Spanish meets Guaraní, Otomí and Quichua: a multilingual confrontation

Dik Bakker
Jorge Gómez-Rendón
Ewald Hekking

Version 21 March 2006

0. Introduction

To date, there are around 7000 languages spoken in the world's 200 countries. According to the survey in Thomason (2001:250f), by far most of these countries have selected one language for their official tongue. The reality, however is that an average of 35 languages are spoken in each of them. For some of the larger countries, the number may run well into the hundreds. Although figures for actual multilingualism are difficult to provide, for most people in the world, multilingualism is simply a fact of social and private life. Probably the most famous example from the literature is the case study by Gumperz & Wilson (1971) of the Indian village of Kupwar. It boasts four languages, two Dravidian (Kannada, Telugu) and two Indic (Marathi, Urdu), which are all used actively in daily life by an important part of the population. As a result of this multilingual situation, all local variants of these languages have taken over structural properties from each other to the extent that a word-by-word translation between any two of the languages yields an understandable utterance. An example from Latin America is the area of the Vaupés River in the northwest Amazonian region where multilingualism is obligatory and marrying someone who belongs to the same language group is considered akin to incest. (Jackson 1974; Aikhenvald 2002)

In general terms, the majority of the around fifty million strong indigenous population in Latin America are bilingual in their native language and either Spanish or Portuguese (Gordon 2005). It is no wonder that especially this area has seen a host of publications[1] about both the social and linguistic aspects of bilingualism. A number of international conferences have been organized on language contact in Latin America, more recently in Amsterdam (2004) and Bremen (2005). A lot of descriptive work has been done especially in the area of loanwords, and to a much lesser extent in the area of morphosyntactic and phonological change. Research has typically been concentrated on the effects of the colonial languages Spanish and Portuguese on a specific Amerindian language, such as Quechua (cf. Hekking & Muysken 1995) and Otomí (Hekking & Bakker 1998; 1999; 2006). Much less frequently, the effects of Amerindian languages on the colonial languages have been studied, as is done by Haboud (1998) for Quechua and Spanish. Much rarer still are studies which take a comparative, multi-contact perspective, where the colonial language is the single source language, i.e. the language from which the borrowings stem, and a number of languages in America, Asia and Europe are the target languages, i.e. the languages which borrow words and other linguistic elements. An example of this is the work of Stolz & Stolz (1996). Next to more or less purely lexical and grammatical studies, many Latin American contact studies concentrate on the social aspects of bilingualism, both from the perspective of the individual speakers and the society at large. Especially the sociolinguistic situation in Paraguay, where both Spanish

and Guaraní are the official languages has generated a lot of studies (cf. Corvalan & Granda 1982; Morínigo 1990).

The bulk of the work in this area is couched in terms of traditional contact linguistics, with a predilection for the borrowing effects on the target (indigenous) language rather than the interference effects on the source language. The main object of interest is the lexicon. Most studies are descriptive-typological and do not work within a specific theoretical framework or from a coherent set of hypotheses concerning the phenomena to expect or not to expect. Also, it is not always clear on what data the observations are based, and what the status of the borrowings is. For instance, it is often not clear from the discussions whether the observed contact phenomena concern real borrowings or mere code switches. Often, no information is available about the level of bilingualism of the speakers who provided the data, and whether we are dealing with spontaneous spoken language or a written variety. This makes it sometimes difficult to interpret the results and compare them to results from other studies. Consequently, it is not always clear to what extent such studies may really deepen our insight into the process of language change through language contact.

The current article reports on an ongoing project in language contact in the Spanish Americas, which aims at repairing several of the shortcomings mentioned above. It is organized as follows. In section 1 we will discuss the major principles and strategies of the project, meant to repair some of the methodological pitfalls. In section 2 we introduce some hypotheses concerning borrowing that may be derived from the principles discussed in section 1. The effects of language contact may become visible only in a large corpus of linguistic data from a variety of sources. The computer is an important tool for the systematic exploration of these data. Section 3 sketches a notational system and some programmes we developed for the processing of digitalized bilingual data. In the rest of the article, we will illustrate our methodology on the basis of three case studies in language contact, all with Spanish as the source language of the borrowing process. It concerns three typologically rather different target languages. Section 4 deals with Guaraní, a Tupi language spoken in Paraguay. Section 5 discusses contact phenomena found in different Quichua[2] dialects of Ecuador. Thirdly, section 6 studies data from Otomí, an Oto-manguean language from Central Mexico. What is possible and not possible in borrowing is very much a matter of tendencies rather than absolute universals. Tendencies can only be borne out by comparison between languages. Therefore, in section 7, we will compare the results from the three languages, and see to what extent any more general conclusions may be drawn.

1. General outlines of the language contact project

As argued above, the methodology of language contact research has been left relatively underdeveloped, and mainly at the discretion of the individual researcher. Therefore, much of the work that has been done in the field so far turns out to be rather incomparable. As a result, the amount of generalization over these results is rather limited and tentative. In order to help repair this situation, we have formulated several points of departure to direct our own investigations in the field. We think they are of a general nature, and may be easily transferred to other language contact studies.

1.1 Borrowing versus shift-induced interference

First, and foremost, this study is about borrowing as opposed to interference caused by shift to another language, in the sense of Thomason (2001:66f). This means that we are interested only in cases where the source of some element is more or less clearly a second language for the speakers whose first language we are studying. Cases of imperfect learning will be left out of consideration as much as possible. Of course, the notion of imperfect learning is not unproblematic in itself. Any version of a target language containing some clear instance of borrowing from some source language not present in the variety of the earlier generations of speakers may be seen as the result of imperfect learning on the side of the speaker in question. This is especially the case if the element borrowed replaces an existing element of the target language, and it may be maintained that the new variety is no longer 'pure'. The distinction between borrowing and interference is crucial, however, since the predictions for the two phenomena are known to be quite different in terms of what may be found in the target language and what not, and the order in which changes may take place. We have tried to cater for the problems involved by requiring that the Amerindian language serve as the dominant instrument for everyday interaction in the speech community under investigation. For the individual speakers who provide our data we require that the Amerindian language is their first language, and that they use it on an everyday basis. In this way we hope to exclude cases of interference, which typically will also be attested in the same communities.

1.2 One source language
Within our project, the source language of the borrowing process is kept constant, in our case Spanish. By taking this language as the only one from which borrowed material is studied, there is as little linguistic variation on the input side as possible. Arguably, for the structural domains of grammar (morphology and syntax) variation between the several varieties of Spanish in Latin America is relatively modest, or even insignificant. As for more superficial variation, i.e. lexical and phonological, potential dialectal differences in the areal varieties of Spanish are controlled by not taking some standard version of Spanish as a point of departure, but rather the Spanish as spoken by the monolinguals in the respective areas. In that sense we fully agree with Van Hout & Muysken (1994), inasmuch as we also include the Spanish of the bilingual speakers of the area whose first language is the Amerindian language under consideration. Our assumption is, however, that there are no fundamental typological differences between the respective varieties of Spanish that would have substantial bearing on the comparison of the respective loan processes.

1.3 Typological variety among target languages
In contrast to the lack of typological variety in the source material of the borrowing process, the target languages should be selected such that, mutually, they manifest considerable typological distance. Only under such conditions typological conclusions of any depth and width concerning the borrowing process may be drawn. This may be done by comparing the differences between the respective sets of borrowed features, to the extent that these are manifest in the first place. The three languages chosen for this article, i.e. Guaraní, Otomí and Quechua seem to fulfil this typological requirement, as will be argued in the respective sections below.[3]

1.4 Culture kept constant

In most language contact studies it is accepted that cultural aspects play a more or less important role in what will and what will not be borrowed. Some authors even assume that 'anything goes' provided that the cultural pressure is strong enough (cf. Thomason 2001). Thus, in order to make the results of two cases of language contact comparable, cultural parameters should be controlled as much as possible. There is no way to do this in any type of realistic experimental setting. By concentrating on Spanish America, we think that a certain amount of cultural unity however is guaranteed. Of course, there are vast differences between the many indigenous cultures of America, and the ways in which they interacted and still interact today with the Spanish-speaking world. These will undoubtedly be reflected in the specific loanwords that may be found as such, and the semantic fields from which they stem. However, with the areal restriction we take into consideration for our data collection, we think that the influence of culture is minimalized. As a result, the differences that we may find between the respective Amerindian languages with respect to borrowing from Spanish may be ascribed to a large extent to typological differences between the borrowing languages.

1.5 Target language: spoken data from a representative group

Language contact is not the contact between two abstract entities such as languages. It is a dynamic process that takes place in the interaction between the speakers of two language communities whose lects are sufficiently different for there to exist a communicative gap. In order for such a situation to arise, the communities involved have to be exocentric in terms of Andersen (1988), i.e. a reasonable proportion of the speakers of one community has a regular contact with those of the other community, and some become more or less bilingual in the process. For a variety of reasons, these bilingual speakers will introduce elements of their second language into their first. Over time, and to the extent that these innovators have a certain amount of authority inside and outside their community, monolingual speakers will copy some of the innovations which will then propagate throughout the community, including learners. This process may or may not be speeded up by an increase in bilingualism within the community both in breadth and depth. As a result, new varieties of the local (target) language may arise, which include aspects of the external (source) language. Apart from so-called catastrophic changes (Ross 2003:177f), when sudden events, such as massive deportation, lead to a complete restructuring of a speech community and its language, language change is a gradual and very complex process, often stretching over several of generations. Therefore, the best way to study the impact of language contact on the languages involved is not by making two snapshots of the target language at two stages far enough apart to reveal large-scale differences. Rather, it should be analyzed on the basis of data produced by a diverse group of individual speakers who find themselves at different stages of the borrowing process.[4] The closest one can get to observing such a process is through longitudinal sampling and in-depth study of stretches of discourse between a variety of relevant speakers, some reappearing in later samples and some new. Since such a complex setup is typically out of the question in most cases, next best is the construction of a corpus of spontaneously spoken samples of the target language by a range of speakers from different age groups, thus simulating time to some extent. Furthermore, the set of speakers should be stratified as best as possible according to a

number of sociolinguistically relevant factors, such as gender, level of education, profession, and contact with the source language community, among others. A sample that is minimally representative in terms of such parameters would amount to between 50 and 80 people. With the typical average text length of around 1000 tokens this would give us a minimum corpus size of 50.000 - 80.000 tokens. Consistent with this view, when we use metaphors such as 'language T borrows element E from language S' this should be read as 'a significant and increasing number of speakers of T uses E, which is originally an element of S only'. E could be a lexical item, a bound morpheme, a specific construction, and so forth. As in most of the literature on borrowing we will use the notions 'significant' and 'increasing' in a rather impressionistic way here. They are in need, however, of a proper definition.

1.6 Source language: from the community itself
Although today the role of external sources - television, film, internet, newspapers - on language is considerable, we think that in most cases the language used in the local communities themselves is the greatest determining factor for language change. Therefore, we collected a corpus of spoken Spanish from the same communities from which we collected the target language samples. Part of the data has been collected from monolingual speakers of Spanish, typically mestizos. However, we also asked some of the native speakers who provided us with the data to contribute a spoken text in Spanish as well. This gives us an instrument to measure their level of bilingualism. Thus, we can relate the proficiency in the mother tongue and the amount of borrowing from Spanish to the proficiency in Spanish, and the diversions from the norm found in it.

1.7 Diachrony
Language change as a result of contact may be a matter of only a short period of time. Media Lengua (Muysken 1994; Gómez-Rendón 2005), a mix of Spanish and Quichua, is an example of a contact language that probably has been created in the span of only two or three generations. In that sense, a well chosen sample which adheres to our sociolinguistic criteria under 1.5 would be good enough to trace rapid changes across generations. However, many other instances of contact-induced language change take more time. Therefore, we think that an adequate database should also contain samples from earlier stages of the target language, preferably also from before the contact era. In most cases, such data is not available, not even in the form of written varieties. In the case of Latin America, however, there exist secondary sources on some of the indigenous languages, such as dictionnaries and grammars written by missionaries (cf. Urbano 1990 [1605]; Ruiz de Montoya 1993 [1640]). When available, we try to employ such sources by lack of primary ones, to the extent of course that they may be considered reliable.

1.8 Regularities and hypotheses concerning borrowing
Not much theory has been developed to date with respect to what may be borrowed and what not. Moravcik (1978) is a classical proposal with respect to the universals of borrowing. However, most of the universals suggested by her, and others found elsewhere in the literature are shown by Campbell (1989) to have counterexamples. This seems to suggest that such universals may be tendencies at best. A second point of orientation is the borrowing scale proposed by Thomason & Kaufmann (1988:74f), and

its revised version in Thomason (2001:70f). This scale measures the extent to which a language has changed as a result of borrowing from another language on the basis of the types of lexical and grammatical elements that are supposed to be or not be borrowed at different stages. Also this scale is not absolute. It provides tendencies which may be influenced by typological differences and which may be overruled by social factors. With the necessary caution, however, some predictions may be derived from this scale as to what types may and may not be found among the borrowed elements in a language. This makes the scale falsifiable to some extent, with the proviso of the overriding social and typological factors. Thirdly, language typology makes certain predictions as to which combinations of features may and may not be found in languages. These so-called universals of language often go back to the original work of Greenberg (1963) on implications between formal features in syntax, morphology and phonology. Most of them were presented as tendencies from the start, have been shown to be not absolute since, or were refined by authors such as Hawkins (1983) in order to make them more reliable. If we take them as more than contingencies, and assume that they represent strong tendencies towards what may or may not be found in a grammar, then certain predictions may be derived from them as far as borrowing is concerned. A final constraint on borrowing, provided by the field of language acquisition, is one on markedness: elements are less likely to be borrowed to the extent that they are marked. This is typically to be taken in the literal sense of (morphological) complexity rather than the metaphorical sense of 'remarkable, unlikely'. Also this claim may be falsified, at least in principle.

1.9 Explanation: the role of theories
Although linguistic theories typically aim at describing linguistic competence (and performance) in the light of the grammar of the first language only, and are not about borrowing, they should be able to make predictions about the nature of the language borrowing process, at least in principle. For syntactic theories of the Chomskyan type, borrowing, both by adult speakers and first language learners, appears to be a real challenge with respect to universal grammar, especially as far as it concerns the structural aspects of the source and the target language. For cognitively and functionally oriented approaches to grammar, such as Construction Grammar (cf. Goldberg 1995; Croft 2001) and Functional Grammar (FG; Dik 1997) borrowing seems to be less of a problem vis-à-vis the major principles. According to such theories, constraints that might be predicted for borrowing at the formal and semantic levels might be overruled by pragmatic or social factors, rendering any absolute prediction untenable. Nevertheless, also in these cases, predictions might be derived, be it that they concern the relative rather than the absolute amount of (several types of) borrowing. In that respect, they more or less naturally coincide with the Thomason scale and with predictions from language typology. More specifically, theories on parts of speech provide a framework from which testable predictions about borrowing may be derived. We will discuss two: the one introduced in Hengeveld (1993), which is syntactically oriented and embedded in FG; and Croft's (2000) cognitively based theory. As another example of a testable framework, Myers-Scotton's (2002) Matrix Language Theory makes predictions about the types of elements one may find in code switches but also in borrowing and other scenarios of language change. Although the current project is about borrowing and not about code switching, the borderline between the two is not always clear. It could be claimed that definitive

borrowing of some element is often preceded by its incidental use in code switches. Studying the correspondences and differences between borrowing and code switching may in fact lead to better definitions of the two related phenomena. Finally, language change takes place as a result of partial or full bilingualism of the speakers of the languages concerned. Studies in language contact should, therefore, take heed of what is known about first and second language acquisition, and of what is known about the bilingual mind.

Although not presented explicitly as a theory on borrowing and shift, Thomason (2001:129f) presents seven what she calls mechanisms according to which contact-induced language change operates. She mentions passive familiarity, first and second language acquisition, code alternation and code switch, negotiation and deliberate decisions. More specific mechanisms mentioned are correspondence rules (phenomena from the source language are projected onto existing phenomena in the target language in a one to one fashion) and simplification (distinctions made in the source language are ignored when a form is introduced in the target language). These are all linguistic environments or practices enhancing language change, and which may throw a light on what kind of changes may take place as a result of language contact and with what probabilities. They are definitely not motivations for change. In principal, all motivating factors are assumed to be social and non-linguistic. However, a few examples of changes are mentioned which might be interpreted as motivated by purely grammatical reasons.[5] As an example may serve the borrowing of the English conjunction *and* by bilingual speakers of Maori. Arguably, its use leads to less complex structures than the Maori system which is based on the preposition *me* 'with'. The latter form also seems to undergo reanalysis leading to its use as a conjunction (Thomason 2001:136).

In typological work it is sometimes argued that a language has copied a specific grammatical strategy or feature from another language. Evidence is typically based on the fact that languages genetically related to the target language do not have such a feature nor is there any indication of internal diachronic development. Comrie (1989) gives an example from Turkish, an Altaic language with mainly infinite subordinate structures, which has borrowed the conjunction *ki* from Persian and uses it to form finite subordinate clauses such as the complement clause in (1) below.

(1) Turkish (Comrie 1989:203)[6]

  Herkes   bilir ki   dünya yuvarlak-tir
  Everyone know CONJ earth   round-COP
  'Everyone knows that the earth is round.'

Another example stems from Estonian. Unlike related Finnish, this language has a passive construction that is arguably borrowed from Russian or German through contact. Finally, according to Dench (1998:37) person clitics in the Pama Nyungan languages Yingkarta and Wajarri are an innovation with respect to the genetically related Australian languages from the north and may be borrowed from unrelated contact languages in the east. Some take this as an example of a purely formal reason for borrowing, i.e. the optimalization of the grammatical system.

In earlier work (Hekking & Bakker 1998; 1999; 2006) some of the principles of borrowing alluded to above were tested on the basis of part of the data then available for Otomí. In section 2 they will be reformulated and extended.

