Guarani Morphology in Paraguayan Spanish: Insights from Code-Mixing Typology

Bruno Estigarribia
University of North Carolina at Chapel Hill

Abstract: In this paper we examine the use of Guarani affixes and clitics in colloquial Paraguayan Spanish. We depart from the traditional view of these as “borrowings,” and instead explore the idea that these phenomena can be integrated within Muysken’s (2000, 2013, 2014) typology of code-mixing. We claim that most of these uses may stem from a strategy of backflagging that bilingual speakers employ to mark their use of a (historical, colonial) L2 (Spanish) with elements from the original community L1 (Guarani). In so doing, however, we identify some problematic areas in applying Muysken’s model to these data. We conclude by showing how this perspective shift in the analysis of Guarani-Spanish contact phenomena can both increase our understanding of this particular contact situation, as well as of code-mixing typology.

Keywords: affix borrowing/préstamo de afijos, agglutinative languages/lenguas aglutinantes, code-switching/alternancia de código, Jopara/jopara, mixed languages/lenguas mixtas

1. Introduction

An oft-noted characteristic of colloquial Paraguayan Spanish is the “intrusion” of particles from Guarani. In (1), we show the occurrence of Guarani (a) interrogative markers; (b) emphatic markers; (c) plurals; (d) imperatives; (e) intensifiers; and (f) diminutives. These particles occur exceptionally frequently in Paraguayan Spanish, with their original morphological form and meaning (Granda 1980). They are assumed to result from interference or borrowing (Gómez Rendón 2008; Granda 1980, 1996; Krivoshein de Canese and Corvalán 1987; Penner, Acosta, and Segovia 2012; Usher de Herreros 1976).

(1)

a. ¿Entendite, pa?
   entendiste = pa
   understood.2sg = Q
  ‘Did you understand?’ (Granda 1980: 62)

b. Así co é la vida
   asi = ko
   es = la vida
   Thus = EMPHATIC is the life
  ‘That’s life.’ (Granda 1980: 63)

c. Vinieron sus amigo cuéra.
   vinieron sus amigo = kuéra
   came.3pl his/her friend = pl.
  ‘His/her friends came.’ (Krivoshein de Canese and Corvalán 1987: 48)
d. Vení na un poco.
vení-na un poco
Come.a2sg-requestative a little
'Come for a second, please' (Krivoshein de Canese and Corvalán 1987: 46)

e. ella no quiere ité a lo negro
ella no quiere-ite a los negros
she not wants-intensifier to the blacks
'she doesn’t like blacks at all' (Granda 1980: 61)

f. en otro lado un bife-imí te cuesta más de un quinientón
en otro lado un bife-i-mí te cuesta más de un quinientón
in another place a steak-dim-dim to. you costs more of a five-hundred
'In a different place a steak will cost you more than five-hundred' (Granda 1980: 60)

Usually, a borrowing has to 1) be morphologically and phonologically integrated in the
system of the borrower language; 2) (often) replace the original item or fill a lexical gap; and
3) be perceived by monolingual speakers of the borrower language as belonging to it, not as
mixed speech (Winford 2010). None of these criteria are consistently verified for Guarani
“grammatical borrowings.” Moreover, borrowings often show semantic change (which Gómez
Rendón (2008) calls resemantization), but the Guarani “borrowings” analyzed here conserve
their original semantic values (Granda 1980; Penner, Acosta, and Segovia 2012).²

We propose a new perspective on these “intrusive” Guarani morphemes. We set aside
the question of whether they are borrowings, and we analyze them instead as code-mixes, produced
by Guarani-Spanish bilinguals when they engage in Spanish-based discourse.³ This shift in
perspective allows us to bring to the analysis tools from code-switching theory. We will classify
these Guarani “morphological borrowings” within the framework of Muysken’s (2000; 2013;
2014) typology of code-mixing processes. All of the examples in (1) satisfy Muysken’s defini-
tion of code-mixing as the appearance of lexical items and overt grammatical features of more
than one language in a single sentence. Nevertheless, our data suggest that further refinements
to Muysken’s theory may be in order, if it aspires to greater empirical coverage. Thus, this new
way of looking at Guarani particles in Paraguayan Spanish allows us to better integrate this
phenomenon within recent advances in language contact research.

Our reasons for employing Muysken’s model are manifold. First, it is a typological model
that aims at maximum generality in classifying code-mixing language pairs. The typology relies
on structural factors, but importantly also on sociohistorical and psycholinguistic parameters
(which will not be discussed here for lack of space). It is our working hypothesis that the model
has sufficient generality to explain our data. This is justified because Muysken (2000) examined
a large database of language pairs (almost 90 languages), representing several different language
families, and in varied language contact situations. Furthermore, a theoretical framework that
recognizes various possible bilingual strategies, using of a plethora of structural and sociolin-
guistic criteria to identify them is well suited to the case of Guarani-Spanish, which displays
many distinct kinds of mixing (Estigarribia 2015).

