Three processes of borrowing: borrowability revisited

Pieter Muysken

There is a tradition of trying to establish borrowability hierarchies and implicational universals of borrowing (e.g., Haugen, 1950; for a summary cf. Van Hout and Muysken, 1994). This effort runs into problems in many bilingual settings: in addition to the standard sets of borrowed nouns and adjectives (which are highly ranked on the predicted hierarchies), there are often recalcitrant sets of borrowed conjunctions and prepositions. In addition, we sometimes find word-internal borrowing processes. I want to argue that these deviant patterns result from the fact that there is not a single borrowing process, just like there is no single code-mixing process. In addition to the familiar pattern of insertion (producing the borrowability hierarchies), there is a pattern of alternation involving interjections and conjunctions, and a pattern of congruent lexicalization involving parts of words.

First, I briefly list the three main types of code-mixing encountered, and in the next section I review some of the work on borrowability hierarchies. The following three sections are dedicated to alternational borrowing, the incorporation of alternational borrowings into the lexicon, and borrowing through congruent lexicalization. I end with some concluding remarks.

Code-mixing

There are several distinct processes at work in code-mixing (Muysken, 1997):

• insertion of material (lexical items or entire constituents) from one language into a structure from the other language;

• alternation between structures from languages;

• congruent lexicalization of a shared grammatical structure with material from different lexical inventories.

The structural interpretation of these notions is illustrated in the trees [1]-[3], where \( A, B \) are language labels for non-terminal nodes (i.e., fictitious markers identifying entire constituents as belonging to one language), and \( a, b \) are labels for terminal, i.e., lexical, nodes, indicating that the words chosen are from a particular language.
[1] insertion

In the situation defined by [1], a single constituent $B$ (with words $b$ from the same language) is inserted into a structure defined by language $A$, with words $a$ from that language. Typically, constituent $B$ is selected by an element from language $A$, and insertion takes place under equivalence between constituents of the two languages.

[2] alternation

In this situation, a constituent from language $A$ (with words from the same language) is followed by a constituent from language $B$ (with words from that language), which is not selected by an element from language $A$. Unspecified is the language of the constituent dominating $A$ and $B$ in [2], and there is no requirement that e.g., the constituent in $B$ is equivalent to an element from language $A$.

[3] congruent lexicalization

Finally, in [3] the grammatical structure is shared by languages $A$ and $B$, and words from both languages $a$ and $b$ are inserted more or less randomly. This last type of code-mixing may characterize e.g., dialect-standard mixing or mixing between closely related languages.
Borrowability hierarchies

The traditional observation, with long roots in language contact research, is that different categories can be borrowed, or are actually borrowed more or less easily. This observation, which had a somewhat shaky empirical base until recently, has received massive support from the work reported in Poplack, Sankoff and Miller (1988). The finding that nouns are the most frequently borrowed element is confirmed for many other language pairs as well (Nortier and Schatz, 1992). For borrowing, constraints can then be formulated in terms of a categorial hierarchy: words of one specific lexical category can be borrowed more easily than those of another. An example of such a hierarchy is [4], partly based on Haugen (1950):


Such hierarchies predict that a noun such as French *automobile* can be borrowed more easily into English than a conjunction such as *que*, and this prediction holds reasonably well in the extreme cases. The problem, however, with a hierarchy such as [4] is that there is no explanation given for the order of the lexical categories in the hierarchies. In addition, there are very striking language-specific deviations, as it turns out.

In Van Hout and Muysken (1994) we explored the possibility of a probabilistic approach to borrowing by comparing two related corpora: a set of Bolivian Quechua folktales collected by Aguiló (1980) and informal Spanish translations of these tales by people from the same villages. The Spanish corpus is taken as the corpus of potentially borrowed elements, the (very numerous) Spanish borrowings in the Quechua folk tales as the actually borrowed corpus. We developed analytical techniques, based on regression analysis, to determine borrowability, i.e., the ease with which a category, a class of lexical items can be borrowed. The analysis is based on two assumptions: (a) the distribution of items in the donor language should be taken into account, to explain why certain items are, and others are not borrowed; (b) the borrowability of a category may result from the interaction of a number of factors. Our research question was framed as follows: Given a donor lexicon \( L_x \) and a recipient lexicon \( L_y \), what is the chance for an item from \( L_x \) to end up in \( L_y \), and what determines this chance?