2. Hypotheses on borrowing

Here we will put forward some hypotheses on borrowing which are in part inspired by work in functional theories of grammar and in part by more general work on contact induced change.

A central issue in the study of language change in general is the extent to which changes are internally motivated or caused by external factors. On the basis of de Saussure's (1916) distinction between *Langue* (the 'inner' language, the system) and *Parole* (the 'outer' language, its use), linguistic theory has sought to base explanations on the former rather than the latter, in the case of Generative Grammar even exclusively. Functionally oriented theories, such as Functional Grammar (FG; Dik 1997) have steered away from this tendency, and placed more emphasis on language use and the role of discourse. We will follow the latter route when looking for factors that might explain the change of languages in the presence of other languages. The functionally oriented hierarchy under (2), proposed earlier in Bakker & Hekking (1999) will serve as a basis.

(2)   Social > Communicative > Cognitive > Formal factors

We may assume that a language that is functioning satisfactorily as a means of communication in some community for a considerable period of time, is a complete system, with its own mechanisms of adaptation, variation and change. Therefore, there seems to be no inherent reason for it to copy any aspect, either form or structure of another language with which it gets into contact at a certain stage. However, languages do, and the most fundamental reason for this seems to be the desire of the speakers of both languages for mutual understanding, be it that this desire is not necessarily the same for both groups, and their individual members. In catastrophic cases, when contact between two language communities increases in a gradual fashion, and the two communities may merge into one, there will arise a certain amount of (partial or full) bilingualism among the speakers involved. This process will be more or less bidirectional in case the position of the original communities is more or less equal. This may lead to the development of a *Sprachbund*, where the languages share a very restricted set of features, but remain mutually non-understandable and do not change typologically in any way. Or we may witness complete bilingualism (or multilingualism) for a large group of speakers, as in the examples of Kupwar and Vaupés mentioned in the introduction, in the long run with rather dramatic results for the languages involved. However, if one community has the edge over the other, economically, socially, politically, as is the case of the languages studied here, bilingualism will generally be unidirectional, and affect only one language – the indigenous one – in a serious way. The results may run from the borrowing of a restricted number of lexical items from specific semantic domains to massive lexical and grammatical borrowing, typological change, relexification, language shift, loss and finally language death.[7]

With social factors forming the basis of language contact, and communication and mutual understanding being the major goals of the speakers concerned, we may

reformulate the hierarchy of (2) as in (3), where we have translated the more general terminology into the corresponding components of the grammar.

(3)   Pragmatics > Semantics > [Syntax-Morphology-Phonology]

Pragmatics covers the communicative aspects of language, such as the types of speech act and finer tuned discourse markers. Semantics deals with the content of the message speakers want to get across. It is precisely with respect to these two central, somewhat more abstract aspects that are crucial for successful communication that languages overlap most. This leads us to the expectation that much of the 'early' borrowing must be found precisely in these two areas of linguistic description. Individual lexical items that fill obvious semantic gaps are easiest to borrow since they enter an open class, for which extension is a routine process in any language, also for speakers above the language acquisition age. Secondly, we expect that languages may quite easily borrow certain discourse markers that are characteristic of the source language, such as connectors, pause fillers etcetera. In this case, gap filling is not a very likely motivation in the strictly functional sense. Rather, the use of such often highly frequent and syntactically peripheral and therefore outstanding markers give the target language some of the flavour of the prestige language, and suggests bilingualism.[8]

At the right hand side of the hierarchy in (3), in brackets, we find the formal components of grammar. In our view of the language system, these components form the (dependent) vehicle of communication rather than its (independent) content. With few exceptions, formal aspects of the source language will typically not be the object of borrowing as such but will rather be potential obstacles to it. Within the formal component of the grammar we distinguish a further subhierarchy going from syntax via morphology to phonology. This subhierarchy runs more or less parallel to the degree to which the functions of the respective subcomponents of the grammar are transparent to the (partially) bilingual speaker of the target language. For instance, we expect that a syntactic feature such as constituent order will turn out to be more transparent and therefore easier to copy than morphological features such as plural or agreement marking. With respect to the phonological component, suprasegmental aspects will be more accessible and therefore borrowable than individual phonemes, which often remain among the last 'fossils' even for semi-bilinguals.

The motivation-versus-constraint interpretation of the hierarchy in (3) operates roughly as follows. Bilingual speakers may be motivated to express a certain functional (i.e. pragmatic or semantic) feature F of source language S, represented by a formal element E_S (a morpheme or construction) in an utterance of target language T. Typically, F is not the only function of E_S, but the most outstanding in the bilingual situation. It may even get an interpretation which is not necessarily the one given to it by native speakers of the source language. If the target language T lacks F, the result is an enrichment of the utterance and, if copied by others, eventually also of the lexicon or the grammar of T as such. If T already possesses an element expressing F, then the result of its use may be a clearer way of expressing that function, because of its unique interpretation, or because element E stands out more than the original form which codes F. Borrowing is easier to the extent that F is a real functional extension of the target language and E_S fits easier into the grammatical structure of T. The more simular both languages are in a typological

sense, the higher the probability that E_S will find its way into an utterance of T, and the grammatical system of T. This typological similarity between S and T should be taken in the obvious synchronic sense but also in the diachronic sense of T being in a state where it might develop an E_S-like element internally.[9]

The above is resumed in the borrowing hypothesis below. This hypothesis is restricted to non-catastrophic, unbalanced language contact situations with a more or less clear distinction between a source and a target language.

Borrowing Hypothesis:
H0. Any element of source language S may be borrowed by target language T. The probability of some element E_S of S to be borrowed by T is dependent on the following linguistic factors:
a. its place on hierarchy (3)
b. its frequency of use in S, and later T
c. the typological distance between S and T on the relevant parameter(s)
d. the (maximum) level of bilingualism in the speech community

From this general hypothesis, several more specific hypotheses may be derived. Note that, since they are framed in terms of probabilities rather than absolute yes/no's, they should be tested by comparing languages of both the same and of a different type. Below, the > sign should be read as 'easier to borrow than'.

H1. pragmatically outstanding elements > non-outstanding elements

      H1.1 discourse marker > topic marker

H2. lexical elements > grammatical elements

      H2.1 open class > half open class > closed class

            H2.1.1 Noun > Verb > Adjective > Adverb > Adposition

            H2.1.2 Adposition > … > Auxiliary > … > Article

      H2.2 free form > clitic > bound form

            H2.2.1 Adposition > Case affix

H3. syntactically simple elements > structures

      H3.1 free lexical element > fixed collocation > noun phrase > construction

In general terms, we follow the view of Functional Grammar (FG; Dik 1997) which acknowledges only nouns, verbs, adjectives and adverbs as part of the lexicon, with their own concrete lexical meaning. All other elements have abstract meanings and are assumed to be part of the grammar. However, we consider adpositions as an intermediate

category in the sense that they may be either lexical or grammatical, depending on their status in the relevant grammar.[10]

As for the typological aspects of the parts of speech a number of proposals have been made in the literature. Some authors take a fundamentally cognitive-semantic position towards their definition. A clear example of this is Croft (2000). On the other end of the scale we find those who define the parts of speech on purely formal grounds, i.e. their role in morphosyntax. Hengeveld (1993) takes this position. Yet others, such as Evans & Osada (2005), take an intermediate stance. In this contribution we will put to the test only one approach, viz. Hengeveld (1993), since some tentative predications may be derived from it with relation to borrowing. Hengeveld's theory only considers the major parts of speech, nouns (N), verbs (V), adjectives (A) and manner adverbs (MAdv). Nothing is said about other adverbs, nor about grammatical elements. The four types distinguished are defined on the basis of their prototypical syntactic behaviour. Thus, V is the head of a predicate phrase, N the head of a referential phrase, while Adv and MAdv typically take the modifying position in predicate and referential phrases, respectively. Languages are ordered on a seven-point scale running from maximum flexibility (type 1) to maximum rigidity (type 7). The 7 language types are defined as in table 1 below.

TYPE	LEXICAL CATEGORIES
1	Contentive
2	Verb, Non-verb
3	Verb, Noun, Modifier
4	V , N , A , MAdv
5	V , N , A
6	V , N
7	V

Table 1. Typology of parts of speech systems (Hengeveld 1993)

Type 1 has only one overarching part of speech, which covers the whole range from V to Adv. In other words, many forms may be used in any of the four syntactic positions, without any extra marking that is indicative of their derivational status. Type 2 languages have an independent category V and a generalized category for the rest. Type 4 shows maximum diversification: all four basic lexical types are attested. Type 5 lacks basic adverbs, and type 6 basic adjectives and adverbs. In terms of this typology, Spanish is basically a type-4 language, i.e. its lexical entities typically specialize for one of the four parts of speech: e.g. its nouns cannot be used as modifiers, as shown in (4a/b); and virtually all adjectives need morphological adaptation in order to be used as adverbs (5).

(4)   a. *Una   casa   piedra   b. Una   casa   de   piedra
           ART house stone   ART   house of   stone
           'a house of stone'   'a house of stone'

(5)   Di-me   la   verdad   sinceramente/*sincero
        Tell-1SG.DAT   ART   truth   honestly/honest

'Tell me the truth honestly'

On the basis of this theory, the following general predictions seem to be warranted. We expect a target language T of type 1-4, which has all four syntactic positions available, to borrow all four types N, V, A and MAdv from a source language S without much constraint. When T is more flexible than S, there are two possibilities: functional adaptation or functional specialization. According to the first, more liberal hypothesis, borrowed elements will be treated as if they belonged to the lexicon of T: e.g., in a type-2 language, borrowed adjectives from a type 3-5 language may be used as heads of referential phrases apart from being used as modifiers. According to the second, less liberal hypothesis, borrowed elements will figure only in their original function. If T is less flexible than S then we only expect specialization among the borrowed elements in the relevant area, e.g. in a type-3 language some [Verb, Non-verb] elements borrowed from a type-2 source language may be used exclusively as heads and others exclusively as modifiers of referential phrases. On the rigid side of the scale, i.e. T languages of types 5-7, we expect to find low numbers of elements from an S language which have an 'unknown' part of speech, and specialization for elements which are borrowed into one of their original classes. E.g. a type-6 language will in principle not borrow a [A, MAdv] element from a type-3 language, and it will borrow [Verb, Non-verb] elements from a type-2 language only in the function of heads of predicate and referential phrases.

All these predictions may be seen as further specifications of hypothesis H2.1.1. There is one striking difference between H2.1.1 on the one hand and Hengeveld's typology on the other hand, i.e. the order of V and N on the hierarchy. Nouns being the most open class, we predict that they will be the first type to be borrowed, and that they will also have the highest frequency among the parts of speech in the list of borrowings. Although no prediction concerning relative frequencies might be derived from Hengeveld's typology, it seems to suggest a very central position for verbs. They are a class of their own in all but type-1 languages.

Apart from parts-of-speech oriented hypotheses, we will also have a look at some of the more structural aspects of language contact and change, as studied by language typology.[11] Some of the well-established parameters that have been discussed in the typological literature over the last 40 years seem to be relatively deeply entrenched in the grammatical system of a language, and take a long time to change, typically many hundreds of years. This leads us to the following general hypothesis on borrowing.

H4. There is a positive correlation between the time a typological parameter takes to change under neutral circumstances, i.e. without a strong external pressure, and the time it takes for it to change in a bilingual situation.

From this, the following more specific hypotheses may be derived.

H4.1 Borrowed elements fall in line with T morpho-syntax, and are easier borrowed when their basic syntactic position in terms of Head-Modifier relations in T is the same as in S. E.g. adpositions will be borrowed in their original syntactic position, and only if it is available in the T syntax. E.g. a postpositional T language may borrow postpositions from S but no prepositions.

H4.2 The frequency of the existing constituent order patterns in T may change in the direction of orders frequently attested in S. However, this will typically not lead to the introduction of new orders, not attested before the contact period. E.g. a VSO language in contact with a SVO language may change its basic order to SVO given that this order is typically among the alternative orders anyway. It will not easily adopt SOV, which is rather rare as an alternative in VSO languages.[12]

H4.3 Languages may borrow elements which express some already existing function more analytically than the original strategy. E.g. when a language marks a possessive relation, it may borrow an adposition expressing possession in S. In actual utterances, both strategies may be combined leading to so called 'doubling'. The opposite, i.e. borrow a more grammaticalized form, is unlikely to happen.

These hypotheses will be tested in the sections 4 through 6, where we discuss the data we collected for the three languages which concern us here. First, in section 3, we will briefly discuss some of the tools we developed for analyzing the data.

3. Notational system and computational tools

As argued above, the empirical basis for any research into contact-related language change should be a relatively large body of spoken language, typically tens of thousands of tokens per language pair, produced by a number of different native speakers. The data are typically collected in recorded fieldwork sessions, by inviting subjects to tell a story either on a freely chosen or a specific topic. The latter can be directed by pictures, as the well-known Pear and Bee stories.[13] The recordings are then transcribed and put into computer-readable form. We chose a solution where the transcriptions are only very partially phonological. For the source language - always Spanish - only the most striking differences from the standard spelling are coded in the protocols. Some examples may be found in (6) below, stemming from speakers of Otomí.

(6)	a. *albañil* 'bricklayer'	albañi
		albañil
		albeni
		albini
	b. *vecino* 'neighbour'	besino
		besinu
		bisino
		bisinu

For the rest, the Spanish tokens are rendered in standard spelling. For the native tokens, a standardized spelling per language is used throughout. For the languages relevant for this article, more details on the collection and the coding of the data will be given in the respective sections below.

The recorded texts are entered into the computer via a standard word processor. The structure of the resulting file is as follows.

```
                  IDENT OF INFORMANT:
                  AGE:
                  LOCATION:
                  PROFESSION:
                  EDUCATION:
                  < further social parameters >
                  TARGET LANGUAGE:
                  SOURCE LANGUAGE:
                  RECORDED BY:
                  TOPIC:
                  $
                  < text >
```

Figure 1. Structure of data file

In the actual text, the loanwords are marked by slashes, as in example (7a) from Otomí, glossed in (7b).

(7)   a. Bí 'mui jar /skwela/ /entre/ 'naha ne hñu ya jeya.

        b. Bí   'mui j-ar   /skwela/ /entre/ 'naha ne hñu ya   jeya.
           PST3 be   LOC-DEF.SG school   between one   and three DEF.PL year
           'I was at school beween one and three years'

On the basis of a text thus coded, a computer programme developed by us compiles frequency tables of all types – i.e. different forms – found for the target language and the source language. Figure 3 is part of the table compiled for the Otomí informant known under the initials AEL, who provided us with a text with a length of 267 tokens (i.e. separate word forms), 16 of which were Spanish words of 11 different types. The table corresponds to the top of the alphabetically ordered list of types per language.[14]

```
Frequencies for informant: AEL


TARGET (Otomi)                SOURCE (Spanish)            informants (35)
ar            7               bentanä        1                  26
bi            2               boi            2
bongu         1               despwes        1
boni          1               frasko         1                   7
boxu          2               i              2                  21
bí           31               kong           2
bötsi        13               nsusya         1
da            2               panal          3                   3
…


Types       71 ( 86.59%)         11 ( 13.41%)                   82
Tokens     251 ( 94.01%)         16 (  5.99%)                  267
TTR        3.54                 1.45                           3.26
```

Figure 3. Lexical overview for one informant (fragment)

For both languages, the type-token ratios (TTR) are given.[15] All borrowed types found in the input text are incrementally added to a dictionary, which also keeps track of the overall number of informants that have used a certain type. These totals are given for

each of the forms in the list at the righthand side. So, of the 35 informants in the corpus, 26 informants have used the form *bentana* at least once.

Apart from containing the overall borrowing lexicon, the dictionary file may be used in several different ways. Firstly, any unmarked form found in a new text that matches a dictionary type may be automatically assigned slashes, marking it as a loanword.[16] Secondly, it may be used to solve both problems created by alternative spellings of source tokens and for lemmatizing inflected forms. The dictionary should then be prespecified as in (8).

(8)   escuela
        escuelas > escuela
        skwela > escuela

If these entries are present in the dictionary, then the forms *escuelas* (plural of Spanish *escuela* 'school') and *skwela* (Otomí pronunciation of *escuela*) will be assigned to the dictionary type *escuela*.

Optionally, the input text may be enriched with parts of speech information for the source language tokens. This is done by extending the form with a one or two letter code for the part of speech in Spanish, plus a code for the syntactic position in which it is found in the target language.[17] Thus enriched, the input sentence of (7) may look as follows.

(7')   Bí 'mui jar /skwelaNHR/ /entrePRL/ 'naha ne hñu ya jeya.

Thus, Spanish *skwela* 'school' is coded as a noun (N) and is functioning as a head of a referential phrase (HR) in this Otomí text. And the Spanish preposition (P) *entre* 'between' is coded as a syntactic relator (RL) in Otomí. Such codes are also represented in the lexicon, and may therefore be assigned automatically by default, i.e. unless explicit coding in the text overrides such a default. A full-blown lexical entry for the form *skwela* may therefore look as in (8') below. Such an entry provides the (canonical) form, the part of speech in Spanish (here N, i.e. noun) and the syntactic positions in which it has been found in the text (here both head and modifier of a referential phrase).

(8')   escuela N (HR, MR)

The programme produces totals per part of speech, both for individual texts and for the whole corpus of texts for a particular target language. The following figure gives a fragment of a part of speech survey for the Spanish loan words found in the complete Otomí corpus.

PoS	SFnc	TYP	TOK
ADJ	– MR	52	188
ADJ	– HP	13	23
ADJ	– HP MR	9	52
ADJ	– HP MP MR	1	4
ADJ	TOTAL	75	257

Figure 4. Total for parts of speech (fragment)

Thus, of the 75 Spanish adjectives found in the collective Otomí texts, 52 were used in their prototypical function of modifiers of the head noun of a noun phrase (MR); 13 were used only as the main predicate of a sentence (head of a predicate phrase, HP); and one was found to perform all three functions.

