Towards a typological profile of the Andean languages

Adelaar, W.F.H.; Lubotsky, A.; Schaeken, J.; Wiedenhof, J.

Towards a typological profile of the Andean languages

In recent years, important progress has been made toward establishing genealogical connections between the Amerindian languages and language families of South America (e.g. Rodrigues 2000). Nevertheless, extreme genealogical diversity is still found among the Amerindian languages spoken along the western fringes of the Amazonian region and in the open areas of South America's southern cone. These areas are adjacent or close to the Andean region where genealogical linguistic diversity, though hollowed out by historical events of the past five centuries, remains as intractable as ever. As a result, the center of gravity of the genealogical linguistic diversity in South America has shifted from the east to the west.

Lack of progress in solving the genealogical puzzle of the South American languages has stimulated researchers to look for typological connections between the Amerindian languages of the subcontinent with the eventual goal to establish typological areas. From the point of view of linguistic typology, however, the diversity that is found in South America is hardly less formidable than from a genealogical point of view. And again, the east and center of the subcontinent seem to offer more possibilities to establish linguistic areas than the west.

It has been a common practice among linguists working on South American languages to make an intuitive distinction between 'Amazonian' and 'Andean' languages on the assumption that there would be two different language types corresponding to these labels. Obviously, this distinction is largely fed by geographical and cultural considerations. If we exclude the southern tip of South

America, where the situation is less clear-cut, there appears to be a wide cultural gap between peoples of the Andes and the Pacific coast, on one hand, and those of the Amazonian lowlands, on the other. Western societies tend to be more complex, with age-old sedentary habits and a highly diversified and technically well developed agriculture. Some of these societies are counted among the great civilizations of the world, whereas for most of the eastern tribes this has never been the case. No matter which perspective one wishes to assume, cultures and languages situated at the edge of the two general areas are likely to share elements of both sides, or else, they will have to be classified as either typically 'Andean' or typically 'Amazonian.'

From a synchronic point of view, a distinction between Andean and Amazonian languages is not entirely out of the question. For instance, Andean languages tend to have elaborate numeral systems, generally conceived on a decimal basis, which have the potential of counting as far as a hundred thousand or even further. Such numeral systems are often paralleled by the use of knotted threads (*quipus*) and abacus-type devices, which reflect a quantitative focus in the organization of society. By contrast, the Amazonian languages are known for having numeral systems of extreme poverty. As a matter of fact, some tribes, such as the Pirahã, have been reported not to count at all (Gordon 2004, Everett 2005). Languages that are closely related to each other, such as Chiriguano and Paraguayan Guaraní, differ by the fact that the former has an elaborate system of numerals, reflecting its Andean contacts, whereas the latter has only four true numerals. Higher numerals exist in Paraguayan Guaraní but were created artificially by missionaries. Early colonial accounts report that speakers of Tupinambá (closely related to the two former languages) used to indicate numbers higher than four by displaying body parts (e.g. 'both these hands, both these feet' meaning 'twenty'; Lemos Barbosa n.d.). Obviously, for languages situated in the interface regions connecting Andean and Amazonian areas, an

elaborate numeral system may betray close or long-standing contacts with the Andean region, whereas the lack of it can be an indication of an Amazonian background. On the other hand, one must also envisage the possibility that numeral systems may be lost or simplified when a non-Amazonian ethnic group moves into an Amazonian environment, especially if we accept the idea that the absence of numerals is cultural (rejection of the concept of counting!) rather than inherited.

It may be significant that Andean numerals higher than two do not seem to be helpful for deep reconstruction goals. (The Chibchan family, which extends into Central America, constitutes an exception.) It looks as if the different numerals that are in use were either borrowed from other languages or invented independently for each language (or shallow language grouping). Even in the Chiriguano case, where the adoption of decimal numerals cannot be old, the source of the actual forms for '6' to '9' is obscure ('5' is designated by the word for 'hand'; Dietrich 1986: 170), although further research may yield an explanation. Whereas the cultural and societal dimensions of Andean numerals are not to be underestimated, their time depth seems to be limited. They have little or no relevance for establishing a typological profile of the Andean languages that takes into account earlier stages of development.