A number of factors are involved. One of the primary motivations for lexical borrowing is to extend the referential potential of a language. Since
reference is established primarily through nouns, these are the elements borrowed most easily. More generally, *content words* such as adjectives, nouns, verbs may be borrowed more easily than *function words* (articles, pronouns, conjunctions) since the former have a clear link to cultural content and the latter do not.

A second explanatory factor to be considered is frequency of lexical items, perhaps both in the donor language and in the recipient language. Furthermore, it is clear from a number of cases that words which play a *peripheral role* in sentence grammar, particularly the grammar of the recipient language: interjections, some types of adverbs, discourse markers, and even sentence coordination markers, are borrowed relatively easily.

A related way to approach the same question is to see to what extent categories are directly implied in the *organization of the sentence*: a verb is more crucial to that organization than a noun, and perhaps therefore it may be harder to borrow verbs than nouns.

The central role of the verb is also reflected in its role in *case assignment*, which may be specific to that verb and idiosyncratic, to different elements in the sentence. This also stands in the way of their being taken from one system to another. Prepositions share this property with verbs, which may hinder their being borrowed.

In addition to these factors deriving from *syntagmatic* coherence (peripherality, structure-building, case-marking), there is also *paradigmatic* coherence. Paradigmatic coherence concerns the tightness of organization of a given subcategory: the pronoun system is tightly organized, and it is difficult to imagine English borrowing a new pronoun to create a second person dual in addition to second person singular and plural.

Often the different elements in the clause are marked on the verb, which may be morphologically quite complex for this reason. Borrowing will imply morphological integration as well, and this often is a hindering factor. A separate dimension then will be *inflection*: agreement (subject/object ... verb and adjective ... noun agreement) and case affixes. We would predict that uninflected elements will be easier to borrow than inflected ones.

In addition to the inflection of the *donor language*, we hypothesize that the inflection of the *host language* plays a role. It may well be easier to incorporate elements into the lexicon that do not have to then become integrated morphologically as well.

Finally, there is the factor of *equivalence* to be considered. Weinreich (1953:61) notes that resistance to borrowing is always a function, not so much of properties of recipient and sources languages by themselves, but of the
difference in structures of the recipient and source languages. In our study a regression analysis was carried out for:
• paradigmaticity;
• equivalence between source and recipient language;
• inflection in the source language;
• inflection in the recipient language;
• status as a function word or not;
• transitivity;
• constituent-peripheral versus -internal position.

Our main results were the following. Paradigmaticity and inflection in the donor language are revealed to be the strongest structural factors in our regression analysis. Lexical content and equivalence do not play a role independently. Frequency also has a (somewhat weaker) effect, while peripherality has an effect, but opposite to what we predicted.

Van Hout and I conclude that the constraints model, operating on the basis of a comparison between a donor language and a recipient language corpus, seems to be a promising way of studying the process of lexical borrowing. The results may be interpreted in such a way as to set up a new hierarchy of borrowability, which would simply result from classifying the individual categories in terms of their value for the factors prohibiting or furthering the borrowing process. However, the results for peripherality were troubling, and may lead to a different analysis of peripheral elements, as I will argue shortly.

Lexical borrowing has been associated with insertional code-mixing, and not without reason. Nouns are the class of elements borrowed par excellence and also the prime example of insertion under categorial equivalence: since all languages in the world have nouns, and nouns can occur in many languages without inflection, they are a natural candidate for borrowing. However, there may also be other kinds of borrowing patterns.