The coding system makes it possible to isolate borrowed lexical elements and bound phonemes in the text. There is also a notational convention for isolating code switches. Any stretch that is analyzed as such may be included in [ ]. By default, these are excluded from the analysis. Optionally, however, they may be added to the totals of the source language. In the analyses below, the code switches are left out of the consideration.

Sofar for the tools we developed. We will now look at the actual data that we collected and analyzed for our three languages.

4. The case of Guaraní

Guaraní is a Tupi language spoken by 3.946.904 (2002) people in Paraguay. Most speakers of Guaraní speak also Spanish with different levels of proficiency. Although Paraguayans take pride in being the only bilingual country in Latin America, bilingualism in Paraguay is neither stable nor symmetrical and the Indian language carries numberless traces of its century-long struggle with Spanish. Based on the investigation of a large corpus of spontaneous speech collected during a three-month period of fieldwork in Paraguay (2004-2005), this section deals with Spanish lexical borrowing in Paraguayan Guaraní (PG).

In this section, we first make the reader familiar with historical and sociolinguistic issues of Guaraní-Spanish contact in Paraguay. The second step is to address the typological classification of PG according to parts of speech and other parameters. In the third section we identify the expected outcomes of Spanish borrowing in PG according to its typological features. Finally we present the statistical results concerning the borrowing of lexical classes and analyze them in the light of the hypotheses described in section 2. Some conclusions are presented in the last section.

4.1 Historic context of the contact between Paraguayan Guaraní and Spanish

Juan Díaz de Solís (1516) and Alejo García (1524) explored the territory of present Paraguay as part of their individual – equally unsuccessful – enterprises of finding easier routes to greater riches. Paraguay itself, however, did not attract the newcomers as other densely populated and economically promising areas of the new empire did. A scattered native population did not offer any important workforce to the Conquerors.[18] Therefore, comparatively few Spaniards settled in the area and were easily outnumbered by the Indian population. Thirty-eight years after its foundation in 1536, Asunción counted two Spaniards for every twenty Indians.

In the context of this demographic unbalance, intermarriage became the best strategy for Spaniards and Indians to create long-lasting bonds that encourage a pacific coexistence. Spaniards used to marry several Indian women at a time and polygamy became a common practice in Paraguay. The resulting racial mixture spread rapidly. Over the years Guaraní speakers – mainly *mestizos* – absorbed most Spanish enclaves in the area.[19] Though Spanish continued to be used for all official transactions, Crown officials

protested that Guaraní was displacing Spanish in the area to the point that even the few remaining descendants of Spaniards preferred to speak the native language with each other (cf. Morínigo 1982). However limited its use, Spanish continued to be associated with the elite's political power throughout the history of Paraguay.

Concurrently with these developments, the Guaraní peoples who had remained outside the Spanish influence were object of individual evangelization enterprises by Franciscan and Jesuit missionaries. In particular Jesuits (1609-1768) developed a unique social organization in their so-called *reducciones* where Indians were encouraged to use only their native language for oral and written communication in their daily lives. Differences between the standardized Guaraní spoken by the Indians in the *reducciones* and the Guaraní spoken by the *mestizos* and Spanish settlers in the towns began to emerge. After Jesuit missionaries were expelled from Spanish America, the Indians of the thirty existing *reducciones* either fled to live in the wilderness or integrated into the mainstream mestizo society and thus made their contribution to the effervescent language pool of the Paraguayan society. The specific relations and influences between the Guaraní as developed in the Jesuit *reducciones* and the Guaraní spoken in the urban areas remain unknown to date (Dietrich 1995:204).

The Guaraní dialect spoken in urban areas became strongly influenced by Spanish both in the lexicon and the grammar. Although many advocates of purism consider this Guaraní (also called Jopara, 'mixture') a mere corruption, it has extended to all sectors of the Paraguayan society and a number of literary pieces have been written in it since the second half of the nineteenth century.[20] The debate about the status of this sometimes-called 'third language' is far from finished. For the purpose of this article, we consider Jopara and Paraguayan Guaraní as synonyms.

When independence from Spain was declared in 1811, Paraguay was an isolated district of the Viceroyalty of Rio de la Plata. However, the new establishment did not imply a new status for Guaraní. Throughout their independent history Paraguayans have shown ambivalent attitudes towards their language: they hold Guaraní in high esteem as symbol of their heritage and past but at the same time associate their language to backwardness and primitivism. This split attitude helps explain why political leaders often took contradictory stands and made divergent decisions on the issue of language use.

Gaspar Rodriguez de Francia, who ruled Paraguay from 1814 until his death in 1840, promoted monolingual Spanish schooling but used Guaraní for all administrative and political issues. His successor Carlos Antonio Lopez, a European-grown progressist and blatant detractor of Guaraní, launched a campaign in 1848 to replace all native family names with Spanish names. With Lopez' death in 1862, his son Francisco Solano became president of Paraguay and led his country to the bloody war with Argentina, Brazil and Uruguay. However tragic the war was in demographic and economic terms, it helped promote Guaraní as the highest landmark of Paraguayan identity. Guaraní was used from the battlefield to the trenches either as a secret code or as the language of folk songs. Francisco López realized the agglutinating potential of Guaraní and organized during the war a *Congreso de Grafía* (i.e. a spelling congress) in which new orthographic rules were set for the language.

The six-year war left a decimated nation at the mercy of the winning powers. And the winners were not interested in preserving any trace whatsoever of the Guaraní culture.

The new puppet regime was appointed on 15 August 1869 with a mass celebrated in the Cathedral by the General Vicar of the Argentinean Army, who preached about "the need to regenerate the Paraguayan people in order to promote their development" and the need to dispel Guaraní from their territory for being "a dreadful creation of ignorance and backwardness" (Trinidad Sanabria 1997). Shortly afterwards, the Argentinean political writer and activist Domingo F. Sarmiento was appointed to conduct an educational reform in Paraguay along the lines of his ideological tenant of the new American civilization, in which a "savage language" had no place (*ibidem*). The linguistic policy of favouring Spanish monolingualism at the cost of the native language was implemented in one way or another for the next half century.

Again, in the 1930's the tragic developments of war showed Guaraní as a centripetal force. The Chaco War between Paraguay and Bolivia motivated the recognition of Guaraní at all spheres of the administration. Politicians, military and religious leaders gave their speeches and harangues in Guaraní while poets composed popular songs in the language. And yet, the liberal, revolutionary and dictatorial administrations that followed did little or nothing to give official status to Guaraní.

The first winds of change came on 15 August 1967 when a new Constitution granted Guaraní the status of 'national language'. But the new status did not have any practical consequences. Only after the 1992 Constitution established the official status of Guaraní vis-à-vis Spanish, the Ministry of Education began to implement a model of Bilingual Bicultural Education according to which every child must be taught in his/her own mother tongue. This implementation has not escaped controversy. For one group of the population the bilingual programme is perpetuating the same structures of oppression on the native language by giving too much space to Spanish borrowings. For others the histories of Guaraní and Spanish are indeed one and it is inevitable that the native language carries all kinds of traces from Spanish (just like Paraguayan Spanish carries traces from Guaraní). At the turn of the new millennium the arena of languages has become the arena of political and cultural issues.

Paraguay is a unique sociolinguistic case in the context of Latin America. But this uniqueness is founded less on its claimed bilingualism than on the fact that Guaraní is the only Indian language in Latin America spoken by non-Indian citizens as their mother tongue. According to the 2002 census[21], Guaraní monolinguals (27%) were significantly more numerous than Spanish monolinguals (6.56%), particularly in rural areas. Interestingly enough, the percentage of bilinguals was only 59%, that is, less than two thirds of the country's population. These figures show that Paraguay is far from being a model bilingual society – if by bilingualism we mean the use of two languages on an equal basis by the whole linguistic community – but might become one in the future if we consider the developments of the last decade[22].

4.2 Typological classification of Guaraní
Guaraní is a flexible type-2 language in Hengeveld's classification of parts-of-speech systems. The language has only two lexical classes: verbs and non-verbs. The class of verbs is clearly identified by the existence of two morphological paradigms as shown in the following examples.

(9)   a. (Che)   a-guata   b. (Nde)   re-mba'apo

1SG 1SG-go	2SG 2SG-work
'I walk'	'You work'

(10)   a. (Che)   ai-pota   b. (Nde)   rei-pota
           1SG   1SG-want   2SG   2SG-want
           'I want'   'You want'

On the other hand, non-verbs may occupy any of the following syntactic positions without any further derivational morphology: head of a referential phrase, modifier of referential phrase, and modifier of a predicate phrase. Some examples illustrate this:

(11)   a. Ko   karai   tuja   b. Che   tuva   tuja
           DEM   man   old   1SG   father old
           'That old man' / 'that man is old'   'The oldness of my father'

(12)   a. Che ro-hayhu   asy   b. Nde   rayhu asy
           1SG 2OBJ-love   intensely   2SG love intense
           'I love you intensely'   'your intense love'

The same lexeme *tuja* is the modifier of a referential phrase in (11a) and the head of a referential phrase in (11b). Likewise, *asy* is the modifier of a predicate phrase in (12a) and the modifier of a referential phrase in (12b).

       The above classification, however, does not show a typical characteristic of Guaraní: the possibility for most lexemes in the language to be used predicatively. This feature is perhaps more evident in the case of so-called quality-attributive verbs (Gregores y Suárez 1967: 138). These lexemes may be used as heads both of predicate and referential phrases as shown in (13a/b):

(13)   a. a-vy'á   ne-recha-rehe   b. a-vy'á   ne-recha-vo
           1-happiness 2.ACC-see-by   1-happiness   2-see-when
           'my happiness of seeing you"   'I am happy to see you'

In similar terms, a predicative reading of (11a) understands *tuja* not as "old" but "to be old". The predicative use of nouns, adjectives and manner adverbs is further illustrated in the following examples:

(14)   a. Pe   *kyse*   puku   b. Che   che-*kyse*
           DEM knife   red   1SG   1SG-knife
           'that red knife'   'I have a knife'

(15)   a. Ajahe'o   *pochy*-rehe   b. Che   che-*pochy*
            1SG-cry   angry-by   1SG   1SG-anger
           'I cry from anger'   'I am angry'

(16)   a. o-mbohovai *mbarete*   b. o-mo   *mbarete*
           3-react   strongly   3-CAUS strongly

      'He reacts strongly'   'He strengthens [it]'

In spite of this evidence, we have not classified Guaraní as a type-1 language (the most flexible as for parts of speech, without a distinction between nouns and verbs). The main reason for our choice is that there clearly exists in Guaraní one class of verbs identified on the basis of morphological distribution (cf. Nordhoff 2004) while no such conclusion may be drawn for any of the other major parts of speech.

      So far for the classification in terms of its parts of speech typology. We will briefly sketch PG in terms of some other typological parameters. From a morphological point of view PG is an agglutinative language. Some authors also characterize PG as a 'typical polysynthetic language' (Trinidad Sanabria 1998). Though this statement remains to be proved in terms of its 'typicality', it is certain that the high degree of synthesis shown by pre-Hispanic Guaraní suffered considerably from the contact with Spanish. Morínigo (1982) offers solid evidence of this and other contact-induced changes. Also, PG is an active language in the sense that the subject of an intransitive verb may be marked either as the subject or the object of a transitive verb (Velazquez-Castillo 2002).

      Constituent order in Guaraní is relatively free, with SVO arguably as the unmarked order in the clause (Gregores & Suarez 1967:182). Within the noun phrase, the order also varies: head-modifier in attributive constructions, modifier-head in possessive constructions. PG lacks gender and number distinctions. Pre-Hispanic Guaraní did not have articles to express definiteness but PG makes an extensive use of Spanish definite article *la* (Gregores & Suarez 1967:144). The functional status of the borrowed article in PG however is not always the same as that of the article in Spanish. Accordingly, we classify Spanish-derived *la* in PG within the overall category of deictics, as it seems to mark not only definiteness but also reference and cohesion (see below section 4.4.2). As additional features worth mentioning here, PG is a pro-drop and postpositional language and shows frequent noun incorporation (Velazquez-Castillo 1995).

4.3 Predictions about Spanish borrowings in Paraguayan Guaraní
Having established that Guaraní is a type-2 language in terms of Hengeveld's part of speech typology discussed in section 2, and that Spanish is type 4, the following predictions can be made as far as borrowing is concerned. We expect that Guaraní will borrow nouns, verbs, adjectives and (manner) adverbs in that order of frequency. If functional adaptation applies, we would find Spanish nouns (N), adjectives (A) and manner adverbs (MAdv) in all relevant functions, i.e. as head (HR) and modifier (MR) of referential phrases and as modifiers (MP) of predicate phrases. If functional specialization applies, we would find N, A and MAdv only as HR, MR and MP, respectively.

      As regards other categories, it is expected that PG does not borrow articles and prepositions from Spanish, as there is no syntactic position for this element in the language. Similarly, from the development of PG towards hypotaxis, it is expected that PG tend to borrow Spanish subordinators and discourse markers provided that they fit in the typological profile of the language. Finally, it should be said that inasmuch as PG allows a predicative use of other non-predicative elements, this preference might be reflected in a similar use of Spanish originally non-predicative borrowings.

In this section we will present our results about borrowing from Spanish in Guaraní and test our findings on the basis of the relevant borrowing hypotheses. Our corpus consists of spoken texts provided by 38 speakers, with a total length of 57.828 tokens. Table 2 contains the figures for Guaraní and for Spanish borrowings in terms of tokens, types and type-token ratio.

N=38	Target Language (PG)	Source Language (Sp)	Total
Tokens	47772 (82.6%)	10056 (17.4%)	57828
Types	9220 (77.0%)	2760 (23.0%)	11980
TTR	0.19	0.27	0.21

Table **2**. Spanish borrowings in PG

It is important to note that in the Guaraní texts we found a considerable amount of code switches, more so than in the corpora of the two other languages. This is probably caused by the fact that a great many speakers of PG are in fact perfectly bilingual. This may have its effect on the amount and type of borrowing.

The influence of Spanish on the PG lexicon becomes evident from the overall figures in the table. Both in terms of tokens and types, the amount of Spanish items represents around one fifth of the corpus. This is just an average, however. The figures for individual speakers range from a mere 5.7% to a maximum of 28.5%. Especially the speakers from urban areas in which Jopara has its stronghold show a relatively high score. Nevertheless, the variation is much less related to age than to geographical area and the division between rural and urban areas is not clear-cut, as it has always been maintained (cf. Garvin & Mathiot 1982:29).

Spanish borrowing in PG is not restricted to the lexicon, however. A parallel development is attested in the morphology and the syntax of the language due to the long-term contact with Spanish (Gomez Rendón fc-a). One illustrative example is today's preference in Guaraní for more analytical constructions that use Spanish subordinators and deviate drastically from the polysynthetic, much less hypotactic model of classical Guaraní. Another contact-induced change – one that was not considered for statistical purposes but certainly exerts influence on the shape of the present language – is the prolific use of code switching strategies.

In order to analyze the borrowing profile in more detail, we classified borrowings according to the parts of speech they belong to in the source language. The figures from table **3** allow us to make some statements about the respective contributions of Spanish lexicon and grammar to PG.

	Verbs	Nouns	Adjectives	Manner Advs.	Others
Tokens	1842 (18.3%)	3738 (37.2%)	747 (7.4%)	95 (0.9%)	3634 (36.1%)
Types	578 (21.5%)	1223 (45.5%)	313 (11.9%)	45 (1.7%)	518 (19.3%)

Table **3**. Spanish borrowings in PG: parts of speech

Table 3 shows a hierarchical ordering of lexical classes in the borrowing process, both in terms of tokens and types, according to the pattern N > V > A > MAdj. This is in fact what

hypothesis H2.1.1 predicts. Furthermore, the contribution of the overall category 'Others' is significant and nearly equivalent to that of borrowed nouns in terms of tokens, and verbs in terms of types. It is precisely here that we find the contribution of Spanish grammatical borrowings to PG. These will be discussed in more detail in section 4.4.2. First we will have a look at major parts of speech.

4.4.1 Major parts of speech

Let us consider now the use of Spanish borrowings in the corpus in order to test the hypotheses of lexical adaptation and specialization. Loanwords were analyzed for their function in PG in order to know which syntactic slots they occupy in the target language. The following are the results for the main lexical classes.

Syntactic Function	Verbs (V)	Nouns (N)	Adjectives (A)	Manner Advs. (MAdv)
	(Tokens)	(Tokens)	(Tokens)	(Tokens)
HP	1842	8	10	
HR		3675	199	
MP			35	95
MR		54	498	
	(Types)	(Types)	(Types)	(Types)
HP	578	7	7	
HR		1214	91	
MP			18	45
MR		37	242	

Table **4**. Spanish borrowings in PG: parts of speech versus function

A first impression of these figures sets aside adjectives and nouns from verbs and adverbs. While V and MAdv are used only in their prototypical positions of head (HP) and modifier (MP) of predicate phrases, respectively, N and A are used almost in all syntactic positions. For N this is only marginally so: 98.3% of the tokens are found in the prototypical HR function; for 93.6% of the types this is the only function. Only 10 Spanish nouns function as MR and 34 as HR plus one other function. They are never found in MP position. For A the distribution is somewhat better. Prototypical use prevails here as well: 66.7% of the tokens are MR; 63.6% of the Spanish adjectives have MR as the only function while 13.7% have MR combined with one or more functions. However, 22.7% are not used in MR function at all, but mainly as HR (85.9%) or MP (11.3%).

As for the borrowing hypotheses, A seems to be the only category that gives some support to our expectations of flexibility and functional adaptation. Of the Spanish nouns, only some 3.6% are used in functions other than the prototypical one. And some of the MR uses might in fact be interpreted as possessive constructions, as in (17) below. The (rare) use of a Spanish noun as a HP is witnessed by (18), where *lado* 'side' is inflected like a verb.