An agglutinative structure with an exclusive or near exclusive reliance on suffixes for all morphological and morphosyntactic purposes has often been mentioned as a typical feature of an alleged Andean language type. Clearly, Aymaran and Quechuan with their 'Turkic' or 'Altaic' structures, have acted as models of inspiration for this view. It is generally agreed that the Aymaran and Quechuan language groups, whether or not related to each other, have gone through a process of structural and lexical convergence, which lasted for a long period of time. It implies that in an initial stage of the contact situation one of

the two ancestral languages may have been remodeled to the extent of losing its inherited typological structure. There is not even a need to assume that the remodeling language in its earlier form would have featured the agglutinative suffixing structure of its daughter language.

When viewed on an Andean scale, however, there does seem to be a drift towards suffixation. One may be inclined to assume that the Andean region diverges in this respect from the rest of the New World, where the favored language type appears to be a mix of prefixes and suffixes (the prefixes often being personal reference markers). It is true that many languages in the Andean region have extensive suffixation and only a few prefixes, if any at all: Barbacoan and Chocoan languages, Páez, Esmeraldeño, Mochica, Uru-Chipaya, Araucanian, Huarpean. At the same time, it can be argued that some languages of the Chibchan family (Chimila, Tunebo or Uwa, possibly Cuna) may have lost the prefix part of a morphological system still present in related languages such as Ika, Kogui, Damana and Muisca. Furthermore, there are other Andean languages that did not follow the trend towards exclusive suffixation: Atacameño, Yahgan, the Chon languages (e.g. Tehuelche) and Cholón (a language of the eastern Andean slopes in Peru) retained the mainstream Amerindian mix of prefixes and suffixes, although suffixation is more elaborate. Araucanian relies on suffixation nearly exclusively, but it has a class of possessive modifiers, which may either represent old personal pronouns, or degrammaticalized prefixes. Kamsá or Sibundoy, a language isolate of the eastern Andean slopes in Colombia, has an elaborate prefix system which seems to mirror the extensive suffixation of Central Andean language groups such as Aymaran and Quechuan.

Interestingly, languages that are rather similar in structure to the ‘proto-typical’ central Andean languages Aymaran and Quechuan are the Jivaroan languages,

located in the pre-Andean Amazonian lowlands. These languages resemble Aymaran and Quechuan structure in many respects, and may be more faithful representatives of the 'Turkic' agglutinative type than many languages presently or formerly spoken in the Andes themselves.

Constituent order is relatively free in Andean languages, although there seems to be a preference for the order in which subject/actor and object precede the verb (SOV). In many languages, including Aymaran and Quechuan, subordinate clauses are strictly verb-final. Languages with a verb-initial constituent order are members of the Arawakan family (Amuesha, Guajiro) of Amazonian provenance. The status of other possible verb-initial languages, such as Esmeraldeño, remains debatable for lack of data. Incomplete descriptive or documentary studies of extinct languages rarely provide all the required information on constituent order, especially if this order is not rigid. A general characteristic of most Andean languages (including Aymaran and Quechuan) is that modifiers must precede the modified in hierarchically organized noun phrases. In some languages, however, adjectives follow the noun whereas other modifiers precede it. Such languages are found in Colombia (Chibchan languages) and in a belt of languages located in northern Argentina and northern Chile (Atacameño, Lule, Quechua of Santiago del Estero). Numerals sometimes align with the adjectives in matters of word order.

As far as phonology is concerned, it is often easier to enumerate the characteristic elements that Andean languages lack, rather than those that they do have. Suprasegmental features such as nasality spread and contrastive tone are common features in the Amazonian region, but are rare in the Andes (not so in the pre-Andean eastern lowlands!). Nasality spread is found in Chocoan. Some of the Chibchan languages of northern Colombia are tonal (Barí, Chimila).