There are several problems with the assumptions behind the insertional view of borrowing and the borrowability hierarchy. Besides the unexpected role of peripherality alluded to above, there are other disturbing facts. First, Hekking and Muysken (1995) compare Otomí and Quechua in terms of their borrowing from Spanish. While in Quechua, the percentage of function words of the total number of borrowings is .06 at the type level, and .13 at the token level, for Otomí the percentage is much higher: .19 and .28, respectively. Even more radical figures can be obtained from Popoloca, another Otomanguean language (Veerman-Leichsenring, 1991). If we contrast the set of borrowed discourse organizers, prepositions and conjunctions, temporal expressions and quantifiers (131 tokens, 42 types, 22 singletons/hapaxes/
nonce borrowings) with the set of content nouns, adjectives and verbs (172 tokens, 79 types, 45 singletons/hapaxes/nonce borrowings), it is clear that the latter set is only slightly larger. Suárez (1983:135) writes in his survey work on the Mesoamerican languages: “The most obvious manifestation of the impact of Spanish is in the grammatical patterns of indigenous languages, and the extent of this type of borrowing runs counter to the traditional view that grammatical words are seldom borrowed.” A general treatment of Spanish borrowed function words is given in Stolz and Stolz (1994).

The deviant pattern results from the fact, I claim, that there is not a single borrowing process, just like there is no single code-mixing process. In addition to the familiar pattern of insertion (producing the Whitney/Haugen borrowability hierarchies), there is a pattern of alternation involving interjections and conjunctions. This alternation-type of borrowing will be further studied in the next section. Insertion is mostly a form of unidirectional language influence, while alternation often goes both ways. Insertion is constituent-internal, alternation is phrase- or clause-peripheral. Congruent lexicalization is word-internal.

**Alternational borrowing**

Conjunctions, prepositions, and discourse-markers are often introduced through alternation rather than insertion. Several arguments may be adduced for this claim.

First of all, sometimes we have *doubling* as in the following examples from Popoloca/Spanish borrowing data encountered in Veerman-Leichsenring (1991):

[5] cùnda nge: thèéná ngu: karru nà para i:ši: me: ... (VL 393)

have-1 that PR-look.for-1ex one car for that then ...

‘We have to look for a car so that then ...’


thus cart-INS-1pl water far for that IMP-drink-1pl

‘Thus we carted the water from afar in order to drink it.’

Here the Spanish preposition/conjunction *para* ‘for’ is combined with the Popoloca conjunction *i:ši:, even though either could have been used (Veerman-Leichsenring, 1991). Doubling in itself suggests a paratactic structure, since what is added does not fulfill an obvious role in the structure.
Second, notice that *para* is external to *i:ši:*, with respect to the complement clause. This external doubling is an extra argument for alternation. The Spanish element is simply added or adjoined to the clause here, and adjunction is always external.

Third, all or most of the borrowed conjunctions and prepositions are not equivalent to Popoloca elements, as becomes clear from Veerman-Leichsenrings grammatical description. Thus, there could not be insertion, since it is not clear what would be the site of insertion, unless the Spanish elements were dramatically restructured.

Unfortunately, nothing is known about the Spanish of the area in which Popoloca is spoken, to see whether Popoloca interjections have been incorporated into the local Spanish or not.

Similar data can be found in Otomi/Spanish language contact (Hekking, 1995:147-186). Hekking has done a very detailed study of a language related to Popoloca, Otomí, as spoken in the Mexican state of Querétaro. The same arguments can be given for the borrowing of Spanish conjunctions and prepositions in Otomí as for Popoloca. Hekking and Muysken (1995) show that the proportion of borrowed function words of all borrowings is much higher for Otomí than for Quechua, as we saw. Furthermore they argue that on the whole the borrowed categories are not present in the same way in Otomí, and it is also clear that there is extensive external doubling with semantically related but structurally dissimilar Otomí elements. In [7] Spanish *pa* precedes Otomí *dige*, both roughly meaning ‘for’.

[7]  När hyokungú bi hoky ‘nar ngú pa dige ar nzóyò *dige pa*
DETsg RC-make-house PRS3 make INDSg house for for DET delegate
‘The mason builds a house for the delegate.’

In the following examples Spanish *como* ‘like’ precedes an Otomí element that either shows up as a suffix or as an independent word:

[8]  Ya tsōho jwei komo-ngu ‘nar nhñe *ngu komo*
DETpl star shine like-like INDSg mirror
‘The stars shine like a mirror.’

