Guaraní-Spanish Jopara Mixing in a Paraguayan Novel

Does it Reflect a Third Language, a Language Variety, or True Codeswitching?

Bruno Estigarribia
University of North Carolina at Chapel Hill
estigarr@email.unc.edu

Abstract

We study the highly idiosyncratic case of Paraguay, the only American nation where an indigenous language has survived as a majority language spoken by the non-indigenous population. Jopara is the name of the commonly used code that mixes Guaraní and Spanish. Characterizations of Jopara in the literature are inconclusive. Some authors call it a variety of Spanish, some a variety of Guaraní, others a new mixed language. The choice of one characterization over the other has important implications for the status of Guaraní vis-à-vis Spanish, especially for Paraguay’s educational and language planning. Here we analyze Jopara as it is represented in the novel Ramona Quebranto (RQ-Jopara). We show that this written code is not a variety of either Spanish or Guaraní, nor a mixed language. Rather, it reflects properties of true code-switching. It displays both insertional and alternational characteristics (Muysken, 2000), as well as evidence of a composite matrix language (Myers-Scotton, 2002). We conclude by suggesting RQ-Jopara fits best a “mixed lect” scenario (Backus, 2003) and discussing generalization to spoken Jopara.

Keywords

codeswitching – language contact – mixed lects – indigenous languages – Guaraní – Jopara
1 Introduction: Paraguay, Guaraní, and Jopara

Spanish coexists with an indigenous language in Paraguay: Guaraní.1 Paraguay is special because it is the only country in the Americas where an indigenous language is spoken by the largely non-indigenous population (under 1% of Paraguayans are ethnically Guaraní) (Gynan, 2001; Lustig, 1996). According to the 2012 census, 77% of Paraguayans speak Guaraní (of which 8% do not speak Spanish) and 73% Spanish (of which 5% do not speak Guaraní). About 69% of Paraguayans are bilingual. Speakers of either Guaraní or Spanish total 82%. Paraguay is clearly an outlier in the Americas vis-à-vis the status of indigenous languages.

Paraguay’s linguistic history is reviewed in almost every publication on Guaraní (Zajícová, 2009; Lustig, 1996; Choi, 2005; K. S. Mortimer, 2012; Gómez Rendón, 2008; Gynan, 2001). In general, it is assumed Guaraní survived in the non-indigenous population because Spain did not have in place a big administration with its concomitant presence of Spanish speakers. The relatively few Spanish males married Guaraní women whose children spoke Guaraní (Morínigo, 1959). Additionally, Guaraní was used repeatedly in Paraguayan history as a rallying symbol in wars against neighbors.2

Of course, the two languages do not coexist as completely separate, “pure” linguistic entities. There is extensive mixing not just as an individual’s choice but also at the societal level (i.e., it is a practice that is sanctioned and entered into by the community). Paraguayans call this practice JOPARA:3

1. a. CHE ND-A-REKÓ-I tanto conocimiento.
   1SG NEG-1SG-have-NEG so.much knowledge
   ‘I don’t have that much knowledge.’

---

1 Henceforth, Guaraní refers to Paraguayan Guaraní [GUG] throughout, specifically to the ‘high’ variety of the language, with few Spanish loans and syntactic calques, known as Guaraniete or ‘true’ Guaraní (e.g. Aiikhenvald, 2012: 95–97).
2 The Jesuits’ conservation of the Guaraní language in the missions can also have played a role, although not all Guaraní speakers joined Spanish-dominated towns in Paraguay after the Jesuits were expelled.
3 Examples are italicized and follow Leipzig Conventions for interlinear morpheme-by-morpheme glosses (http://www.eva.mpg.de/lingua/resources/glossing-rules.php). Guaraní elements are in capitals in mixed language examples; when needed, underlining signals morphemes under discussion. Parentheses mark omission of Spanish phonological segments for clarity (obviously, no prescriptive judgment is intended). Omission of whole Spanish words is generally not indicated. Glosses simplify over details that are not germane.
b. **ND-A-JÚ-I** porque **CHE-RASY**
   neg-1sg-come-NEG because 1sg-sickness
   ‘I didn’t come because I was sick.’

c. **¿Por qué PIKO me decís eso?**
   Why Q me you.say that
   ‘Why are you telling me that?’
   (Zajícová, 2009)

(2) a. **Para peore(s), una noche O-U JULIO PRIMO,**
   For worse, one night 3-come Julio cousin, escaped of police
   ‘What’s worse, one night came Julio’s cousin, running from the police’

b. **MENDA NIKO e(s) jodido porque cuando salí(s) la Iglesia-GUI,**
   Marriage emph is screwed because when you.go.out the church-from
   ‘To be married, I say, is difficult because when you walk out the church’

   **O-JE-creé tu dueño, y te trata como cualquié(r) cosa,**
   3-refl-believe your owner and you treat like any thing
   ‘he believes he is your owner, and he treats you like a thing’

   **HA O-KA’Ú-RÔ, KATU-ETE NE-NUPĀ**
   and 3-drunk-if just-very 2sg-hit
   ‘and when he is drunk, he always beats you up.’
   (Ayala de Michelagnoli, 1989)

To date, structural constraints on Jopara have been relatively understudied. Many different characterizations of Jopara, mainly as a variety of Guaraní, but also as a variety of Spanish, as a whole new third language, or as a non-normal-ized code, are found in the literature. Morínigo (1959) considered all of modern Guaraní to be essentially Spanish in its syntactic and semantic structure, but spoken with Guaraní words. On the other hand, Fernández Guizzetti (1966) defined Jopara as Guaraní with hispanisms, Palacios Alcaine (2008) as Guaraní with extensive influence from Spanish, and Dietrich (2010) as a “popular

---

4 **Primo** is considered a Guaraní word based on Symeonidis (2013) who showed that (Paraguayan) Guaraní has incorporated this and many other family terms from Spanish, replacing the terms from classical Guaraní. Moreover, the order of the possessive phrase **Julio primo** ‘Julio’s cousin’ is possessor-possessum as in Guaraní, not possessum+de+possessor as in Spanish.
spoken Guaraní” with a variable amount of Spanish borrowings.\(^5\) As for the degree of conventionalization of Jopara, Lustig (1996) and Pottier (1970) call it an unstable language mixing and Palacios Alcaine (2005) a non-normalized linguistic complex.\(^6\) More recent research has revived the notion of Jopara as a third language (Bakker, Gómez Rendón, and Hekking, 2008; see also, Boyer, 2010; Rubí, 2011). These authors assume at the outset Jopara refers simply to Paraguayan Guaraní, while concluding that “it is not difficult to side with those who see in today’s Paraguayan Guaraní a ‘third language’ that is neither Spanish nor Guaraní” (p. 197). Such a view was indeed advanced by Melià (1974).

The same authors sometimes offer incompatible characterizations: Boidin (2006a) calls it a transitional pidgin, but elsewhere she claims it is a creole (Boidin, 2006b; see also Pic-Gillard, 2003). Lustig (1996: 22) calls it “a diastratic and situational variety of Guaraní”, and later on (p. 27) a code “with Spanish lexical material adapted to Guaraní morphosyntax.” Certainly, the lack of systematic linguistic descriptions of this mixed code has only compounded the definitional problem (Bakker, Gómez Rendón, and Hekking, 2008; Lustig, 1996).

More recently, Dietrich (2010: 40) considers Jopara simply “spoken Guaraní”. He concludes (as we will here) against a third language or a creole explanation.\(^7\) Dietrich calls Jopara any utterances of Modern Paraguayan Guaraní influenced by Spanish syntax. Hence, in example (3), (a) is Guaraní because it shows a classic Guaraní possessive construction with juxtaposition without a possessive verb, and (b) is, supposedly, Jopara, because areko (equivalent to Spanish tengo) is used in an utterance that is a syntactic calque from Spanish.

\[\begin{align*}
\text{(3) a. } & \text{Che } \text{che-ra’y heta.} \\
& 1\text{SG } 1\text{SG-son much} \\
& ‘I have many sons.’ \\
\text{b. } & \text{Che a-reko heta che-ra’y.} \\
& 1\text{SG } 1\text{SG-have much } 1\text{SG-son} \\
& ‘I have many sons.’ \\
& \text{(Dietrich, 2010)}
\end{align*}\]

\(^5\) Aikhenvald (2012) seems to implicitly agree with this characterization.

\(^6\) “Es evidente, pues, que no puede hablarse del guaraní criollo como una unidad lingüística, sino como un complejo lingüístico sin normalizar.” (It is therefore evident that one cannot talk about creole Guaraní [the author considers this equivalent to Jopara] as being a linguistic unit, but rather a non-normalized linguistic complex. My translation.)