In contrast with the situation of the Amazonian region, vowel systems in Andean languages tend to be simple. Aymaran and Quechuan are renowned for having only three contrastive, qualitatively distinct vowels: *a*, *i (~e)*, *u (~o)*. This system has expanded to a five-vowel system in some of the dialects and in borrowed words. A tri-vocalic system is also found in Amuesha, a pre-Andean Arawakan language that underwent a substantial influence from Quechua (Wise 1976). Other Andean languages normally have five qualitatively distinct vowels (*a*, *e*, *i*, *o*, *u*) or six (the same series augmented with an extra vowel which can be either a non-rounded high back vowel, or a high or mid central vowel). More complex vowel systems are found in the far south of the subcontinent (Chon languages, Kawesqar, Yahgan) and in Mochica, a typologically anomalous language of the Peruvian coast. Other vocalic modifications (length, nasality, voicelessness, aspiration, glottalization) are dispersedly distributed in the Andean region but are not very common.

The use of uvular or post-velar stops (either in contrast with, or instead of their velar counterparts) is widely found in an area comprising the central and southern Andes (Aymaran, Quechuan, Uru-Chipaya, Callahuaya, Atacameño) as well as in the eastern lowlands of the southern cone of South America (Matacoan, Guaicuruan, Vilela, the Chon languages, also the Chilean language called Kawesqar or Alacaluf). Araucanian and Yahgan (and probably Huarpean), which have no uvular stops, are exceptional in this respect. It is interesting to note, that uvular stops occur neither in the Amazonian region, nor in the northern Andes. When moving northward on the map, uvular stops resurface in the Mayan language family of Mesoamerica, as well as in Pacific North America.

Glottalized stops and affricates (or ejectives) are found in roughly the same area as the uvular and post-velar stops, but there are more exceptions. They are found in Aymaran, some Quechuan dialects, Uru-Chipaya, Callahuaya, Atacameño, Matacoan, Vilela and the Chon languages. On the other hand, it is not entirely absent from pre-Andean Amazonia (Itonama, Jebero, Piaroa). Again, Araucanian and Yahgan (and probably Huarpean) are exceptional because they lack glottalized obstruents. When moving northward on the map, glottalized stops resurface in Mesoamerica (in the Mayan languages and in Lenca, for instance).

It is very difficult to make any generalizations about the distribution of aspirated consonants, which are found in dispersed locations of South America (Arawakan languages, Bora, Cofán, Mosetén, Páez, etc.). However, the characteristic three-fold division of plain, glottalized and aspirated stops and affricates is restricted to the southern part of the Central Andes (Aymaran, some Quechua dialects, Callahuaya, Uru-Chipaya, possibly Atacameño). Moving northward, this three-fold division resurfaces in California (in Pomoan).

Retroflex affricates in opposition with alveopalatal affricates are found in several Andean and pre-Andean languages: Quechuan, Jaqaru (Aymaran), Chipaya, Araucanian, Kamsá, Guambiano (Barbacoan), Amuesha and Chamicuro (both Arawakan languages), Gününa Yajich (a Chon language). Given the dispersed locations of these languages it is difficult to establish a pattern.

A palatality contrast is very common in the Andean region (with a number of exceptions), but it is usually restricted to resonants and sibilants. A more fully developed palatal / non-palatal distinction, comparable to Irish Gaelic or

Russian, which also affects other consonants, is found in Amuesha (Arawakan), in Mochica and in Páez.

Some interesting sound contrasts are only found locally: Trilled vibrants have a contrastive status in a number of languages of northern Colombia: in the Chocoan languages, in Damana and Barí (both Chibchan) and in Guajiro (Arawakan). There is an alveolar / interdental contrast in Araucanian. Chipaya and some of the Chon languages (Ona or Selk'nam, Gününa Yajich) have a three-way contrast for sibilants (alveolar-apical-retroflex). Apical sibilants and affricates may have existed in Mochica, although the evidence is not conclusive.

Labial or labio-dental fricatives are rare, as elsewhere in the Americas. They are found in Mochica, Muisca, Páez, Araucanian, in the South Barbacoan languages Cha'palaachi and Tsafiki, in the Aymara dialect of northern Chile, in Kawesqar and in Yahgan. In the pre-Andean eastern lowlands they are found in Betoi, in Huitotoan, in Cofán and in the Matacoan languages (*inter alia*).

