. The details are very complex and not yet entirely understood. What seems beyond doubt, however, is that this happened in at least two places: on the one

3 See Cerrón-Palomino (2003) for more details.
hand, close to the homeland, speakers of the so-called Quechua I varieties of Central Peru interacted with speakers of Jaqaru, which represents a branch of Aymaran languages that is today restricted to a few villages in the highlands near Lima. On the other hand, in southern Peru, the area around Cuzco in particular, the local Quechua variety has not only exchanged a notable amount of vocabulary with Aymara, but has also acquired aspirates and ejectives which are widely attributed to Aymaran influence (cf. Mannheim 1991 on the precise mechanisms by which the contrasts were introduced into Cuzco Quechua). The Quechuan-Aymaran relationship is thus a complex story of multilayered language contact.

3. Language contact and the Northern Peruvian sphere

As we will see now, also the linguistic landscape of Northern Peru has been shaped decisively by language contact. One can observe commonalities between many of the Northern Peruvian languages that make them align with one another, but that at the same time set them apart in particular from Quechuan and Aymaran (Urban 2019a, 2019b). The languages do not generally give the impression of standing in a genealogical relationship to one another given the largely different lexical stocks with different characteristics. Yet, one can, even with the restricted amount of lexical documentation available, detect sets of loanwords that criss-cross coast, highlands, and eastern slopes (Urban 2017, 2019b). These include basic vocabulary items, not thought to be easily borrowed in casual contact situations (Thomason and Kaufman 1988). Concomitantly, one can observe rather clear similarities in root shape canons and phonotactic structures among the languages of Northern Peru which sometimes contrast starkly with the situation in Quechuan and Aymaran (Urban 2019a, 2019b): Northern Peruvian languages rely to varying degrees on monosyllabic roots which are, however, always more prominent than in Quechuan and Aymaran, which both have very strict disyllabic root canons. Northern Peruvian languages also allow word-final plosives more liberally than Quechuan and in particular than Aymaran. Together with the preference for monosyllabic structure and a certain predilection for lexical reduplication, this gives many northern Peruvian languages a distinctly recognizable fingerprint. One can even see how these preferences become activated in the adoption of Quechuan lexical material in a short list of numerals of the Quingnam language (Quilter et al. 2010), which actually constitutes the only dedicated direct documentation of this language that is available (see Urban 2019b for an analysis of placenames, personal names, and lexical localisms in the Spanish spoken in the former Quingnam-speaking area). Applying Thomason and Kaufman’s (1988) borrowing scale post hoc to the northern Peruvian data leads to the impression that Northern Peruvian languages were in non-casual contact with one another for a significant period of time (Urban 2017).

4. Language contact as an overarching point of access to the linguistic history and prehistory of the Central Andes

As the overviews in the preceding two sections have made clear, language contact looms large in the linguistic history of the Central Andes as a whole: there is no way of understanding the history of the Quechuan and Aymaran languages, and in the future possibly their deeper external relations (cf. Emlen
Matthias Urban

210

2017), without paying attention to the language contact phenomena that link them together. And in
Northern Peru, analysis of available data in terms of language contact opens up perspectives on linguistic
interactions and convergence across the different altitudinal belts of the Andes. Therefore, it makes sense
to frame a general linguistic prehistory of the Central Andes not exclusively on the basis of the expansion
of the Quechuan and Aymaran languages (cf. e.g. Heggarty and Beresford-Jones 2010), but to specifically
exploit the contact histories of the individual languages to gain a more fine-grained understanding of
linguistic relationships and thereby past interactions. As Lass (1997: 184–185) puts it beautifully, “[t]
he palimpsest that makes up the observable surface of a language is rarely (if ever) entirely the result
of its own internal history. At least part, either superficial like lexis, or ‘deeper’ in structure will likely
be the scars of encounters with other languages.” The distinct advantage of a contact-based approach is
that it is capable of giving a voice also to the many Central Andean isolates, which are unduly relegated
to supernumerary status in an approach focused on language dispersals: as Campbell (2017) shows, it
is not true that language isolates have to remain entities without histories just because our methods fail
to trace their genealogies. Rather, there is a bundle of techniques available to elucidate the history of
language isolates on par with languages belonging to larger genealogical groupings. Campbell (2017)
mentions, among others, the “[p]hilological study of attestations and historical reports” to achieve a better
understanding of the former extension of language isolates; as far as the Central Andes are concerned,
the value of these types of data has long been recognized and taken into account; in fact, the map in
fig. 1 is based to a not inconceivable extent precisely on evidence of these types. Campbell, however,
also mentions the “[e]vidence from loanwords,” “[l]anguage contact and areal linguistics,” and the study
of “[t]oponyms and other proper names.” Neither of these have been entirely unexplored in research
of the historical linguistics of the Central Andes either, and language contact between Quechuan and
Aymaran has in fact been a central topic ever since, but it is these that promise still further insights into
the dynamics of the development of the linguistic situation in the Central Andes as a whole.

