Contacts between indigenous languages in South America

Pieter Muysken

1. Introduction

this chapter I will try to describe a few aspects of language contact in the history of the languages of the American Indian communities of South America. The topic of contacts between the indigenous languages in South America is vast and almost intractable. This is the case particularly because we still know little about the history of the languages of the continent, in the absence of essential sources of informatin, which include:

- historical sources dating back more than a few centuries
- reliable and complete descriptions for the majority of languages or major representatives of language families
- reliable family trees for a number of linguistic families
- reliable reconstructions of the features of potential ancestor languages

In quite a number of cases, perhaps the majority, we do not know whether a given instance of resemblance between two languages is due to contact or to shared ancestry. This difficult and complex state of affairs is illustrated by the fact that using reliable and tested techniques of historical reconstruction, we can distinguish around a hundred language families for the continent (cf. e.g. Kaufman 1990; Campbell, classification, this volume). At the same time, Joseph Greenberg (1987), using highly contested techniques of mass lexical comparison and relying on impressionistic observations about recurring features, argued that there is only language family. There are a number of "pan-Americanisms" (Kaufman 1990: 26), but they do not allow us to reliably create larger family units.

There are a great many sporadic observations about the contact between indigenous languages in South America in the literature, and few scholars active in the field would deny its importance, but no consistent picture has emerged as yet, nor is there an inventory, let alone a typology of contact phenomena in place.

In this paper I hope to achieve two things: (a) present some of the findings regarding language contact settings involving the indigenous languages in South America, largely (but not entirely) leaving aside here the contact with the colonial languages (Spanish, Portuguese, English, French, and Dutch); (b) sketch some of the most typical types of contact settings known at present, illustrating each one with one or two striking examples. These latter can then serve as proto-types, help-

ing us to discern the specificities and general features of other contact settings in the region as well, which may resemble them to a certain extent.

The specific types or proto-types of language contact that will be discussed in this paper are:

- Sub- and superstrate effects in imperial languages, illustrated with the Quechuan language family
- Symmetrical intensive bilingual contact, illustrated with the Quechuan-Aymaran relationship
- Dispersal languages, illustrated with the Arawakan language family
- Multilingual extended communities: illustrated with the Içana and Vaupés
- Lingua francas in the Tupían family, illustrated with Nheengatú
- Intertwined languages, illustrated with Kallawaya
- Languages resulting from incomplete shift, illustrated with Kokama
- Pluri-directional structural convergence due to prolonged coexistence, illustrated with the Guaporé-Mamoré area
- Shared lexical borrowings from dominant languages.

Some language families are a bit overrrepresented in this survey of proto-types, notably Quechuan, Arawakan, and Tupían. There are two reasons for this. First these families have a wide distribution and complex history, which has meant that members of these families underwent many forms of contact. Second, these families are relatively well-known, which makes it easier to detect contact phenomena involving their members. Surely many more complex contact situations will become unvealed as our understanding of the historical relations between the languages of South America increases. It was difficult to organize the material presented here, since certain special cases (Kallawaya, Kokama, Amuesha) could have been placed in other sections, involving several of the families discussed at the same time.

Before turning to the specific contact types, a number of remarks need to be made about the preconditions for studying language contact in South America.

First, I should mention that incidental word borrowing is something encountered throughout South America, as elsewhere in the world. We find many words shared by individual unrelated or not closely related language pairs; this is the rule rather than the exception, and sometimes the amount of shared, but unequivocably borrowed, vocabulary may be as high as a double digit.

In contrast, structural borrowing is by no means as frequent. In many regions in South America languages are spoken next to each other with widely different structural characteristics. By no means do all languages in the continent or in a specific region necessarily resemble each other structurally. Taking the maps in the *World Atlas of Linguistic Structures* (Haspelmath et al. 2005) as a point of departure, we see quite a bit of structural pluriformity, even though the continent is not very well represented on most maps. Nonetheless, we do find a number of features that recur in a number of the languages of South America:

- complex verbal morphology
- agglutinative morphology
- head marking
- evidentials
- both nominal and verbal classifiers
- possession often marked on the possessed noun
- clause subordination through nominalization

These features may ultimately be due to diffusion or to genetic inheritance, and some are not logically independent.

Second, a methodological remark about the study of language contact in the South American context is in order. While language contact is extremely important if we want to explain the features of a number of South American indian languages, language contact studies can never, repeat never, replace careful historical genealogical research. It will become clear that only if we know a fair amount about the history of a particular language family that we can discuss the specific role of contact at all.

When we are studying the relation between two language isolates, it becomes extremely hard, if not impossible, to separate similarities due to contact from those due to shared ancestry. At this point the only possibility we have involves independently established principles concerning the likelihood that a certain element was transferred from one language to the other. Thus in the lexical domain, resemblances in cultural (e.g. numbers and names of edible foods) and ecological vocabulary (e.g. names of specific plants and animals) may be more likely to be due to borrowing than resemblances in the core vocabulary. In the grammatical domain, resemblances among the discourse markers may be more likely to be due to borrowing than resemblances in the pronominal system.

Third, Taylor (1999) stresses the fact that cultural and socio-economic resemblances between many Amerindian lowland communities at the turn of the second millenium are deceptive. The Amerindian communities were highly diverse in their structural complexity, size, and economic basis of subsistence at the time of the European invations, and only subsequently were reduced to the small and sometimes isolated communities encountered at present. Thus we cannot reason back from the present day sociolinguistic settings in indigenous communities to earlier scenarios of language contact.

Finally, Schwartz and Salomon (1999: 458) emphasize the tremendous amount of ethnic restructuring that has occurred in the wake of the European invasions, and probably long before. Sometimes this involved just a few ethnic groups that were

merged, but in part larger collectivities, as "former ethnic identities were overshadowed by a new collective identity originally imposed from outside". These new collective identities were sometimes associated with linguas francas that were part and parcel of the colonial regime, such as Nheengatú in Brazil and the regional variety of Quechua called Quichua in Ecuador.

Thus the European presence was not only responsible for considerable direct Spanish and Portuguese influence on the indigenous languages, but also for indirect influence on the types of language contact that the indigenous languages underwent in a more general sense. Thus the decision to exclude European influence in this chapter is to some extent an artificial one.