From the point of view of morphological and morphosyntactic categories Andean languages are also highly diverse. One of the most salient characteristics is again negative, namely the absence of Amazonian-type classifier systems (cf. Derbyshire and Payne 1990). Such classifier systems, which are largely based on distinctions of shape, are widely found in the pre-Andean Amazonian region, where they are subject to borrowing and imitation. They are not found in the Andes. By contrast, numeral classifiers of the Chinese or Mayan type exist in Cholón, in Mochica and in Cuna (Chibchan). Some shape morphemes that are reminiscent of Amazonian classifiers have been reconstructed for Proto-Chibchan (Constenla 1988). Timote-Cuica, an extinct language family in the Andes of Venezuela, shows evidence of a Bantu-style class system marked by

prefixes. However, the data for these languages are too poor to permit any conclusive statement.

In most Andean languages, including Aymaran, Quechuan and Araucanian, gender distinctions are not expressed morphologically. As a rule, these languages do not even make a distinction between feminine and masculine deictic pronouns. Surprisingly, two Andean language families, Uru-Chipaya and Chon, do distinguish gender and also exhibit extensive morphological agreement. In the eastern lowlands of South America several language families have gender and gender agreement: Arawakan, Arawá, Boran, Chapacuran, Mosetén, etc. The way gender distinctions are treated is another expression of the extraordinary typological diversity of this subcontinent.

Case marking on noun phrases expressed by means of suffixes or postpositions is common in Andean languages. A surprising exception is Araucanian, which has one single nominal postposition used for a wide array of oblique case relations. It may be a consequence of a general tendency to avoid nominal morphology, which seems to characterize this language. Accusative case marking is found in several central Andean languages (Aymaran, Quechuan, Barbacoan, Páez, also in the pre-Andean Jivaroan languages), but less so in the northern Andes (cf. Constenla 1991) and in the south. Some northern languages, viz. Chimila, Ika (both Chibchan) and Cholón, may use a disambiguating affix that indicates an actor in opposition to an (unmarked) object. It may be tempting to refer to such markers as ergative, but the general structure of the languages in question does not favor such an analysis. True ergative case marking is found in the Chocoan languages, in (Chibchan) Tunebo, and possibly in the extinct Puquina language. Solid ergativity is furthermore attested in the Panoan languages of the Amazonian lowlands of Peru.

The formal distinction between possessed and non-possessed roots is a wide-spread phenomenon in Mesoamerica (Mayan, Uto-Aztecan) and in the eastern lowlands of South America (Arawakan, Cariban, Tupi-Guaraní). It is nearly unknown in the Andean region, except in the Arawakan languages (Amuesha, Asháninka, Guajiro, etc.) established on its eastern slopes. However, such distinction occurred extensively in Mochica, where the difference could be expressed either lexically, or morphologically. Interestingly, Mochica only had non-morphological means at its disposition to indicate the possessor of a noun, whereas most Amerindian languages that distinguish possessed from non-possessed roots indicate the person of the possessor morphologically. The Atacameño language marked possessed roots with a suffix, but there is no evidence of competing roots.

The distinction between transitive and intransitive verbs is formally expressed in some of the northern Andean languages (Ika, Muisca, Cholón). However, in most of the central Andean languages (including Aymaran and Quechuan) the transitive / intransitive distinction hardly plays a role. Like in English, these languages have verbs such as ‘to break’ and ‘to turn’ that can be used both in transitive and in intransitive contexts. This is remarkable considering the high level of explicitness of the morphological categories in these languages.

As many Amerindian languages elsewhere in the continent, some Andean languages encode both the subject and an object in a verb form. Some languages (e.g. Ika) encode different types of objects. In suffixing languages, such as Aymaran, Quechuan, Araucanian and Mosetén, the combined encoding of subject and object may produce highly complex (and sometimes unpredictable) endings to account for all the possible combinations in the different verbal sub-paradigms (tense, mood, nominalization, etc.). Since the distinction between

transitive and intransitive may be fluid (see above), the encoded object often represents an indirect (human) object, rather than a direct object.

The stative-active distinction, which is attested in eastern lowland languages (e.g. Arawakan, Tupi-Guaraní) and in languages of the Gran Chaco (Guaicuruan), has not been found in the Andes. However, it is found in pre-Andean Arawakan languages such as Guajiro.