(17)　ha　upépe katuĩ　　oi-kove ha　ña-hendu　*Kirito*　ñe'ẽ
　　　and　then　sometimes　3-live　and　1PL-listen　Christ　word
　　　'and sometimes they live and listen to Christ's word'

(18) la   che   gente-kuéra   che   rú-gui   o-*lado*
      DEM   1.POS family-PL   1.POS father-ABL 3.side
      'my family sides with my father'

So, there are strong arguments for functional specialization rather than adaptation of Spanish nouns in PG. Possibly, this is due to semantic factors, but we will leave the analysis of the individual word meanings out of consideration here.

 As we have already seen above, borrowed Spanish adverbs are completely inflexible in their use in Guaraní utterances. This lack of flexibility is by no means restricted to manner adverbs. Adverbs of time and place show the same distributional restriction. The other syntactic function of adverbs is the modification of modifiers. Since this use is restricted to a closed class of adverbs that do not show manner, place or time features, we have not included it here. The inflexible behaviour of adverbs might be caused, at least in the case of manner adverbs, by their morphological marking in Spanish, where manner is usually expressed by suffixing *-mente* to the respective adjectives, as with English *-ly*. Thus, for manner adverbs, the hypothesis of functional specialization seems to hold without exception.

 But even for adjectives matters are not so clearcut. When we look at some of them in more detail, the adaptation hypothesis finds even less support. Let us first look at the use of Spanish adjectives with HP function in PG. In such cases, there is always some kind of verbal marking present. In example (19) the Spanish adjective *fanatico* 'fanatic' is marked for third person. And in (20) *provechoso* 'useful' is marked for person and tense.

(19) La   mbo'ehára guaraní   i-*fanático*
      DEM   teacher   Guaraní   3-fanatic
      'That Guaraní teacher is a fanatic'

(20) I-*provechoso*-va'erã pe   i-vida   diaria-pe
      3-useful-FUT   DEM   3.POS-life daily-LOC
      'That will be useful in their daily life'

In the much more frequent use of A as HR, the forms are morphologically unmarked. This is shown in (21) and (22).

(21) Oĩ   avei i-japyté-pe   *colorado* ho'a   va'ekue   *preso*
      be.3   too   3-middle-LOC   red   fall   PST   imprisoned
      'Among them was a red one who also was put in jail'

(22) Umía *kampesino* no-entendé-i
      that peasant   NEG-understand-NEG
      'The peasants don't understand that'

All three Spanish forms appearing in these sentences, as well as the majority of the other ones attested in the corpus, although of adjectival origin, are frequently used as nouns also in Spanish syntactic contexts, without any further formal adaptation. They may well

be borrowed as N rather than A. Therefore, even for A, the case for functional adaptation is not very strong.

4.4.2 Other parts of speech
Apart from the rest of the adverbs the category labelled 'Others' in table 3 contains the following grammatical categories: articles (19.1% of the Spanish tokens), conjunctions (7.5%), numerals (1.7%), discourse markers (0.8%), adpositions (0.5%) and pronouns (0.2%). The prolific use of Spanish articles and conjunctions in PG deserves some attention. Classical Guaraní lacked the category of article and expressed definiteness by other means if necessary, for instance through the use of deictics. The borrowing of the Spanish article in PG has not resulted however in the creation of a new grammatical category. PG uses the Spanish article especially as an independent anaphoric (23), cataphoric (24) or elliptical (25) element for establishing cohesion in the discourse (cf. Gómez-Rendón fc-a). The only two forms used are *la* (feminine; 91.8% of the cases) and *lo* (non-feminine; 8.2%).

(23)   Algunoko   noñe'eiete   *la*   kastellano,
        Some.DEM   NEG.speak.NEG.very   DEM   Spanish$_{(x)}$

        oikerõ   eskuelapente   *la*   ñ-aprende-pa
        3.come.when   school.ALL.only   DEM $_{(x)}$ 1PL-learn-ALL
        'Some don't speak Spanish at all, only when we go to school, we learn it well'

(24)   *La*   oñembo'ére,   *la*   kastellano.
        DEM $_{(x)}$ 3.learn.for   DEM   Spanish$_{(x)}$
        'For them to learn it, Spanish'

(25)   Oĩ   gente-kuera   *la*   noñe'ẽseiva
        there.is   people-PL   DEM$_{(x)}$ NEG.speak.want.NEG.that
        'There are people who don't want to speak it [Spanish]'

The extensive borrowing of the Spanish article in PG, which is second only after N and before V in terms of tokens, is possibly enhanced by the existence of a set of PG demonstratives that are used much in the same way. The Spanish pronouns are either used as a demonstrative with [+distal, -visible] meaning, a shade not present in the PG demonstratives. They are also used as independent elements of an anaphoric character. In turn, these developments are closely related to the incorporation of conjunctions, the second most frequent grammatical class in borrowing. Classical Guaraní (ClG) also lacked the category of conjunctions and showed strong preference for juxtaposition over coordination and subordination. Today, however, PG makes extensive use of Spanish coordinators and subordinators. The heavy borrowing of Spanish articles and conjunctions, with the discourse functions assigned to them in PG, gives support to our hypothesis H1, which favours the borrowing of pragmatically outstanding elements. It goes counter, however, to hypothesis H2, which disfavours grammatical elements, and particularly to subhypothesis H2.1.2, which puts articles at the bottom end of the borrowing hierarchy.

Spanish numerals are relatively frequently used. Although ClG had a five-value number system that is still used in PG, values above 5 are usually borrowings from Spanish. In our data, all values between 2 and 10 are from Spanish (around 41.1% of the borrowed numerals in terms of tokens) and so are another 65 higher values. The evident limitations of the vernacular numeral system resulted in its almost total replacement by Spanish for all practical purposes.

Although adpositions are predicted to be the most likely candidate for borrowing from the category of grammatical items, their number is relatively low in PG. The number of different prepositions used is 12, with an average frequency 4 per preposition. Since ClG is a postpositional language, we assume that hypothesis H4.1 is relevant here. This presents syntax as a potential constraint to borrowing.

Finally, since PG is a pro-drop language in terms of section 4.3, one would not expect the frequent use of personal pronouns in the first place. Borrowing would be further disfavoured by the infrequent use of pronouns in Spanish, which is also a pro drop language. In fact, all 6 types found in the corpus are not personal pronouns but indefinites of the type 'someone'.

4.5 Conclusion

By way of conclusion we can say that the borrowing pattern for the major parts of speech exhibited by Guaraní supports both hypotheses H1 and H2 in terms of the expected type and token frequencies. But although predictions based on parts-of-speech types for both languages suggest functional adaptation rather than specialization, this is not confirmed for nouns and manner adverbs, and only marginally for adjectives.

As far as grammatical categories are concerned, we have seen that there are an unexpectedly high number of Spanish articles. This goes counter to hypothesis H2.1.2, which puts articles towards the bottom of the borrowing hierarchy. Also, since the category of article is absent from PG grammar at all, there might be in fact no syntactic position for articles in the language, according to hypothesis H4.1. However, we also saw that most of the articles were not used in their prototypical function, as definiteness markers in a noun phrase, but as freestanding discourse connectors. This is indicative of the fact that borrowing probabilities should be based not only on the categories of elements in the source language but also on the potential and actual functions they get in the target language. More evidence for this will be found in the case of the other two languages we have studied.

Considering the evidence above, the present profile of the language is not one of extensive borrowing only (i.e. stage 2 on Thomason's scale) but rather one of strong convergence towards Spanish. From this perspective it is not difficult to side with those who see in today's Paraguayan Guaraní a 'third language' that is neither Spanish nor Guaraní (cf. Meliá 1974).

5. The case of Quichua

Ecuadorian Quechua is a language of the Quechua family spoken in Ecuador. Like other Quechua languages, it has been in contact with Spanish since the European conquest. The long and intensive contact between Quechua and Spanish has resulted in a pervasive reciprocal influence. The present section deals with Spanish lexical borrowing in Ecuadorian Quichua (EQ). It explores the statistical results from the investigation of a

corpus of spontaneous speech collected during fieldwork in Ecuador (2003-2004). In this section, we follow the pattern of the previous section on Guaraní. Section 5.1 sketches some historical and sociolinguistic aspects of Spanish-Quichua contact in the Ecuadorian Andes. Section 5.2 addresses the classification of EQ according to its parts of speech system and other parameters. Section 5.3 presents our expectations around borrowings from Spanish in EQ according to the theoretical assumptions of section 5.2. Section 5.4 tests these hypotheses on the basis of the corpus. Some conclusions follow in Section 5.5.

5.1 Historic and sociolinguistic context of contact between Quechua and Spanish
Language contact is part and parcel of conquest and colonization settings such as those that characterize the Andean history before and after the European invasion. The Northern Andes were subject to two successive conquests by the Incas and the Spaniards in less than sixty years. Crossroads for the traffic of goods, languages and people, the Northern Andes featured a great cultural variety by the time the first Europeans arrived. Long-distance traders from Chincha in central Peru introduced Quechua in the present territory of Ecuador around the first centuries of the second millennium (Torero 2003: 93). Chincha Quechua became the source of the present Ecuadorian dialects, with a significant contribution of pre-Inca languages and other Quechua varieties brought by soldiers, officers and immigrants from several areas of the Inca Empire.[23]

Quechua along with ten other native languages were spoken in the Ecuadorian Andes by the time of the Spanish invasion. One hundred years after the foundation of the first Spanish cities in the Northern Andes, most native languages had disappeared or were on the path of extinction. Quichua was the exception. The evangelization of native peoples speaking different languages seemed feasible only to the extent that one language of widespread distribution could be used as *lengua general*, especially if that language was already in use by native peoples. Quichua met these requirements. The Third Lima Council (1583) ordered the translation of Catholic texts into a standardized Quechua resembling very much the Ecuadorian dialects (Adelaar 2004: 183).[24] In the following decades a number of grammars and dictionaries were prepared to help missionary work in the Royal Audience of Quito and other cities of the Viceroyalty of Lima.

The flowering time of Quechua came to an end by the end of the eighteenth century in the frame of the Bourbon reformation, when a series of failed Indian uprisings called the attention of the Spanish Crown to the dangers of Quechua as an instrument of agglutination and revolt. With the birth of the Andean republics in the nineteenth century Spanish remained the only official language supported by the state. Until the late 1970s Quechua was not taught in schools, and Indian pupils used to carry the heavy load of learning subjects alien to their culture in a language they scarcely heard at home. Nowadays a situation of (Spanish) monolingualism imposed on multilingual and multicultural societies prevails throughout the Andes, with the result of an ever-increasing pressure for the Hispanicization of native communities.

Another side of language contact in the Andes is socio-cultural. From the inception of the colonial regime, the Crown furthered the separation of Indians from Spaniards into what came to be known as 'the two republics'. The late seventeenth century witnessed the emergence of the *hacienda* system, based on the appropriation by the ruling elites of vast areas of land at the cost of Indian communities. This establishment contributed to isolating Quechua-speaking peoples from education and

other public institutions except for the Church. The republics founded after the Independence wars of the 1820's did not effect any changes in the situation and Indians remained to a certain extent non-integrated in the nascent nations.

Despite the varied forms of non-integration, Indians always had some type of contact with the Spanish-speaking world. In permanent need of labour for public works and private service, Spaniards and mestizos used – and continue to use – the Indian workforce in urban and rural areas. Indian labour migration became an everyday issue in the life of mestizo towns. Again, the foundation of the Andean republics in the early nineteenth century did not change the scenario. Indian migration from the countryside to the cities was even stronger in the twentieth century. In this context, the emergence of Quechua varieties heavily influenced by Spanish becomes a matter of fact.

5.2 Typological classification of Quechua

Just like Guaraní, Quechua is an example of a verb-nonverb (type 2) language, in which verbal lexemes occupy the syntactic slot assigned to predicate phrase heads, and non-verbal lexemes the other three slots. Examples (26)-(27) from Schachter (1985:17) provide the evidence.

(26)   a. Rikashaka: *hatun*-ta   b. chay   *hatun* runa
           see.PST.1SG   big-ACC   DEM   big   man
           'I saw the big one'   'that big man.'

(27)   a. Rikashaka: *alkalde*-ta   b. chay   *alkalde* runa
           see.PST.1SG   mayor-ACC   DEM   mayor   man
           'I saw the mayor'   'that man who is mayor'

In (26), *hatun* 'big' is both a referential phrase modifier (MR; 26a) and a referential phrase head (MH; 26b). Similarly, *alkalde* 'mayor' is used both as head noun (HR; 27a) and modifier of the head noun (MR; 27b).

Evidence against this classification has been recently presented by Beck (2002: 144ff) and will be discussed in the following. According to Beck, the lack of distinction between nouns and adjectives is not complete because only the latter can be modified by adverbs like *maymi* 'very' as shown in (28):

(28)   Chay warmi   maymi sumak-mi
        DEM   woman very   pretty-FOC
        'That woman is very pretty'

Interestingly, Cole gives an (ungrammatical) example of noun modification precisely with *maymi* (Cole 1985: 99-100), which was elicited by one of us in site as fully grammatical:

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

Beck's argument is meant to point out a semantic distinction between property concepts and entity concepts. Hengeveld's theory of parts of speech does not exclude such a distinction but argues for the existence of one non-specialized lexical class that embrace both concepts. As Beck himself admits: "The existence of a semantic distinction of this type is in itself not enough to establish that there is a parts-of-speech distinction between nouns and adjectives in the lexicon" (2002:144).

Beck's second argument states that noun-noun constructions are better treated as compounds on the basis that nouns acting as attributes of other nouns cannot occur more than once in the same noun phrase – as opposed to adjectival modifiers that may be recursive. As additional evidence Beck mentions that noun-noun constructions themselves may be attributives of other nouns, as in (30), from Cerrón-Palomino (1987:300).

(30)   Hara   chakra   rumi
        corn   field   stone
        'stone of the cornfield'

In (30) *hara* and *chakra* as a whole modify the noun head *rumi*. Readings such as 'field stone of corn' or 'corn field of stone' are not possible. The only possible reading considers the first two nouns as forming an attributive compound just like its English equivalent 'stone of cornfield'. A phonological test would consist in analyzing the stress pattern to know whether *hara chakra* is realized as a single word (with stress only on the penultimate syllable) or two individual words (with double stress). Unfortunately, this phonological information is not available at the moment.

Beck's third argument against the typological classification of EQ as a type-2 language is that property concept words being used apparently as heads of referential phrases as in (31) below are indeed adjectives standing for deleted heads in elliptical constructions. Conclusive proof of this interpretation is, according to Beck, "their reliance on context to supply the identity of a nominal head". Therefore, sentences like (27) would be ungrammatical if out of context.

(31)   Puka-ta   rika
        red-ACC   see.PST
        'he sees the red one'

The claim that contextual reference is always required for the correct interpretation of sentences like (31) is not conclusive either. Colour terms are universally associated to objects and do not exist independently, being to this extent context-dependent in any human language and irrelevant for a noun-adjective distinction.[25] What really matters here is not the contextual dependency of lexemes like *puka* in (31) but the fact that they occupy the syntactic slot of referential heads without any further measures and take nominal morphology, in this case the accusative marker.

Additional evidence for our classification of EQ as a type-2 language comes from the fact that both nouns and adjectives can be used as referential phrase modifiers. Consider the following examples. In (32) *yanka* 'useless' occupies the position of referential phrase modifier and, without any derivation, the position of predicate phrase modifier in (33).

(32)   Kai-ka   yanka   yura   ka-n-mi
        that-TOP   useless plant   be-3-AFF
        'that is a useless plant'

(33)   Kaina   chaupi tuta-kaman   yanka   shuya-ku-rka-ni
        yesterday middle   night-up.to   useless wait-DUR-PST-1SG
        ' yesterday I waited until midnight to no avail'

Similarly, *utka* 'speed' functions as head of a referential phrase in (34) while the same word modifies the predicate *shamui* 'come' in (35).

(34)   utka-ka   rura-shpa   alli-mi   ka-n
        speed-TOP   work-LOC   good-AFF   be.3
        'quickness is good in working'

(35)   utka   huasi-man   shamu-i
        speed   house-ALL   come-IMP
        'come home quickly'

On the basis of the above we conclude that the evidence against classifying EQ as a type-2 language is insufficient. Therefore, we will assume in the following that Quichua is a language that makes no distinction between nouns, adjectives and adverbs.

From a morphological point of view, EQ is an agglutinative language though its morphological profile has suffered from two changes involving simplification with respect to other Quechua varieties: the loss of verb-object agreement and the loss of possessive nominal suffixes. Cole (1982:6) presents the examples in (36a/b) for second-person object agreement in (Peruvian) San Martin Quechua and Ecuadorian Quechua.

	San Martín Quechua	Ecuadorian Quechua

(36)   a. Ñuka-ka   maka-yki   b. Ñuka-ka   kan-ta   maka-ni
           1SG-TOP hit-2OB   1SG-TOP 2SG-ACC   hit-1SG
           ' I hit you'   'I hit you'

As for the loss of possessive nominal suffixes, consider the following examples from Cerrón-Palomino (1987:200):

	Junin Quechua	Ecuadorian Quechua

(37)   a. maki-yki   b. kanpak   maki
           hand-2SG.POSS   2SG.GEN   hand
           'your hand'   'your hand'

The extensive use of pronouns in EQ as compared to other Quechua varieties has, in turn, resulted in the obligatory use of pronouns not only in subordinate but also in main clauses

where other varieties use them only for emphasis. To this exent we can claim that EQ developed from a pro-drop to a non pro-drop language. A parallel development in subordinate clauses in EQ is the increasing tendency to prefer coordination and relativization to the traditional strategy of nominalization, with the result of independent clauses linked by connectors instead of one main clause with embedded noun phrases.

Constituent order is rather fixed in EQ and displays an almost invariable SOV pattern. Within the noun phrase, the order is also fixed most of the time: modifier-head in attributive and possessive constructions. EQ lacks grammatical gender and the lexical category of article to express definiteness. Finally, all adpositions in EQ are postpositions, which many authors classify as members of a large set of inflexional suffixes (Cerrón Palomino 1987: 270ff).