Taking these general points into consideration, I now turn to eight proto-typical forms of contact, illustrated with individual examples.¹

2. Sub- and superstrate effects in imperial languages: Quechua

I will begin by sketching some of the dimensions of sub- and superstrate contact involving Quechua. This is the best known of the South American indian languages associated with a political unit larger than the ethnic group. Many scholars have tried to sketch the language situation in the Andes in the period since around 200 our era, trying to link linguistic, archeological, and ethno-historical evidence.

A core feature of Andean linguistic ecology in the last millenneum concerns Quechua in contact with other languages. A schematic overview of contacts involving Quechua is given in Table 1.

Quechua started out, probably, as a language spoken in the Andean range immediately east of Lima in central Peru. Probably Aymara was spoken in a neighboring area, and there is evidence for very early contacts between the two languages or language families (cf. Section 3). As it expanded throughout the central Andean region, Quechua came in contact with other languages as well, and may have undergone substrate influence from some of them, particularly in Ecuador, and possibly also in Bolivia. At the highpoint of Inca political power, just before the Spanish invasion, Quechua was spoken either as a native language or as a lingua franca from southern Colombia to northern Argentina and Chile. At the same time, languages spoken in the vicinity were influenced by Quechua, both in the pre-colonial period and after the invasion by the Spanish. I will briefly describe the various contact settings involving Quechua here, except for the contacts with Aymara, Puquina, and Kokama, which will be discussed in separate sections, and not discuss the complex relation between Quechua and Spanish, referring the reader to chapter 7 of Adelaar with Muysken (2004). More detailed information on the languages and language families involved likewise can be found in Adelaar with Muysken (2004).

Quechua in contact with	Type of contact	Linguistic effects	
Aymara (cf. Section 3)	Various dominance relations in the course of history – coexistence	Extensive mutual lexical borrowing Aymara derivational affixes in Q Aymara phonological substrate	
Uru-Chipaya	Qas a dominant neighboring language. Some communities have shifted to Q, others (in part) to Aymara	Q loans Possible calque in the pronoun system	
Puquina (cf. Section 6)	Q as a dominant neighboring language, eventually a complete shift	Relexified secret ritual lan- guage Kallawaya	
Shuar and Barbacoan	Q as a dominant neighboring language and as a trading language	Shift: *possible simplification of Q morphology in the lowland varieties *some Shuar and Barbacoan loans in flora/fauna *possible influence in local Q grammar features	
		Maintenance: *some Q cultural loans	
Kokama and Tupinambá (cf. Section 7)	Q as a lingua franca and trading language	Numerous lexical loans; some morphemes	
Cholón	Q as a dominant trading language	Extensive Q lexical borrowing Borrowing of several Q affixes	
Amuesha or Yanesha'	Q as a dominant/trading language	Extensive Q lexical borrowing Borrowing of grammatical elements	
Others, such as Leco and Yurakare	Q trading language	Incidental Q cultural loans	

Table 1. A schematic overview of contacts involving Quechua (Q)

The small Uru-Chipaya language family was widely spoken in hunting/fishing communities in the lake and river basin of the Peruvian/Bolivian altiplano, along the shores of the Titicaca and Poopó lakes and the rivers feeding into them and connecting them. The only place with a viable though small community of speakers is now Santa Ana de Chipaya, department of Oruro. There are a number of lexical

loans from Quechua, including some core vocabulary, some borrowed minor structural features, and the inclusive/exclusive distinction in the pronoun system appears to be calqued on the Quechua pattern. It should be noted, however, that the influence from Aymara, particularly in recent times, is much more extensive (Hannß 2007).

There are a few (in some local varieties perhaps more than a few) Quechua lexical borrowings in Shuar, such as the word for 'corn'; less well known is the possible contribution of Shuar to Quechua local (lowland) flora and fauna terminology.

The now extinct language Cholón was spoken in the upper Huallaga valley in northern Peru, north of the town of Tingo María. While the main source is a single 18th century manuscript, Alexander-Bakkerus (2004) provides a complete modern reconstruction on the basis of the available materials. Cholón has person prefixes, but in other respects it resembles Quechua and Aymara typologically, with a postpositional case system, SOV word order, etc. In addition to a number of Quechua lexical loans, it shares several affixes with Quechua: Ch -*pit* / Q -*pis*/-*pas* 'additive; indefinite'; Ch -(*a*)b' / Q b'a 'only'. Interesting is also that the Cholón system of exclamatives resembles that of Quechua, as the following examples show:

ood!'
0

Even though the actual forms do not always correspond, the principle of construction is similar, suggesting close contacts in an early period.

The Amuesha or Yanesha', speakers of an Arawakan language, orginally lived in a region stretching from the department of Huánuco to the department of Junín, immediately west of Quechua speaking territories in the Andes of central Peru. Since the Cerro de la Sal (Salt Mountain) was located in their area, they had frequent contacts with traders looking for salt from all over. For several centuries contacts must have been intense even if now the Amuesha live further eastward in the Amazonian lowlands. Building on earlier work of Wise (1976), who definitely established both the Arawakan genetic affiliation and pointed to the influence of Quechua on the language, Adelaar (2007) has provided a detailed analysis of the layers of Quechua influence on the language. What makes the Amuesha case particularly complicated is that there is evidence not just of Arawakan and Quechuan linguistic elements, but of a third as yet unknown language as well. The latter is currently being investigated. In contrast with the influence on other languages spoken alongside of Quechua, the influence on Amuesha is not primarily from the Quechua lingua franca associated with the Inca empire, but rather from the neighboring Quechuan dialects that are part of the so-called Quechua I cluster. There is phonological evidence that these loans are quite old. However, there are also a few loans, specifically referring to personal status terms, which must come from the Inca period.

Quechua loanwords cover all word classes and include a striking number (at least sixty) verbs, often with core meanings. Both Quechua noun and verb roots loose their final vowel in Amuesha. There are a few possibly borrowed suffixes in Amuesha, including *-kma* 'precisely', 'always', 'totally', 'of the same sex' (possibly < Q *-kama* 'until', 'distributive'), *-(V:)kop* 'referential', 'benefactive' (possibly Q *-paq* 'benefactive'), *-nya* 'intensifier', 'sequential' (possibly Q *-nya* 'already'), *-Vny* 'desiderative' (possibly Q *-na: / -naya* 'desiderative'), *-V:r* 'stative' (possibly Q *-ra(:)* 'stative'). The metathesis that we find in the Amuesha forms is independently motivated. In addition, there are a number of suggestive structural resemblances between Amuesha and neighboring Quechuan varieties in specific 'non-essential' constructions (Adelaar 2007: 309): the negation system, an apprehensive construction, an applicative reversal construction, subordination, double possessive marking. Altogether, however, Amuesha has not converged structurally with Quechua.

