



UvA-DARE (Digital Academic Repository)

Three languages from America in contact with Spanish

Bakker, D.

DOI

[10.1524/9783050057682.171](https://doi.org/10.1524/9783050057682.171)

Publication date

2012

Document Version

Author accepted manuscript

Published in

Amerindiana: neue Perspektiven auf die indigenen Sprachen Amerikas

[Link to publication](#)

Citation for published version (APA):

Bakker, D. (2012). Three languages from America in contact with Spanish. In J. Sakel, & T. Stolz (Eds.), *Amerindiana: neue Perspektiven auf die indigenen Sprachen Amerikas* (pp. 171-195). Akademie Verlag. <https://doi.org/10.1524/9783050057682.171>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

DIK BAKKER (LANCASTER)

Three languages from America in contact with Spanish

Abstract

Long before Europeans reached the American shores for the first time, and forced their cultures upon the indigenous population, including their languages, a great many other languages were spoken on that continent. These dated back to the original discoverers of America, who probably came from the West rather than the East. This article briefly sketches the arrival of those first Americans, and the languages that they introduced. Then three modern languages that developed out of these are discussed in some more detail, with respect to both sociolinguistic and grammatical aspects: Otomí, Quechua and Guaraní. Finally, an impression is given of what kind of linguistic changes the clash between these languages and Spanish, the official language in many of today's American countries, has brought about.

1. Introduction

This is the story of three languages from America: Otomí from Mexico, Quechua from Peru, and Guaraní from Paraguay. They are just a few of the more than 1,000 indigenous languages that are currently spoken in the Americas. All these languages are the direct descendants of the languages that were spoken by many millions of 'Indians' around the year 1492, when Columbus got sight of the island of Hispaniola, thinking that he was approaching India. So, it must have been the ancestors of the Otomí, the Quechua, and the Guaraní speakers who were the real discoverers of America. According to archeological and biological evidence, these ancestors came from the Siberian steppes, and entered Alaska via what we now call the Bering Strait. The two continents were once connected thanks to the much lower sea level during the last Ice Age. This situation existed maximally from 60,000 to 10,000 years ago. Arguably, those Asian peoples must have crossed before the latter date. Until recently, the oldest reliable evidence of human presence in the Americas was an excavation site near the town of Clovis, in New Mexico (USA), dated by archeologists around 13,500 B(efore) P(resent). However, more recent findings, near Tlapacoya (Mexico), and Monte Verde

(Chile), suggest an earlier date for immigration from the west: around 16,000 BP, or possibly 25,000 BP, or even as early as 50,000 BP.¹ Most of the wrongly named ‘Indians’ are assumed to descend from these first Americans. There could have been more than one group that crossed over to the New World at the earliest stages. But evidence, based both on DNA samples and cultural communalities seems to point towards genetic relatedness among all pre-Columbian inhabitants of America and the populations of East Asia and Siberia.² Those early invaders, now known as Paleoamericans, were hunters, and might have followed their favorite prey across the Bering Strait.³ Some groups of invaders settled, possibly after the larger species of animals that they were hunting became extinct. This happened especially in Meso-America, and the northern part of South-America, where the climate and the vegetation were more welcoming than the plains of North America, encouraging food gathering and the invention of agriculture. Other groups migrated further to the South, eventually spreading over the whole continent.

We must assume that the language, or languages, that the Paleoamericans spoke in the days of their migration were related to the languages of the population of East Asia, with whom these people were also genetically related. Therefore, we might suspect that there would still exist some traces of that relationship in the modern descendants of those languages on both sides of the Bering Strait. However, languages change from generation to generation. And when two dialects of the same language are no longer in contact with each other, they will change autonomously, and in different, and far from predictable directions. After several generations, they will be like two dialects of the same language. After around 1,000 years they will be so different from each other that linguists will consider them as two (related) languages, no longer as two dialects of one language. After more than 5,000 years without contact – equivalent to around 250 generations of speakers – most of the traces that would reveal an historic relationship between two languages, such as an overlap in their every day lexicon, sound correspondences, and certain morphological and word order patterns will have disappeared. Such languages may be as different as today’s Welsh, Lithuanian and Albanian, which have a Proto-Indo-European ancestor in common. After 10,000 years it will be virtually impossible to decide on the basis of language data alone whether they ever had a common ancestor. Further information, such as historical, archeological and genetic evidence is then called for. So, linguistics cannot be the only source, or even a reliable source to decide whether the early Asian immigrants in America came as one group, and shared

¹ Cf. Gruhn (1997:103). Recent studies of mitochondrial DNA and Y chromosome data suggest an initial entry into the Americas around 18,000 – 15,000 BP (cf. Jones 2004).

² It has also been suggested that the earliest inhabitants of the Americas actually came over sea, and were related to Southeast Asians or Polynesians rather than Northeast Asians (Erlandson 2002; Steele & Powell 2002). Others have even speculated about an early European immigration (Stanford & Bradley 2004), or one from the Middle East.

³ This scenario is supported by most authors on the subject, notably Campbell (1997) and Fortescue (1998).

one language, or that there was more than one such invasion. Not even under the (conservative) assumption that the Clovis settlers were the first to arrive, and did so relatively recently. What linguistics can do, however, is cluster the languages into families of related languages that testify of later splits, when the immigrants spread over the entire continent. As a result, we can distinguish today a large number of language families in the Americas. Estimates run from around 100 to 150. This includes some 30 isolates, i.e. languages that are so different from any other language that they are considered to be a language family of their own.⁴ In general, these families cannot be related to one common ancestor language on the basis of linguistic evidence alone, although languages of some families show some resemblance with languages of one or a few other families. In the academic tradition of the study of the languages of the Americas, two approaches to the relation problem may be distinguished, which originate with two scholars from the first half of the 20th century: Edward Sapir and Franz Boas. Very broadly speaking, the followers of Sapir may be associated with a trend to ‘lump’ languages and language families into larger groups. Their method, later refined by Joseph Greenberg (1957; 1987), relies in the first place on the mass comparison of the words that languages employ for everyday notions, such as bodyparts, natural phenomena, and common activities. The most extreme position here may be that of Greenberg’s pupil Merritt Ruhlen (1991), who posits an Amerindian phylum, or superfamily, under which the vast majority of the languages to be found in the Americas are subsumed, implying that they have one common ancestor. The followers of Boas, on the other hand are sometimes characterized as ‘splitters’. They apply the methodology of historical linguistics, which consists of the stepwise reconstruction of language relationships, which relies more on the comparison of language sound and structure. The latter approach is more conservative than the first one in the sense that there is less of a tendency to stipulate higher order groupings. A complication for any approach is that we cannot always be sure whether resemblances between languages have a genetic origin or are the result of the contact between the cultures who speak them. Such contact, if long and intensive enough, may lead to bilingualism and linguistic borrowing between these languages, blurring the initial differences between them, or even leading to one mixed language.

Thus, apart from the largely held assumption that the original population of the Americas came from East-Asia rather than from elsewhere, there is definitely no agreement among scholars with respect to the precise pattern of the immigration, and thus to the origin and the development of their languages. Two groups of languages, however are so different from all the others, as are their speakers, both culturally and genetically, that we must assume that there have been at least two later waves of immigrants, equally from Asia, and probably after the end of the Ice Age. It is therefore likely that these peoples had to travel partially by sea, hopping from island to island.

⁴ The lower number stems from the Ethnologue (Lewis 2009), the higher one from Campbell (1997:94).