5.3 Predictions about Spanish borrowings in Quichua

Since Quichua was classified as a type-2 language, just like Guaraní, we can repeat the predictions for borrowing here. Thus, Quichua will borrow N, V, A and MAdv in that order of frequency. Again, if functional adaptation applies –it did not for Guaraní – we will find Spanish nouns, adjectives and manner adverbs in all relevant functions, i.e. as heads and modifiers of referential phrases and as modifiers of predicate phrases. If functional specialization applies – as it did more or less for Guaraní – we will find N, A and MAdv only as HR, MR and MP, respectively.

In relation to other categories, it may be hypothesized that Spanish borrowings in EQ do not include prepositions and articles, as there is no syntactic position for them in the language. On the other hand, pronouns might be borrowed on the basis that EQ has become a non pro-drop language though this tendency may be disfavoured in the first place by the fact that Spanish is a pro-drop language itself. Finally, coordinators and subordinators are also expected to be borrowed if we consider the recent developments in EQ discussed in section 5.2.

5.4 Spanish borrowing in Quichua: the data

Here we will present our observations about borrowing from Spanish in Quichua and again test our findings on the basis of the relevant borrowing hypotheses. The Quichua corpus consists of spoken texts provided by 25 speakers, with a total length of 79.718 tokens. Table 5 contains the figures for Quichua and for Spanish borrowings in terms of tokens, types and type-token ratio.

N=38	Target Language (EQ)	Source Language (Sp)	Total
Tokens	64620 (81.1%)	15098 (18.9%)	79718
Types	19023 (83.5%)	3768 (16.5%)	22791
TTR	0.29	0.25	0.28

Table **5**. Spanish borrowings in EQ

Looking at these global figures, the influence of Spanish on EQ lexicon is close to 20% of the corpus in the case of tokens. The percentage is somewhat less, however, for types. We will come back to a comparison between the languages in section 7. The figures for the individual speakers range from only 4.0% of the tokens to a stunning maximum of 49.1%. While this percentage is evidence of a strong lexical influence from Spanish, it is by

no means uncommon to each of the Ecuadorian dialects of Quechua. Arguably there exists a correlation between this percentage and some changes in EQ morphosyntax along the borrowing scale proposed by Thomason (2001). EQ uses a few Spanish morphemes productively and shows at the same time syntactic deviations such as inverted possessive constructions and verbs in medial position.

For the respective parts of speech, we found the following numbers.

	Verbs	Nouns	Adjectives	Manner Advs.	Others
Tokens	2672 (17.7%)	8220 (54.0%)	1276 (8.5%)	88 (0.6%)	2842 (18.8%)
Types	600 (15.9%)	2037 (54.0%)	514 (13.6%)	37 (1.0%)	580 (15.4%)

Table **6**. Spanish borrowings in EQ: parts of speech

Again as predicted by hypothesis H2.1.1, there is a hierarchical ordering of lexical classes in the borrowing process, both in terms of tokens and types, according to the pattern N > V > A > MAdj. However, the position of N is much more outstanding than in the case of PG, mainly at the cost of the category 'Others'. The lexical and grammatical borrowings will be discussed in sections 5.4.1 and 5.4.2 respectively.

5.4.1 Major parts of speech
Analysis of the borrowed items in terms of Spanish parts of speech and functions in Quichua syntax gave the overall results in Table 7 below.

Syntactic Function	Verbs (V)	Nouns (N)	Adjectives (A)	Manner Advs. (MAdv)
	(Tokens)	(Tokens)	(Tokens)	(Tokens)
HP	2668	58	16	
HR	3	7618	287	
MP		141	131	88
MR	1	392	841	
	(Types)	(Types)	(Types)	(Types)
HP	596	20	10	
HR	3	1767	128	
MP		29	49	37
MR	1	215	327	

Table **7**. Spanish borrowings in EQ: parts of speech versus function

When we analyze the functional aspects of Spanish borrowings in EQ for the respective parts of speech, we find the following. Verbs are typically used in their HP function only. For adjectives and nouns the situation is quite different. Although the prototypical functions prevail here as well, with 87% (N) and 64% (A) respectively, there are much more than marginal frequencies for the other three functions. Almost 11% of the Spanish nouns are used as nominal modifiers while 24.9% of Spanish adjectives function as nominal heads. A substantial 9.5% of Spanish adjectives operate as verbal modifiers. Let us first look at some examples of the use in EQ of Spanish nouns.

(38)   Chayka   chicha-ka   *aidante*-pak-ka
        then   chicha-TOP   assistant-ALL-TOP

jala-shka   shina   *barro*   manga-pi
       carry-PFT   so   earth   pot-LOC
       'the chicha was then carried in an earthen pot to the assistant's house'

(39)   Ashtawan *juirza*   baila-k   ka-rka
        more   strength   dance-DUR   be-PST
        'they were dancing more intensely'

Example (38) shows the use of Spanish noun *aidante* (from *ayudante* 'assistant') as HR and *barro* 'mud' as MR of the Quichua head *manga* 'pot'. And (39) illustrates the use of the Spanish noun *juirza* (from *fuerza* 'strength') as a manner adverb modifying the predicate head *bailak* 'dance'.

       The next examples illustrate the different uses to which Spanish adjectives are put in EQ other than their prototypical syntactic function of referential phrase modifiers. Thus, the Spanish adjectives *antiwo* 'old' and *pariju* 'equal' are used respectively as head of a referential phrase (40) and modifier of a predicate phrase (41).

(40)   Chasna *antiwo*-ka   ka-rka
        so   old-TOP   be-PST
        'the old (use) was like that'

(41)   Sankuan-ta-ka   paywan   *pariju*   baila-nchik-mi
        Sanjuan-ADV-ka   3SG.INST   equal   dance-1PL-FOC
        'at the Saint John's Festival we dance all together'

Spanish *antiwo* in (40) stands for 'that which is old' and refers to a distant past when some implicit referent was in use. In (41) the adjective *pariju* (from *parejo*, 'equal', 'from both sides') occurs in the function of a predicate phrase modifier, pointing out the manner in which people danced during Saint John's festivals, i.e. men and women danced together on that occasion.

       While Spanish nouns and adjectives are rather flexible in relation to their syntactic functions in EQ, this is again not the case for manner adverbs. There is not a single example of a manner adverb in a syntactic position different from the canonical MP. As suggested already for PG, this might be due to the morphological markedness of Spanish adverbs with the derivational suffix *-mente*. The majority of the borrowed manner adverbs are of this type (29 out of 37 types). The fact that these overtly marked adverbs are used only as predicate phrase modifiers might suggest that bilinguals are aware of the function of the adverbializer *-mente* in the source language and use all corresponding lexemes exclusively in the syntactic slot of predicate modifiers. If this is true, then bilingual processing seems to play a crucial role in the borrowing process and may overrule other factors.

       While 'bare' manner adverbs are relatively few in the corpus (88 appearances), EQ shows preference for complex Spanish constructions in order to indicate manner. Thus, we find prepositional phrases like *de repente* 'suddenly', *a propósito* 'on purpose', *por ley* 'obligatorily', *a la carrera* 'rapidly' and many others acting as adverbials (cf. example 42 below). Counted as single lexical entities, which is probably the way they should be interpreted syntactically, these constructions cover around 1.0% of the borrowed tokens.

Apart from that we find quite a few Spanish adjectives used as predicate phrase modifiers without further measures being taken: 29 types with 141 appearances.

5.4.2 Other parts of speech

Apart from a small set of adverbs (around 2.5% in terms of types and tokens), the category labelled 'Others' in table 6 contains the following grammatical categories: conjunctions (6.9% of the tokens), interjections (1.4%), numerals (0.9%), discourse markers (0.6%), and prepositions (0.5%). Pronouns and articles (0.1%) are very rarely borrowed, two categories which were very frequently borrowed by PG. In other respects, the picture is roughly the same.

There are a total of 17 different Spanish prepositions almost always used as part of fixed adverbial expressions but not as prepositions themselves. Example (42) illustrates this:

(42) ñuka-ka   *por-gusto*   ri-ni   ufia-ngapa
      1SG-TOP   by-pleasure   go-1SG   drink-PURP
      'I gladly go to drink'

Also here, the syntax seems to provide a constraint, not on borrowing as such, but on employing the element in an 'impossible' syntactic slot.

The same motivation as for PG could be given for the presence in the corpus of around 70 different types of numerals from Spanish, i.e. the pragmatics of negotiation and the fact that, in financial transactions, Spanish is often the language of communication.

The use of conjunctions can be easily linked to certain syntactic developments in EQ where nominalization strategies are being replaced by relativization (hence also the introduction of some pronouns, mainly relative and interrogative) as the preferred mechanism of clause linking (cf. Lefebvre & Muysken 1988; Gomez-Rendón fc). Consider the use of the Spanish coordinator *y* 'and' in (43) where traditional Quichua uses simple juxtaposition or the additive suffix on both coordinated elements. Likewise, observe the use of the Spanish relative conjunction *que* 'that' in (44) which connects the otherwise nominalized subordinate clause to the predicate *nini* 'I say':

(43) ñuka ayllu kanchis-pura kan, pichca churi ñuka *y*   ñuka warmi
      1SG family seven-among be.3 five   child 1SG CONJ 1SG wife
      'seven people compose my family: five children, me and my wife'

(44) ñuka ñuka japi-shka-manta   nini *que* ñukanchik
      1SG 1SG   understand-PART-ABL say.1 that 1PL

      llakta-kuna-pi   yachai-kuna-ka   tiya-n
      community-PL-LOC   idea-PL-TOP   exist-3
      'from what I understand I say that there are ideas in our communities'

5.5 Conclusion

Also Quichua gives support to both hypotheses H1 and H2 in terms of the expected relative type and token frequencies. In contrast to Guaraní, however, there is some

support for the functional adaptation strategy in favour of the specialization strategy, be it only for nouns and adjectives. Manner adverbs behave rather inflexible again.

The way Spanish prepositions are employed in Quichua syntax is further support for the fact that both the part of speech of a borrowed element in the source language and the function assigned to it in the target language should be taken into consideration when analyzing borrowing effects.

From the social and historical background in Section 5.1 and the linguistic data we collected, it is safe to identify EQ with the third stage of Thomason's borrowing scale: a fairly high degree of bilingualism among the linguistic community (Büttner 1993:48-49) and lexical borrowing of non-basic and basic vocabulary. Contact-induced changes in the lexicon include the borrowing of a significant percentage (> 20%) of vocabulary, including Spanish conjunctions and numerals. Contact-induced changes in the structure include Spanish-derived SVO word order instead of typical Quechua SOV; subordinate constructions headed by Spanish conjunctions and SVO clauses, instead of typical Quichua participial constructions; and a few Spanish inflectional suffixes of gender and number.

6. The case of Otomí

Otomí is a language from Central Mexico, which since the European conquest has been in contact with Spanish and has undergone a pervasive influence from that language. Below we describe the Spanish lexical and grammatical borrowings in two Otomí dialects spoken in the State of Querétaro, viz. the dialect of Santiago Mexquititlán and the dialect of San Miguel. Both belong to the branch of north-western Otomí, one of the larger variants of Otomí with around 33.000 speakers. Santiago Mexquititlán is located in the southern part of the Querétaro, in the mountains of the neovolcanic axis of Mexico, and is a town with a population of around 15.000 inhabitants. The Otomí spoken in Santiago Mexquititlán is similar to the Otomí dialect of the villages in the north of the State of Mexico. San Miguel is situated in the northern part of Querétaro, in the semidesert of the Sierra Madre Oriental, and is a village with a population of around 700 inhabitants. The Otomí spoken in Tolimán is similar to the Otomí dialect of the Valle del Mezquital in the State of Hidalgo. In both villages the vast majority of the population are Otomís. For our research we base ourselves on a corpus that consists mainly of spontaneous speech collected during our fieldwork between 1993 and 2004.

In section 6.1 we describe the historic and sociolinguistic context of the contact between Otomí and Spanish. In section 6.2 we give a typological classification of Otomí according to its word order, lexical parts-of-speech system, syntactic relations and numerical system. On the base of that classification section 6.3 presents some predictions about what changes may be expected in Otomí as a result of its contact with Spanish. Section 6.4 will discuss what has been attested in our corpus. In 6.5 we will draw some conclusions.

6.1 Historic context of the contact between Otomí and Spanish
Otomí is a central amerindian language, currently spoken by around 220.000 mostly bilingual speakers on the highlands around Mexico City in the States of Mexico, Hidalgo, Querétaro, Puebla, Guanajuato, Tlaxcala, Veracruz and Michoacán. Together with Mazahua, Ocuilteca, Matlatzinca, Pame and Chichimeca it constitutes the Otopame group within the Otomanguean family. It was the native language of the old inhabitants of the

Valley of Mexico and its surrounding valleys. Throughout history its speakers had to confront the Aztecs, Spaniards and Mestizos, speakers of Nahuatl and Spanish, both belonging to other language families. Since the Otomís had to surrender to the Nahuas from the 15th century onwards, there has been a very close contact between the Otomí and Nahuatl languages. During that contact the Nahuas developed a very negative image of the Otomís, which later was passed on to the colonial chroniclers, such as Sahagún (1989 [1557]), whose Nahuatl speaking informants considered the Otomíes "toscos e inhábiles" (coarse and unskillful). The very fact that the word Otomí probably is a derivation of the Nahuatl word *totomitl* 'bird hunter' is an example of the negative image imposed by the Nahuas (Jiménez Moreno 1939).[26] The Otomís like to call their language themselves Ñäñho, Ñöñhö o Ñähñu, i.e. 'he who speaks well' and their language Hñäñho, Ñhöñhö or Hñähñu (Hekking 1995). Since the Otomís were the second most numerous group after the Nahuas on the highlands, the Spaniards were strongly interested in their conversion to catholicism. Despite the fact that the language was considered very difficult because of its complex vowel and consonant system, a spelling system for Otomí was developed and grammars, vocabularies, catechisms and legal documents were written in it. Especially missionary friars of the Franciscan order have studied the Otomí language, above all by Fray Alonso de Rangel, who arrived in Mexico in 1529 and died in 1547. He published the first *Arte i Doctrina Cristiana en lengua Otomí*. Another Franciscan who deserves mentioning is Fray Pedro Cárceres, who wrote his *Arte de la lengua Otomí* in Querétaro around 1580 following the spelling developed by Fray Alonso de Rangel. These spelling rules were improved by Fray Alonso Urbano, who wrote the linguistically important trilingual *Arte breve de la lengua otomí y vocabulario trilingüe, español-otomí-náhuatl* at the beginning of the 16th century. Sociolinguistically very interesting is the *Códice Martín del Toro* from the second part of 17th century. This document aims to show that Pedro Martín del Toro, the leading figure of the text, belongs to a lineage of Otomí nobles who militarily had formed an alliance with the Spaniards. In it there is a constant code switching between Otomí and Spanish (Guerrero 2002). It should be mentioned that the colonial documents written in Otomí are not easily readable since the authors not always distinguish the Otomí phonemes that do not exist in Spanish from those that do, specifically some vowels. Only trained linguists with a good command of Otomí are able to read them, often with difficulty as in the case of the *Códice Martín del Toro*.

After the Independence of Mexico in 1813 major changes occurred in the indigenous community. Officially the indigenous groups were no longer recognized. Many Otomís could not afford their education any more and Otomí was no longer written by the civil authorities, only by a handful of scholars. It was in the nineteenth century that a process of language shift started.

The Mexican Revolution, which lasted from 1911 to 1917 did not result in social change for the Otomí population, nor did it foster more recognition of their language, or stopped the language shift. On the contrary, after a long history in which they have been degraded socially little by little, the Otomís now belong to the lowest social levels of the Mexican society, as is the case for many other indigenous groups in Mexico. They dwell in the most remote and less fertile places on the highlands living on an agriculture of subsistence, reason why many of them have chosen to emigrate to the bigger towns, such as Mexico City, Guadalajara and Monterrey. In the 20th century several attempts have

been made to integrate the indigenous communities in the national community by means of the officially called *Educación Bilingüe* given by indigenous teachers. These, however, often possess a very negative attitude towards their own roots and have a complete lack of knowledge about bilingual education. Consequently, most Otomís are illiterate in their first language and very often have insufficient command of the standard variety of Mexican Spanish. Otomí is only spoken inside informal domains such as the family.

During the last 20 years, because of the construction of roads and schools, the growing influence of the media and the increasing trade and emigration, contact between the relegated Otomís and the Spanish speaking Mestizos has dramatically increased. As a result, a rapid increase in contact phenomena from Spanish is observed in the respective varieties of the Otomí language (Hekking 1995, 2001, 2002; Hekking & Bakker 1998, 1999, 2006). Because of the fact that Otomí is a stigmatized language, only spoken by poor and traditional people, many Otomís do not want to teach the indigenous language to their children any more. This makes Otomí an endangered language, despite its current high number of speakers.

6.2 Typological classification of Otomí

Otomí has many lexical elements which specialize as verbs or nouns. Although nouns may be more or less freely used as the main predicate of a clause, and then may be marked for tense, aspect and person an number of the subject, just like the verbs, only they may be nominal heads of noun phrases and are then obligatorily preceded by an article or a possessive. On the other hand, only verbs may be marked for object, for a beneficiary, a locative or by a reflexivity marker. This is exemplified in (45).

(45) Ar   hyonguu   bi   hyom-bi   ar   nguu   ar   majä.
      DEF.SG bricklayer PST.3 build-BEN DEF.SG house DEF.SG priest
      'The bricklayer built the house for the priest.'