Importantly, we do not expect the application of Muysken’s model to this contact situation
to be without drawbacks. A reviewer notes that this model has mainly been implemented to
analyze switches from a hegemonic language into a minoritized language (for example, Spanish
in Otomi or Spanish in Quechua), not the converse, as is the case for the examples in this paper.
However, a few language pairs in Muysken (2000) defy this generalization. For example, the
analysis of English (hegemonic)/Spanish (minoritized) in New York shows function words from
Spanish inserted into English. In Dutch (hegemonic)/Moroccan Arabic (minoritized), 70% of
discourse marker switches are Moroccan Arabic conjunctions in a Dutch sentence. Crucially for us, this last case is reanalyzed in Muysken (2013, 2014) as backflagging. Hence, it is possible that elements of a minoritized language appear in a hegemonic language via a dedicated mechanism, namely, backflagging as we will propose.

We of course acknowledge that there are other ways of analyzing our data. In section 8, we will examine Guarani switches according to their function (as a reviewer suggests), a dimension that is but weakly represented in Muysken’s model.

2. Preliminary Remarks about Guarani

Guarani and Spanish have a long history of contact since the sixteenth century when a relatively small number of Spaniards occupied the area that is now Paraguay. Historically, this colonization lacked a continuous influx of colonists (Zajícová 2009). Settlers were mostly males, with very few females (Gómez Rendón, forthcoming). Moreover, the area was rather isolated during much of the colonial period. Early on, Perú was much more accessible for the Spaniards via Panama (Gómez Rendón, forthcoming). Later, the Asunción region was always perceived as peripheral with respect to Potosí first and Buenos Aires later (Kleinpenning 2011). Finally, the use of Guarani by Jesuits in the reductions is generally considered of paramount importance in the continuity of Guarani. The Jesuits composed the first grammars and dictionaries of Guarani, many of which survived the Jesuits’ expulsion in 1767 and 1768 (Zajícová 2009).

Guarani is currently estimated to be used by about 80% of the population, of which only under 1% is indigenous Guarani. Paraguayans normally communicate in a mixed code called Jopara. This term is variously used to refer to Guarani with Hispanisms, to Paraguayan Spanish with Guarani influence, or to more extensive Guarani-Spanish code-switching (e.g., Palacios Alcaine 2008; Zajícová 2009). Sentences in Jopara can be either Guarani-based or Spanish-based (Estigarribia 2015; Gómez Rendón 2008). Most Paraguayans would view the data analyzed in this paper as Jopara (see Zajícová 2009 for evidence), hence recognizing them as instances of mixed speech.

Important for Jopara is the fact that Guarani is typologically very different from Spanish. Spanish is an analytic-fusional language: words tend to contain a small number of morphemes, and grammatical morphemes can (and often do) fuse together two or more meanings. In contrast, Guarani is agglutinative: it expresses most grammatical meanings by adding affixes to a base, and these affixes convey a single grammatical meaning. Guarani agglutinative affixes show low degrees of lexical selection, that is, they attach to all bases of a given category, whereas fusional affixes (those from Spanish) show sensitivity to different declension or conjugation classes. These properties make the grammatical affixes of Guarani well-suited for “import” into Spanish (Field 2002; Hickey 2010).

3. Muysken’s Typology of Code-mixing Strategies

Muysken (2000) proposes that code-mixing can come about via three different strategies: insertion, alternation, or congruent lexicalization.

In insertions, content words or small phrases with content heads from an embedded language occur in a clause in a different base/matrix language (Myers-Scotton 2002). Insertions often receive affixal morphology from the base language (morphological integration, 2). The matrix language material before and after the switch is usually grammatically related (nestedness). Insertions tend to happen under categorial equivalence (the inserted embedded language item replaces an equivalent item in the matrix language). Finally, they are usually either predicates or units selected by a predicate.
In *alternations*, we find longer fragments of Language A and Language B, neither of which is nested or integrated inside the other. Such clauses are generally not clearly identifiable as being in either language, and therefore they do not have a base/matrix language (3). Alternation allows for some elements to appear simultaneously in both languages (doubling) whereas insertions replace Language A items with Language B items. Alternation often requires identical word order at the switch site (linear equivalence). Alternational switches tend to be non-nested, and to involve long, complex, or several constituents.

Finally, congruent lexicalization is used when both languages share a single syntactic structure into which items from either language fit (4). Congruent lexicalization often involves short, non-complex functional items, but also switches that are not coherent constituents (called “non-constituent switches”). It often gives rise to a much higher density of switches in a sentence. Typological similarity between the languages in contact is crucial for the viability of this strategy.

Recently, Muysken (2013; 2014) defined a new strategy of backflagging as the use of heritage language L1 discourse markers in L2 discourse (assuming language shift from L1 to L2). Backflagged switches tend to be peripheral to the clause (not participating in predicate-argument structure, for example), non-nested, and to be single, simple forms, usually discourse markers (5). Backflagging is not expected with lexical or functional elements tightly integrated in clausal syntax, and for that reason backflagged switches were originally analyzed as alternations in Muysken (2000).
The strategies are then redefined in the following way (Muysken 2013: 6):

(6) **Insertion**: Use the L1, i.e. the grammatical and lexical properties of the first language, as the matrix or base language.