\(^7\) The exact distinction between creole and mixed language is a matter of some debate (see Matras and Bakker, 2003a) but generally creoles are the product of imperfect learning and show some structures that are untraceable to any parent language. This is not the case of Jopara.
These utterances, however, would be recognized by most speakers as Guaraní (90% according to Zajícová 2009: 78), not as Jopara. In this paper we focus on overt instances of switching between the languages and not on the contact phenomena Dietrich analyses.8

Defining Jopara is not a purely academic issue. The Paraguayan Constitution of 1992 states that “Teaching in the early school career will be in the official language of which the student is a native speaker. Students will be taught to understand and to use both official languages of the Republic likewise.”9

Education in Guaraní could arguably help reduce poverty and bridge class gaps in Paraguay (Nickson, 2009). But what Guaraní? A main cause for the mixed success of teaching in Guaraní is that students’ families and teachers complain that the academic Guaraní taught bears little relation to the Guaraní actually spoken and understood by both children and educators (see, e.g., Choi, 2004: 253). Seventy-one percent of teachers in the Guaraní modality prefer everyday Guaraní to academic Guaraní (Gynan, 2001). To purge Spanish from Guaraní, Academic Guaraní employs many neologisms that that even teachers cannot understand (Caballero, 2008). Mortimer (2006) notes that there is a camp of proponents of Jopara as the language of instruction. This idea is strongly vilified by defenders of purer forms of Guaraní.10 Perhaps both sides are calling different phenomena “Jopara” and this leads to even less agreement.

Here, we analyze a novel, Margot Ayala de Michelagnoli’s “Ramona Quebranto” (1989) which is considered in Paraguay the first novel ever written entirely in Jopara (see details in section 3.1). Because our corpus is an edited written text, and not a spoken variety, we call the language in the novel RQ-Jopara (as a shorthand for “the Jopara found in the novel Ramona Quebranto”), to highlight that we do not assume it is automatically equivalent to colloquial spoken Jopara. We ask ourselves what the linguistic status of this mixing is and how to characterize it from a structural viewpoint. Does this code represent a true third language (e.g. Palenquero, or a mixed language like

---

8 Of course, the question of how to treat syntactic calques from Spanish into Guaraní is a very important one. We are simply highlighting that this is not the focus of our present work.
Media Lengua or Michif), or a variety of Spanish or Guaraní (like Rioplatense Spanish)? The hypothesis we defend is that RQ-Jopara is neither of these, but bonafide codeswitching (cs) showing both bidirectional insertion and alternation patterns (Muysken, 2000; see section 5.6). It also exhibits Spanish Matrix Language (ML, Myers-Scotton, 2002; see section 6.1), Guaraní ML, and composite ML clauses. We conclude in favor of a mixed lect analysis of RQ-Jopara (Backus, 2003; see section 6.2) and discuss possible extensions to an analysis of spoken Jopara.

2 A Brief Sketch of Guaraní

This section is based on Bakker, Gómez Rendón, and Hekking (2008), Gregores and Suárez (1967), Krivoshein de Canese (1983), Palacios Alcaine (2000), Tonhauser (2006), and Velázquez-Castillo (2004). Its purpose is not to give a comprehensive overview of Guaraní (many details are left out; for example, some authors may have different analyses for the morphemes exemplified), but to introduce basic features of Guaraní (in fact, the high variety of Guaraní, known as Guaraníte11) to help the reader evaluate the cs examples.

Guaraní is a member of the Tupí-Guaraní branch of the Tupí linguistic family, genetically unrelated and typologically rather different from Spanish. While Spanish lexical stress is variable, Guaraní monomorphemic roots are oxytone.12 Stress in polymorphemic words falls on the last syllable of the root unless there is non-final a suffix that is lexically specified as stressed. When words are not oxytone, stress is marked orthographically. The language has extensive nasal harmony, which accounts for most allomorphies in the form of affixes. Whereas Spanish has masculine and feminine gender, Guaraní lacks any gender or noun class marking. Classical Guaraní had no definite articles, but modern Guaraní has borrowed Spanish la for singular NPs and lo for plural NPs.13 Plural marking by -kuéra/-nguéra on nouns is optional and used mostly with animates.

As for morphological typology, Guaraní is an agglutinative (4a) and mildly polysynthetic (4b) language. Polysynthesis, however, has been greatly diminished by the contact with Spanish.

11 This high variety tends to eschew most of the Spanish borrowings featured in Gómez Rendón (2007).
12 With a few exceptions (among others túva, ta’yra, and ára).
13 A reviewer notes that this borrowing is not a direct copy of the Spanish, in which plural is las/los and singular is el/la.
Orthographic rules of Guaraní mandate that postpositions are attached if monosyllabic, otherwise written separately. In Ramona Quebranto, application of this rule is erratic, which has implications for what is considered word-internal switching. Unfortunately, there isn’t a consensus in the literature on the affixal, clitic, or free morpheme status of postpositions and sentence markers. Here, we have attached all postpositions to their host following Velázquez-Castillo (2004) and Tonhauser (2011), except those that have sentence scope (interrogative and evidentiality markers) and (h)ina, which Velázquez-Castillo (2004: 1424) considers non-affixal.

(4) a. *nomboguatasevéitapa*
   n-o-mbo-guata-se-vé-i-ta-pa
   NEG-3-CAUS-walk-DES(IDERATIVE)-more-NEG-FUT-Q
   ‘Will he not want to make him walk anymore?’

   b. *ajepojohei*
   a-je-po-johéi
   1SG-REFL-hand-wash
   ‘I wash my hands.’

Many Spanish prepositions have postpositional Guaraní counterparts:

(5) *Ahákuri cherógaguive nderógape kavajuári*
   A-há-kuri che-róga-guive nde-róga-pe kavaju-ári
   1SG-go-PST 1SG-house-ABL you-house-LOC horse-on
   ‘I went from my house to your house on horse.’

Furthermore, Guaraní has two classes of words that serve as head of clausal predicates. One series, called a(i)real, takes the subject prefixes in Table 1.

<table>
<thead>
<tr>
<th>Person/Number</th>
<th>Non-nasal bases</th>
<th>Nasal bases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>a(i)</td>
<td>a(i)</td>
</tr>
<tr>
<td>2SG</td>
<td>re(i)</td>
<td>re(i)</td>
</tr>
<tr>
<td>3</td>
<td>o(i)</td>
<td>o(i)</td>
</tr>
<tr>
<td>1PL inclusive</td>
<td>jà(i)</td>
<td>ña(i)</td>
</tr>
<tr>
<td>1PL exclusive</td>
<td>ro(i)</td>
<td>ro(i)</td>
</tr>
<tr>
<td>2PL</td>
<td>pe(i)</td>
<td>pe(i)</td>
</tr>
</tbody>
</table>
The words in the a(i)real series can only be heads of the clausal predicate. The other class, called “chendal”, can serve as heads of predicates (6a), but also as heads of referential phrases (6b) and as modifier of referential or predicate phrases (6c).

(6)  

a. a-karu  
   1sg-food  
   ‘I eat/ate’

b. karu   guasu  
   food   big  
   ‘banquet’

c. Mitã   karu  
   child   food  
   ‘gluttonous child’

The prefixes this class takes are shown in Table 2.

<table>
<thead>
<tr>
<th>Person/Number</th>
<th>Non-nasal bases</th>
<th>Nasal bases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>che</td>
<td>che</td>
</tr>
<tr>
<td>2sg</td>
<td>nde</td>
<td>ne</td>
</tr>
<tr>
<td>3</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>1pl inclusive</td>
<td>ñande</td>
<td>ñane</td>
</tr>
<tr>
<td>1pl exclusive</td>
<td>ore</td>
<td>ore</td>
</tr>
<tr>
<td>2pl</td>
<td>pende</td>
<td>pene</td>
</tr>
</tbody>
</table>

Subject pronouns also serve as possessive markers (8a). We follow Velázquez-Castillo (1996) and Dietrich (2010) in representing them as prefixes attached to the predicate or possessum. Bare predicates can have present or past interpretations but future must be obligatorily marked (usually with -ta, but also with the probable but uncertain future -ne). The past morphemes -ramo, -kuri, -vaékue, -ra’e, when used, convey different degrees of certainty or proximity.

---

15 See Tonhauser (2011). This author glosses -ta as a prospective aspect/modal and -ne as a modal marker.
Negation is circumfixal $n(d) + ... + (r)i$ (7a,b), or ani ...+ti (7c) for the negative imperative.

(7) a. \textit{Nd-a-japó-i}.
\textit{NEG-1SG-make-NEG}
‘I don’t/didn’t make (it).’

b. \textit{Na-nde-resarái-ri}.
\textit{NEG-2SG-forget-NEG}
‘You don’t/didn’t forget.’

c. \textit{Ani} \textit{e-ju-tí}.
\textit{NEG.IMP} \textit{IMP-come-NEG.IMP}
‘Don’t come.’