The first of these families is generally known as Na-Dene. The people belonging to this group are mainly found in Northwest Canada, with pockets in the United States, along the Pacific coast, and in Arizona. Arguably, their ancestors arrived in America not later than 9,000 years ago. The around 40 languages belonging to the Na-Dene family that are still spoken today may be related to each other on the basis of the vocabulary elements that they share. As an illustration of this relationship, Table 1 shows the forms some of these languages use for five elements from the so-called Swadesh 100 list. This is a list of 100 English words, compiled by Morris Swadesh, all representing an everyday object or activity, and for which every language is supposed to have a word. Moreover, that word is typically not borrowed from another language, and relatively stable over time. Swadesh's wordlist was precisely meant to be an instrument for relating languages genetically.⁵

Language	'eye'	'tooth'	'hand'	'dog'	'path'
Carrier	<i>una</i>	<i>ugo</i>	<i>ula</i>	<i>like</i>	<i>ti</i>
Hupa	<i>na'</i>	<i>wo'</i>	<i>la'</i>	<i>liŋ</i>	<i>tin</i>
Apache	<i>da</i>	<i>wo</i>	<i>la</i>	<i>thini</i>	<i>ikin</i>
Navajo	<i>ana'</i>	<i>awo'</i>	<i>ala'</i>	<i>li</i>	<i>atin</i>

Table 1: Several words from the Swadesh list in some Na-Dene languages

As may be clear from this list, even for someone not specialized in the diachronic development of sound systems and lexicons, these correspondences in form can hardly be the result of sheer coincidence, but must have a historical basis. Furthermore, the Na-Dene languages have some very specific rules for the formation of their rather complex verbal clusters. This is shown in example (1) below, from Koyukon, a Na-Dene language from Alaska with around 100 speakers, according to a census held in 2000.⁶

- (1) Koyukon (Thompson 1996:355)
neel-h-ee-to-de-ts'eyh
 RECP-3-once-FUT-CLF-pinch
 'They will pinch each other once.'

The second exceptional family is Inuit, part of a larger group known as Eskimo-Aleut. This group consists of only 11 languages, spoken in the Arctic area, from Eastern Siberia via Alaska to Greenland. It is the only family present in America that still has more or less clear relatives in Asia. The ancestors of these people entered America probably around 7,000 BP. Culturally, they are very distinct from the Na-Dene and the Paleoamericans. Table 2 shows

⁵ See Swadesh (1955). The lexical information in this table, and some of the tables below stems from the online database of the ASJP project, which attempts to create a framework for large scale classification of the languages of the world, and takes a particularly stable subset of the Swadesh list as a point of departure (<http://email.eva.mpg.de/~wichmann/ASJPHomePage.htm>). See also Holman et al. (2008) and Bakker et al. (2009).

⁶ Definitions for the abbreviations used in the examples may be found towards the end of the article.

the word correspondences for three Inuit languages for the same meanings as used in Table 1 above. Not alone are the correspondences between these three languages striking, but even more so the complete difference from the forms in Table 1, suggesting a considerable time distance between these two groups, if related at all.

Language	'eye'	'tooth'	'hand'	'dog'	'path'
Central Yupik	<i>ii</i>	<i>kəxun</i>	<i>aixaq</i>	<i>qimuxta</i>	<i>tumyaxaq</i>
Inuktitut	<i>iyi</i>	<i>kigut</i>	<i>agak</i>	<i>qimiq</i>	<i>agquti</i>
West Greenlandic	<i>isi</i>	<i>kixut</i>	<i>assak</i>	<i>qimmig</i>	<i>aqqusiaq</i>

Table 2: Several words from the Swadesh list in some Inuit languages

There is linguistic and non-linguistic evidence that relates the Eskimo-Aleut languages to three other language families of Eastern Asia: Uralic, Yukaghir, and possibly also Chukotko-Kamchatkan. This would give support to the Asian origin of at least these people. There is much less evidence for the relationships between Asian languages of today and the other language groupings in America, which therefore must be much farther apart in terms of time.⁷ Typical for the Inuit languages, though not unique even in the Americas, are long and complex words, which may consist of several noun and verb stems at a time, and may contain the information expressed in a whole sentence in some other languages. A flavor of this may be found in example (2), from West-Greenlandic.

- (2) West-Greenlandic (Fortescue 1984:76)
ini-n-nuka-laar-niar-lunga-ana
 room-1POS-go.to-a.little-FUT-1SG-CONT
 'I am going to my room for a while.'

The three-wave theory as discussed above, although definitely not the only one, is the currently most accepted theory with respect to the early population of the Americas. Being there first, the immigrants of the first wave of around 15,000 BP had the area as it were for themselves. They could spread over the whole continent, opposed only by natural obstacles, over time covering the around 20,000 kilometers between Alaska and Tierra del Fuego.⁸ Those of the Na-Dene group, who came some 6,000 years later, found other people there already. They populated the free spaces they found on the wide plains of today's Canada and the USA. Finally, the Inuit, who came another 2,000 years later, did not travel much further than the Arctic zone.

Given their current geographic locations, the three American languages that we will discuss in more detail below most probably all go back to the first wave of immigrants. But there is no purely linguistic evidence for this whatsoever: the last contact between their supposed ancestor languages was simply too long ago for there to be any trace left

⁷ Cf. Fortescue (1998).

⁸ That figure for the distance is more or less as the crow flies. When following today's Pan-American Highway, it would be more like 30,000 kilometers, or three quarters of the circumference of the earth.

of a relationship in the lexicons or grammatical systems of the modern languages. As a result, Otomi, Quechua and Guarani differ as much from each other as they differ from West Greenlandic and Koyukon, and from any language in East Asia for that matter. In fact, many of the ‘first wave’ languages are so different that some authors assume that the differentiation between them must have begun already before the migration to America. A direct consequence of this would be that they must go back to different immigrant groups themselves.⁹ In short, unless we will understand much better than today how languages may change, it will probably not be on the basis of linguistic evidence that we will one day have more detailed information about the earliest population dynamics in the Americas. If so, then it will have to come from biology (genetics), and possibly also from further archeological finds. However, there does not exist a one-to-one relationship between genetic features of a group of people on the one hand and the language they happen to speak on the other hand. The drift of genes within and across human populations and the drift of languages among them are in principle independent from each other, and correspond to different sociological factors. It is therefore not very likely that we will ever know whether it was just one language that gave rise to America’s rich linguistic patchwork – which would justify the postulation of one big Amerind language family in the sense of Ruhlen (1991) - or that it sprung from several languages, be they or their speakers related or not. This does not mean, however, that nothing can be said about the relationships among the close to 1000 languages that do not belong to either the Na-Dene or the Inuit families. Systematic comparison has led to the establishment in the Americas of up to 150 language families and isolates that have been accepted more or less widely among experts. The three languages on which we will focus in the rest of this article each belong to one of these. For any higher orderings and groupings no firm linguistic support has been found so far.

As for our languages, Otomi is part of the Otomanguean family. With a total of around 175 distinct languages, and many more dialects, this is arguably the largest language family of the New World. It once ranged over the whole of Meso-America, although extant languages are now found in Mexico only. A potential candidate for the homeland of this family – the area where Proto-Otomanguean might have been spoken – is the Tehuacán valley in the Mexican state of Puebla. It is in this region that early forms of agriculture were introduced. Traces of this are found in the shared vocabulary of some of the Otomanguean languages, above all in the names for domesticated plants from that area. Specific linguistic traits of the Otomanguean family are vowel nasaliza-

⁹ Campbell (1997:96f) systematically discusses several logical possibilities, from just one group of Paleoamericans migrating to up to 33 consecutive groups, and either assuming a spread scenario, when part of the population stays behind in Asia, or the complete emigration of a whole population, not leaving any trace. The latter scenario might get support from the observation that to date there is three to four times as much linguistic diversity in the New World as in the Old World. And of this diversity, by far most is found in the West, which might be indicative of a greater time depth, and therefore of a West-East direction of the migration.

tion, a rather simple CV syllable structure, and the presence of up to five tones, rare in the Americas. Otomi itself belongs to the Otopame subgroup. Other subgroups of this family are Chinantec, Popolocan, Mixtecan and Zapotecan.