Some languages (Araucanian, Mochica) have well developed morphological passives. Mochica even had a special case marker for agents of passive verbs. Other languages have special verb forms of which the 'passive' status is debatable (Muisca) or peripheral (Cholón). By contrast, many languages, including Aymaran and Quechuan, have no morphological passive at all. Some morphological derivations in specific Quechuan dialects are semantically reminiscent of a passive, but their syntactic status is not that of a passive as generally accepted.

In most Andean languages, as in some pre-Andean languages with an Amazonian background (for instance, the Arawakan languages Amuesha, Ashéninka, Guajiro etc.), the verbal morphology is extremely rich and varied. There is no space to discuss the enormous variety of semantic distinctions that can be encoded in the verbs of Andean languages. We shall mention the case of switch-reference, which assigns different endings to a subordinate verb in an adverbial construction depending on whether their subject is identical with, or different from the subject of the verb to which they are subordinated. Elaborate switch-reference is found in Tsafiki (Barbacoan), in Quechuan, in Jaqaru (Aymaran), in Jivaroan, in Panoan and in Uru-Chipaya. As one can appreciate, the switch-reference belt apparently runs across the Andes, rather than coincide with it. Interestingly, the Barbacoan, Panoan and Uru-Chipaya languages seem

to share formal elements in their switch-reference paradigms which may be coincidental but deserve a closer investigation. Outside the Andean area, switch-reference has been attested in parts of western North America (e.g. in the northern, Numic branch of Uto-Aztecan).

Andean languages such as Aymaran, Quechuan and Uru-Chipaya have a set of morphological nominalizations, which may encode relative tense and which may be combined with case markers to form different types of adverbial clauses and complement clauses. Case markers cannot be attached to finite verb forms, as if often the case in Amazonian languages (Tupi-Guaraní, Cholón). Araucanian, which lacks case markers, also relies on nominalizations to form complement clauses.

Personal reference systems in Andean languages generally reflect the classical Amerindian pattern of 1st person, 2nd person, 3rd person, 1st person inclusive (also known as 4th person). The number distinction is usually the product of a secondary development, although it plays an important role in Araucanian, which distinguishes singular, plural and dual in the pronominal morphology. Originally, the inclusive plural may have referred to the group to which both speaker and hearer belonged. The Aymaran languages exhibit the four-person pattern in its purest form (possibly of all Amerindian languages). In many languages the inclusive has developed into a plain non-singular 1st person (e.g. in Araucanian, in Cholón, and in the Chibchan languages). Others have introduced a secondary inclusive (e.g. Quechuan). Barbacoan languages exhibit a different pattern that separates 1st person from the other grammatical persons.

Aymaran and Quechuan are well known for having strict and compelling systems of evidentials that indicate data source. They may find their expression in the verbal paradigm (Aymaran) or in enclitic elements operating at the

sentence level (Quechuan). Extensive systems of evidentials have been reported for other Andean languages as well, such as Tsafiki (Barbacoan) and Páez, and for pre-Andean languages eastern lowland languages, including Panoan and Tucanoan languages (Aikhenvald 2004). It is possible that other Andean languages (for instance, Araucanian and Uru-Chipaya, for instance) have similar phenomena which still await a systematic comparative interpretation. As for now, it is too early for a definite statement as to which Andean languages distinguish evidential categories and which languages do not.

The preceding pages only give a very general impression of the typological variety to be found in the languages of the Andean region. We must conclude that there is still very little evidence that can be helpful for recognizing and delimiting linguistic typological areas, let alone, an Andean linguistic area that would encompass the entire region. Most conceivable characterizations of Andean languages are negative: Andean languages are predominantly suffixing case-marking languages, which have no prosodic nasality, no tone, no complex vowel systems, no nominal classifier systems (other than numeral), no gender (except for two language families), no stative-active systems and no well developed ergativity (except in one language family and in one language). The most 'exotic' language from an Andean point of view is Mochica. Its unique linguistic characteristics (phonology, numeral classifiers, possessed vs. non-possessed nouns) and its location on the northern coast of Peru are matched by an extraordinary cultural development in the past, which may be indicative of external contact or migration (cf. Shimada 1999). Another language that stands out in many respects is Araucanian (Mapuche). Its lack of nominal morphology, including case, is matched by extensive noun incorporation, another phenomenon otherwise absent from the Andes. As we have seen, it differs from the surrounding languages also from a phonological point of view.