[9]  Yogo’a hingi pa kor ntogebojá, ho gi mpefi, *jangu komo*
komo jangu di pōje?
why NEG-prs2 go with-ps2 ride-iron where Prs2 work like-like Prs1 go-
EXCPl
‘Why don’t you ride your bicycle to work like we do?’
We also have cases of temporal conjunctions like *asta 'until' and *kwando 'when' preceding an Otomi temporal linking element:

[10] Asta nu’bu da zor tsi ‘ye, ja ga pot’i
until when FU3 RC-arrive-Dsg DIM rain, make FU1 sow
‘When the rains come, I will sow.’

when when nothing be DETsg work, sometimes gather DETpl husband
or DETpl man
‘When there is no work at all, sometimes the husbands or men gather
and ...’

Finally, the Spanish complementizer *ke 'that' is doubled sometimes in relative clauses:

[12] ja ‘bowar sei *ná’ä ngi ödi
be be-LOC.CIS DETsg pulque that that Cpr2 ask
‘Here is the pulque that you asked for.’

In all cases, only the order Spanish element/Otomí element occurs, never the reverse. It should be understood that there are other sentences in which either the Spanish or the Otomí elements occur, or in which the linkage is understood from the context and not overtly expressed.

Data from Quechua/Spanish language contact confirm this perspective. A preliminary analysis of the code-mixes in Spanish/Quechua bilingual songs (Escobar and Escobar, 1981) did not show unfamiliar patterns of distribution: frequent switches at the fringe of the sentences, including exclamations, quotes, persons addressed, and within the clause, frequent mixes involving adverbial prepositional phrases. Important from the perspective of this chapter is the phenomenon of doubling, frequent in the corpus of Quechua/Spanish bilingual songs, identified below by their number in the anthology (different from the corpus analyzed in Van Hout and Muysken, 1994).

In three cases Spanish prepositions are borrowed, which may be nearly impossible in spoken Quechua, which has only post-positions and case suffixes. In [13] we find borrowing of *disdi ‘since, from’ in a couplet, and the occurrence of the same preposition *desde in a codeswitched line of the following couplet:
Three processes of borrowing: borrowability revisited

[13] disdi warma-manta from childhood on
yanayuq kanaypaq for me to have a lover
desde su palacio from its palace
gubirnu kamachikamun the government commands

In the first line of [13] disdi doubles with the Quechua case suffix -manta ‘from’. The same thing occurs in [14], where para ‘for’ doubles with Quechua benefactive case -paq.

[14] sipas kahtan nini, manas kuraqtachu I want the girl, not the oldest daughter
para paya-paq-qa, wasiypipas kanmi as for old women, those I have at home

In [15] finally we have the Spanish subordinating conjunction si ‘if’:

[15] si-chus munawanki chayqa if you love me
si-chus waylluwanki chayqa if you care for me
en prueba de tu cariño as proof of your affection
kay kupata tumay drink this cup

In Treffers-Daller (1994) considerable attention is devoted to syntactically unintegrated French discourse markers in Brussels Dutch. In [16a-c] the relevant paradigm is presented. The borrowed adverb pertang ‘still’ does not trigger verb second when it occurs in initial position, as in [16a]. Thus examples such as [16b] do not occur in the corpus. Still, the Dutch equivalent of pertang, toch, does trigger verb second, in examples such as [16c].

[16] a. Pertang ze hadden gezeit dat ik die mocht hebben. Still they had said I could have it. (TD 91)
   b. Pertang hadden ze gezeit dat ik die mocht hebben. [non-occuring]
   c. Toch hadden ze gezeit dat ik die mocht hebben. [constructed]

We can argue that the difference between [16a] and [16c] is due to the fact that toch in [16c] is inserted, and pertang in [16a] simply adjoined.