The second language that I have selected is Quechua. This name is in fact the label for a group of some 45 language varieties that is sometimes seen as a continuum of dialects, since there is a considerable amount of similarity between all the varieties, especially between neighboring ones. Taken together, the Quechua continuum has some eight million speakers, making it the largest extant indigenous language of the Americas. The languages belonging to this family are found over the whole Andes area, from Southern Colombia to Northern Chile. Possibly, Quechua is related to Aymara, another Andean language.¹⁰ The Quechua family is often split up in two subgroups: Central Quechua (or Quechua I), in Central Peru, and Peripheral Quechua (or Quechua II), in the other areas. Generally, these languages are characterized by SOV constituent order, both glottalized and aspirated consonants, and a complex agglutinative morphological system.

Our third language, Guarani belongs to the Tupi family, together with around 75 other languages. These are found in Paraguay, Bolivia, and spread over the Brazilian Amazon area, up to the Guyanas. The greatest variety among these languages is found in Rondônia, in Western Brazil. This is often taken as an indication that this is the homeland of the family, in this case of the Proto-Tupi speakers. These languages are typically postpositional, and have many prefixes on verbs and nouns. They often have ergative alignment. The predecessors of the speakers of Guarani, the largest language of the family in terms of its current number of speakers, must have moved away from this area, first around a thousand kilometers to the southeast, and later even much further to the northeast. An interesting member of this family is Nhegatu, spoken today by 8000 speakers in northwest Brazil, near the Colombian and Venezuelan borders. In the 17th and 18th centuries this language served as a lingua franca in large parts of Brazil, both among native Americans and between them and the Portuguese invaders.

Thus, our three languages share a number of features with languages in the areas surrounding them, testifying of family relationships. But even though it is not impossible that they all have Palaeoamerican for their common ancestor, there is virtually no trace of that left in the modern languages. Just a few elements from the Swadesh list may be used to illustrate this. Table 3 below gives an impression of the formal correspondences between some everyday words for two languages from the same family as opposed to those of the two other families.

Family	Language	'louse'	'tree'	'fish'	'water'	'nose'
Otomanguean	Otomi	<i>t'o</i>	<i>za'a</i>	<i>xwã</i>	<i>dehe</i>	<i>siŋu</i>

¹⁰ Campbell (1995) gives arguments for this relationship. Mannheim (1991), on the other hand, claims that the communalities between the two languages are due to intensive language contact, or more in general to areal tendencies.

	Mazahua	<i>t'oni</i>	<i>za</i>	<i>xmō'ō</i>	<i>ndexe</i>	<i>siŋu</i>
Quechuan	Quechua (Ancash)	<i>uhe</i>	<i>munti</i>	<i>ʃalwa</i>	<i>yaku</i>	<i>seŋgo</i>
	Quechua (Pastaza)	<i>usa</i>	<i>kaspi</i>	<i>ʃalywa</i>	<i>yeku</i>	<i>siŋa</i>
Tupi	Guarani	<i>ky</i>	<i>mata</i>	<i>pira</i>	<i>y</i>	<i>iĩ</i>
	Juruna	<i>kapa</i>	<i>ka</i>	<i>piča</i>	<i>iya</i>	<i>I'ã</i>

Table 3: Several words from the Swadesh list in some Amerindian language families

As the data in the table suggests, there is not much reason to assume that the words for these five basic meanings have a common origin across the three language families. Only the words for ‘nose’ might have a common source for the Otomanguean and the Quechuan languages.

Other language families from the Americas are Salishian, Siouan and Uto-Aztek in North America, Mayan in Meso-America, and Arawak in South-America, to mention just a few of the larger, more well-known groupings.

Most of the pre-Columbian history of the vast numbers of cultures that must have thrived in America before the contact with Europeans is unknown to us. It lives only in archeological finds, and in the oral history and the languages of the cultures that are still extant. Traces of large scale empires with city-like establishments are found only in today’s Peru (e.g. the Inca empire), and in Mexico, Guatemala and Honduras (e.g. Olmec, Zapotec, Maya, and Aztec). Only the Meso-American cultures developed a more or less full-fledged writing system. The Mayan system is possibly the earliest, and certainly the most elaborate of these. It is also the only one for which a reasonable amount of text has survived, despite the fact that the Spaniards have tried to burn all the codices they found. Since the Maya script – at least partially a syllabic system - has been largely deciphered, quite a few details are now known about the pre-contact stages of this culture, which flourished between 250 and 900 AD.¹¹

I will concentrate now on the three languages that we started out with: Otomi, Quechua and Guarani. In section two I will give an impression of their current position in the respective societies. Section three presents a short sketch of the grammars. And in section four we will see what influence contact with Spanish and widespread bilingualism has had so far on each of them.

2. The sociolinguistic situation

The figure of 1,000 mentioned above for the number of languages that are spoken anywhere in the America’s today may suggest a healthy linguistic situation. However, since the arrival of the European conquerors, many of the indigenous languages are in decline. The number of languages at the time of contact is estimated to have been 50% higher, so languages have disappeared at an average rate of around 100 per century, or

¹¹ Cf. Coe & Van Stone (2001).

one per year. More worrying however is the fact that most of these languages have become extinct only in the last century. Especially after 1950 contact between the indigenous cultures and the 'official' culture has intensified, in terms of education, the media, and the infrastructure in general. As a result, most of the speakers of the indigenous languages are bilingual, in English, Spanish, Portuguese, French, Dutch, or some Creole language. People are leaving their native villages at a large scale, to seek employment in the towns where the colonial language is omnipresent. In a multilingual situation, with one language completely dominant everywhere except at home, the native language may no longer be passed on to the next generation. Therefore, without a language community and a living culture to support it, languages may go lost within a few generations, even if the current number of speakers would be an impressive 100,000 or more. In fact, many of the languages that are still extant today are spoken by a small number of people, the youngest of them often being over 50 years of age. It has been suggested that, at the end of the 21st century, most of the native languages of the Americas will not be spoken anymore, and that only the very large ones have a chance to survive the next 4 generations.¹²

I will discuss the sociolinguistic situation with respect to the three languages selected here briefly in the following subsections. A more detailed description may be found in Bakker et al (2008), and the references given there.

2.1 Otomi

Otomi is the native language of some 250,000 people on the highlands around Mexico City. The data we will discuss here stem from the northwestern dialect, with around 33,000 speakers. Together with Mazahua and several other languages, Otomi belongs to the Otopame branch of the Otomanguean family. With around 175 extant languages, this is the largest family in the Americas in terms of linguistic diversity, and it is among the ten largest families in the world.

In the pre-colonial era the Otomis reigned over the Mexican highlands for a long time, but they were subjugated from around 1000 AD onwards by the Aztecs, speakers of the Nahuatl language, who were in control of the area when the Spanish arrived. The word 'Otomi' is probably of Nahuatl origin, meaning something like 'bird hunters'. The Otomis prefer to call themselves *Ñähñu*, i.e. 'he who speaks well', and their language *Hñähñu*. Being marginalized by the Aztecs, a sizeable number of Otomis joined the Spaniards against their enemies, and they regained part of their territory as a result after these had been defeated. Given their relatively central position in the early colonial society, they were a clear target for conversion to Catholicism. From the middle of the

¹² Mithun (1999:2), when discussing Navajo, with 100,000 speakers the most widely spoken and well-studied native language of North America, assumes that all languages of North America will be extinct by the end of this century.

16th century onwards, a number of Spanish missionaries have studied the language, and published dictionaries, grammars, spelling systems, and religious texts in Otomi.¹³ After the Independence of Mexico in 1813 the Otomis lost the special status they enjoyed during the colonial era, and so did their language. Otomi was no longer written by the civil authorities, only by a handful of scholars, and a process of language shift started.