Subjects can be dropped or expressed by the verbal inflections in tables 1 and 2. The free-standing personal pronouns are similar to the chendal series (Table 3).

Whether the first column pronouns express the subject or object depends on the person hierarchy indicating whether A or O are referenced (Jensen, 1997). Word order is mostly pragmatically determined as in Spanish, with possible object/subject drop, even though a basic SVO or OV order is usually identified in the literature (Gregores and Suárez, 1967; but see Tonhauser and Colijn, 2010). The object marker -pe (-me with bases that have a nasal segment) is the same as the locative (with or without movement) marker -pe and it differentially marks animate objects (Shain and Tonhauser, 2010).

\begin{table}[h]
\centering
\begin{tabular}{lcc}
\hline
\textbf{Person/Number} & \textbf{Subject or object} & \textbf{Explicit Object} \\
\hline
1SG & che & cheve \\
2SG & nde & ndeve \\
3 & ha’e & ichupe \\
1PL inclusive & ñande & ñandeve \\
1PL exclusive & ore & orev \\
2PL & peē & peēme \\
\hline
\end{tabular}
\caption{Guaraní free standing personal pronouns}
\end{table}
Possession is marked by \textit{possessor-possessum} juxtaposition, the reversed order from Spanish when the possessor is not a pronoun (8b).

(8)  
\begin{itemize}
\item \textbf{a.} \textit{che-ru}  
\hspace{0.5cm} 1\text{sg}\text{-father}  
\hspace{0.5cm} \text{‘my father’}
\item \textbf{b.} \textit{María memby}  
\hspace{0.5cm} \text{María} \hspace{0.5cm} \text{offspring}  
\hspace{0.5cm} \text{“Maria’s children”}
\end{itemize}

Finally, the Guaraní interrogative markers \textit{pa, pi(k)o} are obligatory and the only means of marking questions. Evidentiality and emphatic markers can be absent. The most widely used emphatic in our corpus is \textit{(ni)ko}, but there are also examples of emphatic \textit{voi}.

To summarize, Guaraní is typologically different from Spanish in important respects: agglutination, postpositions, evidentiality and interrogative marking. Of all these, agglutination is crucial in a contact scenario because it increases the probability of word-internal switching (see e.g., Zajícová, 2010). Such a typological asymmetry commonly entails asymmetries in violations of Sankoff and Poplack’s (1981) Free Morpheme constraint, as Muysken (2000: 76) notes. We will see that mixed words with Spanish bases and Guaraní agglutinative morphology are common in RQ-Jopara. This is most apparent in the verbal domain. Spanish verbs inflect for the categories of thematic vowel, TAM, and syncretic number/person. Guaraní predicates, on the other hand, can be inflected for at least 22 different categories (Canese, 1983: 101–103).

3 \hspace{0.5cm} \textbf{Methodology}

3.1 \hspace{0.5cm} \textbf{The corpus}

Our corpus consists of an excerpt of the novel \textit{Ramona Quebranto} by Margot Ayala de Michelagnoli, considered the first novel ever written entirely in Jopara. We refer to the language in this corpus as “RQ-Jopara” to explicitly acknowledge we are not claiming the novel reflects exactly spoken Jopara, but that it is comparable enough to be relevant for analysis.

To justify our use of this text, we note first that native speakers consider \textit{Ramona Quebranto} to accurately reflect the way people speak (see e.g. Lustig, 1996). Peiró and Rodríguez Alcalá (1999: 153) review it and comment: “Her novel \textit{Ramona Quebranto} (1989) deserves special attention because it is
the first narration to reproduce the real Paraguayan language... the Jopara, a mix of Spanish and Guaraní." This is not trivial, since previous authors (including the great Augusto Roa Bastos, arguably the best Paraguayan writer to have ever lived) attempted the use of Jopara and were heavily criticized (Rodríguez-Alcalá, 1955: 33).

In addition, fictional texts are routinely employed in linguistic analysis of both monolingual and bilingual grammars. Relevantly, in her analysis of direct object drop in Paraguayan Spanish, Choi (2000) illustrates with data from literary works, assuming those data are not at odds with spoken Paraguayan Spanish. Likewise, Zajícová (2010) studies Spanish borrowings in Guaraní and Guaraní borrowings in Spanish on the basis of newspaper texts that, while not being fictional, are an “effort to reproduce the popular language.” (p. 185). This author even cites Ramona Quebranto when discussing nasalization of the postposition -pe (p. 191). An example of use of a bilingual written corpus is provided by Thomas and Sayahi (2012: 265) who use medieval texts to characterize Al-Andalus bilingual speech. They state that “the language use in the kharjas [bilingual closures of Hispano-Arabic poems] most probably approximates that of the lower class of Al-Andalus.” They furthermore justify their study noting that their literary corpus shows evidence the authors (of the poems) attempted to imitate natural speech. | This latter characteristic is especially true of Ramona Quebranto. Specifically, the novel chronicles the miseries of the “Chacariteños”, the inhabitants of a poor neighborhood in Asunción, la Chacarita. Although situated right at the bottom of the Presidential Palace, this neighborhood is constantly flooded by the Paraguay River surges. In this context the use of Jopara acquires maximum authenticity, as the code

---

16 “Su novela “ramona quebranto” (1989) merece una atención particular por ser la primera narración que reproduce el lenguaje real paraguayo... el jopará, mezcla de español y guaraní.” My translation. Additionally, see Francisco Pérez-Maricevich’s prologue to Ramona Quebranto.

17 “Pero cuando el escritor abandona la ciudad y va al campo, y cuando sus personajes deben expresarse en el idioma guaraní desvirtuado que hablan [...], Roa pone en los labios de aquéllos una jerigonza que no es [...] ni el guaraní de los guaraníes ni el guaraní de los paraguayos, ni el español que se habla en el Paraguay. El resultado es un feo pastiche que acaso pueda pasar por algo auténtico sólo para quien desconozca en absoluto la expresión oral del campesino paraguayo.” (“But when the writer leaves the city for the countryside, when his characters must express themselves in the adulterated Guarani they speak [...] Roa puts in their mouths a jargon that is not the Guarani Indians’ Guarani, nor the Paraguayans’ Guarani, nor the Spanish spoken in Paraguay. The result is an ugly hodgepodge that might perhaps sound authentic only to those who are completely ignorant of the speech of Paraguayan peasant.” My translation.)
employed by these Chacariteños both in their dealings with the Spanish-speaking upper classes (“la patrona” or boss), as well as among themselves.

Finally, when we analyze a Spanish novel, we consider it to be in a different register from spoken Spanish, but it is still Spanish. Likewise, a written text in Jopara is also a reflection of the author’s competence in Jopara. As Callahan (2004: 100) notes, the charge of inauthenticity leveled at written CS is “a by-product of describing written codeswitching solely from the perspective of its oral counterpart... Although writing may not exactly mimic oral conversation, it is not fundamentally different from speech”. The issue for us is in the end not the authenticity of the language in Ramona Quebranto, but perhaps its representativity, namely whether the use of only one text from one author obscures the range of variation found in this code. We acknowledge these limitations for our analysis. In this paper’s conclusion, we show briefly that the main phenomena analyzed here are also found in online blogs. We recognize that further work with spoken sources is needed to confirm or refute the views we espouse.18

A last important claim this paper makes is that we can apply tools from theories of spoken CS to this particular kind of textual CS. While recognizing that texts need different tools, Sebba (2012) suggests that a theory of textual CS should subsume theories of spoken CS, not that the latter simply do not apply. Callahan (2004), for example, successfully applied the Matrix Language Frame model (MLF, Myers-Scotton, 2002) to written data.

3.2 Coding methodology

We converted to Unicode a freely available online pdf version of the novel, prepared by the Centro Cultural de la República El Cabildo (www.cabildoccr.gov.py). A native Paraguayan speaker, with formal instruction in both Guaraní and Spanish (but no Linguistics training, and blind to the purpose of the study), checked it for typos, then coded each word in the novel as (1) Guaraní, (2) Spanish, (3) either (mainly proper names, about two dozen), (4) mixed (with bound morphology from both languages), and (5) undecidable (where typos and/or context made it impossible to assign the word to Guaraní or Spanish; less than a dozen cases).19 We then prepared the first 15 pages of the corpus in CHAT format (http://childes.psy.cmu.edu/manuals/CHAT.pdf) for

---

18 The author has recently (fall of 2013) completed 20 sociolinguistic interviews in Jopara. Such data will provide corroboration or rebuttal of the proposals in this paper.