In addition to doubling, externality, lack of equivalence, and lack of syntactic integration we can also consider directionality. This can be illustrated in several bilingual data sets. Treffers-Daller (1994:99-104) looks both at French mixing in Brussels Dutch and at Dutch mixing in Brussels French. Here an asymmetry becomes apparent. In Brussels French only 0.8% of the nouns is of Dutch origin (28 out of 3419 tokens), while for interjections
this is 2.5% (63 out of 2499 tokens). Interjections are in fact the only type of Dutch element which represents more than one percent of the corresponding category in Brussels French. In Brussels Dutch, the reverse holds. French loans constitute 17.7% of the Brussels Dutch nouns (2329 out of 13179), and French interjections 6.8% (496 out of 7258) of the Brussels Dutch set. So within Brussels there is a tremendous asymmetry (Treffers-Daller, 1995): the influence of French on Dutch is much more massive than the reverse. Nonetheless, this asymmetry affects nouns much more than interjections. If we assume that nouns are borrowed through insertion and interjections through alternation, it is clear that insertional mixing is unidirectional and involves a matrix/non-matrix asymmetry, while alternational mixing is bi-directional.

Similar effects can be observed in Nortier (1990). Consider cases such as [17], where an Arabic conjunction occurs in a Dutch sentence:

[17] Ik ben een dokter *wella* ik ben een ingenieur.          (Nortier 142)
I am a doctor or I am an engineer.

The following Arabic conjunctions occur in Dutch sentences:

[18] walakin  ‘but’
9la-heqq-a$  ‘because’
wella  ‘or’
be-l-heqq  ‘but’

We also have many Arabic interjections, which Nortier (1990:124) terms extra-sentential mixes:

[19] ... mineraalwater, *ze9ma*, als cadeau
     ... mineral water, so to speak, as present

[20] Maar de tijd die gaat toch voorbij, *fhemti*?
     ‘But the time, it still goes by, understand-2?’

Elements in extra-sentential mixes include:

[21] fhemti(ni)  ‘you (PL) understand?’  14  (Nortier 124)
u/wa dak-ś-śi  ‘and/or so’  5
la  ‘no’  5
iwa  ‘well’  4
We also have borrowing of conjunctions in the other direction:

[22]  tebgĩ teفهم eh terwijl hadĩk l-mas’ala ma-tehtaż-§ (Nortier 125)  
2-want 2-understand uh while that DET-question 3-be-not-necessary  
‘You want to understand uh, while that question is not necessary.’

Comparing the number of introduced Arabic discourse linkers in [21] to the number of Dutch ones in [23], we find it is roughly the same number in both directions:

[23]  ja ‘yes’ 15  
nee ‘no’ 7  
nou ‘well’ 6  
hè ‘isn’t’ 5

For content words a dramatically different picture emerges, as can be seen in section (a) of Table 1. Almost all single major category elements borrowed are Dutch:

Table 1: Directionality of single word switches in Moroccan Arabic (MA) / Dutch (DU) codeswitching (based on Nortier, 1990:Table 7.15)

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>%MA&gt;DU</th>
<th>%DU&gt;MA</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Major</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adjectives</td>
<td>26</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>verbs</td>
<td>11</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>numerals</td>
<td>4</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>nouns</td>
<td>286</td>
<td>99.7</td>
<td>.3</td>
</tr>
<tr>
<td>(b) Minor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pronouns</td>
<td>6</td>
<td>83.3</td>
<td>16.7</td>
</tr>
<tr>
<td>adverbs</td>
<td>41</td>
<td>70.7</td>
<td>29.3</td>
</tr>
<tr>
<td>prepositions</td>
<td>8</td>
<td>62.5</td>
<td>37.5</td>
</tr>
<tr>
<td>(c) Discourse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>conjunctions</td>
<td>20</td>
<td>30</td>
<td>70</td>
</tr>
</tbody>
</table>

The category ‘minor’ (b) probably presents a mixed case, and needs much further analysis. However, for conjunctions (c) a bi-directional pattern is evident. A number of explanations can be given for the use of discourse markers from another language. I will review some of these here.
Hamel (1995:158), analyzing Otomi/Spanish bilingualism (but in a different region from Hekking) claims that language shift affects different domains at different speeds, and distinguishes the domain of cultural patterns and procedures from that of discourse structures and finally, from that of linguistic codes and structures. Given that either cultural patterns and procedures or discourse structures could shift first (depending on the circumstance), there is a potential explanation for the frequent occurrence of Spanish discourse markers in Otomí: once we adopt that idea, we could say discourse linkers belong to a domain (that of discourse structure) affected earlier by language shift, than that of the sentence itself. Hence they are in Spanish before the rest of the sentence. There are several problems with such an explanation:

• there are cases where the intrusion of discourse markers from a different language does not correspond to language shift, like the case described by De Rooij (1996) involving Shaba Swahili with French discourse markers;

• there are cases where language shift occurs but discourse organization patterns, even in the language of wider communication, remain modeled on indigenous cultural practices;

• it is not obvious that in the case of e.g., Moroccan Arabic/Dutch bilingualism, the use of discourse markers corresponds to a specific cultural orientation or discourse structure. Recall that the use of alien discourse markers was bi-directional.

