

5-1-2016

# A variationist perspective on Spanish-origin verbs in Paraguayan Guarani

Josefina Bittar Prieto

Follow this and additional works at: [http://digitalrepository.unm.edu/ling\\_etds](http://digitalrepository.unm.edu/ling_etds)

---

## Recommended Citation

Bittar Prieto, Josefina. "A variationist perspective on Spanish-origin verbs in Paraguayan Guarani." (2016).  
[http://digitalrepository.unm.edu/ling\\_etds/4](http://digitalrepository.unm.edu/ling_etds/4)

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Linguistics ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact [disc@unm.edu](mailto:disc@unm.edu).

Josefina Bittar Prieto  
*Candidate*

Linguistics  
*Department*

This thesis is approved, and it is acceptable  
in quality and form for publication:

Approved by the Thesis Committee:

Christian Koops, PhD, Chairperson

Naomi Shin, PhD

Rosa Vallejos, PhD

**A VARIATIONIST PERSPECTIVE ON SPANISH-ORIGIN VERBS  
IN PARAGUAYAN GUARANI**

BY

JOSEFINA BITTAR PRIETO

B.A. SPANISH  
UNIVERSIDAD NACIONAL DE ASUNCIÓN, 2011

THESIS

Submitted in Partial Fulfillment of the  
Requirements for the Degree of

**Master of Arts  
Linguistics**

The University of New Mexico  
Albuquerque, New Mexico

May, 2016

## **DEDICATION**

Para Guille y Tati.

## ACKNOWLEDGMENTS

I would like to thank Dr. Christian Koops, Dr. Naomi Shin, and Dr. Rosa Vallejos for helping me not only in the process of writing my thesis but also in the development of my research.

También agradezco a toda mi familia, especialmente a mi abuela Polaca, que siempre hace lo posible por aclarar mis dudas con la predisposición y el cariño que solo puede tener una abuela. A Israel Pedrozo, por abrirme las puertas de gente maravillosa. A las voces anónimas del corpus de esta tesis, por regalarme su tiempo y sus historias. A Celia Candado, por ser una fantástica consultora lingüística.

Tati y Guille: agradecerles es poco y dedicarles esta tesis también. Así que van en los agradecimientos y en la dedicatoria. Gracias por el apoyo de todos los días.

# **A VARIATIONIST PERSPECTIVE ON SPANISH-ORIGIN VERBS IN PARAGUAYAN GUARANI**

by

Josefina Bittar Prieto

B.A., Spanish, Universidad Nacional de Asunción, 2011

M.A., Linguistics, University of New Mexico, 2016

## **ABSTRACT**

Paraguay presents a unique case of societal bilingualism in Latin America. Almost 90% of its population speaks the indigenous language Guarani, and half of the population is bilingual in Spanish and Guarani. However, contemporary Guarani has been described as an unpredictable, ad-hoc mix of Guarani and Spanish. This idea is also reflected in the local term *jopara* ‘mixture’, which is used to refer to the variety by its speakers. A particularly salient contact feature are loan verbs, i.e. Spanish-origin roots used with Guarani morphology. To gain an empirical perspective on the presence and motivation of Spanish loan verbs in Guarani, 35 sociolinguistic interviews were conducted with Guarani speakers who live or work in Asunción (Paraguay’s capital).

The initial hypothesis was that, due to a tendency towards bilingualism in urban areas, younger speakers would use Spanish-origin verbs more frequently than older speakers when speaking Guarani. However, an initial count of all verb tokens used by the two oldest and the two youngest speakers (average age: 77 vs. 21; total tokens: 2930) showed only a modest increase in the use of Spanish-origin verbs (an 8% rise). Next, the degree to which some verb meanings have competing Spanish-origin and Guarani-origin forms was quantified to test the hypothesis of ongoing relexification. Of the 331 different verb meanings expressed in these four interviews, very few have broadly equivalent and

therefore interchangeable forms. The vast majority of the meanings were expressed categorically, i.e. using only a Guaraní or a Spanish-origin form.