The Mexican Revolution (1911–1917) did not lead to social change for the indigenous population. On the contrary, their degradation continued. Today, the Otomis belong to the lowest social levels of the Mexican society, as is the case for many other indigenous groups. They dwell in the most remote and less fertile places on the highlands, living from agriculture of subsistence. Many have migrated to the bigger towns, such as Mexico City, Guadalajara and Monterrey.

Several attempts have been made in Mexico to integrate the indigenous communities in the national processes by means of a bilingual education scheme, so far without much success. Most Otomis are illiterate in their first language. Today, Otomi is only spoken within informal domains such as the family, while Spanish has become the language in all other domains, leading to a high degree of bilingualism. In the near future, it may be expected that increasing globalization and stigmatization will push Otomi, and many of the more than 100 indigenous languages and dialects of Mexico closer to extinction.

2.2 Quechua

With over eight million native speakers, the Quechua family is the largest of the indigenous American families in terms of language users. Some 45 varieties, each typically called Quechua followed by a regional indication, may be found over a large stretch of the Andes, from Ecuador via Peru to Bolivia, with pockets in Colombia, Argentina and Chile. The largest variety is spoken around Cusco, in Peru, with around 1,500,000 speakers, a quarter of whom are monolingual. Many varieties of Quechua have less than 5,000 speakers, and some are on the brink of extinction. As stated above, there is a lot of similarity among the Quechua languages, especially in the lexicon. Neighboring versions are often mutually understandable, but those at greater distances from each other are not. The variety that I will look at in more detail below, Chimborazo Quechua, is from Ecuador, and has a total of around one million speakers, many of them monolingual. The Ecuadorians call their language Quichua rather than Quechua.

When the Spanish invaders arrived in the Northern Andes, around 1530, the so-called Incas, ancestors of today's Quechua people, had just completed the expansion of their empire. They started out from their basis in Cusco around 1450, and subjected the surrounding civilizations to their rule. In less than 80 years they established the largest empire of pre-Columbian America, which covered the Andean area from today's Ecu-

¹³ Cf. Urbano (1990[1605]).

dor to Bolivia. In the process, they exported their language, Quechua, much in the way Latin spread over southwest Europe, via soldiers, officials, tradesmen and immigrants. In their turn however, the Incas were rapidly subjected by the Spanish, mainly as a result of internal feuds, the superior European weapons, and imported diseases.

Although a number of very different languages must have been spoken within the boundaries of the Inca Empire, most of them had disappeared 100 years after the Spanish invasion. This was partially due to the role of Quechua, which had become a *lingua franca* in the area. The colonial powers built up on this by ordering that, for the purpose of spreading the Catholic faith, texts should be translated into a standardized variety of Quechua.¹⁴ Dictionaries and grammars were produced precisely for this purpose. However, when it became apparent that the common language had grown into a unifying factor for the suppressed native peoples, the Spanish abandoned their multilingual policy. Apart from this, the indigenous population saw its place in society being reduced as a result of the 18th century hacienda system under which large amounts of land were expropriated by the colonial newcomers. As in the case of Otomi discussed above, the position of Quechua and other indigenous languages was marginalized even more after the liberation of the Spanish Americas, and the birth of the modern republics around 1825. Being the creation of the descendants of the colonizers rather than the original population, these states introduced Spanish as the only official language. Only after 1980 was Quechua introduced in primary education. Arguably, this is too late to stop further hispanicization of the indigenous population, resulting from the large scale depopulation of the countryside and massive migration to the monolingual cities. For some Quechua communities in Ecuador, who live literally halfway between both cultures, this has had an enormous impact on their language. The permanent bilingualism of those who divide their time between family life in a traditional village and professional life in the modern city, has led to the development of *Media Lengua* – or half way language – a mix of Quechua grammar and Spanish lexicon.¹⁵

2.3 Guarani

Guarani is part of the Tupi family, with around 75 languages one of the most widespread families of the Americas. Together with three other languages – Aché, Kaiwá and Xetá – it forms a subgroup of the Tupi-Guarani branch. Paraguayan Guarani, or Avanye'e, is the major variety, boasting 4.6 million speakers. Four more varieties – Ava, Mbyá, Eastern and Western Guarani – which have only between 7,000 and 35,000 speakers each, are found in Paraguay, Bolivia and Brazil.

Although the area of modern day Paraguay was visited by Spanish explorers as early as 1516, these did not find it interesting enough to settle in any numbers. The town of

¹⁴ Cf. Adelaar (2004:183).

¹⁵ Cf. Muysken (1994).

Asunción counted one European on every ten Indians around 1600. Polygamy with indigenous women became the norm for the Spanish men, and Guarani remained the everyday language for all – Indians, mestizos and Spaniards alike – despite the fact that Spanish was the official language. Also the evangelization process was rather different from elsewhere in the Americas. In this area, the Jesuits developed so called *reducciones* (reductions), economically self-sustaining village-like communities, where thousands of Indians lived under the supervision of priests. The obvious goal was evangelization, however without affecting the local culture, including the language. A specific kind of standardized Guarani developed here, until the Jesuits were expelled, in 1768, and the reductions abolished and destroyed. Further developments of the language took place in the cities, and foremost in Asunción. It is as yet unclear what influence the Guarani of the reductions has had on the urban varieties.

After the independence, in 1814, Paraguay developed rather in isolation from the surrounding republics, which were much more open for external – European, North American – influences. Different Paraguayan regimes held different attitudes towards Guarani. Some considered it a symbol of national unity. This is especially prominent in periods of war with the surrounding countries. Other regimes saw the language as a sign of backwardness. But even the periods of extreme suppression have not been able to displace the language. Quite the opposite, Guarani was given the status of national language in 1967, next to Spanish. Over time, in the larger towns a variety of Guarani developed with many traces of Spanish. An extreme version of this, called Jopara, is a kind of mixed language, or rather a speech attitude of bilinguals who, in certain speech situations make constant code switches between Guarani and Spanish.¹⁶ In the smaller, rural communities a more ‘pure’ variety of Guarani may be found. A recent census in Paraguay established that 59% of the population of Paraguay is bilingual in Guarani and Spanish, while 27% is monolingual in Guarani, and a mere 7% monolingual in Spanish.

3. Three Amerindian languages, three different grammars

In this section I will have a look at the three languages from a typological perspective. Only those aspects will be illustrated that will be relevant for section 4, where I will explore the influence that contact with Spanish has had on them. No attempt is made to give a representative sketch of the languages in any way. For that, descriptive grammars should be consulted, or more specialized studies on specific phenomena, and above all of course native speakers.¹⁷

¹⁶ Cf. Dietrich (2010).

¹⁷ Cf. Hekking & Andrés de Jesús (1984) for Otomí; Cole (1982) for Imbabura Quechua; and Gregores & Suarez (1967) for Guarani.

3.1 Otomi

Like many other Otomanguean languages, Otomi is a tone language, which distinguishes between a high, a low and a rising tone. The basic word order pattern of classical Otomi is VOS. Only when in focus, the subject may be fronted. Compare examples (3a) and (3b) below.

- (3a) *mi=ñä-wi* *ár to* *ar* *Xuwa*
 PAST.3=speak-with his mother-in-law the Juan
 ‘Juan spoke with his mother-in-law.’
- (3b) *ar Xuwa mi=ñä-wi* *ár to*
 the Juan PAST.3=speak-with his mother-in-law
 ‘It was Juan who spoke with his mother-in-law.’

The morphological type is fusional, with tense+person markers affixed to the verb, either as a prefix or a suffix. Otomi has articles, both definite and indefinite ones; (3a,b) above and (4b) below give examples of these. As typical for a V-first language, Otomi has prepositions rather than postpositions. However, there are only a few of these, among them *dige* ‘with respect to’, which has a very general meaning and wide application, as exemplified in (4a,b). Most relations of noun phrases within a clause, typically coded with a preposition in Spanish, remain unexpressed.