19 Even though the chendal prefixes are glossed here as attached to their bases, some authors (Guarani, 2008; Lustig, 2005) consider them free standing words. As a conservative measure, we have not included them in the counts or examples of mixed words.
This subcorpus excludes sentences attributed to one character (la “patrona”) who always speaks Spanish, without cs. The clan command used to identify switches is combo +s[@1^@2^@4^@2^@4^@2^@4^@4^@2]+f.

analysis with the CLAN suite (http://childes.psy.cmu.edu/manuals/CHAT.pdf, MacWhinney 2000), following the guidelines set in Barnett et al. (2000). The excerpt contains 272 sentences (522 clauses). Mixed sentences (inter-and intra-clausally) number 186 (68%), containing 504 switch points. Table 4 gives the distribution of word and morpheme tokens in this corpus. Morpheme counts are crucial given the agglutinative nature of Guaraní, but the general distribution of the languages is similar whether one counts words or morphemes.

### Table 4

<table>
<thead>
<tr>
<th>Word tokens</th>
<th>Spanish</th>
<th>Guaraní</th>
<th>Mixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1146</td>
<td>798</td>
<td>261</td>
<td>87</td>
</tr>
<tr>
<td>70%</td>
<td>23%</td>
<td>8%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Morpheme tokens</th>
<th>Spanish</th>
<th>Guaraní</th>
</tr>
</thead>
<tbody>
<tr>
<td>1680</td>
<td>1273</td>
<td>587</td>
</tr>
<tr>
<td>68%</td>
<td>32%</td>
<td></td>
</tr>
</tbody>
</table>

4 RQ-Jopara does not represent a language variety

4.1 Definition of “language variety”

Dialectologists define LANGUAGE VARIETIES or DIALECTS mainly by mutual intelligibility. For example, Chambers and Trudgill (1998) and Trudgill (1998) consider some degree of mutual intelligibility an important factor in determining whether two observed codes are in fact varieties of the same language. Additionally, if speakers consider the code a variety of a particular standard, and any standardizing changes tend to that standard (heteronomy), this also indicates the code is a variety of that standard (Chambers and Trudgill, 1998: 9).

Even though speakers generally use the appellation “Guaraní Jopara”—suggesting Jopara is a variety of Guaraní—some Paraguayans define Jopara as primarily Spanish, as Kallfell (2011) reports. Moreover, compliance with the intelligibility criterion fails, as we demonstrate next.

---

20 This subcorpus excludes sentences attributed to one character (la "patrona") who always speaks Spanish, without cs. The CLAN command used to identify switches is combo +s[@1^@2^@4^@2^@4^@2^@4^@4^@2]+f.
4.2 Lack of mutual intelligibility

RQ-Jopara is clearly not mutually intelligible with monolingual Spanish. Consider the following passage and its translation into Spanish (Lustig, 1996):

\[\begin{align*}
\text{—¡TE-RE-HO \ PARAGUAY-PE E-MBA'APO, CHE-MEMBY!} \\
\text{HORTATIVE-2SG-go Asunción-LOC 1MP-work 1SG-child} \\
\text{‘Go work in Asuncion, my daughter!}
\end{align*}\]

\[\begin{align*}
\text{CHE A-PYTÁ-TA KOKUÉ-PE.} \\
\text{1SG 1SG-stay-FUT farm-LOC} \\
\text{I will stay in the farm.}
\end{align*}\]

\[\begin{align*}
\text{Mientra(s) tengo salú(d), via^{21} atendé(r) por tu criatura,} \\
\text{while I have health I will take care for your creature} \\
\text{As long as I am healthy, I will take care of your child,}
\end{align*}\]

\[\begin{align*}
\text{Dame nomá(s) SAPYAPYA un poco de PIRAPIRE} \\
\text{give me just from time to time a little of money} \\
\text{just give me from time to time a little money.}
\end{align*}\]

\[\begin{align*}
\text{MANDYJU cosecha HA soja OPA... no hay-ETE porvení(r).} \\
\text{cotton harvest and soy finished no there is true future} \\
\text{The cotton and soy harvest is over... there is really no future.}
\end{align*}\]

\[\begin{align*}
\text{¡Si te quedá(s) aquí NE-MO-MEMBY-JEVÝ-TA PE bandido HÍNA!} \\
\text{if you stay here 2SG-CAUS-child-again-FUT that rascal PROG} \\
\text{If you stay here, that rascal will get you pregnant again!}
\end{align*}\]

\[\begin{align*}
\text{—ND-A-HA-SÉ-I NIKO, mamita, CHE KO A-VY'A KO'ÁPE.} \\
\text{NEG-1SG-go-want-NEG EMPH mom-DIM 1SG EMPH 1SG-happy here} \\
\text{‘I don’t want to go, mom, I am happy here.}
\end{align*}\]

\[\begin{align*}
\text{A-JAHE'Ó-TA MANTE A-HEJÁ-VO CHE-valle-PE.} \\
\text{1SG-lament-FUT only 1SG-leave-as^{22} 1SG-valley-LOC} \\
\text{I will just cry when I leave my village.}
\end{align*}\]

---

21 Colloquial contraction of voy a (similar to “going to” > "gonna").
22 Tonhauser (2006) glosses this as -AT, and calls this a “cotemporaneity connector.”
“El Jopara sigue anclado más en la estructura del guaraní que en la del español, y ciertas nociones básicas de guaraní son la clave indispensable para entrar en el laberinto del Jopara.” My translation.

As Lustig (1996: 21) states, “Jopara remains more anchored on the Guaraní structure than on the Spanish structure, so that a basic knowledge of Guaraní is an indispensable key to penetrate the labyrinth of Jopara.”

On the other hand, if speakers of Guaraní understand RQ-Jopara it is because most Guaraní speakers also have some Spanish proficiency. But RQ-Jopara per se is not mutually intelligible with Guaraní. First, Spanish words are roughly four times as frequent as Guaraní words, as table 4 above shows. Guaraní speakers have to have a good command of the Spanish lexicon to understand RQ-Jopara. Second, many sentences in RQ-Jopara are almost exclusively Spanish:

---
This extensive use of Spanish is highly inconsistent with the idea that Jopara is plain spoken Guaraní or a Guaraní with hispanisms. Furthermore, Hekking, Bakker, and Gómez Rendón (2010) find only 17.4% Spanish borrowings in their Modern Guaraní corpus, a much lower figure of Spanish tokens than what we find in Ramona Quebranto, another difference between RQ-Jopara and spoken Guaraní.

Finally, for some passages it is not clear that they would be intelligible to speakers of either monolingual Guaraní or monolingual Spanish:

(11) ¿Cómo PQ anda ese tu problema? Quiero-ETE que sea(s) feli(z)...
    ‘How is it going with that problem of yours? I really want you to be happy…’

(12) Lo(s) rico(s) NIKO son iguale(s) nomá(s) que de nosotros.
    ‘Rich people are just like us.’

Mamita dice que alguno no e(s)-ETE bueno con pobre(s)
    ‘Mom says some are not really good to the poor’

This passage shows alternating Spanish and Guaraní sv agreement, possessive nps sometimes fully in Spanish (“la e(s)cuelita de mi hijo”, “su galpón”), sometimes fully in Guaraní (“che membykuéra”), and composite nps (“veterano-KUÉRA Guerra Chaco”) with Spanish lexical items, Guaraní plural, Spanish word order but omission of the Spanish marker de, much like in Guaraní where
possession doesn’t require a marker. Are we in Spanish or in Guaraní here? We conclude that RQ-Jopara is neither a variety of Spanish nor of Guaraní.

5 RQ-Jopara is not a Mixed Language

5.1 Definition of “Mixed Language”

Mixed languages are languages with a split between lexicon and grammar, commonly (but not always) morphosyntax from language 1 (the grammaticizer language) and lexic from language 2 (the lexifier language). Matras and Bakker (2003b) identify the following additional prototypical characteristics of mixed languages: They are native linguistic codes endorsed as unmarked choice in a community; they have emerged from situations of full bilingualism; they show unclear parentage; and they are conventionalized, not on-the-spot mixing. Of all these, conventionalization is the only feature common to all claimed cases of mixed languages. Codes that are not conventionalized and that rely on a speaker’s active knowledge of both languages are not mixed languages. Likewise, Muysken (2008; chap. 16) highlights the stability (even partial) of mixed languages, and explicitly excludes languages which have undergone extensive lexical borrowing, and bilingual code-mixing.

One of the best-known, clearest examples of a mixed language is Media Lengua. Muysken (1997: 365) describes it as essentially Quechua with relexification, that is, replacement of all Quechua stems, including all core vocabulary, by Spanish stems. Furthermore, Media Lengua “is not made up on the spot every time it is spoken” (p. 366). Quechua personal pronouns and interrogative words have been fully replaced by Spanish forms. On the other hand, grammatical elements by and large come from Quechua. This exemplifies the necessary presence of a split in a mixed language, where each contributing code takes over a specific subset of the grammar and lexicon. Michif is another well-known example in the Americas, where the split consists of French taking over the noun domain and Cree the verb domain. We will see that no similar split can be found in RQ-Jopara.