Two of the attested variable verb meanings ('know' and 'live') were coded in all 35 interviewees and tested for a possible correlation with demographic data and internal factors. It was found that the age of the speaker did not have the hypothesized effect. The amount of tokens of the Spanish-origin form of these verbs did not increase in the younger generation. A re-examination of the initial verb count showed that the difference in preference for Spanish-origin forms between the oldest and the youngest speakers was, instead, found in the low-frequency verbs. Thus, verb meanings like 'live' and 'know', which are mid-frequency, are not undergoing change. This resistance to change is even more evident in high-frequency verbs, which are all categorical Guaraní-origin stems, with no attested Spanish-origin counterparts.

Overall, these results suggest that to the degree that Spanish-origin verbs are replacing Guaraní verbs, only low-frequency verbs are affected. Also, the process is slower than generally assumed. The data analyzed here indicate that the incorporation of Spanish-origin verbs into the language for the most part does not represent a loss of Guaraní verbs, but rather serves the purpose of lexical expansion. Finally, the low variability in the expression of verb meanings helps discredit the idea that spoken Guaraní is a random, on-the-spot mixture of Guaraní and Spanish.

## TABLE OF CONTENTS

<i>1. Introduction</i> .....	1
<i>2. Theoretical background</i> .....	2
2.1. Language contact in Paraguay .....	2
2.1.1. Increase of bilingualism .....	2
2.1.2. Spanish influence in Guarani .....	3
2.2. The object of this study: ‘bilingual verbs’ or ‘loan verbs’ .....	4
2.2.1. The motivation and (social) distribution of Spanish-origin verbs in Guarani .....	7
2.2.2. Semantic domains of loanwords .....	11
2.3. Apparent-time variationist studies .....	13
<i>3. The study</i> .....	15
3.1. Hypotheses .....	15
3.2 Methodology .....	16
3.2.1. Data collection .....	16
3.3. Data analysis .....	19
3.3.1. First hypothesis .....	19
3.3.2. Second and third hypotheses .....	21
3.3.2.1. The first category: Spanish-origin forms only .....	22
3.3.2.2. More observations on semantic domains .....	26
3.3.2. The second category: Guarani-origin forms only.....	27
3.3.3. The third category: variable verbs .....	29
3.3.3.1. The initial list: tokens from the 4 interviews .....	29
3.3.3.2. The new list of variable verbs: tokens from 12 interviews .....	30
3.3.3.3. Two variable verbs: ‘know’ and ‘live’. Tokens from 35 interviews .....	38
3.3.3.3.1. Know and live: Results .....	39
3.3.3.3.1.1. Internal factors .....	39
3.3.3.3.1.2. External factors .....	41
3.3.3.4. Variable verbs and the third hypothesis .....	44
3.3.4. Frequency and Spanish-origin verbs .....	45
<i>4. Discussion and conclusion</i> .....	46
4.1. First hypothesis .....	46
4.2. Second hypothesis .....	47
4.3. Third hypothesis .....	48
4.4. Conclusion.....	49
<i>5. References</i> .....	52
Appendix 1: list of verb meanings in the corpus of 4 speakers .....	56
Appendix 2: corpus of Guarani speakers .....	64



# **A variationist perspective on Spanish-origin verbs in Paraguayan Guarani**

## **1. Introduction**

This thesis explores Spanish-origin loan verbs in present-day spoken Guarani. These two languages, Spanish and Guarani, have been in contact in what is now Paraguay for almost five centuries. According to the most recent census data, half of Paraguay's population is now bilingual (Paraguay, 2003). As a result of this intense and continuous co-occurrence, both Spanish and Guarani have experienced contact-induced change in all linguistic levels. This study specifically looks at verbal borrowings from Spanish into Guarani, a phenomenon that has been virtually unexplored, from a variationist perspective.