Thus differential shifting patterns can at best only provide a partial explanation.

A second type of explanation, given by Hekking and Muysken (1995), likewise lacks sufficient generality. We suggested that Spanish discourse markers truly fill gaps in Otomí, which tends to leave relations between clauses and arguments very implicit. Again, this explanation does not get very far with some of the other bilingual settings discussed, and even in Otomí the pattern of borrowing is more pervasive than could be predicted from gap-filling.

A third explanation, given by De Rooij (1996), is more satisfactory. De Rooij argues in his extensive case study of Shaba Swahili/French code-mixing in Lubumbashi, Zaire, that code-mixing has an important contrastive function and as such functions as a contextualization cue. Since discourse markers have the same function, we will often find code-mixes involving discourse-markers. De Rooij illustrates his argument with a detailed analysis of the alternate use of French *que* and Swahili *asema*, showing that *que* is used almost as frequently as its Swahili counterpart in Swahili bilingual discourse,
but has been reinterpreted as syntactically peripheral, parallel to *asema*. In this way, it escapes being a potential side of code-mixing constraint violations.

**Integration of elements borrowed through alternation**

Given the relatively free and peripheral character of alternation-style code-mixing, we may wonder how elements introduced through alternation are integrated. This is a subject which needs much more serious study.

The Quechua data show evidence of grammaticalization of adverbs such as *seguro* ‘certain’ and *acaso* ‘perhaps’ as subordinators or evidential markers:

[24]  
\[
\begin{align*}
\text{sigrur} & \text{ manan}a \text{ mamayqa kanchu} \\
& \text{ certain that I have no mother any more?}
\end{align*}
\]
\[
\begin{align*}
\text{sigrur} & \text{ taytayqa manan}a \text{ kanchu} \\
& \text{ certain that I have no father any more?}
\end{align*}
\]

[25]  
\[
\begin{align*}
\text{akasu} & \text{ nuqapaq mansana phaltanchu} \\
& \text{ as if I lack apples}
\end{align*}
\]
\[
\begin{align*}
\text{akasu} & \text{ nuqapaq sultira phaltanchu} \\
& \text{ as if I lack girls}
\end{align*}
\]

In [26] a Spanish form *awir* (<Sp. *a ver*) ‘let’s see, lit. to see’ is used as a sentence introducing element:

[26]  
\[
\begin{align*}
\text{chukchachaykita t’ipiykukuy} \\
& \text{ pull your little hair}
\end{align*}
\]
\[
\begin{align*}
\text{awir} & \text{ manachus nanasunki} \\
& \text{ let’s see if it does not hurt}
\end{align*}
\]

Possibly *sigrur* parallels Quechua *-mi* ‘affirmation, certainty’, *akasu* Quechua *-cha*, and *awir* Quechua *-chus*. In this case, we would have grammaticalization along the lines of already existing categories in the language.

For Otomi the situation is somewhat more complicated. Compare the small Otomí corpus analyzed in Hekking and Muysken (1995) with the Quechua corpus from Aguiló (1980) analyzed in the same paper. In Table 2 some relevant contrasts are given:
Table 2: Spanish function words in the Otomí corpus and the Quechua corpus analyzed in Hekking and Muysken (1995) (types in parentheses, * forms only in fixed expressions)