Jopara has emerged from a situation of extended and extensive (if not actually full) bilingualism, and it is the unmarked choice in the community. In addition to this, Jopara can be considered an in-group language, just as Meakins (2012) shows for Gurindji Kriol, if one views bilingual Paraguayans as an in-group relative to monolingual Spanish speakers. If RQ-Jopara reflects spoken Jopara, one could be tempted to conclude it exemplifies the use of a mixed language. But the crucial split (with replacement of whole parts of a language A by a language B) and conventionalization that define mixed languages fail in RQ-Jopara’s case.
5.2 Evidence for Guaraní as Grammaticizer and Spanish as lexifier

For Jopara to be a mixed language we should find that its morphosyntax comes from one language, its vocabulary from another, or that the language shows a similar clear split (Bakker and Muysken, 1995). First, note that there is a strong asymmetry favoring Guaraní morphosyntax in the corpus. Guarani bound morphology often occurs with Spanish lexical bases and phrases (Table 5), but not the other way around.

These data are consistent with Spanish being a lexifier language in a mixed-language scenario. For example, in the corpus we find O-JE-creé ‘he believes himself’, but not the converse *se JEROVIA (jerovia ‘believe’), with Spanish reflexive marking on a Guarani base; O-vendé ‘he sells’, but not *él ÑEMU/-a/-e/-Ø (ñemu ‘sells’), with any of the possible 3SG Spanish endings; A-pensá ‘I think’, but not *yo MO’Ã (mo’a ‘think’); nuestro casita-PE ‘in our house(s)’, but not *en ÑANDE RÓGA(kuéra).24

We mentioned above that Guaraní is agglutinative and therefore provides more opportunities to add its affixal morphology to Spanish bases. In some of the cases cited above, Spanish affixation is not likely because suffixes are associated with particular verb classes (that is, they have a high degree of lexical selection; Muysken, 2000: 76). It is possible that ñemu or mo’a do not take Spanish person marking because the particular Spanish allomorph depends on the verb’s conjugation, something that is not equivalent to the Guaraní system. But the future tense marker of Spanish does not show allomorphy and yet it does not appear affixed to Guaraní words either. Moreover, Spanish has inflectional suffixes with little or no lexical restriction, for example, the plural suffix. Yet plurals never appear affixed to Guaraní bases in RQ-Jopara (consistent with findings by Zajícová, 2010).

5.3 Rethinking the Status of Spanish Insertions.

What complicates a putative ‘relexification’ explanation for RQ-Jopara is that Spanish insertions in the novel have two origins. The first process is the historical process of lexical borrowing. Modern Guaraní has indeed been relexified to some extent by Spanish (see, for example, Gómez-Rendón, 2007). Many of the mixed words found in RQ-Jopara can be argued to be integrated borrowings. For example, in the most common type of mixed single-word construction, with a subject prefix accompanying a Spanish verb (the 13 subject-verb agreement cases identified above), the Spanish base almost invariably carries the native oxytone stress of Guaraní, even when the Spanish verb is paroxytone in the correspondent TAM+person/number combination:

24 Zajícová (2010) also finds the same asymmetry in texts from newspapers.
<table>
<thead>
<tr>
<th>Type</th>
<th>Number of cases</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject cross-reference</td>
<td>13</td>
<td>Ojecréé, oñeinundá, ovendé, oñeconchavá, aeplicá, apensá, adeconfiá (2x), opena, adecansá, oganá, osolucionátá, apregunté</td>
</tr>
<tr>
<td>Locative -pe</td>
<td>12</td>
<td>Insiernope, Sol de Américape, casitape, guerrape, calle Palmape, centrope, peteĩ tiendape, portónpe, Ríope, cárcelepe, teléfonope, cóntrape</td>
</tr>
<tr>
<td>Superlative -ete</td>
<td>12</td>
<td>No quierete, quierete, no sabemoete, no hayete, no e(s)ete (x6), no valeiete, no me enojoete</td>
</tr>
<tr>
<td>Negations</td>
<td>7</td>
<td>Ndanegái, nañanimái, no entendéi, no malratái, ndaigutoi, no valeiete, ani renegáti</td>
</tr>
<tr>
<td>Friendly imperative, polite request -{(mi)+na}</td>
<td>6</td>
<td>Miramíná, ecambiamíná, ecuchamíná, acompañaamíná, usamíná, camínána</td>
</tr>
<tr>
<td>Chendal subject 3 -i</td>
<td>5</td>
<td>Isinverguenzo, idesubicada, itríte, iviva, imáníá</td>
</tr>
<tr>
<td>Reflexive -je/-ñe-</td>
<td>5</td>
<td>Ojecréé, oñeinundá, oñeconchavá, jegustaha</td>
</tr>
<tr>
<td>Plural -kuéra</td>
<td>4</td>
<td>Hijokuéra, veteranokuéra, militarkuéra, sobrainkuéra, membykuéra, poriahukuéra</td>
</tr>
<tr>
<td>Future -ta</td>
<td>2</td>
<td>Osolucionátá, picáta</td>
</tr>
<tr>
<td>Nominal temporal marker -kue</td>
<td>2</td>
<td>Ranchokue, lavanderakue</td>
</tr>
<tr>
<td>Focus operator -nte (“only”)</td>
<td>2</td>
<td>Éllante, propaganda governamentalnte</td>
</tr>
<tr>
<td>Nominalizers/Adjectivizers -va, -ha</td>
<td>2</td>
<td>Ganává, jegustaha</td>
</tr>
<tr>
<td>Imperative e-</td>
<td>2</td>
<td>Ecambiamíná, ecuchamíná</td>
</tr>
<tr>
<td>Locative -re (“for”)</td>
<td>1</td>
<td>Sentimientore</td>
</tr>
<tr>
<td>Animate object marking -pe</td>
<td>1</td>
<td>Jesúpe</td>
</tr>
<tr>
<td>Ablative -gui</td>
<td>1</td>
<td>Iglesiaguí</td>
</tr>
<tr>
<td>Forceful imperative -ke</td>
<td>1</td>
<td>Bañáteke</td>
</tr>
</tbody>
</table>
Se cree > OJEcreé
Se inunda > OÑEinundá
Vende > Ovendé
Conchaba > OÑEconchavá
Explico > Aeplicá
Desconfía > Adeconfiá
Descanso > Adecansá
Ganan > Oganá

In some cases, even the root’s allomorphy is lost:

Pienso > Apensá

This extended borrowing creates a first stratum of mixing that is evident in the high number of Spanish insertions in the corpus. Such borrowings are present even in Paraguayan Guaraní monolingual-intended discourse, where Spanish roots have sometimes replaced Guaraní ones (Gómez-Rendón, 2007). RQ-Jopara has, however, other nonce borrowings, evidence of bonafide cs. One possible case is opena, where the oxytone stress is not marked so we assume it has paroxytone stress. Apregunté is also likely cs and not a borrowing because it bears the Spanish 1 sg preterite marking -é, together with the Guaraní 1sg active person marker a-. Hence, the occurrence of insertions in RQ-Jopara reflects processes both diachronic and synchronic.

5.4 Evidence for Spanish as Grammaticizer

Yet, we cannot conclude that RQ-Jopara is a mixed language with Guaraní morphology and Spanish lexemes because there are exceptions to the Spanish lexeme-Guaraní affix pattern. Spanish feminine gender can occasionally be marked on Guaraní lexemes, as in VÝR-a ‘stupid-fem’ instead of Guaraní výro ‘stupid-masc/fem’, perhaps via reanalysis of the final /o/ as Spanish masculine -o.

Also, Spanish derivational morphology can appear, creatively, on a Guaraní lexeme. Such is the case of OvyreáTA, formed from Guaraní výro via the Spanish verbalizing suffix -ear, with the non-conventionalized meaning ‘do stupid things’.25

Furthermore, Spanish morphemes are not limited to content words. The corpus also includes sentences where most of the items, including lexical ones, are in Guaraní, but some lexical and function words appear in Spanish (14b):

---

25 Thanks to Maura Velázquez-Castillo for identifying this example in the novel.
These Spanish words have equivalents in “pure” Guaraní: “otro” is ambue, “pero” is (agã) katu or jepe, “que” is the suffix -ha. Matras (2003) points out that in mixed languages one component consistently provides the inflectional morphology (in particular on the verb), while the other may provide a variety of other lexical or even morphological resources. While it is true that borrowing of function words like pero ‘but’ is quite common crosslinguistically and does not suffice to rule out an interpretation of RQ-Jopara as mixed language, identical inflectional material should not come from both languages. Yet, inflectional and clause building material does come from both Guaraní and Spanish in RQ-Jopara. For example, we find clauses combining Guaraní (1sg) and Spanish (1sg past) verbal inflection in (15a) and use of the complementizer -ha in (15b):

(15) a. Sí, le Apregunté AKUE.
    Yes to-her 1SG-Lashed pst
    ‘Yes, I asked her.’

    b. A-jurá AKUE compromiso A-cumpli-TA-HA
       1SG-swear pst compromise 1SG-keep-FUT-COMP
       ‘I swore that I would keep that promise’

Moreover, our corpus shows clauses whose grammatical structure is that of the presumed donor language (Spanish), and the inserted items are those of the presumed recipient language (Guaraní). Here we show insertions of Guaraní items in a Spanish-ML clause (16a-c), or Guaraní verbs inserted in a Spanish sentence, giving rise to a composite ML clause (16d):

(16) a. ¿vo(s) PIKO no sabé(s) que KUIMBA'E e(s) como gallo?
    you Q no you.know that man is like rooster
    ‘Don’t you know that man is like a rooster?’
This is the same conclusion Zajícová (2009: 86), based mostly on the relatively low degree of Spanish relexification in Modern Guaraní varieties. 