This thesis starts with a theoretical background on language contact in Paraguay, loanwords, and variationist studies in minority indigenous languages. Then, it continues with a description of the study, which includes the hypothesis, a description of data collection process, and an analysis of the data found in the corpus. The analysis is divided into the types of verbs that were found in the corpus: categorical Spanish-origin forms, categorical Guarani-origin forms, and verbs with variable forms. In the last part of the data analysis, the production of the variable verbs is correlated with speakers' demographics and language-internal factors.

Finally, a discussion of the results and the limitations of this study are presented with ideas for future research on Spanish loan verbs into Guarani.

## **2. Theoretical background**

### **2.1. Language contact in Paraguay**

#### *2.1.1. Increase of bilingualism*

Paraguay is well-known for its unique linguistic situation in Latin America: almost 90% of its population speaks the indigenous language, Guarani. Half of the population in the country is bilingual in Spanish and Guarani (Paraguay, 2003). Although, traditionally, Guarani has been associated with rural areas, and Spanish, with urban areas (Rubin, 1968), this split does not fully represent Paraguay's current linguistic reality. In the 1960s, Garvin and Mathiot (1968) already spotted what they called 'the urbanization of Guarani.' They claimed that the efforts that were being made to standardize the Guarani language reflected the urbanization of its speakers. It is unclear, however, whether Garvin and Mathiot (1968) were conveying that rural parts of Paraguay were becoming more urban or whether speakers of Guarani in the countryside were migrating to the city, and, thus, acquiring an 'urban culture.' Many decades later, Jinny Choi (2005) claimed that the standardization of Guarani –i.e. its recognition as an official language of Paraguay and its implementation in the school (K-12) system– was partially responsible for the increase of bilingualism in Asunción (the capital of Paraguay) that she attested through her survey-based research. Choi (2005) compared surveys of language use from 1990 and 2000 taken by people who live in the capital and detected an increase in the use of both languages in different types of interactions. She found that people were using both languages instead of just Spanish or just Guarani. For instance, while in 1990 30.0% of the participants responded that they spoke both Guarani and Spanish with their siblings, in 2000, 42.2% gave this answer. This increase is visible in most types of interaction (e.g.

with the parents, with the workmates at work, etc), and it is accompanied by a decrease in percentages of the answers ‘Spanish’ and ‘Guarani.’ Choi also claims that Guarani may be used more in the city because of rural-urban migration. While in 1982 42.8% of the population was living in urban areas, in 2002 this percentage increased to 56.7% (Paraguay, 2003). That is, rural-urban migrants might be bringing Guarani to the cities. Also based on census data, Gómez Rendón claims that societal bilingualism –that is, the co-occurrence of two languages in a society– “appears to be on the way to be attained in the next twenty years as the diatopic gap between P[araguayan] G[uarani] (rural) and Spanish (urban) is gradually bridged with a bilingualism spreading nation-wide” (forthcoming:3).

### *2.1.2. Spanish influence in Guarani*

What implications does this tendency towards bilingualism have for the structure of the languages and, in particular, for the structure of Guarani? It has been proposed that, due to the intensification of the contact with Spanish, Guarani has changed due to both lexical and grammatical borrowings (Gómez Rendón, forthcoming). This highly ‘Hispanicized’ Guarani has been referred to as *Jopara* (the Guarani word for ‘mixture’) by many scholars (e.g. Palacios Alcaine, 1999; Gómez Rendón, forthcoming) who have embraced this widespread term used by Paraguayans to refer to the ‘mix’ of Guarani and Spanish in a very broad sense. Moreover, the latest publications on Paraguayan Guarani make a distinction between ‘traditional Guarani’ (a variety that only exists in the written form) and ‘Paraguayan Guarani’ or ‘Jopara’, the only spoken form of Guarani (Dietrich, 2010). Some have even proposed that Jopara is used differently by every speaker and that it is a somewhat random and on-the-spot mixture between Guarani and Spanish (Palacios