**Congruent lexicalization**: Produce structures and words which share properties of L1 and L2.

**Alternation**: Use universal combinatory principles, procedures by which fragments from different languages can be combined independently of the grammars involved.

**Backflagging**: Use as much as possible of the L2 (i.e., the grammatical and lexical properties of the second language, as the matrix or base language).

When viewed this way, backflagging turns out to be not a kind of alternation, but a mirror image of insertion. However, insertions tend to involve content items tightly integrated in the grammatical frame of the matrix language, whereas backflaggings occur with discourse markers from an L1 which are only loosely integrated in the grammatical structure of the sentence in L2. That is, backflaggings, unlike insertions, tend to happen at clause boundaries and in non-nested contexts. Moreover, a non-structural difference between the two strategies is apparent in the definitions above: for a switch to be diagnosed as backflagging, the switch must belong to the original community language in a case of language shift, and be used as a marker of ethnic identity.

4. Data Sources

For this study, we collected data from two written sources and a movie. We chose them because they employ Jopara instead of Standard Paraguayan Spanish (which is by definition devoid of any Guarani mixing).

Our first source is the novel *Ramona Quebranto* by Paraguayan author Margot Ayala de Michelagnoli (1989). This novel tells the struggles of Ramona, an impoverished Paraguayan from a very poor neighborhood in Asunción. The authenticity and linguistic representativity of Ramona Quebranto has recently been argued for in detail in Estigarribia and Wilkins (forthcoming). Several researchers have used it as a corpus for studying Jopara (see, e.g., Estigarribia 2015; Gómez Rendón 2008; Kallfell 2011; Zajícová 2009).

Our second source is also fictional, the Paraguayan film *Siete Cajas/7 Boxes* (Maneglia and Schembori 2014). This film tells the story of a pushcart porter, Víctor, in the Mercado 4 (Market number four) in Asunción. Víctor receives a proposal to transport seven boxes of unknown contents in exchange for an important sum of money and thus becomes involved in a serious crime. Linguistically, the film is important because it takes place mostly in Jopara.

Finally, our third and last source of data comes from social media. An advantage of social media data over literary data is that the former is naturalistic (produced spontaneously in a real interaction in a typical communicative setting, as opposed to rehearsed or elicited).

We culled data from the Facebook page Paraguay Oficial and from the Facebook feed for the Paraguayan newspaper *ABC*. We also looked at the Revista Arsenio Eríco, a digital newspaper from the Paraguayan community in New York, and we looked in the comment threads from the page of Paraguay’s *Diario Popular* as well. Additionally, we examined a Paraguayan forum about cars called Motores.com.py and a general forum for Paraguayans, La Jaula. Finally, we looked at the blog ndersasore.com, a personal blog with jokes, viral content, and memes. We supplemented all these data with Google searches for specific Guarani morphemes in Paraguayan webpages.

We focused on a subset of bound morphemes from Granda (1980) because bound morphemes *a priori* violate word-internal switching constraints, which may in turn explain the tendency to classify them as borrowings. Additionally, bound morphemes are of theoretical relevance, since their analysis within the parameters of Muysken’s typology is not clear (see section 7).\(^5\)
5. Guarani Particles in Paraguayan Spanish as Code-mixes

We classified our Guarani bound morphemes in three different groups: Guarani particles with sentential scope (interrogative and emphatic markers); Guarani word-level affixes (imperatives, intensifier, diminutive, and nominal past); and the Guarani plural clitic which has phrasal scope. This classification highlights commonalities and differences in the use of these morphemes and their relation to different code-mixing strategies. (We discuss a different classification, suggested by a reviewer, in section 8.)

5.1 Guarani Second-position Clitics with Sentential Scope

Guarani questions have the same intonation and word order as declaratives. They are marked by interrogative second position clitics, the most common being pa and piko (7) (Ayala 1996). Similarly, Guarani has a number of evidential or attitudinal morphemes. One of the most frequent is the second position clitic =ni(k)o, =ko, =ningo, or =ngo, marking truthful narrative mood (Guarania 2008) and emphasis (Ayala 1996) (8).

(7) Kola ha Peru pa opytu’u.
   [Kola ha Peru]=pa o-pytu’u
   Nicolas and Pedro=q 3-rest
   ‘Are Nicolas and Pedro resting?’

(8) Kuehe niko omba’apo gueteri.
   Kuehe=niko o-omba’apo gueteri
   Yesterday=EMPHATIC 3.ACTIVE-WORK still
   ‘Yesterday, yes, he still worked.’ (Guarania, 2008: 40)

The usage of both these particles in Spanish is well-attested in our corpus. Interrogative markers are used in yes/no questions, wh-questions, and embedded questions (9–12).