<table>
<thead>
<tr>
<th></th>
<th>Otomí</th>
<th>Quechua</th>
</tr>
</thead>
<tbody>
<tr>
<td>coordinators</td>
<td>26 (6)</td>
<td>119 (14)</td>
</tr>
<tr>
<td>discourse markers</td>
<td>15 (9)</td>
<td>4* (3)</td>
</tr>
<tr>
<td>prepositions</td>
<td>36 (11)</td>
<td>17+2* (4)</td>
</tr>
<tr>
<td>subordinators</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is clear that overall the Quechua corpus contains many more coordinators and discourse markers than the Otomí one (even though for nouns the discrepancy is even greater: 98 Spanish nouns in the Otomí corpus, versus 823 in the Quechua corpus). However, for prepositions (not really borrowed into Quechua) and subordinators (twice as frequent in the Otomí corpus) a different picture obtains. This sharp difference can be related, perhaps, to the fact that Otomí is VO, like Spanish, and Quechua OV. Thus Spanish prepositions and subordinators can be incorporated much more easily into Otomí, even where the category was lacking in the original Otomí. Here clause linking was mostly indicated through parataxis and particles expressing oblique grammatical relations such as benefactive (corresponding to the Spanish preposition *pa*) and comitative (corresponding to Spanish *con*) were incorporated into the verb. Typological factors may thus facilitate integration of alternated elements.

Congruent lexicalization

When we turn to congruent lexicalization, it may not be easy to find patterns of borrowing that resemble it, since congruent lexicalization generally involves consideration of a string of words. Still, there are some word-internal borrowing phenomena that result from shared (word)grammar. I will give two examples, one from English - German mixed compounds (Clyne, 1967:34-5), and one from Aymara - Quechua affix borrowing (Adelaar, 1986; Muysken, 1988; Van de Kerke, 1996).

In Australia, German immigrants will often form bilingual compounds, headed (Williams, 1981) either by a German (most common 24 types listed) or an English word (7 types listed):
The predominance of German headed compounds reflects the fact that German is the matrix language in this bilingual corpus. Several cases of English headed compounds are based on very specific English compounds, like greengrocer and chain store. In any case, it is clear from the examples and their gloss how close German and English are in this respect. The similarity of the compounding pattern in the two languages makes it plausible to regard this as an example of borrowing through congruent lexicalization. The bidirectionality of the process points in the same direction: both German-headed and English-headed compounds occur.

A second example of borrowing through congruent lexicalization comes from the contact between Quechua and Aymara (Adelaar, 1986), e.g., in the Quechua of Puno, Peru. These two highly agglutinative Andean languages have been spoken in the same regions for over thousand years, but they present an enigma for genetic and historical linguists in that their basic vocabulary and many of their affixes are clearly distinct, making a common ancestor unlikely. At the same time they have borrowed extensively over time and show an uncanny resemblance in their grammatical structure. In several areas, affixes have been borrowed between them. An example from Chumbivilcas Quechua (Muysken, 1988) is given in [28].

[28] llank’a - naqa - yu - ni - n
  work try INT 1s AF
  ‘I am certainly trying to work.’
Here the affix \textit{-naqa-} has been borrowed into Quechua from Aymara, but it occupies a slot that is easily available for affixes in either language. The similarity of patterning makes affix borrowing through congruent lexicalization quite easily possible.

Adelaar (1986) shows that the borrowed Aymara affixes \textit{-thapi-} and \textit{-t’a-} trigger deletion of the last vowel of the Quechua base they are attached to:

\begin{align*}
[29] & \quad \text{tiy - thapi - chi - ...} & \text{(full form tiya-)} \\
& \quad \text{live together CAU} \\
& \quad \text{‘permit that they live together’} \\
[30] & \quad \text{pas - t’a - ku - ...} & \text{(full form pasa-)} \\
& \quad \text{pass suddenly RE} \\
& \quad \text{‘pass suddenly’}
\end{align*}

This deletion rule does not exist in Quechua, which underscores the analysis given of these cases as congruent lexicalization. The Aymara morphonemic system is active at the same time as the Quechua system. Van de Kerke (1996) documents the pervasive influence of the Aymara morphological structuring in Bolivian Quechua.

Conclusions and discussion

The different borrowing patterns reflect the tri-partite division proposed here between different styles of code-mixing and provide independent support for them. At the same time, the tenor of this chapter has been that there is no watertight division between code-mixing and borrowing: any mixed type of item that can be interpreted as a lexical unit could potentially be borrowed.

References