Finally, in relexified mixed languages, words from the lexifier language replace words from the other language, but this is not the case in RQ-Jopara. RQ-Jopara does not show anything remotely close to the replacement of all core vocabulary we see, for example, in Media Lengua. For instance, the Spanish word hombre appears in (17a) but Guaraní kuimba’e is also used (cf. 15a). In (17b), the Guaraní word memby ‘son’ appears, but in (17c) it is Spanish hijo ‘son’.

(17) a. Hombre KO e(s) recurso.
   Man EMPH is resource
   ‘Men are (certainly) a resource’

b. para nue(s)tro(s) hijo-KUÉRA
   for our son-PL
   ‘for our children’

c. ¡CHE-MEMBY-KUÉRA trabaja(n) má(s) fácil aquí!
   1SG-son-PL work:3PL more easy here
   ‘My children (find) work more easily here!’

To sum, even though what exact kind of lexical or functional material comes from each contributing code in a mixed language is crosslinguistically somewhat variable, what cannot happen is what we see in our corpus, where the same inflections, function words, or lexical items can come from either language. We conclude, then, that RQ-Jopara does not have a clear grammaticizer and a clear lexifier language.

26 This is the same conclusion Zajícová (2009: 86), based mostly on the relatively low degree of Spanish relexification in Modern Guaraní varieties.
5.5 **Evidence against Conventionalization**

It should already be clear from the preceding discussion that RQ-Jopara is not a conventionalized code. Many sentences (like 16 above) have a Spanish morphosyntactic frame, while others (like 14 above) have a Guaraní morphosyntactic frame. The same lexical content can appear alternatively in Spanish or in Guaraní. Additional evidence comes from sentence pairs standing in a paraphrase relation, which shows choice and not conventionalization. In (18), the reportative phrase “as X says” appear in Spanish in (a) and in Guaraní in (b):

(18) a. *E(s)tá fuera de lugar, *como dice mi comadre*.  
‘She is out of line, as my bosom buddy says.’

b. *Tiene que aguantá(r) PETEĨ KUIMBAE “toda la vida”, HEI-HÁ-ICHA PAI*.  
‘She has to stand one man her whole life, like the priest says.’

Example (19) is a question answer sequence where the content “did you ask her” appears in Guaraní in (a) and in Spanish in (b):

(19) a. —*Re-porandu-ma piko ra’e ichupe?*  
2SG-question-already Q PST 3.OBJ  
‘Did you ask her already?’

b. —*Sí, le A-pregunté AKUE*  
Yes to-her 1SG-I.asked PST  
‘Yes, I asked her.’

Example (20) shows the Spanish discourse marker *nomás* and the Guaraní counterpart *katu* both used as embedded language (Myers-Scotton, 2002) items:

(20) a. *¡ÓĨ KO nomá(s) PETEĨ problema!*  
there.is EMPH just one problem  
‘Only there is one problem!’

---

27 *Comadre* is used by a woman to designate her child’s Godmother. It has no equivalent in English. *Comagre* is a dialectal variant here, meaning “bosom buddy”.

28 In (b), additionally, the Guaraní past morpheme *akue* doubles the Spanish preterit inflection -é.

29 Tonhauser (2013) analyzes *katu* as a contrastive topic marker.
Evidence from Alternation

As a last piece of evidence against a mixed language analysis note that, beyond the insertions discussed above, RQ-Jopara displays ALTERNATIONAL CS (Muysken, 2000). Whereas insertions are usually nested single words or constituents, alternations are switches of generally longer stretches of discourse, or a “succession of fragments in language A and B in a sentence, which is overall not identifiable as belonging to either A, or B.” (Muysken, 2013: 5). RQ-Jopara presents alternations involving Guaraní discourse markers (21a), Spanish adverbials and discourse markers (21b,c), and longer fragments where the two participating languages are on a more “even footing” (21d,e).

(21)  a. **MENDA NIKO e(s) jodido**
Marriage EMPH is screwed
‘Marriage is indeed difficult’

b. **sin falta OI-KE-SE NDE-KOTÝ-PE.**
without fail 3-enter-DES 2SG-room-LOC
‘Without fail, he wants to come into your room.’

c. **¡OĨ KO nomá(s) PETEĬ problema!**
there.is EMPH just one problem
‘Only there is one problem!’

d. **¡MA(E)RÃ PIKO CHE-MOMBE’U si no podemo(s) remediá(r)!**
Why Q 1sg-tell if no we.can remedy
‘Why are you telling me if we can’t fix it!’

e. **te trata como cualquié(r) cosa, HA O-KA’Ũ-RÔ, KATU-ETE NENUPÁ.**
you treats like any thing and 3-drunk-when just-very you-hit
‘He treats you like any (old) thing, and when he is drunk, he beats you up.’

Backus (2003) claims that alternational CS cannot give rise to mixed languages (see McConvell and Meakins, 2005, for discussion). The presence of alternation, then, argues against a mixed language classification.

Hence, since RQ-Jopara does not show a clear split, does not show conventionalization, and it does show alternation, we conclude it does not represent...
a mixed language. In the next section we will argue for a CS and mixed lect analysis.

6 RQ-Jopara as Codeswitching Mixed Lect

We propose that RQ-Jopara is best analyzed as CS. The evidence in sections 4 and 5 argues against a variety or a third language analysis. This section shows RQ-Jopara fits a composite CS scenario with a range of structural variation. This, in turn, justifies calling Jopara a “mixed lect” reflecting a continuum of CS varieties.

6.1 RQ-Jopara and a Composite Matrix Language

An important question in determining the structural properties of RQ-Jopara is whether it responds to constraints of classic CS. In classic CS, as defined within the MLF and the 4M models (Myers-Scotton, 2002; Jake, Myers-Scotton, and Gross, 2002), the grammatical structure of each clause comes from one of the participant languages. Which language this is can vary in different corpora of the same varieties, or more rarely within clauses in the same corpus. The crucial principle of these models is the recognition of the asymmetry of participating languages in classic CS. The language that determines word order and supplies outsider system morphemes that build clausal structure (and whose presence is required by an element outside of their maximal projection) is called the matrix language (ML).30 All other morphemes can and often do come from the embedded language (EL). Outsiders, according to Myers-Scotton (2008), are primarily the morphemes that indicate argument structure and agreement relations within a clause (e.g., subject-verb agreement, case markers).

Applying these criteria, we find that RQ-Jopara has Guaraní ML clauses, Spanish ML clauses, and composite ML clauses. In the following examples, subject-verb agreement and case/grammatical function markers are underlined. Clauses with Guaraní ML are, for example:

(22) a. Sin falta, OI-KE-SE NDE-KOTÝ-PE.
   ‘Without fail, he wants to come into your room.’

---

30 Technically, the ML is an abstract construct and not one of the participating languages, but the distinction is of no consequence here.
b. ¡Q-HEKA otro KUÑA AVEI!  
   3-seek other woman too  
   ‘He seeks other women too!’

c. Q-JE-creé tu dueño  
   3-refl-believe your owner  
   ‘He believes he is your owner’

d. HA'EK-sinvergüenzo  
   3 3-shameless  
   ‘He is shameless’

e. ND-Á-negá-I VOI que CHE-gu(s)tá  
   NEG-1SG-deny-NEG EMPH that 1SG-like  
   ‘I do not deny that I like (him)’

f. CHE-vecina Q-HECHA Q-JEROKY Petrona-NDIVE clu(b) Sol de América-PE  
   1SG-neighbor 3-see 3-dance Petrona-com club Sun of America-LOC  
   ‘My neighbor saw him dance with Petrona at the Sol de América club.’