Its phonology, unusual both from an Arawakan and a Quechuan perspective (but partially resembling that of Cholón in the vowel elision rules in verb stems, and its partly untraceable lexicon suggest that the speakers of Amuesha originally spoke a different language and only later became Arawakanized. Needed is a detailed study of the grammatical system of Amuesha, the other Arawakan languages, and all neighboring languages, including Cholón, to see whether further insights can be gained.

Finally in languages such as Yurakare, borrowings are limited to elements such as *atalipa* 'chicken', *kuchi* 'pig' and *michi* 'cat' (the latter two themselves a borrowing from Spanish).

3. Intensive bilingual contact: the early Quechua-Aymara relation

The relation between Quechua and Aymara, or more appropriate between the Quechuan and the Aymaran languages, has long been an issue of often heated discussion. Orr and Longacre (1968) and, using more principled arguments, Campbell (1995) have argued for a common origin. Following Adelaar with Muysken (2004: 34–36), Heggarty (2005), and McMahon et al. (2005), I will assume that the evidence for a separate origin of Quechua and Aymara and intensive borrowing is stronger than that supporting a common origin. The (striking) evidence for the genetic link includes:

- the phoneme systems of the two language families are similar enough to allow for the reconstruction of a common proto-system (Orr and Longacre 1968);
- disregarding the large number of later borrowings between branches of the two families in both directions, about 20% of the core vocabulary is shared;
- the morpho-syntax of both languages shows an uncanny resemblance on the level of the structural and semantic organization of the grammar (Cerrón Palomino 1994).

However, other factors militate against postulating a genetic unit:

- the phonotactic patterning and very specific vowel deletion rules characteristic of the Aymaran languages as a group are not found in Quechua;
- while 20% of the words in the core vocabulary correspond between the two language families, 80% do not at all;
- this is all the more surprising since the 20% words are very similar if not identical in both language groups;
- there are very specific structural and semantic correspondences, but these do not extend to the actual shapes of the grammatical morphemes, which are all different;
- core parts of verbal inflection do not correspond;
- the semantic fields covered by a group of 150 specific culturally relevant lexical items do not overlap in the two language families.

Altogether, the scenario that best fits the data is that two possibly unrelated or only very distantly related languages coexisted in the same area, most likely central Peru, for a long time and profoundly influenced one another. One of the two probably was restructured in fundamental ways under the influence of the other language, and was remodeled on the basis of its morpho-syntactic patterns without taking over the actual grammatical morphemes associated with these patterns. Given its overall more complex phonotactics and more regular verbal and nominal person inflection, it is more likely that Aymara provided the model on the basis of which Quechua was restructured.

Apart from the early contacts that have affected both families in their entirety, there has been intensive subsequent contact between individual branches of the Quechuan and Aymaran families. The most striking effects are found in southern Quechua varieties. The Quechua of Cuzco and Bolivia probably has adopted the aspirated and glottalized stops in word initial position from Aymara (Mannheim 1991). Quechua dialects in the area of Arequipa and Puno (Peru) have adopted several Aymara verbal suffixes, inclusive of the accompanying vowel reduction rules (Adelaar 1987). In many areas of southern Peru and Bolivia there have been processes of language shift in rural communities from Aymara to Quechua (mostly) or from Quechua to Aymara.

4. Dispersal languages: Arawakan

A language family such as Quechuan was spoken in a more or less continuous area (interspersed, to be sure with pockets of speakers of Aymara, Puquina in southern Peru and Bolivia, and a few other languages in northern Peru and southern Ecuador). In contrast, the Arawakan languages are dispersed across a very wide region, spreading from Belize to Paraguay. Typically, many Arawakan languages or language clusters (such as the Campa branch) are surrounded by non-Arawakan languages. Ethnohistorically, the ancestors of the speakers of the Arawakan languages were associated with the archeologically defined Saladoid culture in the Orinoco basin (Allaire 1999: 696), ultimately introduced there perhaps from the Amazon basin.

For all Quechuan languages and language varieties, it can be unambiguously determined that they belong to the same family (the possible exception is the specialized ritual language Kallawaya (cf. Section 7). For Arawakan, this also holds on the level of basic verbal morphology, but the structural differences between the varieties are greater. Some Arawakan languages have become very different from Proto-Arawakan, at least morpho-syntacitcally.

Language	Arawakan in contact with	Type of contact	Linguistic effects
Island Carib	Kariña / Kariña pidgin	Male conquest of female community; use of Kariña pidgin features	Gender differences in language use
Tariana	Eastern Tukanoan	Exogamous bilingualism	Extensive grammatical borrowing, cf. Section 5
Resigaro	Bora, Ocaina	Close contact in terms of dependency relations	Borrowing of nominal morphology and classifiers
Mawayana	Cariban	Language attrition in a minority setting	Borrowing of grammatical distinctions from majority language
Amuesha	Quechua	Long-term interethnic trading contacts	Extensive lexical and grammatical borrowing, cf. Section 2
Palikur	Cariban and unknown other languages	Possibly shift to Arawakan	Grammatical borrowing
Kokama	Tupinambá	Imposition of and shift to Tupinambá in a complex multi-ethnic setting	grammatical restructuring and borrowing, cf. Sec- tion 8
Wãnsöhöt (Puinave)	Makúan	Shift to a Makúan lan- guage by Arawak speakers	Grammatical restructuring

Table 2. A schematic overview of contacts involving Arawakan languages

I will discuss these settings one by one.