(9) no tenes pio nada de ideas para esto?
    no tenes=pi(k)o nada de ideas para esto
    no you.have=q nothing of ideas for this
    ‘You don’t have any ideas for this?’ (Marcelo 2011)

(10) ¿E cierto pa que Julio e tu novio?
     es cierto=pa que Julio es tu novio
     is=true=q that Julio is your boyfriend
     ‘Is it true that Julio is your boyfriend?’ (Ramona Quebranto, henceforth [RQ])

(11) como pio guarani no vas a entender?
     como=pi(k)o guarani no vas a entender
     how=q guarani no you.go to understand
     ‘what do you mean, you don’t understand Guarani?’ (WhatsApp Paraguay 2013)

(12) Solo queremos sabé si vive o se murió pa.
     solo queremos saber si vive o se murió=pa
     only we.want know if lives or died=q
     ‘All we want to know is if he is alive or dead.’ [RQ]
The spoken data from *Siete Cajas* reveal that the interrogative markers are not strictly necessary in Paraguayan Spanish, because Spanish question intonation is already present (13, 14).

(13) ¿Para qué\(5\) pa\(5\) queré(s)\(5\) saber?
For what\(=q\) you want know
‘Why do you want to know?’ (*Siete Cajas*, henceforth [7C])

(14) No queré\(s\)\(5\) pi(k)o\(5\) que te \(5\) pase a buscar . . .
No you want\(=q\) that to you I pass to seek
‘Don't you want me to come get you . . .?’ [7C]

The Guarani emphatic marker -(ni)ko (and its allomorphs) also occurs in our data, phonologically attached to a variety of categories (15–18).

(15) no me hace niko falta.
no me hace=niko falta
‘I don’t need him.’ [RQ]

(16) ella\=ngo es tu amiga
 ella=ngo es tu amiga
‘She is your friend.’ [RQ]

(17) cuando alguien está siendo electrocutado nio no podés tocarle!!
cuando alguien está siendo electrocutado=niko no podés tocarle
when someone is being electrocuted=EMPHATIC no you can touch him
‘When someone is being electrocuted you cannot touch him!’ (Ross Saucedo 2014)

(18) A quien se\(5\)ngo le tengo que vendé ese
A quien sea=ngo le tengo que vender ese
‘I must sell that to whoever’ [7C]

5.2 Guarani Word-level Affixes

In this category, we find Guarani imperative affixes, the intensifier suffix -ete/-ite, the diminutiv suffix -mi, and the nominal past suffix -kue. We consider them word-level affixes because, except for the intensifier suffix, each affix attaches to a particular word class, and because they modify words, not entire phrases.

Basic imperatives are conveyed in Guarani via the verb prefix e- for 2sg and pe- for 2pl. They are supplemented by suffixes that convey different degrees of imperative meaning (19): the forceful imperative -ke (glossed FORCEFUL), the requestative -na (REQUESTATIVE), for polite, mitigated requests, and the pleading imperative -mi (MITIGATIVE), with affectionate nuances (Guarania 2008).

(19) a. Ehecháke
E-hecha-ke
IMP.2SG-look-FORCEFUL
b. Ehechána
E-hecha-na
IMP.2SG-work-REQUESTATIVE
'Work, please'

c. Ehechami
E-hecha-mi
IMP.2SG-work-MITIGATIVE
'I beg you to look'

All of these imperative suffixes occur in colloquial Paraguayan Spanish, attached to the Spanish imperative verb forms (20). This can be perhaps interpreted as a kind of “doubling” (but see note 7 for more information).

(20) . . . que caminána así, que bañáteke.
que camíná-na así que bañá-te-ke
that walk.IMP-REQUESTATIVE thus that take.a.bath.IMP-FORCEFUL
‘Walk this way, take a bath.’ [RQ]

(21) Tomaneque el jugo y come la ensalada
Tomá-ke el jugo y comé la ensalada
Drink.IMP.me-FORCEFUL the juice and eat.IMP the salad
‘Drink your juice and eat your salad’ (De Paula 2012)

(22) papa eso es en Argentina, lee en la ensalada
pa-pá eso es en Argentina lee-na lo que pusé
dad that is in Argentina read.IMP-REQUESTATIVE it that I.put
‘That's in Argentina, dude, please read what I wrote’ (Enzomas 2013)

(23) Atendémeke bien ne mita'i
Atendéme-ke bien ne-mita’i
pay.attention.to.me.IMP-FORCEFUL well 2SG.INACTIVE-child-DIM
‘Pay close attention to me, child.’ [7C]

The second word-level affix we found in our corpus is the Guarani intensifying suffix -ete/-ite (Eng. ‘very’ or ‘at all’) (24).

(24) Ndaikuaaiete mbaèvetete.
Nd-ai-kuaa-i-ete mbaève-te
NEG-1SG.ACT-know-NEG-INTENSIFIER nothing- INTENSIFIER
‘I don’t know anything at all.’ (“Diputado Carlos Soler” 2009)

This suffix occurs mostly with Spanish verbs, but also with adverbs and nouns, intensifying their degree (25–28).

(25) Quiero ete que sea feli
Quiero-ete que sea[s] feli[z]
want-1SG-INTENSIFIER that be.SBJV.2SG happy
‘I want (so much for) you to be happy’ [RQ]
YA CONOZCOITE YA! . . . jajaja
ya conozco-ite ya!
already I know-intensifier already
‘Now I know (you), hahaha!’ (Fabioa 2012)

ahoraite degenero el laburo
ahora-ite degenero el laburo
now-intensifier I degenerate the work
‘I’ll play hooky from work this minute . . .’ (Nova 2012)

ajaj este si es paraguayoite jajajaj
este si es paraguayo-ite
this one yes is Paraguayan-intensifier
‘Haha this one is a real Paraguayan hahaha’ (Jepy 2010)

The diminutive affix -mi from Guarani occurs in Spanish much less frequently (see section 8), with an affectionate, not literal reading (29, 30).