References:

Adelaar, Willem F.H., with the collaboration of Pieter C. Muysken (2004) *The Languages of the Andes*. Cambridge University Press, Cambridge U.K.

Aikhenvald, Alexandra Y. (2004) *Evidentiality*. Oxford University Press, Oxford-New York.

Alexander-Bakkerus, Astrid (2005) *18th Century Cholón*. Ph.D. dissertation Leiden University.

Cerrón-Palomino, Rodolfo M. (in press) *El chipaya. La lengua de los hombres del agua*. Lima.

Constenla Umaña, Adolfo (1988) Indicios para la reconstrucción de clasificadores en el sintagma nominal protochibcha (1). *Filología y Lingüística* 14:2, pp. 111-8. Editorial de la Universidad de Costa Rica, San José C.R.

Constenla Umaña, Adolfo (1991) *Las lenguas del Area Intermedia: Introducción a su estudio oral*. Editorial de la Universidad de Costa Rica, San José (C.R.).

Derbyshire, Desmond C., and Doris L. Payne (1990) Noun classification systems of Amazonian languages. In: Doris L. Payne (ed.) *Amazonian Linguistics. Studies in Lowland South American Languages*, pp. 243-71. University of Texas Press, Austin.

Dietrich, Wolf (1986) *El idioma chiriguano.* Instituto de Cooperación Iberoamericana, Madrid.

Everett, Daniel L. (2005) Cultural constraints on grammar and cognition in Pirahã. *Cultural Anthropology* 46:4, pp. 621-46.

Fernández Garay, Ana V. (1998) *El tehuelche. Una lengua en vías de extinción.* Estudios Filológicos, Anejo 15. Valdivia (Chile).

Gordon, Peter (2004) Numeral cognition without words: Evidence from Amazonia. *Science* 306 (15 Oct. 2004), pp. 496-9.

Lemos Barbosa, A. (n.d.) *Curso de Tupi Antigo.* Livraria São José, Rio de Janeiro.

Mithun, Marianne (1999) *The Languages of Native North America.* Cambridge University Press, Cambridge U.K.

Rodrigues, Aryon Dall'Igna (2000) 'Ge-Pano-Carib' x 'Jê-Tupí-Karib'; sobre relaciones lingüísticas prehistóricas en Sudamérica. In: Luis Miranda Esquerre (ed.), *Actas del I Congreso de Lenguas Indígenas de Sudamérica* (Lima, 4-6 August 1999), vol. 1, pp. 95-104. Universidad Ricardo Palma, Facultad de Lenguas Modernas, Lima.

González de Pérez, María Stella, and María Luisa Rodríguez de Montes, eds. (2000) *Lenguas Indígenas de Colombia. Una visión descriptiva.* Instituto Caro y Cuervo, Bogotá.

Sakel, Jeanette (2004) *A Grammar of Mosetén.* Mouton de Gruyter, Berlin – New York.

Shimada, Izumi (1999) Evolution of Andean diversity: Regional formations (500 B.C.E.-C.E. 600). In: Frank Salomon and Stuart B. Schwartz, *The Cambridge History of the Native Peoples of the Americas. Volume III: South America.* Part 1, pp. 350-517. Cambridge University Press, Cambridge U.K.

Smeets, Ineke (1989) A Mapuche Grammar. Ph.D. dissertation, Leiden University.

Valenzuela Bismarck, Pilar M. (2000) Ergatividad escindida en wariapano, yaminawa y shipibo-konibo. In : Hein van der Voort and Simon van de Kerke (eds.), *Essays on Indigenous Languages of Lowland South America* (Contributions to the 49th International Congress of Americanists, Quito, 1997). ILLA Series no. 1, pp. 111-28. CNWS, Leiden.

Wise, Mary Ruth (1976) Apuntes sobre la influencia inca entre los amuesha : Factor que oscurece la clasificación de su idioma. *Revista del Museo Nacional* 42, pp. 355-66. Lima.

Zamponi, Raoul (2003) *Betoi*. Languages of the World/Materials 428. Muenchen, Lincom Europa.