- (4a) *di ñä-he* *dige ma* *bñi-he* *Maxei*
 we speak-1PL.EXCL about our trip-1PL.EXCL Querétaro
 ‘We speak about our trip to Querétaro.’
- (4b) *'nar jä'* *pwede da du* *dige-r* *t'ete*
 IND.SG person may FUT.3 die through-DEF.SG witchcraft
 ‘A human being may die through witchcraft’

Although Otomi has several coordinators and subordinators at its disposal, the prevailing form for both coordination and subordination is asyndetic juxtaposition at the clause level.

Apart from lexical elements which specialize as a verb or a noun, Otomi has very few words that have only an adjectival application. To my knowledge, this is the complete set: *t'olo*, *t'uku* ‘small’; *hogi* ‘good’; *bente* ‘poor, unfortunate’; and *bet'o* ‘older’. For all other nominal modification, either nouns or verbs are used. (5) below gives an example of the former. That *goda* ‘blind’ is a noun rather than an adjective is shown by the definite article preceding it. In (6), a verb is used as a nominal modifier, witness the past participle prefix.

- (5) *ar tsat'yo ar goda*
 DEF.SG dog DEF.SG blind
 ‘The blind dog.’

- (6) *ar n-t'axu 'bitu*
 DEF.SG PP-be.white sheet
 'The white sheet.'

3.2 Quechua

Quechua is a typical V-final language, with SOV main clause order (7), and postpositions, which are suffixed to the nominal (8).¹⁸

- (7) *kanun-ka ilkimas rundin tuka-ria-n*
 today-TOP only pan.flute play-DUR-PRES.3
 'Today they play only the pan flute.'
- (8) *kitu-man ri-rka-ni chay-pi trabaja-ngapaj*
 Quito-to go-PAST-1 that-in work-SBJNC
 'I went to Quito to work there.'

Morphologically it is an agglutinative language, with a large number of suffixes, as can be seen in the examples above, and in (9) below.

- (9) *miku-chi-wa-shka-rka-ngui*
 eat-CAUS-1-PERF-PAST-2
 'You had fed me.'

Quechua has a word class specific for verbal use (verb), and one for words that can function both as a nominal head (noun) and as a nominal modifier (adjective). An example of the latter is given in (10). In (10a) *hatun* is used as a noun, while in (10b) it functions as an adjective.

- (10a) *rika-sha-ka: hatun-ta* (10b) *chay hatun runa*
 see-PAST-1SG big-ACC DEM big man
 'I saw a big one.' 'That big man.'

The flexibility of this word class is demonstrated in (11), where a typical nominal, *duktur* 'doctor' is modified by an adverb.

- (11) *Chay warmi maymi duktur-mi*
 DEM woman very doctor-FOC
 'That woman is a real doctor.'

Quechua has no articles.

¹⁸ For that reason, some authors analyze them as case markers (cf. Cerrón-Palomino 1987). Given enough time, this is likely what at least some of them might develop into anyway.

3.3 Guarani

Guarani is a SVO language, as shown in example (12).

- (12) *tuvisava o-gweraha-vai orereta*
 president 3-manage-badly country
 ‘The president manages our country badly.’

Unlike most SVO languages, it has postpositions.¹⁹ (13) gives an example.

- (13) *a-jahe'o pochy-rehe*
 1SG-cry anger-by
 ‘I cry from anger.’

Guarani is a fusional language, and is mainly prefixing, as can be gathered from the examples above and below. In terms of its parts of speech system, it is very flexible, even more so than Quechua. Not only are there a large number of words that have nominal as well as adjectival use, as demonstrated in (14a,b). Many others are ambivalent between noun and verb, as shown in (15a,b).

- | | |
|--|---|
| <p>(14a) <i>ko karai tuja</i>
 that man old
 ‘That old man.’</p> | <p>(14b) <i>che tuva tuja</i>
 I father old
 ‘My father’s old age.’</p> |
| <p>(15a) <i>a-jahe'o pochy-rehe</i>
 1SG-cry anger-by
 ‘I cry from anger.’</p> | <p>(15b) <i>che che-pochy</i>
 I I-anger
 ‘I am angry.’</p> |

There are no articles in Guarani.

4. What has contact with Spanish done to them?

As may be clear from the short historical sketches in section 2, all three languages that concern us here have been in close contact with Spanish. Otomi enjoyed some special status until the beginning of the 19th century, but has lost that completely since. Until the middle of the 20th century, it was spoken mainly in isolation. However today, contact of the Otomi speakers with the outside world is very intensive, thanks to recent developments in society. As a result, most of them are bilingual, and use Spanish outside the realm of the family. Education beyond the primary level is in Spanish, and so are the media in Mexico. These aspects hold as well for Quechua, be it that this language never had any official status in the Ecuadorian and Peruvian societies, and con-

¹⁹ In the WALS database (<http://wals.info/>, and see Haspelmath et al. 2005), only 20% of the SVO languages are postpositional.

tact with the Spanish speaking world started much earlier on, and has been more intensive since. Guarani has had a higher status than both of the other two languages throughout the last two centuries, culminating in its recognition as the ‘other’ national language of Paraguay. However, it has always been in close contact with Spanish, which in its turn has always been the language of the upper echelons of society, and the ‘modern world’, especially in the cities, but not only there.

Given the different position and roles in the respective societies, and in communicative situations of Spanish on the one hand and Otomi, Quechua and Guarani on the other hand, we may expect that there has been mutual influence, and that this influence will have mainly gone from ‘dominating’ Spanish to the ‘dominated’ indigenous languages. Thus, we expect to find traces of Spanish in all three of them, lexical as well as grammatical ones. In order to test this assumption, three sets of spoken language were collected in situ by Bakker et al. (2008). In each case, a number of native speakers were involved stemming from different categories in terms of age, gender, educational level, and profession. In order to test whether borrowing is dependent on certain aspects of the grammars of the source and target languages of the process, data were collected from two different, non-contiguous dialects of each of the languages. These data collections were transcribed, and entered into a database. Subsequently, all Spanish elements were marked, and a computer program was developed to explore the database. Table 4 gives a global overview.

	Otomi	Quechua	Guarani
Number of respondents	59	25	38
Number of dialects	2	2	2
Corpus size (tokens)	110,541	79,469	57,828

Table 4: Data collected for the three languages

What I will be interested in here is whether there are qualitative and quantitative differences between the three languages in terms of what they have borrowed from Spanish. If there would be significant differences, the next step would be to try and explain these on the basis of the contact history, as well as the typological characteristics of the three languages in their relation to Spanish. Table 5 shows the language features I have selected for this discussion.

	Spanish	Otomi	Quechua	Guarani
Word Order	SVO	VOS/SVO	SOV	SVO
Adposition Type	Prepositions	Prepositions	Postpositions	Postpositions
Articles	DEF/IND	DEF/IND	-/-	-/-
Parts of Speech	V/N/A	V/N/-	V/N+A	V+N+A

Table 5: Relevant features of the four languages

Now let us first have a look at the overall borrowing figures in terms of words. Of the total number of tokens – i.e. individual word occurrences in the text – in the respective subcorpora, Quechua has the highest number of borrowings with 19.0%, or almost one in every five words of running conversation. Guarani comes second with 17.4%. And Otomi ends up last with 14.1%, slightly less than one in seven words. Although these percentages are not dramatically different, they are nevertheless statistically significant. The sociolinguistic facts seem to give an explanation for this. Quechua has been in contact with Spanish virtually from the earliest days of the invasion, with a relatively high proportion of bilinguals as a result. Guarani has been spoken in an environment with a much lower amount of Spanish speakers, and has had spells of a relatively high status, when purism prevailed. For Otomi, intensive contact started only half a century ago, with until recently much lower levels of bilingualism.