A special case are the Garifuna, descendants of the Island Carib. Their history is a complex one, and involves Arawaks, Caribs, and descendants of African slaves. The Arawak Indians left Guyana, Surinam and Venezuela in the second century A.D., settling in the Greater Antilles. The Carib Indians left Orinoco Delta in the thirteenth century and conquered the Lesser Antilles. From the mixture of the Carib and Arawak the well-known but now extinct language Island Carib resulted. The descendants of the original mixed communities formed new communities on Saint Vincent. There a number of marooned slaves were adopted into the community. Finally the ensuing group, the Black Carib, settled off the coast of Honduras and then spread into Belize and Nicaragua. Hoff (1995: 50) argues that the contacts between the Kariña speaking Caribs and the Igneri speaking Arawaks took place in Kariña pidgin rather than full Kariña, which survived until the 20th century along the northern coast of the South American mainland. Characteristic of this pidgin was the use of formal marking on the verb of the transitive (*s*- prefix) / intransitive (*n*- prefix) distinction, which survived in Island Carib:

- (2) a. *amoré s-ineri touna* you TR-drink water 'You are drinking water.'
 - b. *amoré* **n**-oboüi you INTR-come 'You are coming.'

The verb forms are in fact frozen nominalized Kariña verbs, which are used as main verbs in the pidgin.

Contacts between Mawayana and languages of the Cariban family such as Trio and Waiwai date from the last 150 years, as described by Carlin (2007). Remnants of Mawayana speaking groups were incorporated into a Trio (and partially also Waiwai) speaking village and became bi- or trilingual. As a consequence, Mawayana has adopted a number of obligatory Cariban categories: a first person plural inclusive/exclusive distinction, the category of nominal past, the category affective (pity), frustrative (in vain, almost, partly), and similative (to be like, but not quite). In contrast, it has lost its Arawakan gender system and the classifier system. Carlin argues (2007: 330) that there was "a clear resistance to the transfer of actual morphological forms but not to the transfer of structural categories". The morphological material required for the new categories is generally taken from the original languages.

Pa'ikwaki or Palikur is spoken by slightly over a thousand people in the border area of State of Amapá, Brazil, and French Guyana. The first historical records after contact with the Portuguese situate the Palikur slightly to the north of the mouth of the Amazon River. Since they became involved in conflicts between the Portuguese and the French colonial powers in Cayenne, they were forced northward and subject to bad treatment by the Portuguese. The language has undergone a number of grammatical changes, without a clear source as far as is known at this moment.

Payne (1985) has established that the genetic classification of Resígaro is squarely within the Arawakan language family, in spite of the many elements shared with Bora that Allin (1975) had discovered. Aikhenvald (2001) has further analyzed the considerable influence from Bora on the language, pointing to the predominance of borrowed Bora classifiers in Resígaro. This influence is currently being further studied by Seifart (2006) and Wloczkowska (2006), partly on the basis of new fieldwork data. In Resígaro 24 % of the core vocabulary is of Bora origin, while there is no borrowing in the other direction. Resigaro has adopted a two tone contrast, a phonemic glottal stop, and syllable structure restrictions from Bora. In the pronominal system, it has adopted the inclusive/exclusive distinction in the first person plural through a Bora pronoun. Of the 56 classifiers in Resígaro, only 8 or 9 have an Arawakan etymology, and 36 are from Bora. Striking and highly unusual is that borrowing is limited to the nominal domain, and involves a high proportion of the nominal affixes and the pronouns. Except for animal names and nouns that can be used as classifiers, almost no other lexemes have been borrowed. Semantically, the domain into which there has been borrowing concerns unitization (through class markers), number, and quantity.

Wãnsöhöt (Puinave) has been studied by Girón (2007). He confirms the relation between Wãnsöhöt and Makúan, but suggests as a likely hypothesis to account for the extensive grammatical restructuring and non-Makúan vacobulary in Puinave an earlier process of shift to a Makúan language by Arawakan speakers. However, much more detailed comparison with Arawakan and other languages is needed to trace the possible roots of this language.

5. Multilingual extended communities: The Içana and Vaupés

Starting with the work of Arthur Sorensen (1967), it has been noted that the Içana and Vaupés river basin in Northwest Amazonia, on the border of Colombia and Brazil, is a region with extensive multilingualism and language contact. The results of this have been documented in detail by many scholars, including Jackson (1983) and Gomez-Imbert (1996) from the perspective of the Tucanoan languages, Aikhenvald (1996, 1999b, 2002) for the Arawakan languages, and Epps (2007) for the Makúan languages. The most likely scenario is that originally the relevant part of the region was inhabited by the Makúans, who were forest-dwelling semi-nomadic hunters, and that later the Tukanoans and the Arawakans arrived, in that order. These latter groups lived along the rivers and were agriculturalists. The Makúans remained in a subordinate position and their language and culture had very

little social prestige. They did not intermarry with the other two groups, but had extensive economic interchange relations with them. The Arawakans and the Tucanoans intermarried, due to exogamy restrictions, so that many children grew up multilingually. However, all three groups maintained their languages as separate entities, at least at the level of the lexical shapes and the morphemes (except for the existential verb *ni*-, which has spread across members of all three families); there is little lexical borrowing. The Arawakan language directly influenced by Tucanoan is Tariana, while of the Makúan languages, Hup has been affected, but a slightly more distant language, Dâw, much less and Nadëb not at all.

In the processes of structural and semantic change which has lead to the emergence of a linguistic area in the Icana and Vaupés river basins the Tucanoan languages have been the source of innovations in the Arawakan and Makúan languages. In other words, the change was unidirectional. Change involved a number of domains of the language. In phonology, nasalization, a particular pitch accent system, and a number of of segmental features have spread from the Tucanoan languages (Aikhenvald 1999b: 394–396; Epps 2007: 272–273). In the lexicon, particular features of the Tucanoan classifier system have spread to Tariana, while the Makúan language Hup is developing a Tucanoan-like classifier system, with inanimates classified on the basis of shape, and animates classified in terms of gender. Likewise, a particular organization of the numeral system has also been adopted by Hup and Tariana, as well as a split number system (based on animacy) and associative plural. A striking case of diffusion concerns the complicated Tucanoan fiveway evidential system, into which tense, peson, and number markings are embedded. In addition, there is evidence for the spread of additional tense distinctions marked by particles. While sometimes the actual morphological realization differs in the three language families, there is also evidence of the spread of verbal compounding patterns, including the expression of causativity. In the case marking and alignment systems, Hup and Tariana have adopted typical Tucanoan features as well, e.g. in the system of animacy-based differential object marking, where human objects are always marked, animal objects depending on the degree of individuation, and inanimates are not marked. While the Makúan languages originally probably had a system of noun incorporation, it appears to be lost in Hup, and it does not occur in Tucanoan or Tariana either. All three languages frequently show verb final word order, again a feature spread from Tucanoan. The list given here, partly based on Epps (2007), could easily be extended.