Clauses with Spanish ML, however, are also common:

(23) a. ¡Y qué PIKO le32 vampo(s) (a) hacé(r)!  
   And what Q to.it we.go (to) do  
   ‘And what are we going to do?’

b. Él KO tiene di(s)culpa  
   He EMPH has excuse  
   ‘He has an excuse’

c. PATI dice que e(s) pecado mortal  
   Priest says that is sin mortal  
   ‘The priest says that it is a mortal sin’

31 Interestingly, Spanish sinvergüenza ‘shameless’ is invariable for gender. Here it is used with the masculine ending -o suggesting that there is at stake here a reanalysis process specific to Spanish morphology, and not to Guaraní.

32 This le is not a referential pronoun, but it is part of the lexical entry for the expression, therefore it does not help determine the ML of the clause.
Guaraní-Spanish Jopara mixing in a Paraguayan novel

Journal of Language Contact 8 (2015) 183-222

33 Guaraní shows mixed marking. Arguments of the verb can be referenced on the verb itself (head-marking) but also marked via postpositions (dependent-marking).

33 However, in RQ-Jopara such mixed marking is entirely optional. We have seen many examples where sv agreement is Guaraní. It is comparatively more difficult to find examples with Spanish positional argument marking. There are some Spanish ablative de in the corpus (subcategorized-for by the verb sacar, ‘to take off/out’):

(24) a. Cuando salí(s) la iglesia-GUI
when you.exit the church-from
‘when you walk out of church’

b. Vamo(s) sin falta infierno-PE
we.go without fail hell-LOC
‘We go to hell for sure’

c. Vemo(s) nue(s)tro casita-PE día a día entra agua
we.see our house-in day to day enters water
‘We see that there flows water into our house day after day’

As McConvell (2002: 345) states “dependent-marking noun-coding languages retain nominal grammar from the old language after verbal grammar has turned over to the new language.” However, in RQ-Jopara such mixed marking is entirely optional. We have seen many examples where sv agreement is Guaraní. It is comparatively more difficult to find examples with Spanish positional argument marking. There are some Spanish ablative de in the corpus (subcategorized-for by the verb sacar, ‘to take off/out’):

(25) a. ¿De dónde PIKO ella saca todo eso?
From where Q she takes all that
‘Where does she get all that from?’

b. porque no saca su(s) ojo(s) de mi RETYMA
because no he.takes his eyes of my leg
‘because he does not take his eyes off my legs’

d. Porque e(s) JUKY
because is salt
‘because he is likeable’

Yet other clauses show evidence of a composite ml. These are mostly cases where the subject-verb agreement is Spanish, but argument marking is Guaraní:
Directional marking is usually omitted, or expressed by the Guaraní postposition -pe. There are rare cases of directional Spanish a in the corpus (again, subcategorized for by the predicate):

(26) ha(s)ta que llega a lo(s) techo(s).

‘until it reaches the roofs’

Therefore, argument marking is not conventionalized either. In sum, although most clauses have a clear monolingual ML, there is non-negligible evidence of a composite ML in the corpus, the most frequent case being a split Spanish SV agreement/Guaraní argument marking.

6.2 RQ-Jopara as a mixed lect

Given the patterns reviewed in the previous sections, we consider RQ-Jopara a mixed lect. Backus (2003) defines mixed lects as registers of immigrant language that show a great number of insertional CS and frequent alternational CS. Mixed lects, therefore, admit much more structural variation that mixed languages. Furthermore, mixed lects are any bilingual speech that is unmarked in a community (Backus, 2003: 238). Gómez Rendón (2008; section 7.2) reviews evidence that Jopara is the “default” code in Paraguay. Jopara is “the colloquial variety of Guaraní spoken by most Paraguayans” (p. 208), and “Jopara is the variety in which Paraguayan Guaraní is realized in daily communication.” RQ-Jopara is also the default code used by Ramona Quebranto characters (except for the upper class characters that use Spanish).

Finally, mixed lects do not require a high level of conventionalization. In fact, my reading of Backus (2003) is that they may require no conventionalization at all. Vocabulary replacement does not occur as much as vocabulary addition and alternational codeswitching is allowed. Both characteristics are true of RQ-Jopara.

Mixed lects are a useful conceptualization of the concept of a “continuum of varieties”. From this perspective, RQ-Jopara does not represent a single variety, but as a spectrum of varieties between the Guaraní and Spanish poles.

Gómez-Rendón (2008: 209) seems to advance this view when he describes Guarañol and Castení. Guarañol is a term created by Melià (1986: 244) and

---

As a reviewer notes, Spanish is technically historically the immigrant language in Paraguay, even though it has prestige and is politically dominant. Given Guaraní’s unique sociohistorical situation, it is not surprising that conventional categorizations of contact situations only apply with reservations.
can be interpreted as Jopara with Guaraní ML. Castení (or “Guaraniticized Spanish”), on the other hand, would refer to Jopara with Spanish ML. According to Gómez Rendón, Castení is mostly an urban variety (i.e., Asunción) produced by speakers with low socioeconomic status and low education. Conversely, Guarañol (or “Hispanicized Guaraní”) is the better documented “elsewhere” variety. Zajícová (2010) presents a similar view when she talks about both castellano Jopara and guaraní Jopara.

RQ-Jopara presents data consistent with this proposal. First, Castení helps explain the frequency distribution of words in RQ-Jopara: all the characters (except Ramona’s “patrona”, whose data are not included here since she never codeswitches) are chacariteños from a lower socioeconomic class neighborhood in Asunción. For this reason, they are expected to produce Castení, that is, to produce basically a Spanish discourse “peppered” with Guaraní. Furthermore, the corpus being a novel favors the occurrence of Castení, given that Guaraní is generally excluded from this genre (but not from poetry or theater).

Yet, Guarañol does occur, in particular in segments like the one in (27), involving a conversation between two chacariteños, Ramona and Severino:

(27) —¡ANÍ-KE! Dio te va ca(s)tigá(r) por tu malo pensamiento 
    neg-imp God you goes punish for your bad thought 
    ‘Don’t! God will punish you for your bad thoughts
    y va caé(r) por cabeza de tu hijo. 
    and goes fall for head of your son
    and will fall on your children’s head.

    Nunca má(s) diga eso bla(s)femia, 
    never more say that blasphemy
    Never say those blasphemies again,

    lavá tu boca con JAVÔ, o te echo de mi rancho. 
    wash your mouth with soap or you I.chase from my house
    wash your mouth with soap or I will kick you out of my house.’

    -Eso KATU e(s) lo que yo e(s)tá queriendo, HÍNA. 
    That but is the that I is wanting prog
    ‘But that’s what I want.
Cuando más te renegas, más te quiero, ¿REI-KUAÁ-PA, CHE-paloma?
The more you protest, the more I like you, understand, my dove?

NDE-JUKY HA PE NDE-RESA PÖRÄ CHE-MBO-TAROVA-ITE.
2sg-salt and that 2sg-eye good 1sg-caus-crazy-true
You are cute and those pretty eyes of yours drive me really crazy.

—¡CHÁKE! OÚ-TA LA CHE-compañero HA O-MEʻẼ-TA NDÉVE
careful 3.come-fut the 1sg-mate and 3-give-fut 2sg.obj
‘Careful! My partner will come give you

tu merecido HÍNA porque e(s) celoso por mí.
your due proq because is jealous for me
what you deserve because he is protective of me.

— MÁVA PIKO O-MOMBEʻÚ-TA ICHUPE?
who Q 3-tell-fut 3.obj
‘Who is going to tell him?’

— HA CHE!
and 1sg
‘Well, I!’

— AÑETE? ND-A-GUEROVÍ-A,
Truth NEG-1sg-believe-NEG
‘Really? I don’t believe you,

porque KATU-ETE O-VYR-EÁ-TA ÑANDE-CÓNTRA-PE.
because just-very 3-stupid-vblz-fut 2sg.pl-opposition-loc
because he will do stupid things against us.’

— AÑA-MEMBY-RE! NDAIPÓRI NDE-ganá-VA NDÉVE.
devil-child-loc there.is.not 2sg-win-habitual 2sg.obj
‘Son of the devil! There’s no winning with you.’

— ANI RENEGÁ-TI, Ramonita.
NEG.imp protest-NEG.imp Ramona-dim
‘Don’t protest, Ramonita.’
ND-O-VALÉ-I  RE-luchá  NDE-sentimiento-RE.
NEG-3-be.worth-NEG  2SG-fight  2SG-feeling-LOCATIVE
It’s not worth it to fight against your feelings.’

NE-MANDUÁ-NE  CHE-REHE.
2SG-remember-FUT.PROB  1SG-LOCATIVE
You will no doubt remember me.’

This fragment begins in Castení and then as it becomes more charged with sexual innuendo, immediately switches first to Guarañol, then to Guarani. This reflects the increase in intimacy that is accompanied by a switch to the basilect in this context.