So, Otomí is a rigid language in terms of Hengeveld (1993) beyond dispute. The major question is whether Otomí has adjectives. Many words that would translate to adjectives in Spanish and English, are nouns in Otomí, since they are preceded by an article. An example is *'ra'yo* 'new'. Others such as *dathi* 'ill', *txutx'ulo* 'small' and *johya* 'happy' are of a verbal nature since they are obligatorily preceded by verbal proclitics such as the continuative aspect marker (Lastra 1997:30f). However, yet another group with typical adjectival meanings such as 'thin', 'thick', 'bitter', 'sweet', 'cold', 'warm', 'beautiful', 'ugly', 'good', 'high', 'low' and colour words are found in bare form in an adnominal position, and are therefore considered to be adjectives by a number of linguists (cf. Ecker 1952; Hekking & Andrés de Jesús 1984; Voigtlander & Echegoyen 1985; Lastra de Suárez 1992, 1994, 1997; Andrews 1993; Bartholomew 2004). [27] This is exemplified in (46) below (from Lastra 1997:61).

(46) a. Ar   'bo   'yo   b. Ar   t'axi   nguu
         DEF.SG black dog   DEF.SG white house
         'the black dog'   'The white house'

Although it is generally acknowledged that there are less adjectives in Otomí than for example in Spanish or English, there seems to be enough reason to assume that there is at least some specialization for the A category. We also find some elements which operate as manner adverbs in Otomi. Examples are found in (47a/b).

(47)   a. Bi   ma   *'nitho*   ndi metu̱-we   ár   'be̱hñä
             PST.3 leave *quickly* 3   tell-1.INCL   POSS.3SG   wife
             nä'ä   xki   thogi
             what   PSTPRF.3   happen
           'He left quickly in order to tell his wife what had happened.'

         b. Gi ñuni *njante*, gi jänt'i   hingi   ñähi gi ñuni komo-ngu gar   mi'ño.
            2   eat   *slowly* 2   swallow NEG.2 chew   2   eat   like-like   3   coyote
            'Eat slowly, you eat without chewing, like a coyote'

However, they are about the only examples of such use. Typically, a verbal construction is found where English or Spanish would use an adverb, as in (48).

(48)   Ma   da̱   hingi   handi *xí hño*   ja-r   'be̱xui.
        POSS.1 eye   NEG   see   3   good   LOC-DEF.SG   darkness
        'I can't see very well in the darkness'

This leads us to the categorization of Otomí as a type 5 language with the extra stipulation that the class of adjectives is only semi-open in the sense that it does not allow easy extension.

Just like the other Otomangue languages, Otomí is a tone language. It distinguishes three tones: a high, a low and a rising tone. The basic constituent order is VOS, a common trait of Otomangue as well (Suárez 1983; Yasugi 1995).

From a morphological point of view, Otomí has a rather complicated synthetic structure on the lower syntactic levels, in particular the noun phrase and the verbal complex. But at the same time Otomí has an analytical structure at the sentence level and is a language with asyndetic compounding or juxtaposition of constituents, with very few explicit markers of the semantic or syntactic relations between these constituents, such as prepositions, coordinators, subordinators and relatives, as can be seen in example (49).

(49)   Nu'bya di   ne   ga   tso̱ni Nxu̱ni.
        today   1   want FUT.1   arrive Morelia
        'Today I want to arrive at Morelia'

Some relations are marked on the verb, such as the beneficiary in (45), and the suffix *-wi* in (50) which marks accompaniment.

(50)   Ar   Xuwa   mi   ñä-wi   ár   to.
        DEF.SG   Juan   IMPRF.3   speak-ACMP POSS.3SG   mother-in-law
        'Juan spoke with his mother-in-law'

Otomí does not have a full class of adpositions. However, the particles *dige*, *ja*, *ir nge* and *ngu* mark a variety of relations between noun phrases and other parts of the sentence. Their exact interpretation depends heavily on the context. See the use of *dige* in (51a/b):

(51)   a. Di   ñä-he   dige ma   b<u>o</u>ni-he   pa Maxei.
          1   speak-1PL.EXC about POSS.1   trip-1PL.EXC   to Querétaro
          'We speak about our trip to Querétaro'

       b. 'Nar   jä'i   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'

Otomí has several coordinators (*ne* 'and', *'nehe* 'also') and subordinators (*wa* 'or', *jange* 'therefore', *ngetho* 'because', *ngu* 'as', and *nu'b<u>u</u>* 'if', 'when') at its disposal. However, the prevailing form for both coordination and subordination of clauses is asyndetic juxtaposition.

In Classical Otomí a relative clause, which is always in a postnominal position, generally is attached to a main clause according to the gapping strategy without any form of connection (cf. Comrie 1989:147f). It is possible to relativize on any part of the relative clause, that is to say not only on the Subject, the Direct or Indirect Object, and the possessor in a possessive construction, but also on the accompaniment, the beneficiary, the instrumental and the locative.

Like many Mesoamerican languagues Otomí uses a vigesimal system for its numerals, a system that according to Bartholomew (2000) was adopted by the Nahuas, but not by the Spaniards. Otomí numerals often function as verbs or as nouns. When used as a noun they are preceded by the nominal clitic *ar* and when used as a verb, as shown in (52), they are accompanied by verbal proclitics and suffixes.

(52)   Di   hñu-he.
       1   three-PL.EXC
       'We are three'

6.3 Predictions about Spanish borrowings in Otomí
With Otomí a type 5 language and Spanish a type 4, we expect easy borrowing for classes V and N under functional specialization, i.e. V in HP function and N in HR function. Adjectives will probably be borrowed, only as MR. It remains to be seen whether this will happen at a large scale, since the class of adjectives seems to be only semi-open in Otomí. Spanish manner adverbs will not be borrowed since the language lacks the category and a real MP slot.

For the other categories, it may be expected that prepositions will be borrowed, since a slot for them seems to be present in the form of particles with a corresponding function. The same goes for coordinators and subordinators. Furthermore, we expect that many speakers will use Spanish numerals for values higher than 10 since the vigesimal system is unpractical in the hispanic world. Finally, we assume that we will quite frequently find Spanish discourse markers, typically as first elements of a clause.

6.4 Spanish borrowing in Otomí: the data

For Otomí we collected data from native speakers of two rather different dialects, 31 from Santiago Mexquititlán (OS) and 28 from Tolimán (OT). In both cases, all speakers were interviewed on the basis of a series of questions about the sociolinguistic and cultural situation in the two communities. Furthermore, the same informants were asked to make a translation of 88 Spanish sentences into Otomí. In all we collected a corpus of 110.541 tokens. Table 8 gives an overview for the two dialects. In these totals, no distinction is made for the two types of elicitation.

N=31	Target Language (OS)	Source Language (Sp)	Total
Tokens	61693 (87.5%)	8841 (12.5%)	70534
Types	5344 (81.0%)	1255 (19.0%)	6599
TTR	0.09	0.14	0.09
N=28	Target Language (OT)	Source Language (Sp)	Total
Tokens	33277 (83.2%)	6730 (16.8%)	40007
Types	2672 (72.2%)	1029 (27.8%)	3701
TTR	0.08	0.15	0.09
TOTAL (N=59)	Target Language (O)	Source Language (Sp)	Total
Tokens	94970 (85.9%)	15571 (14.1%)	110541
Types	8016 (77.8%)	2284 (22.2%)	10300
TTR	0.08	0.15	0.09

Table **8**. Spanish borrowings in OS, OT and O

These is an overall borrowing rate of around 14% in terms of tokens, with the OS variant slightly lower (12.5%) and the OT variant slightly higher (16.8%). In both cases, the percentage for the types is considerably higher. This is probably due to some amount in variety in the spelling of the Spanish loanwords, not fully catered for by the lexicon we constructed. For the individual speakers the figures range from 6.7% (OS) and 10.7% (OT) to 20.1% (OS) and 26.0% (OT). The differences betwee the two dialect may be explained on the basis of the fact that Tolimán is a very small community indeed, with very intensive contact with the Spanish speaking world.

When we look at the two data sources, we find that the amount of borrowing is less in the spontaneous data than in the translation data. This is shown in table 9. We restict ourselves to numbers of tokens, and make no differentiation for the two dialects.

N=59	Target Language (O)	Source Language (Sp)	Total
Interviews	52447	7623 (12.7%)	60070
Translation task	42523	7948 (15.8%)	50471
Total	94970 (85.9%)	15571 (14.1%)	110541

Table **9**. Spanish borrowings in Otomí for two types of elicitation

In Otomí not only contents words are borrowed, but also a high amount of function words. The adoption of some of these function words makes that some synthetic constructions in the Otomí grammar are changing into analytical constructions. For example, the Spanish loanword 'ko' is replacing the Otomí suffixes *-hu̱, -wi, -he* and *-'be* (Hekking 1995). Furthermore it has been observed that basic constituent order VOS is

frequently replaced by SVO, the basic order of Spanish (Lastra de Suárez 1994, 1997; Hekking 1995). Tone does not seem to have been affected so far.

Looking in more detail at the parts of speech of the lexical borrowings, we find the following totals.

	Verbs	Nouns	Adjectives	Manner Advs.	Others
Santiago					
Tokens	400 (4.5%)	3478 (39.3%)	158 (1.8%)	28 (0.3%)	4777 (54.0%)
Types	118 (9.4%)	583 (46.5%)	43 (3.4%)	10 (0.8%)	501 (39.9%)
Tolimán					
Tokens	344 (5.1%)	2856 (42.4%)	137 (2.0%)	5 (0.1%)	3388 (50.3%)
Types	115 (11.2%)	505 (49.1%)	53 (5.2%)	4 (0.4%)	352 (34.2%)

Table **10**. Spanish borrowings in O: parts of speech

As predicted by hypothesis H2.1.1, also for Otomí borrowing takes place according to the pattern N > V > A > MAdj. The contribution of the overall category 'Others' is very significant indeed. In terms of tokens it surpasses even the total of the four major categories. In terms of types only nouns have a higher percentage. As can be seen from table 10, there is no striking difference between the two dialects. For most analyses we will therefore aggregate the data of all 59 respondents.

Apart from separate lexical entities, we have found a huge amount of Spanish phrases, which are not to be considered as code switches but as composite or frozen borrowings, because they appear in our corpora several times and for several informants. Most of them are noun phrases such as: *barryo kinto* 'neighbourhood five', *el beintisinko de julyo* 'the twenty fifth of July', *la mera verdad* 'the pure truth' and *rweda de fortunä* ,'wheel of fortune'.
We will discuss some details in terms of lexical and grammatical borrowings in sections 6.4.1 and 6.4.2, respectively.

6.4.1 Major parts of speech
For the borrowings belonging to the four major parts of speech in Spanish we found the functional distributions given in table 11.

Syntactic Function	Verbs (V)	Nouns (N)	Adjectives (A)	Manner Advs. (MAdv)
N=59	(Tokens)	(Tokens)	(Tokens)	(Tokens)
HP	725	37	120	
HR		6275	11	
MP			1	31
MR		7	162	
	(Types)	(Types)	(Types)	(Types)
HP	191	24	53	
HR		880	4	
MP			1	10
MR		3	25	

Table **11**. Spanish borrowings in O: parts of speech versus function

As for the other languages, verbs are borrowed as the second largest of the four lexical classes, and exclusively in the prototypical function of head of a predication. Spanish verbs are always borrowed in the third person singular present (the stem form) and may then be accompanied by any Otomí verbal proclitic or suffix. Both the bare stem form is found, as in example (53a) and the diphtongued version some verbs get in Spanish when the corresponding syllable gets stress (53b).

(53) a. Bi   ma bi   *konta*-wi yá   kompañera nu-'u̠
          PST3 go PST.3 tell-lINC POS.3 friend   DEM-3PL
          o̠t'-ar   wela.
          do-DEF.SG grandmother
          'She visited her girlfriends to tell them what her grandmother had done'

      b. Nu-ge   bwelo-ga̠   mi   *kwenta* ke hä mi   ting-ar
          DEM-DEM grandfather-1 IMPRF.3 tell   that yes IMPRF.3 find-DEF.SG
          oro   'ne-r   plata.
          gold also-DEF.SG silver
          'My grandfather used to tell that he used to find gold and also silver'

Nouns are borrowed in high numbers, and overwhelmingly as heads of referential phrases. This is in full compliance with the rigid type assigned of Otomí, with its specialized noun class. Spanish nouns are always adopted by Otomí in the singular, and but be accompanied by any Otomí nominal proclitic or suffix.

   That we find Spanish adjectives, though not in large numbers, is also expected on the basis of the class 5 categorization of the language. What is striking, however, is that Spanish adjectives, unlike V and N do not specialize for their prototypical function, i.e. MR. Quite a large number (40% of the tokens, 64% of the types) function as predicates. They are accompanied by verbal (53a) or nominal (53b) morphology, often according to the treatment the corresponding vernacular word would get.

(53)   a. ... nu'bya dige-'bya   ya   bi   m-*bibo* ...
               now   with.respect.to-now already PST.3 PST.3-smart
            '... nowadays they became smart ...'

       b. ya   na *pobre* i   hinti   mi   pe̠ts'i nada.
           DEF.PL INT poor and nothing IMPRF.3 have   nothing
           'They were very poor and had nothing.'

When used as a nominal modifier in Otomí, Spanish adjectives appear in their bare form, typically with the male suffix –*o*.[28] Interestingly, they may be found in prenominal position, following Otomí syntax, but also postnominally, following Spanish syntax.

   Finally, the borrowing of several manner adverbs, though not very frequent, goes counter to our predictions. The borrowing of adverbs is definitely not uncommon in Otomí. A further 663 Spanish adverbs (tokens; 61 types) of other categories (Time, Place) are found in our corpus.

If we subtract the lexical elements found under Other in table 10 above, we are left with a large number of grammatical borrowings: 7502 of 15.571 borrowed tokens, i.e. 48.2 %. By far the most frequent category among these are Spanish prepositions. We found 3305 tokens for this category, or 21.2%, placing this category second after the nouns. The majority of these are used in their function of relating a noun phrase to another constituent. Around a third of the prepositions are markers of a subordinate clause. In fact, these may be derived not from the Spanish prepositions as such but from the complex Spanish subordinators consisting of a preposition followed by the general subordinator _que_ 'that', as in _porque_ 'because' and _para que_ 'so that'. An example of this use is found in (54).

(54) Ñä   nts'e̲di, _pa_ da   'yo̲de xi hño ya   jä'i.
      speak strong   for FUT.3 hear 3   good DEF.PL person
      'Speak loud, so that the people hear you well'

After the prepositions follow coordinators such as _i_ 'and', _o_ 'or' and _pero_ 'but' (7.5%), and discourse markers such as _pues_ 'well' and _este_, a demonstrative used as a continuation marker (6.5%). The fourth category are subordinators like _aunque_ 'although', _porque_ 'because' and _de que_ 'so that' (6.1%, 40 types). Relatively few pronouns are found, of four different subtypes: personal, impersonal, interrogative and relative (0.5%). Finally, we found 42 different numerals, totalling to 0.9% of the Spanish tokens. All values under 10 are among them. Coordinators, subordinators, discourse markers, numerals and pronouns are all used mainly or exclusively for their prototypical functions. Spanish articles, certainly a category in Otomí, are not attested in our corpus.

To sum up, most predictions concerning grammatical borrowing got support from the actual data, the use of lower numerals being an exception.

6.5 Conclusion

As in the case of both Guaraní and Quichua, borrowing hypotheses H1 and H2 find some support in terms of the relative token and type frequencies, be it to a somewhat lesser extent. This is mainly due to the high frequency of borrowed prepositions, not only in terms of tokens but also in terms of types. An explanation may be that they Otomí has several particles which function like prepositions, and has in fact a syntactic slot for that category. This is in line with hypothesis H4.1. Given that Otomí leaves many relationships between noun phrases and other constituents unmarked, especially at the clause level, the wide variety of Spanish prepositions may provide an extra resource in this respect. The same may be said about the category of manner adverbs.

An interesting phenomenon is the fact that borrowed adjectives are used not only as modifiers in noun phrases, but also as heads of predications. In terms of types this is even the prevalent use. Per word type, however, the function is almost always unique. Three types are used both as MR and HP. These are _blanco_ 'white', _viejo_ 'old' and _rico_ 'rich'. Like other Spanish adjectives, they are regularly found in 'headless' noun phrases, and may therefore bee analyzed as nouns rather than adjectives, or as homonymes for that matter. So we may conclude that by far most adjectives which are borrowed specialize,

be it not necessarily for the prototypical function. This is not what the part of speech theory would predict.

The basic constituent order, VOS in Classical Otomí, is being replaced by SVO, which is the basic order of Spanish, a marked though existing order of the classical language.

Finally, given the huge number of Spanish function words that are borrowed, such as prepositions, coordinators and subordinators, with the concomitant restructuring of subordinate clauses, and the ongoing change in basic clause order, we estimate that Otomí is situated somewhere between point 2 and 3 on the borrowing scale of Thomason (2001).

7. Comparing the results

In the previous sections we analyzed the contact data we collected for three individual, typologically quite different languages, viz. Guaraní (PG), Quichua (EQ) and Otomí (O). As observed earlier, much of what happens in language contact situations follows the same more or less general patterns. The differences are quantitative rather than qualitative. Therefore, comparison is an absolute necessity. Therefore, in this section we will compare the results of the preceding sections in the light of the hypotheses of section 3, and see to what extent the typological differences between the languages throw any light on the differences we may find. Let us first have a look at the overall number of borrowings. Table 12 displays the totals of tokens borrowed from Spanish found in the corpora of the three languages.

	Quichua (EQ) N=25	Guaraní (PG) N=38	Otomí (O) N=59
Source	15098 (18.9%)	10056 (17.4%)	15571 (14.1%)
Target	64620 (81.1%)	47772 (82.6%)	94970 (85.9%)

Table 12. Comparison between borrowed tokens

Although the differences seem to be slight on first view, they turn out to be significant on a χ^2 test at the 0.5% level. This would imply that in terms of language use, the presence of Spanish is greatest in EQ, somewhat less in PG and least in O. It is interesting now to compare the figures for the speakers with the lowest and the highest percentages of Spanish tokens for the three target languages. This is done in table 13.