Altogether it is clear that there has been systematic and profound, mostly unilateral, grammatical diffusion in the Içana and Vaupés river basins, but that it generally did not involve the transference of lexical items or language shift. Crucial is the link between ethnic (Tucanoan, Arawakan, Makúan) identity and language. In Tucanoan-Tariana relations this is instantiated through linguistic exogamy, and in Tucanoan-Hup relations through a long-term trading dependency relation.

6. Lingua francas in the Tupían family

The large Tupían language family is spread over much of the Amazon area, branching out into the Chaco with members of the well-known Tupí-Guaraní branch also spoken in Paraguay, Bolivia, and northern Argentina. While most of the languages in the family remained small languages of fairly isolated ethnic groups, a few members of the Tupí-Guaraní branch developed into linguas francas as a result of the Portuguese and Spanish colonization. The two best known ones are Nheengatú and Guaraní; the latter now functions as a national language in Paraguay.

Nheengatú (also called *língua geral* of Amazonia, or *lingua brasilica*) originated in the 17th century in what are now the states of Pará and Maranhão, as lingua franca on the basis of Tupinambá lexicon but with strong grammatical influence from Portuguese, also due to internvention by Jesuit missionaries (see also Moore et al. 1994). The original mixture was termed *ie'engatu* 'good language'. Around 1700 it was spoken in a large area in Brazil, as a contact language between whites and indians, but it lost some its support with the expulsion of the Jesuits in 1758.

Currently there are around 8,000 speakers in the area of the Upper Rio Negro, as well as in adjacent territories of Colombia and Venezuela. In one municipality, São Gabriel da Cachoeira, it has official status and is taught in schools. The first grammar of Nheengatú was written in 1556 (published in 1595) by father José de Anchieta. Its sister language in the colonial period was Língua Geral Paulista (in the state of São Paulo), a lingua franca which is now extinct. A text fragment from the Río Negro (cited from Taylor 1985) is given in (3) (with Portuguese items non-italicized):

(3) yepé viagem paa pedro umunhã festa iwáka-kiti yepé viagem paá pedro u-muñá festa iwàka-kití DT.IN time CIT Pedro 3s-make feast sky-DIR They say that once Peter gave a feast in heaven.

ápe paa jabuti *paa unhee*~: "*ti maye asu à-kiti aá-pe paá jabuti paá u-ñee ti mayé a-sú*

aá-pe paá jabuti paá u-ñee ti mayé a-sú aà-kití that-LOC CIT jabuti CIT 3s-say NEG like 1s-go that-DIR Then the jabutí (tortoise) said: There is no way for me to go there.

ixe aputái amaã nhaã festa ixé a-putári a-maã ñaã festa I 1s-want 1s-see that feast I want to see that feast. *ti aríku sepepu awewèu-rã ti a-rikú se-pepú a-wewèu-ráma* NEG 1s-have 1p-wing 1s-fly-FIN I do not have wings to fly.

àpe(-te) paa usasa icompadre urubu aà-pe(-té) paá u-sasá i-compadre *urubu* that-LO-EMP CIT 3s-pass 3p-friend urubu At that moment his friend, the urubú (vulture), passed by.

"eh compadre!" unhee~ paa, "ti sera repodéi rerasu ixé asu-rã iwàka-kiti amaã festa

"eh compadre u-ñee paá ti será re-podé+ri re-rasú
eh friend 3s-say CIT NEG INT 2s-can 2s-take
ixé a-sù-ráma iwàka-kití a-maã festa
I 1s-go-FIN sky-DIR 1s-see feast

"Eh friend!" he said, "couldn't you take me up to the sky so that I can see the feast?"

"ah!" paa unhee~, "anhu~ resu kwá violão-kwàra-kiti" ah paá u-ñee añu re-sú kwaá violão-kwàra-kití ah CIT 3s-say only 2s-go this violin-hole-DIR "Ah!" said <the urubú (vulture)>, "you can only go inside of the violin."

It is clear that a considerable amount of morphology has been retained in Nheengatú, but that it is quite transparent and regular.

While Nheengatú currently only has the status of a local language of a few groups in the upper Amazon, Guaraní has become a very important national language. Melià (2003), based on a Strasbourg thesis from 1969 by this expert on the language, reconstructs the development of Guaraní as the lingua franca of Paraguay in terms of the efforts by Jesuit missionaries to "reduce" the hitherto dispersed bands of indians into structured mission settlements, and similarly, to reduce the alien tongue of the infidels in terms of a writing system and systematic grammatical descriptions and ecclesiastical written materials. In Melià (1992) the continuities are stressed between the different varieties of Guaraní from the colonial past to the present day.

7. Intertwined languages: The case of Kallawaya

In the case of language intertwining, substantial portions of two languages are paired together, typically lexicon from one language and grammar from another one. The most well-known case of such a language involving to South American Indian languages is Kallawaya, a specialised ritual language, now almost extinct, spoken by healers from Charazani in northern Bolivia. Recent analyses are given in Muysken (1997; 2009), Adelaar with Muysken (2004: 350–362), and Muysken (2009). In this language the structure of Quechua is combined with lexical elements from Puquina as well as from other languages in the region; there are also some lexical elements of unknown origin. Kallawaya in its current form is best seen as a case of paralexification (Mous 2003): the special lexicon of Kallawaya only appears when the people from Charazani perform healing rituals (and even there the data presented by Rösing [e.g. 1990] suggest that oftentimes healing rituals are performed with Quechua lexicon). In ordinary language use, only Quechua lexicon appears.

Although Puquina was once important enough to be rated as the third *lengua* general 'general language' of the Andes in the early colonial period, it rapidly lost its prominent status and is now extinct. It was spoken in and around the Peruvian/Bolivian altiplano, mostly in the area between Arequipa and Lake Titicaca, as far as can be ascertained from place names and ethnohistorical records. It appears that the Puquina-speaking region was overrun and split up, in pre-conquest times, by Aymara. Little is known of Puquina; the elements in its pronominal system suggest a relationship with the Arawakan language family, but lexically this has by no means been established. There are a number of Quechua and Aymara lexical loans in Puquina, and possibly some Aymara words are of Puquina origin.