No tengo ni plata para guiso
No tengo ni plata para guiso- mi!
no I have nor money for stew-dim
‘I don’t even have money for a little stew!’ [RQ]

jaja y pongan ahi agua y pasto para los camellos
y pongan ahi agua- mi y pasto para los camellos
and put there water-dim and grass for the camels
‘Hahaha, and put a little water and grass for the camels there’ (La Caballota, n.d.)

Finally, unlike Spanish, Guarani nouns take temporal markers. One of them is preteritive -kue (here preteritive), meaning ‘ex-‘ or ‘former’ (31).

che róga kue
che-róga-kue
1sg.inactive-house-preteritive
‘my former house’ (i.e., because I don’t own it anymore, or I don’t live there, or it is not a house anymore)

Preteritive -kue is found in our corpus, although not very frequently (see section 8) (32, 33).

este eraa mi auto kueee!!!!
este era mi auto-kue
this was my car-preteritive
‘this was my car.’ (Gio777 2010)

asi era mi novio kue
asi era mi novio-kue
thus was my boyfriend-preteritive
‘My ex-boyfriend was like that’ (Anónimo 2011)
5.3 Switches Involving the Guarani Phrase-level Plural Clitic

Traditional grammars consider the Guarani plural =\textit{kuéra} a separate word, but it is a phrasal clitic: it attaches to the last word in the phrase it pluralizes; it shows morphophonologically-conditioned alternations (=\textit{nguéra} is used when the last word of the phrase is nasal, as in (34)); and it cannot appear alone or move to a different place in the sentence. It is not an affix because it attaches to different word classes and it pluralizes the whole phrase.

(34) \textit{óga moroti nguéra}
\textit{óga moroti=nguéra}
house white=pl
'(the) white houses'

Use of \textit{=kuéra} in Spanish sentences is well attested in Ramona Quebranto (35), but we only found one example in social media (36). In \textit{Siete Cajas =kuéra} appears only in sentences that have Guarani as a matrix language, never in Spanish sentences.

(35) \textit{yo hablo a mi sobrina kuéra}
yo hablo a mi sobrina=\textit{kuéra}
'I talk to my niece=pl
'I talk to my nieces.' [RQ]

(36) \textit{Hoy no vas a salir afuera con esos tus amigo kuera.}
 hoy no vas a salir afuera con esos tus amigo=\textit{kuera}
today no you.go to go.out outside with those your friend=pl
'Today you will not go out with those friends of yours.' (De Paula 2012)


This section provides evidence that most of these mixes emerge via backflagging. This conclusion holds across the different morphemes we examined, except for plural =\textit{kuéra}, which we will examine separately and for which the evidence is inconclusive (see Table 1 below for a summary of the analysis). We also highlight what we believe are shortcomings of this typology in accounting for our data.

Let us consider first the evidence for and against insertion. The interrogative and emphatic markers considered here are often nested, since they appear in second position in the clause. Additionally, the material before and after the marker is often related by subject-verb agreement as in (37), or the material after the marker is an argument of the material before, as in (38).

(37) \textit{Nelson ko llegó tarde.}
Nelson=\textit{ko} llegó tarde
Nelson=EMPHATIC arrived late
‘Nelson came late.’ [7C]

(38) \textit{¿Viste pa cómo te miró Jim?}
\textit{viste=pa cómo te miró Jim}
you.saw=Q how to.you he.looked Jim
‘Did you notice how Jim was looking at you?’ [7C]

However, nestedness is not required and is purely accidental. Moreover, the mixed-in morphemes are functional items (not content words), are not selected by a predicate, and
furthermore they are syntactically optional. Additionally, these Guarani morphemes do not have equivalents in Spanish (except diminutive -mi and plural =kuéra). This lack of categorial equivalence actually precludes an analysis in terms of insertion (Muysken 2000: 108).

Turning now to alternation, lack of categorial equivalence is an argument in favor of an alternational analysis. Additionally, alternations are not selected by a predicate, and they are peripheral (not belonging to the core marking of the clause). Moreover, discourse markers are often borrowed via alternation (Muysken 2000: 106), and emphatic markers at least could fall under that definition.

However, alternations are generally non-nested, and involve long, complex, or several constituents. They tend to happen under linear equivalence in sentences where it is difficult to identify a matrix language. None of these criteria is verified in these cases, since there is no linear equivalence at the point of the switch, and the sentences are clearly Spanish sentences with well-delimited, small Guarani switches. We conclude, then, that alternation is not a plausible scenario.