Alcaine, 1999:30; Kallfell, 2006:334). Hedy Penner (2014) has criticized the idea of a random mixture by claiming that if the only variety of Guarani spoken today is a mixture of Guarani and Spanish, or Jopara, then where would speakers get the input of a non-mixed variety of Guarani? In fact, many studies on Paraguayan Guarani have commented on the difference between ‘pure’ Guarani, existent only in the written form, and ‘Paraguayan’ and ‘Jopara’ Guarani, the spoken form of Guarani (e.g. Dietrich, 2010; Gómez Rendón forthcoming). This distinction is irrelevant for this study, as only the spoken language is being studied. Also, the term ‘Jopara’ is avoided because it is rather problematic, as there is no consistency in scholars’ –or speakers’– definition of this word. Throughout this study, both ‘Guarani’ and ‘Paraguayan Guarani’ will be used interchangeably to refer to the variety of Guarani that is spoken by non-indigenous people in Paraguay. ‘Spanish’ or ‘Paraguayan Spanish’ will be used to refer to the variety of Spanish spoken by people in Paraguay, either natively or as a second language.

## **2.2. The object of this study: ‘bilingual verbs’ or ‘loan verbs’**

According to Wichman and Wohlgemuth, “[f]ew studies have been devoted to the topic of borrowed verbs” (2008:89). However, two typological efforts to classify these verbs have been recently made: one by Pieter Muysken (2000) and the other by Wichman and Wohlgemuth (2008).

Muysken claims that a language can borrow a verb from another language using one of these strategies:

a. *Insertion*: the donor language verb is used in the place of a native verb. The ‘new’ verbs can have an adapted form or not. One of the examples Muysken gives is *ge-zoom-d* in Dutch. In this case, the English verb *zoom* is imported directly (Muysken, 2000:184).

b. *Adjunction*: the donor language verb is adjoined to a helping verb. It is not a complement of the helping verb and thus, the adjoined verb is not required for the syntactic completeness of the sentence. For instance, from Tamil

Indian women-e avaa discriminate panna-ille

AC they do-NEG

‘They don't discriminate (against) Indian women.’ (Sankoff, Poplack, and Vanniarajan 1990: 80, qtd in Musyken, 2000:206)

c. The new verb is a nominalized complement of a causative helping verb in a compound. For instance, from American Portuguese: *fazer o spoil* (‘to spoil’) (Muysken 2000:207).

d. The borrowed verb has the infinitive form and it is a complement of the auxiliary. It does not have to be right next to the auxiliary. The new verb is an infinitive and it complements a native auxiliary. For instance, from Moroccan Arabic Dutch

ka-t-dir m a-hum voetballen?

DUR-2±do.IMPF with-3pl football

‘Do you play football with them?’ (Muysken 2000:216)

Muysken calls the result of all the strategies above ‘bilingual verbs.’

Wichman and Wohlgemuth (2008) also identify four patterns in the integration of loan verbs into the recipient language:

1. *The light verb strategy*, in which a verb with a meaning similar to ‘do’ in the native language is used with the loan verb.
2. *Indirect insertion*, in which a native affix is used to ‘accommodate’ the loan verb, direct insertion and paradigm transfer.
3. *Direct insertion*, in which the loan verb does not change morphologically but is rather borrowed as “root-like, infinitive-like, imperative-like, inflected for third person or nominalized [...]” (2008:99)
4. *Paradigm transfer*, in which the new verb is borrowed not only with its stem but with some of its verbal morphology.

If we were to group these two classifications, Wichman and Wohlgemuth (2008)’s light verb strategy would correspond to three of Muysken (2000)’s categories (b, c, d, above), all of which use a native verb (usually ‘do’) to accommodate the loan verb. Likewise, Muysken’s insertion strategy corresponds to two of Wichman and Wohlgemuth’s strategies: direct and indirect insertion. The paradigm transfer strategy, much less typical according to Wichman and Wohlgemuth, is not contemplated in Muysken’s typology.