	Quichua (EQ)	Guaraní (PG)	Otomí (O)
Mimimum	4.0%	5.7%	6.7%
Maximum	27.0%	28.5%	26.0%
Mean	18.9%	17.4%	14.1%
SD	8.92	6.42	3.97

Table 13. Minimum and maximum Spanish tokens

So we see that, although the borrowing percentages are slightly different for the three languages, the minimum and maximum percentages do not differ much. The values for the standard deviation suggest that in the case of Otomí the borrowing percentages are somewhat closer to the mean than for the other two languages.

Now let us look at the figures for the individual parts of speech. In table 14 below we put the figures for the categories which make a more than marginal contribution to the total of the borrowings for at least one of the languages. Note that under the category Adv we have aggregated all adverbs, including MAdv.[29] Outlying values are in bold.

	Quichua (EQ)	Guaraní (PG)	Otomí (O)	OS – OT
Art	0.0%	**19.4%**	0.0%	0.0 – 0.0%
Adpos	0.5%	0.5%	**21.2%**	17.8 – 25.8
V	17.7%	18.3%	**4.8%**	4.5 – 5.1%
DsMrk	0.6%	0.8%	6.5%	7.2 – 5.7%
A	8.5%	7.4%	1.9%	1.8 – 2.0%
N	**54.4%**	37.2%	40.7%	39.3 – 42.4%
Subj	1.6%	4.6%	6.1%	6.8 – 5.3%
Conj	6.9%	4.4%	7.5%	9.1 – 5.3%
Adv	3.4%	2.4%	4.5%	4.6 – 4.3%
Complex	1.0%	0.1%	0.0%	0.0 – 0.0%
REST	6.3%	5.0%	6.8%	8.9 – 4.1%

Table 14. Comparison of most frequent parts of speech

What is immediately clear from this table is that there are vast differences between the contributions of the respective parts of speech to the total. The real outlying language is Otomí; the differences between Quichua and Guaraní, though considerable, are less striking. This is in accordance with the typological correspondences and differences between the three languages. That the borrowing behaviour of Otomí speakers is not due to purely local processes gets support from the fact that the figures for the Santiago (OS) and Tolimán (OT) dialects, as given in the rightmost column, converge to a high extent. The only significant differences between the two are the extremely high number of prepositions in OT and the high number of conjunctions in OS. We think therefore that it is safe to assume that the differences are mainly due to the differences in the grammatical make up of the target languages.

The most extreme case are the articles borrowed by Guaraní, and already mentioned in section 4.4.2. They may fill a gap in the sense that Classical Guaraní does not have articles. However, it is unlikely that this is in fact the case. The same can be said of Quichua, for which we find not a single example for this category. It turns out that they are mainly used as a freestanding anaphor. In case of their use within a noun phrase, they may often be interpreted pragmatically, as demonstratives indicating stress or contrast. Only in a restricted number of cases they merely code definiteness.$$$$

Second are the adpositions, which are borrowed abundantly by Otomí, but hardly by EQ and PG. They fall in step with the syntax of O, which has some adverbs with the function and in the syntactic position of prepositions. So Spanish prepositions may fill a functional gap in a language that on thw whole leaves nominal relations unmarked and implicit. In EQ and PG, both postpositional, the borrowing of prepositions is arguably hampered for purely structural reasons.

The third most striking difference is the very low number of verbs borrowed by Otomí. It is not clear to us why this is the case. All three languages, like Spanish have a specialized category for verbs. And the part of speech type of O (high specialization) is in many respects closer to that of Spanish (maximum specialization) than the other two

languages (both highly flexible). As we will see below, there seems to be no 'compensation' for this lack of borrowing from the other lexical categories, N and A.

The next most remarkable difference between the categories in table 15 is the use, in Otomí, of Spanish discourse markers, mainly the particle *pues* and the demonstrative *este*. Both are found very frequently in all varieties of spoken Mexican Spanish as well. They appear only very infrequently in the EQ and PG corpora, but they are much less characteristic of the colloquial Spanish of both Ecuador and Paraguay.

The following exception is again for Otomí. The language borrows a very reduced number of adjectives in comparison to the other languages. EQ and PG have a flexible class for the typical nominal and adjectival functions, and might therefore be flexible with the acceptation of both N and A. Otomí specializes for these categories, but there is only a restricted, semi-open class of adjectives in Otomí itself. Most of the meanings for which there would be an adjective in Spanish, are expressed by a verb or a noun in Otomí. This may explain why there are apparently restrictions on borrowing them.

One of the categories for which Quichua is an outlier is the high percentage of nouns that it borrows. Well over half of the total loans belong to this category, while both PG and O have around 40%. A high percentage of nominal borrowings may be indicative of an early stage of language change through contact. However, the overall amount of Spanish borrowings of both lexical and grammatical nature, which is the highest of the three languages under investigation, is strongly indicative of a later stage. Below we will contrast the amount of nouns borrowed only with two other lexical types, N and S, and see whether there may be an explanation for the relatively high number of Spanish nouns in Quechua.

Finally, Quichua has a rather low percentage of subjunctors among the borrowed tokens. The number of different types, nine, is the same as for Guaraní, though. All these are rather frequent in colloquial Spanish.

On the prefinal row in table 14 we find the total for complex borrowings. Only for Quechua do we find a number of these. By far the most of these are of the type found in () below, i.e. Spanish preposional phrases with a clearly fixed meaning.

(55)   a lo mejor 'at best'
         a lo primero 'firstly'
         a tiempo 'in time'
         a veces 'sometimes'
         a ver 'let us see'
         de acuerdo 'okay'
         de repente 'suddenly'
         de todas maneras 'anyhow'
         por ejemplo 'for example'
         por lo menos 'at least'

If we look at the three major parts of speech V, N and A in isolation, and assume for the sake of argument that these are in 'competition' for the same semantic space in all languages, then we see a remarkable difference between the three languages again. The columns in the middle of figure 15 show that the relative proportions of coverage for Quichua and Guaraní are relatively close to each other. This relativizes a bit, though not

entirely the high percentage of nouns for Quichua in table 14 above.[30] However, for Otomí the figures are vastly different, with a strong overrepresentation of N at the cost of V and A. On the basis of these figures we may tentatively conclude that Otomí attracts referential material while Quichua and above all Guaraní tend to interpret borrowings more predicatively. The columns on the righthand side make it clear that for Quichua and Guaraní [V, N, A] represent 70-80% of overall borrowing, while in the case of Otomí they cover less than 50% of the total. So, also from this perspective EQ and PG differ strongly from O.

Part of Speech	RELATIVE PERCENTAGES			ABSOLUTE PERCENTAGES		
	EQ	PG	O	EQ	PG	O
N	67.6%	59.1%	85.9%	54.4%	37.2%	40.7%
V	22.0%	29.1%	10.1%	17.7%	18.3%	4.8%
A	10.5%	11.8%	4.0%	8.5%	7.4%	1.9%
TOTAL	100.0%	100.0%	100.0%	80.6%	72.7%	47.4%

Table 15. Proportion major parts of speech

Now let us aggregate the figures for the individual parts of speech to totals for lexical and grammatical borrowings. We use the term 'lexical' for the classes V, N, A and MAdv plus all other types of adverbs that we found. The rest of the elements will be subsumed under 'grammatical'. Table 16 provides an overview, again in terms of tokens, and rounded off to whole percentages.

	Quichua (EQ)	Guaraní (PG)	Otomí (O)
Lexical	84%	65%	52%
Grammatical	16%	35%	48%

Table 16. Comparison in terms of lexical versus grammatical

These differences are even more remarkable than the individual figures of table 14 and 15. They reflect the mirror image of the order that we get when we rank these languages in terms of number of tokens borrowed. As indicated above the major cause of the high percentage for Otomí is the high number of Spanish prepositions. It could be argued that these elements are on the borderline between lexical and grammatical, or that some are more lexical and some more grammatical. When we add, for the sake of argument the prepositions to the lexical categories the figures for grammatical borrowing would drop to 15.5% (EQ), 34.2% (PG) and 26.9% (O), respectively. As we have already seen in table 14, the high number of grammatical borrowings in Guaraní is mainly caused by the Spanish articles *la* and *lo*. They make up 55.9% of the total of this category. Quichua and Otomí do not borrow articles at all. When we would take out articles altogether from the calculations the total percentages for grammtical borrowing would be 15.5% (EQ), 18.4% (PG) and 26.9% (O). These two manipulations would still leave the languages in the reverse order of the one reflecting the total number of borrowings in table 12. It seems to be obvious that these two measures – overall percentage of borrowings and division lexical:grammatical - are not necessarily highly correlated. Whatever may be the precise meaning of this, the figures presented in the tables in this section must lead to a

reconsideration of several of the hypotheses presented in section 3. We think that they also have implications for general borrowing scales such as the one in Thomason (2001). Possibly, factors such as level of bilingualism, and amount of code switching, briefly mentioned in the section on Guaraní, should be included as factors.

To round of this section, we will compare the functions in which the lexical categories are used in the respective languages. As has already been pointed out in the sections on the individual languages, borrowed V and MAdv are virtually completely fixed, in all three languages, to their prototypical functions of HP and MP, respectively. Below, in table 16, we find the relative uses for the categories N and A.

	Quichua (EQ)	Guaraní (PG)	Otomí (O)
NOUN			
HP	0.7%	0.2%	0.6%
HR	92.8%	98.3%	99.3%
MP	1.7%		
MR	4.8%	1.4%	0.1%
ADJ			
HP	1.3%	1.4%	40.8%
HR	22.5%	26.8%	3.7%
MP	10.3%	4.7%	0.3%
MR	66.0%	67.1%	55.1%

Table 17. Relative distribution over functions

For the nouns the differences are rather small. As to be expected, Otomí is the least flexible language of the three, and close to total rigidity in its application of Spanish nouns. But one would have expected more flexibility indeed for the other two languages, which do not have a specialized class of nouns. An explanation may be that nouns are borrowed in the more specific nominal meaning they have in Spanish rather than the more flexible meaning that non-verbal lexemes might have in EQ and PG.

For the adjectives, the picture is as one would expect for EQ and PG: a rather high flexibility. The second most frequent use is as the head of a referential phrase, in line with non-verbal elements of these languages themselves. The picture for the – relatively few – Spanish adjectives borrowed by Otomí does not seem to support the assumption made above that this language has a tendency to interpret its borrowings referentially rather than predicationally. Around 41% of the borrowed adjectives are used as predicates.

A final remark should be devoted to code switching. As indicated above we have distinguished between code switches and 'real' borrowings. The former have been left out of the calculations. However, there are considerable differences between the numbers of code switches in the three corpora. This is shown in table 18.

	Quichua (EQ)	Guaraní (PG)	Otomí (O)
Tokens	3103	10500	1717
Percentage of real borrowings	20.6%	95.8%	11.0%

Table 18. Number of tokens in code switches

For Otomí the figure is rather low. However, in the case of Guaraní adding the code switches to the borrowings would virtually double the figures. This is strongly indicative of the bilinguality of the Paraguayan language community.

8. Conclusions
In this contribution we have compared the borrowing phenomena in three typologically different languages as a result of contact with the same source language. On the basis of data collected according to a more or less uniform strategy we tested several hypotheses concerning borrowing which were derived from a functionally oriented linguistic theory on parts of speech and from the recent literature on language contact.

Hypothesis 1, about pragmatically more or less outstanding markers, has gotten strong support above all from Otomí. Also in the other languages, we found only discourse markers at the extra-sentential level. No intra-sentential topic or focus marker were found in anywhere.

Hypothesis 2, which predicts that more lexical than grammatical elements should be found, at least at the borrowing stage of the three languages concerned, was supported by Quechua and Guaraní, but only marginally by Otomí, unless we consider adpositions as lexical rather than grammatical elements. When we look at the respective subhypotheses, we observe the following. There is a strong tendency for open classes to be borrowed more easily than closed classes. This is the case for both lexical (H2.1.1) and non-lexical (H2.1.2) classes, with two notable exceptions. If borrowed at all, as for Otomí adpositions rank among rank at the top of the grammatical categories, as predicted, and definitely not at the bottom of the lexical categories. In that sense they fare better as grammatical categories. And articles, which are expected to rank at the lower end of the grammatical hierarchy, if borrowed, as in the case of Guaraní score higher than any of the other grammatical categories. But the hypotheses are applied under the assumption of 'all things being equal'. However, things are not equal: we have seen that constraints originating from typological differences between the languages may explain several of the discrepancies observed, be it not all of them. Particularly the heavy borrowing of the Spanish article in Guaraní is puzzling. Possibly, the part of speech type of the target language rather than that of the source language plays a crucial role in the explanation of such cases. Hypothesis H2.2, i.e. that the amount of boundedness of source elements plays a role in their being borrowed, has not been considered here at all. Although we have quite some information on this from our corpus, which suggests that especially Quichua has relatively many mixed complex forms, space prevents us from going into details on this here.

Hypothesis H3 predicts that we would find many more separate lexical elements than complex structures. This is indeed the case, overwhelmingly so, especially when we disregard code switches. Only in Quichua do we find a fair number of construction-like borrowings. For most of them it may be assumed that they are borrowed as unanalyzable entities.

As far as the hypotheses under H4 are concerned, we can say the following. H4.1, about head-modifier relationships finds support from the large amount of prepositions borrowed by Otomí as opposed to the very low numbers of Spanish prepositions which made it into both postpositional languages. As far as main clause orders are concerned (H4.2), we observed that Otomí shifts from VOS towards SVO order, both the basic order

in Spanish, and a variant of Otomí. Also Quechua, rather rigidly SOV, shifts towards SVO, a rare order though not an ungrammatical order in Pre-spanish Quechua. This is more or less in line with the hypothesis. The basic order for adjectives in Spanish is postnominal, with prenominal order rather infrequent and marked for most adjectives. Still, this does not prevent Quechua, with prenominal adjective ordering, to borrow relatively more adjectives than Guaraní. In this case, the hypothesis finds no support from typology. Finally, H4.3 finds support from many 'doublings' in especially Otomí, where the borrowed form is typically higher on the grammaticalization scale than the native element.

The predictions made on the basis of Hengeveld's part of speech theory found support in a more or less general sense. The borrowing characteristics in terms of the four central types V, N, A and MAdv is rather similar for Guaraní and Quechua, both type 2, and quite different for Otomí, type 5. However, when we look in more detail, quite a few discrepancies remained, especially the borrowing of relatively many nouns and adverbs and the low number of verbs by Otomí and the extremely high number of nouns by Quichua. Our major conclusion is that, although most of the hypotheses get support to some extent, the formal conditions have to be specified in much more detail to be of real value. Since they are mainly based on hypotheses from the literature on borrowing, more specifically the work of Thomason & Kaufmann (1988) and Thomason (2001), our observations on these three examples of borrowing must have repercussions for some generally accepted insights. Obviously, in order to confirm these insights, and derive a robust theory on borrowing, a more refined analysis of the data than we have made here is called for. In that future work, more attention should be paid to the role of code switching and the level of bilingualism of both the language communities and the individual informants.

Abbreviations
1	first person
2	second person
3	third person
ACC	accusative
ACMP	accompaniment
ADIT	additive
ALL	allative
COM	commitative
CONJ	conjunction
COP	copula
DEF	definite
DEM	demonstrative
DIM	diminutive
DIREV	direct evidence
EXC	exclusive
FUT	future
GEN	genitive
GER	gerund

HP	head of a predicate phrase
HR	head of a referential phrase
IMPRF	imperfective
INC	inclusive
IND	indefinite
INFER	inference
INST	instrumental
INT	intensifier
INTERP	interpelative
LIM	limitative
LOC	locative
MP	modifier of a predicate phrase
MR	modifier of a referential phrase
NEG	negative
OBJ	(in)direct object
PERF	perfect
PL	plural
PSTPRF	past perfect
POSS	possessive
PRES	present
PROL	prolative
PST	past
REP	reportative
SG	singular
SUB	subordinator
SUPL	superlative
TOP	topicalizer.

References

Adelaar, Willem, in collaboration with Pieter Muysken (2004). *Languages of the Andes.* England: Cambridge University Press.

Aikhenvald, Alexandra (2002). *Language contact in Amazonia*. Oxford: Oxford University Press.

Aitchison, J. (1981). *Language Change: Progress or Decay.* London: Fortuna.

Andersen, H. (1988). 'Centre and periphery: adoption, diffusion and spread'. In J. Fisiak (ed.) *Historical dialectology*. Berlin: Mouton de Gruyter, 39-85.

Andrews, H. (1993). 'The Function of Verb Prefixes in Southwestern Otomí' *The Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics* 115. Dallas: SIL.

Bakker, D. & E. Hekking (1999). 'A functional approach to linguistic change through language contact: the case of Spanish and Otomí'. Working Papers in Functional Grammar 71, 1-32.

Bakker, D. & A. Siewierska (2002). 'Adpositions, the Lexicon and

Expression Rules'. In Ricardo Mairal Usón & María Jesús Pérez Quintero (eds) *New Perspectives on Argument Structure in Functional Grammar*. Berlin: Mouton de Gruyter, 125-78.

Bartholomew, D. (2004). 'Notas sobre la gramática.' In Luis Hernández Cruz et al. (eds.). *Diccionario del hñähñu (otomí) del Valle del Mezquital (Hidalgo)*. Vocabularios Indígenas, no. 45. México, D.F.: Instituto Lingüístico de Verano.

Beck, David (2002). *The typology of parts of speech systems: the markedness of adjectives*. New York: Routledge.

Campbell, L. (1989). 'On Proposed Universals of Grammatical Borrowing'. In H. Aertsen & R.J. Jeffers (eds) *Papers from the 9th International Conference on Historical Linguistics*. Amsterdam: John Benjamins, 91-109.

Cerrón-Palomino, Rodolfo (1987). *Lingüística Quechua*. Biblioteca de la Tradición Andina. Cusco: Centro de Estudios Rurales Andinos Bartolomé de las Casas.