Not surprisingly, given the typological distance between Guarani and Spanish, the features defining congruent lexicalizations are rather inconsistent with all switches (except perhaps the plural). Although congruent lexicalization often involves short, non-complex, functional items, and is consistent with switches not being constituents, the label “congruent lexicalization” explicitly highlights the idea of congruent syntactic structures of both languages as engines driving rather dense intra-sentential code-switches. That is, congruent lexicalization mostly requires linear and categorial equivalence at many switch sites; these are both absent in the Guarani-Spanish contact case. As Muysken (2014: 245) reminds us, congruent lexicalization occurs when speakers’ chosen strategy consists of the production of “structures and words that share properties of L1 and L2.” The fact that the Guarani morphemes are optional, not syntactically required in the Spanish sentences in which they appear, also argues against congruent lexicalization.

The fourth strategy in Muysken’s model is backflagging. There are a few arguments against a backflagging analysis. Backflagging switches (like alternations) tend to happen at major clause boundaries. Whereas some second-position clitics are switched at clause boundaries, and the other morphemes can appear often at the end of a clause, this is accidental, not a necessary occurrence. Also, Muysken (2014) states (albeit without argument) that linear equivalence favors a diagnosis of backflagging, just as it does for alternation. However, no linear equivalence is present here, except in the case of plural =kuéra, which appears in the same position in which one would find a Spanish plural suffix.

Nevertheless, in favor of this strategy we note that our Guarani switches are single, simple, well-defined items, exactly as required in Muysken (2013; 2014). Furthermore, they are peripheral because they are not selected by a predicate and do not participate in the core structure of the clause. We conclude, then, that backflagging is the code-mixing strategy that best fits our cases of use of Guarani morphology in Spanish sentences, except for the plural =kuéra, to which we turn now.

6.1 The Case of Plural =kuéra

Just as we argued above for the other Guarani morphemes, the plural clitic =kuéra is a single, simple, functional item, favoring a backflagging account. However, the plural is not peripheral to the clause, since it is required by the semantics of selected arguments and can enter into agreement relationships with other constituents. Plural as a grammatical category belongs to the core sentential syntax, not to a “looser” discourse syntax.

Furthermore, plural switches happen under categorial equivalence, since =kuéra has the same semantic value as the Spanish plural (Granda 1980). Linear equivalence is likewise verified in most cases since =kuéra cliticizes to a word that would receive Spanish plural marking in Standard Spanish.
The combination of categorial equivalence, linear equivalence, and non-peripheral status, plus the fact that {$=\text{k}u\text{era}$} is a functional item, strongly suggests a congruent lexicalization analysis. Yet, such an analysis is inconsistent with the lack of dense, back-and-forth switching in our sentences, which have Spanish as clear base (or matrix) language, and with the typological distance between Guarani and Spanish, since congruent lexicalization occurs when the languages in contact are typologically very close. We conclude that, given the evidence presented here, it is not possible to decide whether Guarani plural markers appear in Paraguayan Spanish via backflagging or via congruent lexicalization. Table 1 summarizes the analysis. In the next section, we will provide a critical reassessment of Muysken’s model in light of our data.

7. A Critical Reassessment of Muysken’s Model

We believe our Guarani-Spanish contact data brings to light some issues in applying Muysken’s typology to the analysis of bound morpheme switches. First, note that Muysken (2000) does not treat mixed words and function word switches uniformly. Lexical switches from a non-agglutinative language that are integrated in the matrix language clause via agglutinative morphology are straightforwardly analyzed as insertions. On the other hand, switches of functional items into an agglutinative language clause are analyzed as alternations (see 39). Neither of these are similar to Guarani switches into a Spanish clause and there are no other examples in Muysken (2000; 2013; 2014) on which to model the analysis of Guarani “grammatical borrowings.”


<table>
<thead>
<tr>
<th>Popoloca</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mē tʔayá-šaː=ná ndaʔ? hík pari išiː kʔúe-kʔtá=ná</td>
<td>‘Thus we carted the water from afar in order to drink it.’</td>
</tr>
</tbody>
</table>

Furthermore, Muysken’s typology accords a great deal of importance to whether a switch is a constituent or not. But are bound morphemes single constituents even though they are smaller than a syntactic constituent? Or are they non-constituents? We chose to interpret “non-constituent” as “ragged” mixing, where a switch straddles constituent boundaries, as do Muysken (2000) and Lipski (2009). Counting bound morphemes as single constituents of course favors an insertional or backflagging analysis.

Another problem concerns the interpretation of “morphological integration,” difficult in the case of bound morphology switches. Morphological integration is diagnostic of insertion. For bound morpheme switches, it is clear that some sort of morphological operation is taking place, but the bound morpheme which constitutes the switch can hardly be said to be morphologically integrated in the sentence in the intended sense, since it does not receive any morphology itself. And the bound morpheme does not help integrate the roots to which it attaches, since the sentence itself is in the language of the roots. Thus, we decided to consider morphological integration not applicable. As a result, both an insertional and a congruent lexicalization analysis are disfavored.

The application of the notion of “nestedness” is also problematic because for morphological switches nestedness is accidental, not necessary. Does that mean that nested instances of a given clitic/affix switch should be considered insertions, while non-nested instances of the same item possibly not? This doesn’t seem to make sense: the same grammatical morpheme is likely to be switched via the same strategy in all cases. This is a particular analytical conundrum for grammatical words, which have higher token frequency than lexical words, entailing that each morpheme has a high probability of displaying different accidental properties when switched.