On top of their classification based on more than a sample of 80 languages, Wichman and Wohlgemuth propose a *Loan Verb Integration Hierarchy*: “light verb strategy < indirect insertion < direct insertion < paradigm transfer” (2008:108). They hypothesize that the more intense the degree of bilingualism of the speakers, the more the chosen strategy will tend to go further to the right, and thus, that the strategy a language uses to loan verbs is not entirely based on the syntactic properties of the languages involved. Although testing this hypothesis is beyond the scope of this research, it is interesting to note that the case

studied here might support the claim. Paraguay has an intense bilingualism and the loan verb strategy in Guarani is the direct insertion strategy, which will be explained below.

The mechanism that Guarani uses to incorporate verbs from Spanish has been described as follows (Gregores and Suárez, 1967): the Spanish infinitive form minus the final *-r* is borrowed (for instance, *leer*, ‘to read’, would be borrowed as *lee*). Then, the speakers add the Guarani inflectional morphemes (for example, to *lee* they add *o*, which indexes the third person. The result is *olee*, which means ‘He/She reads’). This pattern follows Muysken’s (2000) insertion strategy: an imported stem takes the native affixes; or, in terms of Wichman and Wohlgemuth (2008), the direct insertion strategy. To my knowledge, there is no evidence that other strategies are used or have been used in the past. Also interesting to test Wichman and Wohlgemuth (2008)’s hierarchy would be a study of Guarani-origin verbs in Paraguayan Spanish, which has not been done to this date.

### *2.2.1. The motivation and (social) distribution of Spanish-origin verbs in Guarani*

Muysken (2000) claims that the formation of this or any other type of ‘bilingual verbs’ serves the purpose of vocabulary extension or indicates lexical loss. Lyle Campbell defines ‘loanword’ as a “lexical item (a word), which has been ‘borrowed’ from another language, a word which originally was not part of the vocabulary of the recipient language” (2004:63). He claims that the two main motivations for borrowing are “need and prestige” (2004:64). The first motivation could be paralleled with Muysken’s (2000) suggestion that bilingual verbs arise to expand the native language’s vocabulary: there are new concepts that need to be expressed with new words. The second motivation, prestige,

might be associated with what Muysken (2000) considers the other nature of bilingual verbs: lexical loss. In the latter case, foreign verbs replace native verbs, perhaps because of the prestige associated with the donor language. Another motivation for loanwords in general which has been proposed is a lexical gap in the speakers (e.g. Silva Corvalán, 1994), that is, the inaccessibility of certain words in language A makes bilingual speakers borrow words from language B while speaking language A. In other words, the borrowing happens when the speaker can't access the verb in the target language. This phenomenon is different from the incorporation of verbs due to a lexical need. The lexical gap motivation presupposes that both languages spoken by the bilinguals have a form for the target meaning. The lexical need motivation, on the other hand, assumes that one language lacks the form for the target meaning. One interesting case study of loan verbs that ruled out the lexical gap motivation was Wilson and Dumont (2015)'s research on the *hacer* + English verb construction in the Spanish of New Mexico. The authors claimed that the lack of pauses before the *hacer* + English verb construction and the fact that there was intra-speaker variation were evidence against the lexical gap motivation.

Whether loan verbs in Guarani are an indicator of vocabulary expansion or lexical loss remains to be determined, as no prior studies of these loan verbs exist. The influence of Spanish on Guarani, however, has caught the attention of several scholars. In a recent corpus-based study, Gómez Rendón (forthcoming) quantified the percentage of Spanish loanwords in Paraguayan Guarani. Overall, he found that 17.4% of the total tokens and 23% of the total word types were of Spanish-origin. Dividing the borrowings into major parts of speech, he found that nouns were the most frequently borrowed word class (37.2%), "followed by verbs (18.3%), adjectives (7.4%) and manner adverbs (9.9%)"