The research project “The Language Dynamics of the ancient Central Andes,” hosted by the
University of Tübingen and funded by the German Research Foundation’s (DFG) Emmy-Noether
Program, seeks to explore these in close dialogue with evidence from archaeology and anthropology.
For instance, archaeological findings show that goods were exchanged within the Central Andes across
short and large distances from early on in prehistory, and likewise, the evidence indicates also early and
continuing interactions across ecological zones, in particular with the moist and hot lowlands of western
Amazonia (cf. e.g. Paulsen 1974, Lathrap 1973). At the same time, however, organic material is poorly
preserved in the climatic conditions of the rainforest. This leads to a situation in which trade routes and
patterns of exchange often cannot be pinned down with ultimate certainty. Through the analysis of lexical
borrowing, linguistics has the potential to help by establishing a second, and logically independent, pillar
of evidence for pre-Columbian exchange within and across the Central Andes. In fact, linguist Johanna
Nichols (as quoted in McConvell and Evans 1997: 13) has likened a loanword whose borrowing history
can be traced to the excavation of a pot which has the fingerprint of the previous owner still attached.
Therefore, members of the project are currently investigating possible lexical similarities in the languages
of the Central Andes and adjacent parts of Western South America, with a targeted focus on items and
goods that have been traded or exchanged in pre-Columbian times.

Of particular interest for the understanding for the language dynamics of the ancient Central
Andes is also the situation on the eastern slopes of the Andes. Extant studies (e.g. Muysken 2012a: 239–
Language Contact and the Linguistic Prehistory

240 on Cholón, Adelaar 1996 on Amuesha, Pache 2018 on the languages of Bolivia) suggest that speakers of languages of the eastern slopes were engaged in complex and multifaceted interactions with neighbors from the highlands that is both reflected in the loanword record as well as their areal-typological profile; likely such interactions also took place with the lowlands in the east (as e.g. suggested by Torero 2002: 161 for Cholón). Interactions were not only such that they reflect imperial Inca influence, but also bespeak more long-term, likely egalitarian, relations. Project members will seek to determine more clearly just what the linguistic record says about the nature of highland-lowland interactions.

Not only are, as Lass’s (1997) metaphor of the palimpsest suggests, languages often layered entities (cf. Aikhenvald 1996: 4–7), but many regions of the Central Andes have also seen successions of different languages at different points of time in history and prehistory. Spanish is now spoken in many areas of the Central Andes where formerly indigenous languages were spoken. Previously, Quechua must have replaced many local indigenous languages as it expanded from its homeland in Central Peru. Language shift, of course, can be conceived of as an extreme outcome of language contact, rendering also the succession of languages in geographical space relevant to a contact-based approach to the linguistic prehistory of the Central Andes. To deepen the understanding of the patterns of language replacement and shift and to reach layers of linguistic prehistory before prehistoric language shifts to Quechua, the study of toponymy plays an important role. Toponymic research can advantageously be complemented by investigating the lexicon of the replacing language (often, a local variety of Quechua) for possible remnants of the original language (so-called “substrate” vocabulary) as well as by paying attention to local personal names. Through the triangulation of such types of evidence, Urban (to appear) was able to suggest that a language related to Cholón has likely been replaced in the Chachapoyas area of the northern Peruvian Andes. A further application of an approach centered on the investigation of toponymy but supplemented by other types of data concerns the existence of a linguistic layer related to Aymaran in much of present-day Peru which Heggarty and Beresford-Jones (2010) associate with the so-called Early Horizon in the Central Andean cultural chronology. However, while indeed toponyms that can potentially be explained through Aymaran have been noted (Cerrón-Palomino 2000), more systematic research efforts are necessary to corroborate the idea of a former presence of Aymara especially in Northern Peru and, if successful, to determine the relationship of the language from which these toponyms remain to the Aymaran languages as we know them. Possibly, this is one of the earliest linguistic layers that can be reached in this linguistically and culturally complex region of South America.

References


Adelaar, Willem F. H. (2012b) “Modeling convergence: Towards a reconstruction of the history of Quechuan–Aymaran interaction.” Lingua Vol. 122 (5); 461–469.