6.3  RQ-Jopara is not a Fused Lect
A last question is whether RQ-Jopara shows characteristics of a fused lect, rather than a mixed lect.35 We don’t believe this characterization would be appropriate. FUSED LECTS are stabilized mixed varieties, with reduced variation and increased structural regularity compared to codeswitching and language mixing (Auer, 1999: 310). We saw that RQ-Jopara is characterized by extensive variation, with little structural regularity except for that that pertains to either the Guaraní or the Spanish grammar.36 Some parts of the grammar of a fused lect are fixed, and speakers have no choice but to use a particular language for certain constituents (p. 321). Again, RQ-Jopara seems to allow more freedom in constituent structure, and even in lexical choice than what a fused lect allows.37

Secondly, fused lects can only be of the insertional, not alternational, type. Section 5.6 above documents the existence of alternation in RQ-Jopara. Moreover, unlike language mixes and CS, fused lects are neither locally nor globally meaningful. But we showed above (section 6.2 especially) that choices in Ramona Quebranto can be interpreted as meaningful. Whether they would be meaningfully locally (as in CS) or more globally (as in language mixing) is an issue that this corpus cannot help us decide. Finally, Auer (1999) also argues that language mixing is more dense that mere CS and that insertion and

35 I thank an anonymous reviewer for raising this possibility and prompting me to address it.

36 Zajícová (2009: 77) states that “in the case of Jopara, it is precisely difficult to talk about any degree of homogeneity.” (“...en el caso del Jopara, es justamente dificil hablar de homogeneidad alguna.” My translation.)

37 The sole possible exception is the almost obligatory use of Guaraní interrogative markers in RQ-Jopara: out of 40 interrogatives in our corpus, only 4 do not have an interrogative marker. The putative obligatoriness of interrogative marking clearly deserves more study, especially with data from spoken Jopara.
alternation seem to converge in language mixing. In that sense RQ-Jopara does resemble language mixing. Future research should consider the possibility that spoken Jopara might match more closely Auer’s definition of language mixing than that of cs or mixed lects.

7 Conclusion

We conducted here an analysis informed by language mixing theory and codeswitching theory of the Jopara portrayed in the novel Ramona Quebranto (RQ-Jopara). This is a useful step towards a structural characterization of Jopara, given the amount of contradictory opinions about this code. We acknowledged in section 3.1 that RQ-Jopara cannot be taken automatically to represent spoken Jopara. However, the same phenomena highlighted in this paper (unless otherwise indicated, cf. Table 5) occur in blog and social media posts in Jopara, lending more authenticity to our corpus.38

For example, (28) shows Guaraní sv agreement on Spanish bases, Guaraní negation on Spanish lexemes, the Spanish words pero and otro (cf. ex. 14) inserted in a Guaraní clause, and modification by the suffix -ete (cf. ex. 11, 12).

(28) Pe guaranime ehaíacue apilla pero umi otro catu ndaleeiete voy jajaja

PE GUARANÍ-ME (R)E-HAÍ-VAEKUE A-pillá pero UMI otro

that Guaraní-loc 2sg-read-past 1sg-understand but that other

‘What you wrote in Guaraní I understood but the other things

KATU ND-A-leé-1-ETE VOI

just NEG-1SG-read-NEG-very EMPH

I don’t read them at all.’


Example (29) shows use of the Guaraní interrogative marker pí(k)ø in a Spanish question (cf. ex. 25a)

(29) para que pio existen los mensajes internos?

Para qué Pí(K)ø existen los mensajes internos?

For what Q they.exist the messages internal

‘What are private messages for?’

38 I thank an anonymous reviewer for suggesting this source of data for comparison.
Alternating use of Spanish plurals and Guaraní plural -kuëra with Spanish bases (cf. ex. 13) is shown in example (30).

(30) ANIKETI los padres y les llevan a sus hijo-KUERA!
NEG.IMP the parents and them they.take A their son-PL
‘Would that parents not take their children (there).’

The examples in (31) exemplify insertions of emphatic (nī)ko (cf. ex. 12, 23b).

(31) a. Gachita, vos NIKO sos una diosa CHE mamá.
Gachita you EMPH are a goddess 1SG mom
‘Gachita, you are indeed a goddess (=beautiful), hot mama’

b. Gracias lorito, pero eso KO entra en otra categoría
thanks little.parrot but that EMPH enters in another category
‘Thank you, Lorito, but that falls in a different category.’

Example (33) shows insertion of Guaraní lexical items in Spanish sentences (cf. ex. 16d).

(33) por tu culpa mi TAKARA’A casi me sale por la garganta
by your fault my gizzards almost me exits by the throat
‘because of you I almost had my stomach in my mouth.’
Finally, alternational cs is shown in (34) (cf. ex. 21).

(34) Así son las cosas, CHE TAVY-ETE NDE-REHE!

thus are the things 1sg crazy-very 2sg-locative

‘That’s the way things are, I am crazy about you!’

These data support our contention that RQ-Jopara is comparable enough to Jopara to provide preliminary insights into its structure. But of course, a detailed structural analysis of naturalistic Jopara is needed to assess empirically the proposals made here.

The analysis we presented shows that RQ-Jopara is indeed cs and a mixed lect in the sense of Backus (2003). None of the extant classifications of Jopara in the literature are applicable to RQ-Jopara. RQ-Jopara is not a pidgin (Boidin, 2006b), since it does not show the grammatical simplification that classically accompanies pidgins. Neither is it a creole (Boidin, 2006a): it is not normalized or fixed; it does not meet Bakker’s (2003) requirement that creoles present structure that is not traceable to any substrate. Lack of normalization also argues against calling RQ-Jopara a “third language” (Bakker, Gómez Rendón, and Hekking, 2008).

Word and morpheme counts suggest Spanish dominance in RQ-Jopara. So does the presence of Guaraní insertions in Spanish clauses. But the affixal asymmetry in the whole corpus and the presence of Spanish insertions (especially integrated Spanish verbs with Guaraní morphosyntax), both indicate Guaraní dominance at a more abstract level. The presence of bidirectional insertion can explain contradictory definitions of Jopara. Cases of insertion of Spanish words in a Guaraní-ml give RQ-Jopara the air of a “Guaraní with hispanisms” (Fernández Guizzetti, 1966), whereas the presence of Guaraní insertions in a Spanish-ml look like a Spanish spoken in Guaraní (Morínigo, 1959). These two extremes, combined with the presence of longer stretches of Guaraní alternating with longer stretches of Spanish also justify characterizations of
RQ-Jopara as “an unstable language mixing” (Lustig, 1996; Pottier, 1970), or a “complejo lingüístico sin normalizar” (non-normed linguistic complex, Palacios Alcaine, 2005). The observed lack of conventionalization makes it clear is that we are not in the presence of a third language. Characters in the novel do not show a consistent usage, a criterion Meakins (2012; section 3.1.2) adduces to identify an autonomous third language. Likewise, the lack of a consistent grammar-lexicon split (or similar) also argues against a mixed language analysis.

In sum, RQ-Jopara shows (1) a very fluid recruitment of Guaraní and Spanish as alternating MLS in bilingual CPSs and (2) alternations of Guaraní and Spanish monolingual CPSs in bilingual sentences. This yields the continuum of varieties in Figure 1:

The data from the novel Ramona Quebranto supporting a continuum view at the interactional level are consistent with the observation that there exists also a societal continuum, with the two poles of Guaraní and Spanish monolingualism, and degrees of bilingual competence in between (Gynan, 2001: 65).

We would like to conclude by pointing out the usefulness of terms like Guarañol and Castení to describe particular kinds of Guaraní-Spanish CS, and suggesting to reserve the term Jopara as an overall term for the mixed lect derived from Guaraní-Spanish CS. The notion of a continuum of varieties is consistent with both the concept of a mixed lect and the idea that RQ-Jopara is, first and foremost, CS. Structurally, Castení would be the production of Guaraní-Spanish CS patterns which are more easily understandable by monolingual speakers of Spanish, where most of the language content is Spanish and the Guaraní content is somehow not crucial for interpretation. Guarañol is, on the other hand, the production of Guaraní-Spanish CS patterns which would be more easily understandable by putative monolingual speakers of Guaraní, where most of the language content is Guaraní and the Spanish content is somehow not crucial for interpretation or it is fully integrated as a...
borrowing in monolingual Guaraní. Jopara is the production of any Guaraní-Spanish cs. One important question that remains for future research is whether children actually target Jopara in acquisition, a situation that, according to Meakins (2012) strengthens claims of autonomy for a mixed code.

References


---

39 Implicit here is the difficult issue of what counts as Guaraní nowadays, if the high variety (Guaraníeté) referred to in most typological and much linguistic literature (for example, in Canese, 1983; Gregores and Suárez, 1967; Zarratea, 2002) or the low variety described for example by Gómez Rendón (2007) of Kallfell (2011).