(Gómez Rendón, forthcoming:7). Although these numbers offer a general idea of the influence of Spanish on Guarani, they do not provide information of the social distribution of these loanwords, let alone a diachronic view on the status of these borrowings. However, even with very scarce empirical evidence, these Spanish-origin verbs have been contrasted with Guarani-origin verbs to compare two different ‘varieties’ of Guarani: the ‘standard’ and the ‘Jopara’ (Palacios Alcaine, 1999). In her *Introducción al estudio de la lengua y la cultura guaraníes*, Palacios Alcaine presents sample sentences of how the use of verbs differs in standard Guarani and Jopara.

“Standard” Guarani: (1) Nde kyvy o-hai peteĩ kuatione’ë. (1999:30)

Your brother 3SG-write one letter.

Jopara: (2) Nde kyvy o-escribi peteĩ kuatione’ë. (1999:30)

Your brother 3SG-write one letter.

According to Palacios Alcaine’s examples, in ‘standard’ Guarani, the Guarani-stem (*hai*) is preferred to express the meaning ‘to write’, while in ‘Jopara’ the Spanish-stem (*escribi*) is preferred. However, these examples do not come from a corpus. They are sentences created by the author to illustrate the imagined ‘varieties’ of Guarani. Gómez Rendón, on the other hand, illustrates the use of loan verbs in Paraguayan Guarani with corpus data. But he does not contrast it with a sentence in which the same verb meaning is expressed with a Guarani-origin stem. In his example, the loan verb is *rrekohe*, which is borrowed from Spanish *recoger*: ‘to collect’.

(3) ha'é-nte o-rrekohe-paité-va'erã mandyju

3SG-only 3SG-collect-EMPH-OBLG cotton

'She was the only one who had to collect cotton' (forthcoming:10)

Gómez Rendón does not explicitly claim that these Spanish-origin verbs distinguish a variety of Guarani from other varieties. However, he claims that they are a feature of Paraguayan Guarani or Jopara Guarani (he considers them synonyms) and that Paraguayan Guarani is “different from Jesuitical Guarani and Indigenous Guarani” (forthcoming:2). Therefore, although he is not directly stating it, he might be suggesting that these verbs were not part of the Jesuitical Guarani lexicon, “the language spoken by Indians living in the Jesuitical missions from 1610 to 1767” (forthcoming:1), or the contemporary Indigenous Guarani lexicon, “the language spoken nowadays by ethnic groups such as Paí Tavytera, Mbya, Ava and Ache” (forthcoming:1).

If the incorporation of loan verbs from Spanish in Guarani has increased with time, as seems to be suggested by Gómez Rendón (forthcoming), it might be possible to visualize this change. However, as mentioned earlier, there aren't any studies that focus specifically on loan verbs in Guarani. A source that can indicate the time-depth of some borrowings is the book by Marcos A. Morínigo *Hispanismos en el Guarani* (1931), which offers an inventory of Spanish-origin words that were present in Paraguayan Guarani at the time it was published. In the book, which is based on observation, 185 out of the 1,176 entries are verbs; this corresponds to 15.73% of the total loanwords.

Previous research on Spanish-origin loanwords into Guarani provides evidence for the claim that Spanish influence on Guarani is, at the same time, an old and ongoing phenomenon. In other words, Guarani will continue to incorporate new Spanish-origin words into its lexicon. Thus, as an ongoing phenomenon, this external-induced change should be visible in today's speakers of Guarani.

### *2.2.2. Semantic domains of loanwords*

The World Loanword Database has classified loanwords according to the effect they have on the lexicon of the recipient language (Haspelmath, 2009). If the loanword is entering the language to express a new meaning, then it is considered an insertion. These insertions are said to be cultural borrowings, which go along the “need” motivation proposed by Campbell (2009). If the loanword is replacing an existing word or if it coexists with the existing word, the borrowings are said to be core borrowings.