Our conclusion is that Muysken’s framework is perhaps not completely adequately equipped to deal with switches that are bound morphemes as opposed to independent words or larger
Table 1. Count of features favoring each code-mixing strategy, applied to Guarani “grammatical borrowings”

<table>
<thead>
<tr>
<th></th>
<th>Insertion</th>
<th>Alternation</th>
<th>Congruent Lexicalization</th>
<th>Backflagging</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Most markers</strong></td>
<td><strong>In favor</strong></td>
<td><strong>Against</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. single/simple item</td>
<td>1. no categorial equivalence</td>
<td>1. single/simple item</td>
<td>1. single/simple item</td>
</tr>
<tr>
<td></td>
<td>2. not selected</td>
<td>2. functional</td>
<td>2. functional</td>
<td>2. functional</td>
</tr>
<tr>
<td></td>
<td>3. peripheral</td>
<td>3. peripheral</td>
<td>3. peripheral</td>
<td>3. peripheral</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. no categorial equivalence</td>
<td>4. no categorial equivalence</td>
<td>4. not selected</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. not selected</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Plural = kuéra</strong></td>
<td><strong>In favor</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. single/simple item</td>
<td>1. linear equivalence</td>
<td>1. single/simple item</td>
<td>1. single/simple item</td>
</tr>
<tr>
<td></td>
<td>2. categorial equivalence</td>
<td></td>
<td>2. functional</td>
<td>2. functional</td>
</tr>
<tr>
<td></td>
<td>3. linear equivalence</td>
<td></td>
<td>3. categorial equivalence</td>
<td>3. linear equivalence</td>
</tr>
<tr>
<td></td>
<td>4. syntactically required</td>
<td></td>
<td>4. linear equivalence</td>
<td>4. syntactically required</td>
</tr>
<tr>
<td></td>
<td>5. enters into agreement relations</td>
<td></td>
<td>5. enters into agreement relations</td>
<td>5. enters into agreement relations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1. not always at clause boundaries</td>
<td>1. not always at clause boundaries</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. not always at clause boundaries</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. enters into agreement relations</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. functional</td>
<td>1. single/simple item</td>
<td>1. single/simple item</td>
<td>1. single/simple item</td>
</tr>
<tr>
<td></td>
<td>2. not selected</td>
<td>2. functional</td>
<td>2. functional</td>
<td>2. functional</td>
</tr>
<tr>
<td></td>
<td>3. not always at clause boundaries</td>
<td></td>
<td>3. not always at clause boundaries</td>
<td>3. not always at clause boundaries</td>
</tr>
<tr>
<td></td>
<td>4. enters into agreement relations</td>
<td></td>
<td>4. enters into agreement relations</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
chunks. In spite of this, we believe our data provide empirical support for the newly-defined strategy of backflagging. We propose ways in which this strategy could be refined in the next section. Then, in section 8, we consider an alternative functional analysis.

7.1 Empirical Support for Backflagging and a Proposal for Redefinition

Muysken (2000) presented a wealth of data from different languages to define the three strategies of insertion, alternation, and congruent lexicalization. Backflagging, however, was motivated in Muysken (2013) by theoretical considerations of symmetry in the model, rather than forced by the need to adequately engage empirical data. Examples of backflagging were originally treated as alternations, which means that the bulk of the empirical data they represent was already satisfactorily accounted for in the original model.

We believe that the data presented here provides some empirical support for backflagging. Our examples demonstrate that backflagging is in reality distinct from alternation, and rather close to insertion. But, in order to make this notion more useful, we believe it should be redefined. First, the requirement that backflagging switches be discourse markers must be abandoned. Next, backflagging should always be considered a distinct possibility when switches are single functional items, in particular bound morphemes. This, in turn, means abandoning the requirement that backflagged switches be clausePeripheral, since many bound morphemes participate in core syntactic relations. Likewise, backflagging should be considered when there is no linear or categorial equivalence for a switch since this is the only strategy that does not require either.

To summarize, we propose the following structural criteria to re-define backflagging:

1) Properties of the clause showing code-mixing:
   a) it has a clear, identifiable base language

2) Properties of the switched material:
   a) it is a single, simple item
   b) especially if it is a functional morpheme OR a discourse marker

3) Assessment of equivalence:
   a) linear equivalence is absent
   b) categorial equivalence is absent

We believe that backflagging will be a particularly viable strategy when bilingual clauses comprise a non-agglutinative matrix language like Spanish and an agglutinative embedded language like Guarani, since in this case the potential for categorial and linear equivalence will be reduced.

8. A Functional Analysis of Guarani Grammatical Morpheme Switches

As highlighted in the previous section, one of the main difficulties faced by our typological analysis is that all of the markers we studied are more tightly morphosyntactically integrated in the sentence than what backflagging predicts. This suggests that perhaps the impact of morphosyntactic properties in a complete analysis should be weighed against a consideration of the functions these items have. To that end, we reclassified our Guarani morphemes in Table 2 according to semantic/pragmatic function.