When we make a breakdown in percentages of the major parts of speech – Verb, Noun and Adjective – that the borrowed words have in Spanish, we find the following. I will give the relative percentages of the tokens rather than the absolute ones, since these three word categories can be seen as in competition for the same semantic space with each other. This is hardly the case for the more grammatical elements that are borrowed, such as prepositions and articles, which will be discussed separately below. Therefore, comparing the mutual contributions that these three major categories make in relation to just each other seems to be the more interesting angle. The following percentages were found in the corpus.

	Quechua	Guarani	Otomi
Noun	68%	59%	86%
Verb	22%	29%	10%
Adjective	11%	12%	4%
Total	100%	100%	100%

Table 6: Relative percentages of major parts of speech borrowed

In all three cases nouns are the most often borrowed category. This does not come as a surprise: nouns are among the first items to be borrowed. They belong to the largest of all word classes, are relatively easy to identify in a context, are often stressed in speech, and are morphologically not very complex in most languages. Their meanings are often concrete, and they may be adopted by a language community along with hitherto unknown objects or concepts. These may be the reasons why they are by far the largest contingent for Otomi, the language with the least intensive exposure to Spanish. Quechua and Guarani have considerably lower percentages of nouns, and have borrowed considerably more words from the other two major categories. The fact that Quechua has a higher percentage of nouns than Guarani may be caused by its even longer exposure to contact, during which the borrowing of elements from this largest and most open class has simply continued, while that of the more restricted classes slowed down. An in depth study of the

precise meanings of the words borrowed could throw more light on this aspect.²⁰ Interestingly, Guarani seems to be borrowing relatively more verbs as opposed to adjectives: the proportions are 2.4:1 for Guarani and 2:1 for Quechua, respectively. Possibly, this could be related to the fact that Guarani is a SVO language, just like Spanish. This would then make a verb easier to access for a speaker of Guarani than for a speaker of Quechua with its SOV perspective, and given that verbs are much more formally complex than nouns in the languages studied here. Note also that Guarani has the most flexible of parts of speech systems of all three languages. This may make identification by meaning rather than function in Spanish easiest for Guarani natives. The nature of the part of speech system may also help explain why for Otomi both the overall figure of adjectives borrowed, and its proportion to verbs is so low. This language does not have an open class of adjectives, and typically resorts to a noun or verb to modify nominals.

Let us now move on to the minor parts of speech. In this case we will take the absolute percentages, since some of these seem to be rather remarkable, as Table 7 shows.

	Quechua	Guarani	Otomi
Preposition	0.5%	0.5%	20.2%
Definite article	0.0%	19.4%	0.0%
Conjunctions	7.7%	7.6%	12.4%

Table 7: Absolute token percentages of some minor parts of speech borrowed

The differences of the figures between the languages are considerable, especially for the first two categories, for which the percentage is very high for one language, and (close to) zero for the other two. This begs for an explanation.

Starting with the prepositions, we see that more than one in five of the borrowed tokens in Otomi belong to this category. In fact, it is the second largest loan category after the nouns, which cater for 40.7% overall. Our tentative explanation is that Otomi is a prepositional language, be it that the number of prepositions is rather small. This means, however, that native speakers of this language will have little difficulty with the identification of prepositions in Spanish, which has a very large number of them, several of them being highly frequent.²¹ It also means that Otomi has a syntactic slot to insert them in. Since most of the relations at the phrase level are not expressed in classical Otomi, but have to be inferred from the context, insertion of a preposition makes a relation more explicit. Guarani and Quechua, on the other hand, borrow only very few

²⁰ Haspelmath & Tadmor (2009) present a database with information on loanwords for around 1600 lexical entries, which could give support to such an exercise. Of the three languages discussed here, only Otomi is represented in this collection (cf. Hekking & Bakker 2009).

²¹ In the 100 million word Spanish corpus of Brigham Young University (Davies 2002), around 13.5% of the tokens are prepositions, with the top three *de* 'of' (6.8%), *a* 'to' (2.4%), and *en* 'in' (2.2%). Among the prepositions borrowed by Otomí, *con* 'with', *para* 'for', and *de* 'of' are the most frequent ones found in the corpus of Bakker et al. (2008).

prepositions from Spanish. The ones that I found in the corpus are typically part of a fixed expression that is borrowed as a whole, as in example (16) from Quechua.

- (16) *ñuka-ka por-gusto ri-ni uña-ngapa*
 1SG-TOP by-pleasure go-1SG drink-PURP
 ‘I like to go to drink’

Other examples are *a lo menos* ‘at least’, *de repente* ‘suddenly’, and *por ejemplo* ‘for example’ (in the Quechua corpus), and *a lo mejor* ‘perhaps’, *de lado* ‘on the side’, and *en cambio* ‘in return’ (in Guarani). Both languages have a sizeable set of postpositions, which are suffixed to the noun or nominal expression that they are relating to the rest of the clause. So, there is no natural syntactic slot for prepositions, nor does there seem to be much functional need for them.

The borrowing of Spanish definite articles shows the same, rather dramatic differences. Again, one language, in this case Guarani, borrows them at a very large scale: around one in five loans in the corpus is an article. Quechua and Otomi, on the other hand, borrow none at all. Otomi has both definite and indefinite articles, and uses them frequently. They have more or less the same function as the ones in Spanish, so there does not seem to be much functional gain in borrowing them. But both Guarani and Quechua lack articles. Still, the former borrows them at a large scale, while for the latter none were found. A possible explanation is the following. When we look at the actual use of the borrowed definite articles in Guarani, then it is not so much definiteness versus indefiniteness that they seem to code. They either function as demonstratives in a noun phrase headed by a noun, coding [+remote, –visible], an extension of the function of the native demonstratives, which only code a tripartite spatial dimension. Or they appear as independent constituents, as anaphoric elements, in positions that would be left empty in the classical language, which is ‘pro drop’. The overall effect is mainly pragmatic, in the sense of further distinguishing the topical from the focal elements of the utterance. There seem to be no native forms in Guarani that have the function of marking topics. Quechua, on the other hand, has a rather explicit system for the topic – focus distinction. The suffix *-ka* marks the topic and the suffix *-mi* the focal element of a sentence. Compare (17a) and (17b).

- (17a) *ñuka tayta-ka alpa-ta-mi yapu-n*
 1SG father-TOP land-ACC-FOC plow-3
 ‘It is the field that my father plows.’
- (17b) *ñuka tayta-mi alpa-ta-ka yapu-n*
 1SG father-FOC land-ACC-TOP plow-3
 ‘It is my father who plows the field.’

Since definiteness and topicality are closely related, there seems to be less functional motivation within the Quechua system to import definite articles.