As said, the structure of Kallawaya is almost entirely that of the local Quechua. The following example (cited from Oblitas Poblete 1968: 44) illustrates this:

(4) č'ana-či-rqa-yki isna-pu-na-yki-pax
 call-CAU-PA-1S.2O go-RS-NOM-2S-BEN
 'I had you called so that so you can go.'

The non-italicized elements are not from Quechua, but from Puquina, while all other elements are regular Quechua affixes.

While the verbal system by and large follows Quechua rules, in the nominal system a number of deviations occur. First of all, in some sources the markers for second and third person appear to be reversed from the Quechua system. Second, plural marking is not consistent, suggesting closer similarity to the Puquina system. Third, the way pronominal forms are used is also fairly close to Puquina, as far as can be ascertained.

8. Languages resulting from incomplete shift: Kokama

A number of languages in South America appear to be the result of incomplete second language acquisition in a process of shift. One example of such a language is Kokama-Cocamilla (also known as Kukama-Kukamiria or Kokama), an endangered language spoken in the Peruvian Amazon (provinces of Loreto and Ucayali) by about 1500 people. According to Cabral (1995), the origin of Kokama must go back to the late pre-Columbian periuod, when a group of Tupinambá speakers migrated in the late 15th century from the Atlantic coast inland to the upper Amazon, and came in contact with speakers of one or more other languages, possibly Arawakan. Cabral argues that the large number of lexical elements shared with Tupinambá, coupled with an almost complete absence of shared morphological and grammatical features, and a number of phonological changes untypical of the Tupían family as a whole suggests that Tupinambá was learned as a second language, albeit imperfectly, by other groups, and that that Kokama cannot properly be classified as Tupían but rather has a mixed origin. Kokama morphology is extremely reduced, in comparison with that of Tupinambá. Tupinambá multi-morphemic words correspond to Kokama single morphemes (see also Vallejos Yopán 2010). Examples include:

(5) a. *Tupinambá Kokama* a?e-pe aepe that-LOC there
b. *e-i-pek epeka* 2IM-3-open open

There is no allomorphy in Kokama, while the few corresponding Tupinambá forms have allomorphic variants. Vallejos (2005) argues, on the basis of new fieldwork data, that the suffixes in Kokama are all simple concatenative elements, and cannot be separated into derivational and infectional:

 (6) yaepe-tsui ajan animaru-pura-tu-anu ipu-ka there-ABL DEM animal-FOC-AUM-PL make.sounds-REI 'Afterwards, those big animals make noises again.'

(Vallejos 2005: 8)

The Kokama lexicon, as stated, is primarily of Tupinambá origin, but also contains elements from Portuguese, Spanish, Arawakan, Nheengatú, and Panoan origin. There are also a substantial number of Quechua words, including plant and animal names, verbs, adverbials, and numerals. Very interesting is that the Spanish verbs in Kokama appear with a Quechua perfective morpheme:

 (7) regala-ška 'give' lea-ška 'read' sufri-ška 'suffer'

This morpheme appears to serve as a way of integrating Spanish verbs into the language, and may be the remnant of an earlier Quechua-based pidgin used in the upper Amazon (Crevels and Muysken 2005). In Kokama there is no case marking on arguments and no verbal argument marking – subject and object are indicated through word order only. Clauses marked for tense are SVO or OSV, while clauses marked for aspect are mainly SOV. In the paradigms for person, number, and demonstratives there is speaker-oriented gender marking. The following examples show the contrast between Ko-kama (8b) and Tupinambá (8a) (Cabral 1995: 170) (PL = punctual locative):

- (8) a. syé yára o- I- me?éŋ piná isé -be my lord 3- 3- give hook 1 PL 'My lord gave the hook to me.'
 - b. *maría yúme manipyára ta cúpe* Mary give hook 1 to 'Mary gave the hook to me.'

As further fieldwork data become available, the precise grammatical features, lexical roots, and morphological properties of this language can be investigated.

9. Pluri-directional structural convergence due to prolonged coexistence, illustrated with the Guaporé-Mamoré area

Reporting on still ongoing research, Crevels and Van der Voort (2007) argue that the Guaporé-Mamoré area, comprising parts of the Brazilian state of Rondonia and the adjacent Amazonian region of Bolivia, show many signs of linguistic convergence, in addition to shared cultural traits, between a great many unrelated or only very distantly related language families. The Guaporé and Mamoré are two great rivers of the Southwestern Amazon region. It is one of the world's linguistically most diverse regions, with over 50 languages representing seven different stocks (Arawakan, Chapacuran, Macro-Ge [Jabutían], Nambikwaran, Panoan, Tacanan, and Tupían) and 11 genetic isolates. In spite of the fact that these languages diverge enormously at the lexical level, they do seem to share a considerable number of structural features. In addition, we find considerable morphological borrowing in the Brazilian part of the zone. However, the structural features do not clearly distinguish the region from surrounding zones, as far as can be established using fairly standard gross characteristics, such as head-marking, nominal number, gender, evidentiality, classifiers, verbal classification, asymmetrical morphology, subordination through nominalization, and switch reference.

10. Shared lexical borrowings from prestige languages

While the phenomenon of lexical borrowing as such needs no special mention here, common as it is, the phenomenon of shared borrowings is strking and frequent in the region. What it involves is the adoption, in parallel or in a chain, of the same set of words, generally from dominant prestige languages.

The example given in Adelaar with Muysken (2004: 500–501) involves the Quechua word for 'chicken', *ata-wal^spa*, which spread through 35 pre-Andine Amazonian languages, ranging from *óta* in Ticuna to *wa-tawah* in Amarakaeri. It is highly unlikely that all these languages borrowed this word from Quechua one by one. It is much more likely that word spread in a chain.

Less extreme cases involve the spread of Aymara numerals like *pataka* 'hundred', which were borrowed into Araucanian, and then into Allentiac, Tehuelche, and Gününa Yajich, in a process of chain borrowing. Similar cases involve other Quechua and Aymara numerals, generally above 'three' or 'four', which have spread into a series of pre-Andine languages, possibly also sometimes in processes of chain borrowing.