The illocutionary force/speech act markers, the propositional attitude markers, and the diminutive affix as a marker of speaker expressivity have propositional scope: their contribution affects the whole proposition and lies outside of truth-conditional calculations. The remaining morphemes all have scopes that are narrower than the proposition and all make truth-conditional contributions. Importantly, we believe that this difference has an impact in the differential frequencies with which these markers appear in Spanish.
For example, Table 3 counts occurrences of these morphemes in the first 25 pages of the Ramona Quebranto corpus (about 40–45% of the whole novel), containing 272 sentences/522 clauses. Items with propositional scope account for 92% of all occurrences of bound morphemes in Spanish clauses.

Moreover, in the Spanish sentences in Siete Cajas, there are 47 examples of interrogative =pa/=piko, 29 examples of imperative -ke/-na, 27 examples of emphatic =ko/=niko, and only one example of the other cases listed in (1). In Table 2, we cross-classify our morphemes according to morphosyntactic position and semantico-pragmatic function, and we boldface the most frequently switched subgroups. As we can see, whether an item is a word-level affix or a sentential or phrasal clitic does not determine how frequently it will be used in Spanish in our corpus. What does predict high frequency of use is propositional scope, that is, expression of a speech act or of speaker propositional attitude.

Why would this be the case? In the case of the emphatics, we can surmise that speakers may “feel the need” for expressing a distinction that doesn’t exist in the language of the sentence (Agnes Bolonyai, personal communication, 2011). Hickey (2010: 155), for example, explains the emergence of a second person plural pronoun in Irish English and the co-opting of West African uma ‘you.pl’ in Caribbean English, from the existence of such a difference in Irish and West African languages. This explanation is applicable to some extent to the Guarani intricate grammaticalized system of imperatives and to the Guarani interrogative particles, neither of which exists as a grammatical category in Spanish. The domain of markers with propositional scope is highly grammaticalized in Guarani. In Spanish, however, it is not. Spanish has no dedicated category of grammaticalized evidentials or emphatics (although of course other forms can be co-opted for this use). It has no morphological marking of interrogatives, relying mostly on intonation (and word order for wh-questions). It does not have a grammaticalized intensifier,
relying instead on lexical means. Even though the imperative is morphological in Spanish, the extended, graded paradigm of imperatives that we see in Guarani is absent. On the other hand, the other markers we studied do have grammaticalized counterparts, and therefore pressure to use the Guarani morphemes in Spanish may be lower.

9. Conclusion

We have provided a new perspective on the use of Guarani morphology in Paraguayan Spanish. The extant literature has treated them simply as "borrowings", without any clear argumentation in favor of this view and without linking it to a theory of language contact and its outcomes. We have argued here that these phenomena can be integrated within an established typology of code-mixing. In particular, we propose that most of these uses stem from a backflagging strategy that bilingual speakers use to mark their discourse in a community L2 with elements from the original community L1. In so doing, we have highlighted ways in which this perspective can both illuminate our understanding of Guarani-Spanish Jopara, and help us refine existing typologies of code-mixing.

ACKNOWLEDGMENTS

The author would like to acknowledge Agnes Bolonyai, Rakesh Bhatt, Ad Backus, Pieter Muysken, and Mily Crevels for their helpful discussions.

NOTES

1 Glosses follow Leipzig glossing conventions (https://www.eva.mpg.de/lingua/resources/glossing-rules.php). The first line presents the example as found in the corpora or literature. The second line, if needed, attaches all bound morphemes and follows the official orthography, following Estigarribia (forthcoming, a). Guarani particles are boldfaced when mixed in Spanish sentences.

2 Krivoshein de Canese and Corvalán (1987) give cases of resemantization involving lexical items. They explicitly exclude meaning changes for grammatical items.

3 Future research should assess whether these markers are used by monolingual speakers of Paraguayan Spanish, which would favor a borrowing analysis. But even then, Winford (as well as Estigarribia, forthcoming, b, and Muysken 2000), does not see a clear-cut difference between the processes of borrowing and the outcomes of classic code-switching.

4 The 2012 census counted 60,930 Paraguayans in six Guarani ethnicities (Guarani Occidental, Aché, Ava Guarani, Mbya, Pái Tavyterá, Guarani Ñandeva) (STP/DGEEC 2012). The total population of Paraguay was 6,600,284 (DGEEC 2012). See also Palacios Alcaine (2008).

5 See Estigarribia (forthcoming, b) for the implications our analysis has on constraints on mixed words in code-switching theory.

6 That is, "omissible in Spanish without loss of grammaticality." Inserted items tend to be necessary for the grammaticality of a sentence because they are selected by a predicate or because they participate in core syntactic relations.

7 Strictly speaking, these imperative switches and the interrogatives in yes/no questions, are not cases of doubling, since these suffixes do not have morphological counterparts in Spanish. In wh-questions, the Guarani wh-words co-occur with the interrogative markers, hence the occurrence of the interrogative markers with Spanish wh-words does not constitute doubling either. It is likely that examination of a wider array of doubling or quasi-doubling cases would yield important insights, but such an undertaking falls outside the scope of this paper.

WORKS CITED

Maneglia, Juan Carlos, and Tana Schembori. (2014). Siete Cajas/7 Boxes. Film.