Comrie, B. (1989). *Language Universals and Linguistic Typology*. Oxford: Blackwell.

Corvalan G. & G. de Granda (eds.) (1982). *Sociedad y lengua: bilinguismo en el Paraguay*. Asunción: Centro Paraguayo de Estudios Sociológicos.

Croft, W. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: Chicago University Press.

Croft, W. (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Dench (1998). Australian lgs $$$$

Dietrich, Wolf (1995). 'El español del Paraguay en contacto con el guaraní. Ejemplos seleccionados de nuevas grabaciones lingüísticas'. In Klaus Zimmermann (ed.): *Lenguas en contacto en Hispanoamérica: Nuevos enfoques*. Frankfurt am Main: Vervuert. 203-216.

Ecker, L. (1952). *Compendio de gramática otomí: Introducción a un diccionario otomí-español*. México, D.F.: Instituto Nacional de Antropología e Historia, Secretaría de Educación Pública.

Escobar, A.M. (2004). 'Bilingualism in Latin America'. In T.K. Bhatia & W.C. Ritchie (eds) *The Handbook of Bilingualism*. Oxford: Blackwell, 642-661.

Evans, N. & T. Osada (2005). 'Mundari: The myth of a language without word classes'. *Linguistic Typology* 9-3, 351-390.

Filppula, M. (2003). 'The quest for the most 'parsimonious' explanations: endogeny vs. contact revisited'. In : R. Hickey (ed.), 161-173.

Garvin, P. & M. Mathiot (1982). 'La urbanización del idioma Guaraní. Problema de lengua y cultura'. In G. Corvalán y G. de Granda (eds.): *Sociedad y Lengua: Bilingüismo en el Paraguay*. Vol 1. Asunción: CEPAG, 27-41.

Goldberg, A. (1995). *Constructions : A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

Gómez-Rendón, J. (2005). 'La media lengua de Imbabura'. In Pieter Muysken and Hella Olbertz (eds), *Lenguas en contacto en los Andes*. Berlin: Vervuert Hispanoamericana, 39-57.

Gómez-Rendón, J. (fc-a). 'Spanish grammatical borrowing in Paraguayan Guaraní: an overview'. Manchester Project on Grammatical Borrowing.

Gómez-Rendón, J. (fc-b). 'Spanish grammatical borrowing in Imbabura Quechua: an overview'. Manchester Project on Grammatical Borrowing.

Gordon, R.G., Jr. (ed.) (2005). *Ethnologue: Languages of the World, Fifteenth edition*. Dallas, Tex.: SIL International. Online Version: http://www.ethnologue.com/.

Guerrero (2002). $$$

Gumperz, J.J. & R. Wilson (1971). 'Convergence and creolization: a case from the Indo-Aryan/Dravidian border'. In D. Hymes (ed.) *Pidginization and Creolization of Languages*. Cambridge : Cambridge University Press, 151-167.

Haboud, Marleen (1998). *Quichua y Castellano en los Andes Ecuatorianos: los efectos de un contacto prolongado*. Quito: Ediciones Abya Yala.

Hawkins, J.A. (1983). *Word Order Universals*. Cambridge: Cambridge University Press.

Hekking, E. (1995). *El Otomí de Santiago M: desplazamiento linguïstico, préstamos y cambios grammaticales.* Amsterdam: IFOTT.

Hekking, E. (2001). 'Cambios gramaticales por el contacto entre el otomí y el español'. In K. Zimmermann & T. Stolz (eds), 127-151.

Hekking, E. (2002). 'Desplazamiento, pérdida y perspectivas para la revitalización del hñäñho'. In Y. Lastra de Suárez & N. Quezada (eds), 221-248.

Hekking, E. & D. Bakker (1998). 'Language shift and Spanish content and function words in Otomí'. In B. Caron (ed.) *Actes du 16e Congres International des Linguistes*. Oxford: Elsevier Sciences.

Hekking, E. & D. Bakker (1999). 'El otomí y el español de Santiago Maxquititlán: Dos lenguas en contacto'. *Foro Hispánico* 13, 45-74.

Hekking, E. & D. Bakker (2006). 'Problemas en la adquisición de una segunda lengua: El Otomí frente al Español'. In C. Chamoreau & J. Lastra (eds) '$$$'. Sonora (Mex.): Universidad de Hermosillo, xxx-yyy.

Hekking, E. & Andrés de Jesús, S. (1984). *Gramática Otomí.* Querétaro (México): Universidad Autónoma de Querétaro.

Hekking, E. & P. Muysken (1995). 'Otomí y Quechua: una comparación de los elementos prestados del español'. In Klaus Zimmermann (ed.) *Lenguas en contacto en Hispanoamérica: nuevos enfoques*. Frankfurt am Main: Vervuert.

Hengeveld, K. (1993). *Non-verbal predication: theory, typology and diachrony.* Berlin: Mouton de Gruyter.

Hengeveld, K., J. Rijkhoff & A. Siewierska (2004). 'Parts-of-speech systems and word order'. *Journal of Linguistics* 40.3, 527-570.

Hickey, R. (ed.) (2003). *Motives for Language Change*. Cambridge: Cambridge University Press.

Hout, R. van & P. Muysken (1994). 'Modelling lexical borrowability'. *Language Variation and Change* 6-1, 39-62.

Jackson, J. (1974). 'Language identity of the Colombian Vaupes Indians'. In R. Bauman and J. Sherzer (eds) *Explorations in the Ethnography of Speaking.* New York: Cambridge University Press.

Jiménez Moreno, W. (1939). 'Origen y significación del nombre otomí'. In *Revista Mexicana de Estudios Antropológicos* III, 62-68.

Lastra de Suárez, Y. (1992). *El otomí de Toluca.* México, D.F.: Instituto de Investigaciones Antropológicas, UNAM.

Lastra de Suárez, Y. (1994). 'Préstamos y alternancias de código en otomí y en español.'

In C. Mackay & V. Vazquez (eds) *Investigaciones Lingüísticas en Mesoamérica.* México: Universidad Autónoma de México.

Lastra de Suárez, Y. (1997). *El otomí de Ixtenco.* México, D.F.: Instituto de Investigaciones Antropológicas, UNAM.

Lastra de Suárez, Y. & N. Quezada (eds) (2002). *Estudios de Cultura Otopame* 3.3. Mexico: Instituto de Investigaciones Antropológicas.

Meliá, B. (1974). 'Hacia una tercera lengua en el Paraguay'. *Estudios Paraguayos* 2, 31-71.

Meliá, B. (1992). *La lengua guaraní del Paraguay: historia, sociedad y literatura*. Madrid: MAPFRE.

Milroy, J. (2003). 'On the role of the speaker in language change'. In R. Hickey (ed.) 143-157.

Morínigo, M. (1990). *Raíz y destino del Guaraní.* Asunción: Universidad Católica.

Moravcik, E. (1978). 'Universals of Language Contact'. In Joseph Greenberg: *Universals of Language, Vol I, Method and Theory*. USA: Stanford University Press, 95-122.

Muysken, P. (1985). 'Contactos entre Quichua y Castellano en el Ecuador', in S. Moreno (ed.): *Memorias del Primer Simposio Europeo sobre Antropología del Ecuador.* Quito: Abya Yala, 377-472.

Muysken, P. (1994). 'Media Lengua'. In P. Bakker & M. Mous (eds) *Mixed Languages: 15 case studies in language intertwining.* Amsterdam: IFOTT, 207-211.

Muysken, P. (2000). 'Semantic transparency in Lowland Ecuadorian Quechua morphosyntax'. In *Linguistics* 39-5, 873-988.

Myers-Scotton, (2002). *Contact linguistics: bilingual encounters and grammatical outcomes.* New York: Oxford University Press.

Nordhoff, Sebastian (2004). *Nomen/Verb Distinktion in Guaraní.* MA thesis Universität Köln.

Palancar, E.L. (fc). 'Property concepts in Otomí: a language with no adjectives'. International Journal of American Linguistics.

Parker, Gary (1963). 'La clasificación genética de los dialectos quechuas'. *Revista del Museo Nacional* 32, 241-252.

Schachter P. (1985). 'Parts of speech Systems'. In Timothy Shopen (ed.) *Language typology and syntactic description* 1. Cambridge: Cambridge University Press, 3-61.

Ross, M. (2003). 'Diagnosing prehistoric language contact'. In R. Hickey (ed.), 174-198.

Ruhlen, M. (1991). *A Guide to the World's Languages*. London: Edward Arnold.

Ruiz de Montoya, A. (1993 [1640]). *Arte de la lengua guaraní.* Asunción: Centro de Estudios Paraguayos "Antonio Guasch".

Sahagún, B. de (1982). *Historia general de las cosas de Nueva España*, 5ª ed. Angel María Garibay (ed.). México: Porrúa.

Saussure, F. de (1916). *Cours de linguistique générale*. Paris: $$$.

Sasse, H-J. (1992). 'Theory of language death'. In: M. Brenzinger (ed.) *Language Death*. Berlin: Mouton de Gruyter, 7-30.

Siewierska, A. (1998). 'Variation in major constituent order'. In A. Siewierska (ed.) *Constituent Order in the Languages of Europe*. Berlin: Mouton de Gruyter, 475-552.

Song, J.J. (2001). *Linguistic Typology: Morphology and Syntax*. London: Longman.

Stolz T. & C. Stolz (1996). 'Funktionswortentlehnung in Mesoamerika. Spanish-amerinindischer Sprachkontakt'. *Sprachtypologie und Universalienforschung* 49:1, $$$-$$$.

Suárez, J. A. (1983). *The Mesoamerican Indian Languages.* Cambridge: Cambridge University Press.

Thomason, S.G. (2001). *Language Contact. An Introduction*. Edinburgh: Edinburgh University Press.

Thomason, S.G. & T. Kaufmann (1988). *Language Contact Creolization, and Genetic Linguistics.* Berkeley: University of California Press.

Trinidad Sanabria, L. (1997). 'La evolución histórica del Guaraní'. http://www.datamex.com.py/Guaraní/marandeko/trinidad_evolucion_historica_del_idioma_Guaraní.html.

Trinidad Sanabria, L. (1998). *Polisíntesis Guaraní. Contribución para el conocimiento tipológicos de esta lengua amerindia*. Asunción: Intercontinental Editora.

Urbano, A. (1990 [1605]). *Arte breve de la lengua Otomí y vocabulario trilingüe.* R. Acuña (ed.). Ciudad de México: Universidad Nacional Autónoma de México.

Velazquez-Castillo, M. (1995). 'Noun incorporation and object placement in discourse: the case of Guaraní'. In P. Downing & M. Noonan (eds) *Word Order in Discourse*. New York: John Benjamins, 555-$$$.

Velazquez-Castillo, M. (2002). 'Grammatical relations in active systems: the case of Guaraní'. In *Functions of Language* 9:2, 133-167.

Voigtlander, K. & Echegoyen, A. (1985). *Luces contemporáneas del otomí: Gramática del otomí de la sierra.* México, D.F.: Instituto Lingüístico de Verano.

Yasugi, Y. (1995). *Native American Languages. An areal-Typological Perspective*. Osaka: National Museum of Ethnology.

Zimmermann, K. (1992). *Sprachkontakt, ethnische Identität und Identitätsbeschädigung: Aspekte der Assimilation der Otomí-Indianer an die hispanophone mexikanische Kultur.* Frankfurt: Vervuert.

Zimmermann, K. & T. Stolz (eds) (2001). *Lo propio y lo ajeno en las lenguas austronésicas y amerindias: procesos interculturales en el contacto de lenguas indígenas con el español en el Pacífico e Hispanoamérica.* Frankfurt: Vervuert-Iberoamericana.

---

[1] For a recent overview of bilingualism in Latin America, see Escobar (2004).

[2] We follow here the convention of calling Quichua the Ecuadorian dialects that form a branch of the Quechua language family. Ecuadorian Quichua is classified as part of Quechua II (Torero 1964) or Quechua A (Parker 1963). The dialects of this subfamily share a number of traits with other varieties from Southern Peru, Bolivia and Argentina (Adelaar 2004: 185ff).

[3] Obviously, for reasons of comparison there should also be languages with more or less the same typological characteristics. The scope of the current study prevents us from including these. For one of the languages, however, we will briefly look at the situation in two different dialects.

[4] Cf. Milroy (2003) for more extensive argumentation on this view.

[5] One might claim that any change in a language and its grammar is eventually motivated by the fact that it is meant to improve its communicative function. This position would deny any autonomy to grammar. Croft (199x) has coined the term extreme functionalism for such a view, and argued convincingly against such a position for monolingual models of grammar. We would like to leave some room for grammar internal explanations of change as well, at least as a theoretical possibility.

[6] Abbreviations used in the glosses and elsewhere are found at the end of the article.

---

[7] Cf. Sasse 1992:19f for a complete scenario leading to language death. All these stages of bilingualism are reported for languages in Spanish America in Grimes & Grimes (200$). Ruhlen (1991) lists $$ extinct languages from the area. This is not to say that all have been driven into extinction as a result of contact with Spanish.

[8] Another analysis that has been given to the borrowing of discourse markers is that it is indicative of the restructuring of the discourse in the direction of the source language. Others, like Zimmermann (1992: 299-305), explain Spanish borrowing as based on the prestige that this language has with respect to the target language.

[9] Thomason (2001:93) rightly warns for jumping to conclusions about borrowing too easily in such cases. In the same vein, Comrie (1991:03) talks about the 'quite widespread myth' concerning the typological stability of languages both diachronically and in the light of language contact that 'must be explored'.

[10] Cf. Bakker & Siewierska (2002) for an analysis of adpositions within the framework of FG.

[11] Cf. Song (2001) for a recent introduction.

[12] Compare universal Ux in Greenberg (1963). According to Siewierska (1998:493), from a sample of 171 languages, 63% of the VSO languages have SVO as an alternative order, and only 13% (also) SOV.

[13] We have collected these for part of the languages discussed here. However, in order to keep the three corpora comparable, we have left them out of our analyses below.

[14] Note that the notion type is used in its most simplistic way, i.e. the pure form, disregarding morphological complexity or homonymy. In other words, no lemmatization has taken place. This has a great influence on the interpretation of this measure, especially when comparing languages of different morphological types. Below, we will mainly be concerned with numbers of tokens.

[15] The programme gives several other statistics, e.g. TTR2, a type-token ratio where the number of types is divided by the square root of the number of tokens. This figure is somewhat more stable and less dependent on the actual text length (cf. van Hout & Muysken 1994).

[16] Both devices are optional, and should be applied with caution. Due to the inherent ambiguity of lexical material no device as simple as the ones discussed here might ever take such decisions with near 100% certainty. All results should be checked on hindsight. E.g. the pattern match procedure scores around 90-95% security for Otomí, but considerably lower for Quichua. This may be due to the amount of overlap between the two phonological systems involved, to the extent that they are reflected in the spelling systems.

[17] Rather than providing a straightforward part of speech code for the target language as well, we decided to code its role in the target language syntax. As will be explained below, in section 3, it is not always the case that certain parts of speech should be distinguished for the lexicon of target language involved.

[18] Three Guaraní-speaking groups originally inhabited the territory of present Paraguay: Kario, Tovatĩ and Guarambaré Indians were the first Guaraní to contact the Spaniards to be finally absorbed in the mixing process. Itatĩ, Tapé and Paranaguá Indians had their first contact with Europeans only through the Jesuit missionaries and were perhaps who most suffered from Spanish exploitation, as they were reluctant to mixing. Finally, the Mbya group had only occasional contact with Europeans and thus preserved their culture and language to a great extent well into the twentieth century (cf. Trinidad Sanabria 1997; Meliá 1992).

[19] The absence of a caste system in Paraguay's emerging colonial society left ample space for interracial and intercultural practices.

[20] Some of the first literary works written in Jopara date from 1867 when the satirical journals *Cabichu'í* and *Cacique Lambaré* were printed during the war against Argentina, Brazil and Uruguay.

[21] For full data on both censuses, visit the website www.dgeec.gov.py.

[22] With respect to the 1992 census, Guarani-Spanish billingualism shows an increase of over 10% and a corresponding decrease in Guarani monolingualism of 12%.

[23] This mosaic of languages in the Northern Andes might have influenced on the configuration of Ecuadorian Quechua as a koineized variety and could shed light on the tendency of some Ecuadorian dialects to show higher degrees of Spanish lexical borrowing. Borrowing in some dialects can be massive and lead to the origin of mixed languages such as Media Lengua, reported only in Ecuador (Muysken 1985; Gomez-Rendón 2005).

[24] The fact that Quechua was initially standardized with evangelization purposes and so disseminated by non-native speakers (Spanish clergymen) may be an important factor when it comes to explaining the presence of Spanish elements in all the varieties of Quechua. For a discussion of this and other relevant

factors that may have led to the rapid spread of Quechua in the territory of present Ecuador, see Muysken (2000).

[25] Typically, dictionaries define colours with reference to objects known in the community's sociocultural context. See, for example, the Webster dictionary's definition for 'white', which resort to metaphors such as intensity of light, racial groups and the like.

[26] The hypothesis that Otomí is a derivation of the two Otomí words *otho* 'without' and *m<u>u</u>i* 'house' is another example of the contempt in which the Otomís were held.

[27] Palancar (fc) gives some arguments to consider Otomí as a language without adjectives.

[28] Even the Spanish adjectival quantifier *cada* 'each' which does not inflect for gender is found several times in the 'masculine' form *cado*.

[29] There does not seem to be any correlation between the numbers of manner adverbs on the one hand and other adverbs on the other hand. Cf. Q: 88 MAdv + 432 other; G: 95 MAcv + 142 other; O: 33 MAdv + 663 other

[30] In could be argued that in comparisons such as the ones in tables 14 and 15 only categories should be compared which are indeed in competition for some semantic or grammatical space. However, certain categories may imply each other morposyntactically, such as articles and nouns, and therefore may have a positive or negative influence on their overall distributions among borrowings.