More complex are cases of Spanish borrowings for domestic animals such as *misi* 'cat' (< Sp. miche) and *khuchi* 'pig' (< Sp. *coche*) which occur in a similar or identical way in a host of South American indian languages. Were they first borrowed into Quechua and then spread from there to other languages, or were they borrowed in parallel in a number of different languages?

Even more complex are cases of borrowing of Spanish conjunctions. Stolz and Stolz (1996) show that the languages of Meso-America by and large have borrowed the same conjunctions and discourse markers all the time. It is not exactly clear how to interpret this finding. Is it chain borrowing? Are there common typological constraints involved? Are the sociolinguistic settings so similar that this is to be expected?

11. Conclusions

It is too early to draw any firm conclusions from the above survey. There is no doubt, however, that language contact, both in the period before the European invasions and afterwards, has been very intense. In some cases this has lead to complex patterns of restructuring and to languages which are difficult to classify genetically using standard techniques.

In the coming years, more solid knowledge about the make up and common linguistic properties of individual linguistic families will lead to further insights about those members of a family that do not fit into the pattern, and must have undergone extense restructuring from the outside.

References

- 1987 Aymarismos en el quechua de Puno. *Indiana* 11: 223–231.
- Adelaar, Willem F. H.
 - 2007 The Quechua impact in Amuesha, an Arawak language of the Peruvian Amazon. In: Aikhenvald and Dixon (eds.), *Grammars in Contact. A Cross-Linguistic Typology*, 290–312. Oxford University Press.

Adelaar, Willem F. H. with Pieter C. Muysken

- 2004 *The languages of the Andes.* Cambridge: Cambridge University Press.
- Aikhenvald, Alexandra Y.
 - 1996 Areal diffusion in Northwest Amazonia: The case of Tariana. *Anthropological Linguistics* 38: 73–116.
- Aikhenvald, Alexandra Y.
 - 1999a The Arawak language family. In: Dixon and Aikhenvald (eds.), *The Amazonian languages*, 65–106. Cambridge: Cambridge University Press.

Aikhenvald, Alexandra Y.

1999b Areal diffusion and language contact in the Içana-Vaupés basin. In: Dixon and Aikhenvald (eds.), *The Amazonian languages*, 385–416. Cambridge: Cambridge University Press.

Aikhenvald, Alexandra Y.

- 2001 Areal diffusion, genetic inheritance, and problems of subgrouping: A North Arawak case study. In: Dixon and Aikhenvald (eds.), *Areal diffusion and genetic inheritance. Problems in comparative linguistics*, 167–194. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y.

2002 Language contact in Amazonia. Oxford: Oxford University Press.

- Aikhenvald, Alexandra Y. and Robert M. W. Dixon (eds.)
- 2007 Grammars in Contact. A Cross-Linguistic Typology. Oxford University Press.
- Aikhenvald, Alexandra Y. and Diana Green
 - 1998 Palikur and the typology of classifiers. *Anthropological Linguistics* 40: 429–80.
- Alexander-Bakkerus, Astrid (ed.)
 - 2007 *Pedro de la Mata: Arte de la lengua cholona (1748).* Iberoamericana, Madrid u. Vervuert, Frankfurt.

Allaire, Louis

1999 Archeology of the Caribbean region. In: Salomon and Schwartz (eds.), *The Cambrige history of the native peoples of the Americas*. Volume 3. *South America*, 668–733. Cambridge: Cambridge University Press.

Allin, Trevor R.

1975. A grammar of Resígaro. Unpublished doctoral dissertation, University of St. Andrews.

Cabral, Anna Suelly

1995 Contact-induced language change in the western Amazon: The non-genetic origin of the Kokama language. PhD dissertation, Faculty of Arts and Sciences, University of Pittsburgh.

Campbell, Lyle

- 1995 The Quechumaran hypothesis and lessons for distant genetic comparison. *Diachronica* 12: 157–200.
- Campbell, Lyle
 - 1997 *American Indian languages: The historical linguistics of native America.* Oxford: Oxford University Press.

Carlin, Eithne B.

- 2007 Feeling the need: The borrowing of Cariban functional categories into Mawayana (Arawak). In: Aikhenvald and Dixon (eds.), Grammars in Contact. A Cross-Linguistic Typology, 313–332. Oxford University Press.
- Cerrón Palomino, Rodolfo
 - 1994 *Quechumara: estructuras parallas del quechua y del aimara*. Vol. 42: Cuadros de investigación. La Paz, Bolivia: Centro de Investigación y Promoción del Campesinado.
- Crevels, Mily and Hein van der Voort
 - 2007 The Guaporé-Mamoré region as a linguistic area. In: Pieter Muysken (ed.), From linguistic areas to areal linguistics, 151–179. Amsterdam: Benjamins.
- Crevels, Mily and Pieter Muysken
 - 2005 La influencia léxica de las lenguas amerindias en el español del Oriente Boliviano y Peruano. In: V. Noll and H. Symeonidis (eds.), *Sprache in Iberoamerika*. *Festschrift für Wolf Dietrich zum 65. Geburtstag.* 179–201. Hamburg: Buske.
- Dixon, Robert M. W. and Alexandra Y. Aikhenvald (eds.)

1999 The Amazonian languages. Cambridge: Cambridge University Press.

- Epps, Patience
 - 2007 The Vaupés melting pot: Tukanoan influence on Hup. In: Aikhenvald and Dixon (eds.), *Grammars in Contact. A Cross-Linguistic Typology*, 267–289. Oxford University Press.
- Girón Higuita, José Maria
 - 2007 Gramática del wansohot (puinave). Doctoral dissertation, Free University Amsterdam.

Gomez-Imbert, Elsa

1996 When animals become "rounded" and "feminine": Conceptual categories and linguistic classification in a miltilingual setting. In: Gumberz, John J. and Stephen C. Levinson (eds.), *Rethinking linguistic relativity*, 438–469. Cambridge: Cambridge University Press.

Greenberg, Joseph H.

1987 *Language in the Americas*. Stanford: Stanford University Press.

Hannß, Katja

2007 A grammar of Uchumataqu. Doctoral dissertation, Radboud University Nijmegen. Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie (eds.)