The third minor category that we will have a look at are conjunctions, i.e. coordinators and subordinators. Here, according to Table 7 the three languages seem to be in agreement: all borrow considerable numbers of conjunctions, be it that Otomi borrows even more of these than the other two languages. Let us try to find an explanation for this. Both classical Otomi and classical Guarani do have a small set of coordinators with a more or less general meaning. Some are used frequently, others less so. In the Otomi corpus there are 1708 occurrences of *ne* ‘and’, 153 of *wa* ‘or’, and 93 of *pe* ‘but’. Nevertheless, almost all speakers use the Spanish loans *y* ‘and’ (200 instances), *pero* ‘but’ (188), and *o* ‘or’ (188). The Guarani corpus counts no less than 3036 instances of native *ha* ‘and’, but only 32 of *tera* ‘or’, and there is no equivalent for ‘but’. On the other hand there are 225 instances of Spanish *pero* ‘but’, 32 of *o* ‘or’, but only 6 of *y* ‘and’. Finally, Quechua speakers make very frequent use of the native suffix *-pash* ‘and’, which may be attached to the final element of a noun phrase or a clause. There are no forms for ‘or’ and ‘but’. However, we find 301 instances of Spanish *pero*, 217 of *y*, and 124 of *o* used by the vast majority of the speakers in the corpus. A further point, already observed in the short grammatical sketches of section 3, is that in all three cases, and in contrast to Spanish, the classical languages leave many connections without overt marking. So it seems that these borrowings literally fill a gap, by making the nature of the connections more explicit, in the way Spanish does, by using the borrowed coordinator, especially when none is available in the language itself. This kind of borrowing strategy may have a rather pragmatic background. Coordinators like ‘and’, ‘but’ and ‘or’ are typically located at the periphery of utterances, and therefore also serve the purpose of turn holders in conversation very well. In this respect it is interesting to note that in all three subcorpora there are quite a few instances of Spanish *o sea* ‘that is to say’, a very colloquial element of spoken Spanish, that precisely serves that very same pragmatic purpose.

As for subordination, classical Otomi and Guarani often resort to asyndetic coordination where a language like Spanish (and English) would have a subordinate clause marked by some kind of subordinator. The typical Quechua strategy is nominalization in such cases. Although these languages do have markers, either neutral (‘that’) or with different modal shades of subordination (cause, reason, purpose, concession), the latter often in the form of adverbs, they are rare in comparison to Spanish, and are used only very infrequently. They are, however borrowed from Spanish, and are rather popular in their use. The most frequently borrowed subordinator is *porque* ‘because’, used by 87% (Guarani), 76% (Otomi), and 40% (Quechua) of the speakers in the corpus, respectively. Semantically unmarked Spanish *que* ‘that’ is highly frequent in Otomi, but less so in the other languages. Many subordinators of Spanish are a combination of a preposition and *que*: *hasta que* ‘until’, *para que* ‘in order to’, *porque* ‘because’, *sin que* ‘without’. These are all present in the Otomi corpus, sometimes occurring with and sometimes without the *que* form. As such, they strengthen the presence of the corresponding prepositions in this language. As for a motivation of this borrowing, we may

assume that, just like the borrowed coordinators, their use makes the relation between the corresponding clauses more explicit, while in the classical language it has to be inferred. Their systematic introduction in the three languages, however, may lead to some syntactic restructuring with respect to the classical formats, resulting in (marked) subordination being the norm rather than the original (asyndetic) coordination. This would mean a typological shift in all three cases.

A final point in this section concerns morphology. In neither of the three subcorpora is there systematic borrowing of Spanish bound morpheme, tight to a native lexical element. Among the few examples to be found are those in (18) below, each used by just one speaker.²² Both are derivations based on a Quechua noun and the Spanish agentive suffix *-ero*.

- (18) a. *huasipunguero* ‘villager, farmer’ < *huasipungo* ‘village’
 b. *warminero* ‘womanizer’ < *warmi* ‘woman’

If this lack of morphological loans would be a diagnostic for the advancement of borrowing in terms of the well-known scales (cf. Thomason & Kaufmann 1988, Thomason 2001), then all three languages would still be at a relatively early stage. What we do find in all three languages, however are Spanish loans that are morphologically complex in the sense that they are borrowed with an affix. Table 8 below gives an overview of the three most frequently occurring affixes, the first one inflectional, coding plural, and the others derivational, coding agentive and diminutive, respectively.

Spanish	Quechua	Guarani	Otomi
<i>-s</i> (PL)	160 (tokens)	7 (tokens)	8 (tokens)
<i>-dor</i> (AG)	22 (types)	10 (types)	1 (types)
<i>-ito/-illo</i> (DIM)	6 (types)	3 (types)	1 (types)

Table 8. Some Spanish suffixes found on borrowed elements

Overall, the figures in Table 8 seem to confirm the differences on the borrowing scale for the three languages that we already observed with respect to the major parts of speech above, with the highest numbers for Quechua, and the lowest for Otomi. The plural markers are by far the most frequent ones in all cases. For Quechua, 7 out of the 160 instances also had the native plural marker *-kuna*. And almost all nouns that have the Spanish plural marker on borrowings in this language also occur without it in the corpus. Furthermore, in most contexts in Quechua with Spanish *-s* but without the native plural marker, it seems to be clear that plural was indeed intended. This does not seem to be the case for Guarani and Otomi, where we do not find them accompanied by native plural markers at all. As for the other two types of suffixes in Table 8, we may assume that, from the perspective of the speakers involved all these word forms are units, and not really derived morphologically complex entities. Their low frequencies and the absence in the corpus of the corresponding bare stem forms are an indication of this. So possibly, only the plural

²² These, and some other examples discussed below are from Bakker & Hekking (forthc).

marker in Quechua might be in the process of making its way into the other language, be it for the time being on borrowed items only. All these cases, however, may function as so many Trojan horses, vehicles that import foreign morphological material into another linguistic environment where it will be ready one day to appear also on a native lexical element.

As indicated in Table 4 above, where an overview of the corpus was given, the data was collected from two different dialects, which are not in contact with each other, and which also might have slightly different contact histories. If the structure of the languages involved would not play a significant role in borrowing, then we could expect that different dialects might show considerably different outcomes even if they would borrow from the same language. However, the two dialects that were selected for each of the languages show very similar results, considerably better than chance could ever explain. The totals in Table 9 below may be convincing enough to give support to this assumption. The figures in brackets present the results for the two dialects with respect to the categories for which there are significantly different figures between the three languages. Note that the sizes of the subcorpora for pairs of dialects are not necessarily equal. The percentages are rounded up.

	Quechua	Guarani	Otomi
Noun	68% (70–59)	59% (53–74)	86% (84–88)
Verb	22% (24–15)	29% (27–33)	10% (8–10)
Adjective	11% (11–13)	12% (11–14)	4% (4–4)
Preposition	1% (1–1)	1% (1–0)	20% (18–26)
Definite article	0.0% (0–0)	19.4% (23–12)	0.0% (0–0)

Table 9: Breakdown for pairs of dialects

So it seems rather safe to conclude that certain differences in the qualitative and quantitative aspects of the borrowing process for the languages may indeed be related to differences in structural features of each of them, at least to some extent.

Let me conclude this section by observing that our three languages seem to be far from affected deeply by borrowing at the moment. However, not all inroads are necessarily visible at this kind of superficial inspection. Observations may be hampered by our lack of knowledge of the classical languages. And the figures are consistently higher for younger and more educated speakers than for the other informants. But even if the process has progressed in fact much further than it seems on first sight, borrowing should also be seen as a strategy precisely to adapt a language to the present requirements. As such it might lead to survival rather than to demise.

5. Conclusion

This ends our story of three languages from America: Otomi from Mexico, Quechua from Peru, and Guarani from Paraguay. Or rather, a tiny fragment of such a story. No paper account of something as complex as a language can come anywhere near completeness. This holds for the linguistic system itself, in the more technical sense of a finite set of elements and rules that operate on them, the lexicon and the grammar. It holds even more for the intricate dynamics of the use of such a system within a language community, in conversation, and the way the system it is handed on to the next generation of speakers.

We have seen that there must have been many thousands of languages in the Americas long before the arrival of the Europeans, most of which will never be known to us. Their history is longer than that of the languages currently spoken in Europe, probably at least twice as long. As a result, and by lack of much linguistic and non-linguistic evidence, we don't know for sure whether the majority of these languages go back to one ancestor language, Paleoamerican, and would therefore form one large Amerindian language family. We do know that there are at least two relatively small groups of languages which are fundamentally different, and probably originate from later invasions: Na-Dene and Inuit.