- 2005 World Atlas of Linguistic Structures. Oxford: Oxford University Press.
- Heggarty, Paul
 - 2005 Enigmas en el origen de las lenguas andinas: aplicando nuevas técnicas a las incógnitas por resolver. *Revista Andina* 40: 9–80.
- Hill, Jonathan D. and Fernando Santos-Granero (eds.)
 - 2002 *Comparative Arawakan Histories: Rethinking Language Family and Culture Area in Amazonia.* Champaign-Urbana, Ill.: University of Illinois Press.

Hoff, Berend J.

1995 Language contact, war, and Amerindian historical tradition: The special case of the Island Carib. In: Neil Whitehead (ed.), *Wolves from the sea*, 37–60. Leiden: KITLV Press.

Jackson, Jean

1983 *The Fish People: Linguistic exogamy and Tukanoan identity in Northwest Amazonia.* Cambridge: Cambridge University press.

Kaufman, Terrence

1990 Language history in South America: What we know and how to know more. In: Doris L. Payne (ed.), Amazonian linguistics: Studies in lowland South American languages, 13–74. Austin, Texas: University of Texas Press.

Launey, Michel

2003 *Awna parikwaki: Introduction à la langue Palikur de Guyane et de l'Amapá.* Paris: IRD Editions.

Mannheim, Bruce

- 1991 *The language of the Inka since the European invasion*. Austin: University of Texas Press.
- McMahon, April, Paul Heggarty, Robert McMahon and Natalia Slaska
 - 2005 Swadesh sublists and the benefits of borrowing: An Andean case study. In: April McMahon (ed.), *Quantitative Methods in Language Comparison. Transactions of the Philological Society* 103, 147–169. Oxford: Blackwell.

Melià, Batomeu

1992 *La lengua guaraní de Paraguay. Historia, sociedad, y literatura.* Madrid: Editorial Mapfre.

Melià, Batomeu

- 2003 *La lengua guaraní en el Paraguay colonial.* Asunción, Paraguay: Centro de estudios paraguayos "Antonio Guasch".
- Moore, Denny, Sidney Facundes and Nádia Pires
 - 1994 Nheengatu (Língua Geral Amazónica), its history, and the effects of language contact. In: Margaret Langdon (ed.), *Survey of California and other Indian languages. Report 8: Proceedings of the Meeting of the Society for the Study of the Indigenous Languages of the Americas and the Hokan-Penutian Workshop*, 93–118. University of California, Berkely.

Mous, Maarten

2003 *The making of a mixed language: The case of Ma'a/Mbugu*. Amsterdam: Benjamins.

Muysken, Pieter

1997 Callahuaya. In: Sarah G. Thomason (ed.), *Contact languages: A wider per*spective, 427–447. Amsterdam: Benjamins.

Muysken, Pieter

2009 *Kallawaya*. In: Mily Crevels and Pieter Muysken (eds.), *Lenguas de Bolivia*, vol. 1, Ámbito andino (pp. 147–168). La Paz: Plural editores.

Oblitas Poblete, Enrique

1968 *El idioma secreto de los incas.* Cochabamba: Los amigos del libro.

Orr, Carolyn J. and Robert E. Longacre

1968 Proto Quechumaran. *Language* 44: 528–555.

Payne, David L.

1985 The genetic classification of Resigaro. *International Journal of American Linguistics* 51: 222–231.

Rösing, Ina

- 1990 Introducción al mundo callahuaya: curación ritual para vencer penas y tristezas, Volume I: Introducción y documentación. Cochabamba: Los amigos del libro.
- Salomon, Frank and Stuart B. Schwartz
 - 1999 *The Cambridge history of the native peoples of the Americas.* Volume 3: *South America.* Cambridge: Cambridge University Press.
- Schwartz, Stuart B. and Frank Salomon
 - 1999 New peoples and new kinds of people: Adaptation, readjustment, and ethnogenesis in South American indigenous societies (colonial era). In: Salomon and Schwartz (eds.), *The Cambrige history of the native peoples of the Americas.* Volume 3: *South America*, 443–501. Cambridge: Cambridge University Press.
- Seifart, Frank
 - 2010 The Principle of Morphosyntactic Subsystem Integrity in language contact: a case study on morphological borrowing in Resígaro (Arawakan). Max Planck Institute for Evolutionary Anthropology, Leipzig. Manuscript.
- Sorensen, Arthur P.
 - 1967 Multilingualism in the Northwest Amazon. *American Anthropologist* 69: 670–684.
- Stolz, Thomas and Christel Stolz
 - 1996 Funktionswortentlehnung in Mesoamerika. Spanisch-amerindischer Sprachkontakt. *Sprachtypologie und Sprachuniversalienforschung* 49: 86–123.
- Taylor, Anne Christine
 - 1999 The western margins of Amazonia from the early sixteenth to the early nineteenth century. In Salomon and Schwartz (eds.), *The Cambridge history of the native peoples of the Americas.* Volume 3: *South America*, 188–256. Cambridge: Cambridge University Press.
- Taylor, Douglas
 - 1958 The place of Island Carib within the Arawakan family. *International Journal of American Linguistics* 24: 153–156.

Taylor, Douglas

1977 A note on Palikur and Northern Arawakan. *International Journal of American Linguistics* 43: 58–60.

Taylor, Gerald

- Apontamentos sobre o nheengatu falado no Rio Negro, Brasil Amerindia 10: 5–23.
- Vallejos, Rosa
 - 2005 Entre flexión y derivación: examinando algunos morfemas en Kokama-Cocamilla. *Memorias del congreso de idiomas indígenas de Latinoamérica-II.* 27–29 October. University of Texas at Austin.

Vallejos Yopán, Rosa

2010 A Grammar of Kokama-Kokamilla. Unpublished Doctoral Dissertation, University of Oregon.

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: 355–366.

Wise, Mary Ruth and Harald Green

1971 Compound propositions and surface structure sentences in Palikur. *Lingua* 26: 252–280.

Wloczkowska, Carolina

2006 Entlehnung von Bora ins Resigaro. Paper, University of Bochum.

The abbreviations used in this chapter are:

- ABL ablative
- AUM augmentative
- CIT citation
- DEM demonstrative
- DIR directional
- DT determiner
- EMP emphatic
- FIN purposive
- FOC focus
- IM imperative
- IN indefinite
- INT interrogative
- INTR intransitive
- LOC locative
- NEG negative
- PL plural
- Q Quechua
- REI reiteration
- TR transitive