We then made a jump ahead in time of many thousands of years, and had a look at the linguistic situation shortly before the European invasion. There must have been a rich variety at that stage, considerably greater than the concurrent situation in Europe, with at least 1500 different languages belonging to over 100 families. We had a closer look at three of the largest languages, each from a different area. We briefly discussed the histories of these languages, in as far as these are known, and the place they hold in today's societies. We also sketched some aspects of their grammars, putting them in typologically different classes.

Finally, we had a look at borrowing, more specifically from the perspective of the three languages and Spanish, the colonial language and the official language of the post-colonial Latin American countries where they are spoken. It is accepted wisdom that borrowing as such is a sociolinguistic process, and typically goes from a language of 'power' to a socially and politically underlying one. However, our database data suggests that what actually happens, and what does and does not take place in a concrete borrowing process, is at least to some extent dependent on the typological similarities and differences between the languages involved. The fact that the borrowing behavior for the two dialects seems to be quite similar in all cases gives extra support to such a conclusion.

The indigenous languages of the Americas lose ground in a rapid fashion. Arguably, 500 have gone lost since Columbus arrived, and this process is gathering momentum since the last half of a century. At this pace, probably more than 95% of the American languages extant today will be extinct by the end of the 21st century, possibly with the exception of the largest ones, among which we might count some form of Quechua and Guarani. It remains to be seen then how these survivors would compare to their ances-

tors of today, and yesterday, assuming that they will be affected further by borrowing. Jopara and Media Lengua may be foreshadowing this process. On the other hand, these languages, and their speech communities may turn out to be more robust than they looked so far.

Abbreviations

1	first person	FUT	future tense
2	second person	IND	indefinite
3	third person	PAST	past tense
ACC	accusative	PERF	perfect
AG	agentive	PL	plural
CAUS	causative	POS	possessive
CLF	classifier	PP	past participle
CONT	contemporary mood	PRES	present tense
DEF	definite	PURP	purpose
DEM	demonstrative pronoun	RECP	reciprocal
DIM	diminutive	SBJNC	subjunctive
DUR	durative	SG	singular
EXCL	exclusive	TOP	topic
FOC	focus		

References

- Adelaar, Willem, with Muysken, Pieter (2004): *Languages of the Andes*. Cambridge: Cambridge University Press.
- Bakker, Dik; Gómez-Rendón, Jorge & Hekking, Ewald (2008): Spanish meets Guaraní, Otomi and Quichua: a multilingual confrontation, in: Stolz, Thomas; Bakker, Dik & Salas Palomo, Rosa (eds.), *Aspects of language contact*. Berlin: Mouton de Gruyter, 165–238.
- Bakker, Dik; Müller, Andre; Velupillai, Viveka; Wichmann, Soeren; Brown, Cecil H.; Brown, Pamela; Egorov, Dmitry; Mailhammer, Robert; Grant, Anthony Grant & Holman, Eric W. (2009): Adding typology to lexicostatistics: a combined approach to language classification, in: *Linguistic Typology* 13-1, 167-179.
- Bakker, Dik & Hekking, Ewald (forthc): Constraints on Morphological Borrowing: Evidence from Latin America.
- Campbell, Lyle (1995): The Quechumaran hypothesis and lessons for distant genetic comparison, in: *Diachronica* 12, 157-200.
- Campbell, Lyle (1997): *American Indian Languages: The Historical Linguistics of Native America*. Oxford: Oxford University Press.
- Cerrón-Palomino, Rodolfo (1987): *Lingüística Quechua*. Cusco: Centro de Estudios Rurales Andinos Bartolomé de las Casas.
- Coe, Michael D. & Van Stone, Mark (2001): *Reading the Maya Glyphs*. London: Thames & Hudson.
- Cole, Peter (1982): *Imbabura Quechua*. London: Croom Helm.
- Davies, Mark (2002): *Corpus del Español* (100 million words, 1200s–1900s). Available online at <http://www.corpusdelespanol.org>.

- Dietrich, Wolf (2010): Lexical evidence for a redefinition of Paraguayan Jopara, in: STUF 63(1), 39–51.
- Erlandson, Jon M. (2002): Anatomically modern humans, maritime voyaging, and the Pleistocene colonization of the Americas, in: Jablonski, Nina G. (ed.), *Memoirs of the California Academy of Sciences*. San Francisco: University of California Press, 59–92.
- Fortescue, Michael (1984): *West Greenlandic*. London: Croom Helm.
- Fortescue, Michael (1998): *Language relations across Bering Strait*. London: Cassell.
- Greenberg, Joseph (1957): *Essays in Linguistics*. Chicago: University of Chicago Press.
- Greenberg, Joseph (1987): *Language in the Americas*. Stanford: Stanford University Press.
- Gregores, Emma & Suarez, Jorge A. (1967): *Colloquial Guarani*. The Hague: Mouton.
- Gruhn, Ruth (1997): Language classification and models of the peopling of the Americas, in: McConvell, Patrick & Evans, Nicolas (eds.), *Archeology and linguistics*. Oxford: Oxford University Press, 99–110.
- Haspelmath Martin; Dryer, Matthew S.; Gil, David & Comrie, Bernard (2005): *The world atlas of language structures*. Oxford: Oxford University Press.
- Haspelmath, Martin & Tadmor, Uri (2009) *World Loanword Database*. Munich: Max Planck Digital Library, 2158 entries. Available online at <http://wold.livingsources.org/vocabulary/>.
- Hekking, Ewald & Andrés de Jesús, Severiano (1984): *Gramática Otomi*. Querétaro: Universidad Autónoma de Querétaro.
- Hekking, Ewald & Bakker, Dik (2009): Loanwords in Otomi, an Otomanguean language of Mexico, in: Haspelmath, Martin & Tadmor, Uri (eds.) *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: Mouton de Gruyter, 897-917.
- Holman, Eric; Wichmann, Soeren; Brown, Cecil; Velupillai, Viveka; Müller, Andre & Bakker, Dik (2008): Explorations in automated language classification, in: *Folia Linguistica* 42(2), 331–354.
- Jones, Peter N. (2004): *American Indian mtDNA, Y Chromosome Genetic Data, and the Peopling of North America*. Boulder: Bauu Press.
- Lewis, M. Paul (ed.) (2009): *Ethnologue: languages of the world*, Sixteenth edition. Dallas: SIL International. Available online at <http://www.ethnologue.com/>.
- Mannheim, Bruce (1991): *The language of the Inka since the European invasion*. Austin: University of Texas Press.
- Mithun, Marianne (1999): *The languages of native North America*. Cambridge: Cambridge University Press.
- Muysken, Pieter (1994): Media Lengua, in: Bakker, Peter & Mous, Maarten (eds.) *Mixed languages: 15 case studies in language intertwining*. Amsterdam: IFOTT, 207–211.
- Ruhlen, Merritt (1991): *A guide to the world's languages*. Stanford: Stanford University Press.
- Stanford, Dennis & Bradley, Bruce (2004): The North Atlantic ice-edge corridor: a possible Palaeolithic route to the New World, in: *World Archaeology*, 36(4): 459–478.
- Steele, D. Gentry & Powell, John F. (2002): Facing the past: a view of the North American human fossil record, in: Jablonski, Nina G. (ed.), 93–122.
- Swadesh, Morris (1955): Towards greater accuracy in lexicostatistic dating, in: *International Journal of American Linguistics* 21, 121–137.
- Thomason, Sarah G. (2001): *Language contact. An introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah G. & Kaufmann, Terrence (1988): *Language contact creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thompson, Chad (1996): The Na-Dene middle voice: an impersonal source of the D-element, in: *International Journal of American Linguistics* 62(4), 351–378.
- Urbano, Alonso (1990[1605]): *Arte breve de la lengua Otomi y vocabulario trilingüe*. Ciudad de México: Universidad Nacional Autónoma de México.