According to Haspelmath (2009), the motivation for these core borrowings is difficult to determine, however, he agrees with Campbell (2009) in that the prestige of the donor language might play a role. Therefore, from a variationist perspective, cultural borrowings are categorical. They are always expressed with the donor-language-origin word. Core borrowings, on the other hand, either show a quasi-categorical pattern or are variable one.

Based on a sample of 41 languages, the Loanword Typology Project created a list of items that are borrowing-resistant: the Leipzig-Jakarta list of Basic Vocabulary, and confirmed the idea “that words with culture-free meaning are less likely to be borrowed” (Tadmor, Haspelmath, & Taylor, 2010:233). The authors grouped the loanwords into

semantic fields and ranked them according to the percentage of loanwords they represented in the combined database. The three semantic fields that were most likely to be borrowed were: *religion and belief* (41.2%), *clothing and grooming* (38.6%), and *the house* (37.2%). The three that were less likely to be borrowed were: *the body* (14.2%), *spatial relations* (14.0%), and *sense perception* (11.0%).

In his inventory of Spanish-origin loanwords in Guarani, Morínigo (1931) also divided the items into the semantic fields. In the following table the name of the semantic field and the number of verbs per field, in parenthesis, is in the left column and an example is provided in the right column:

<b>Morínigo's categories</b>	<b>examples</b>
<i>material actions</i> (33)	<i>falta</i> ('lack'); <i>topa</i> ('find')
<i>relationships</i> (24)	<i>avisa</i> ('let know'); <i>ofrece</i> ('offer')
<i>accidents and qualities</i> (17)	<i>nace</i> ('be born'); <i>lastima</i> ('hurt')
<i>military and weapons</i> (17)	<i>ataca</i> ('attack'); <i>defende</i> ('defend')
<i>material activities</i> [food ingestion-related] (16)	<i>ayuna</i> ('fast'); <i>cena</i> ('have dinner')
<i>cattle raising</i> (11)	<i>amansa</i> ('tame'); <i>galopea</i> ('gallop')
<i>commerce</i> (11)	<i>cobra</i> ('collect'); <i>vende</i> ('sell')
<i>school activities</i> (9)	<i>lee</i> ('read'); <i>multiplica</i> ('multiply')
<i>clothing</i> (7)	<i>teje</i> ('knit'); <i>dibraza</i> ('wear a costume')
<i>games and superstitions</i> (6)	<i>aposta</i> ('bet'); <i>adivina</i> ('guess')
<i>religion and funerary practices</i> (5)	<i>adora</i> ('adore'); <i>roga</i> ('beg')

<i>parts of the body</i> (4)	<i>denuca</i> ('break the neck')
<i>medicines</i> (3)	<i>purga</i> ('purge')
<i>smoking</i> (2)	<i>pita</i> ('smoke')
<i>carpenter's utensils</i> (2)	<i>serrucha</i> ('saw')
<i>tailor's utensils</i> (2)	<i>planchea</i> ('iron')
<i>maritime transportation</i> (2)	<i>rema</i> ('row')
<i>kitchen</i> (1)	<i>bartea</i> ('wash down')
<i>food</i> (1)	<i>nseba</i> ('add grease')

Table 1. Morínigo (1931)'s semantic domains of Spanish loan verbs in Guarani

As can be seen, most of these categories belong to cultural domains. However, the category with the most verbs, *material actions*, is a rather vague category, as it contains verbs like: *accommodate, get used to, hurry, cross, deliver, paint, answer, destroy*, etc, which not only have very different meanings among each other but also appear to be culture-free and, thus –as proposed by Tadmor, Haspelmath, & Taylor, 2010– unlikely to be borrowed. What is to be determined is which of these Spanish-origin verbs coexist with Guarani-origin verbs with the same meaning and which ones do not have Guarani-origin equivalents, and whether semantic domains can predict which verb meaning fall in which categories.

### 2.3. Apparent-time variationist studies

Although spoken by most of Paraguay's population, Guarani is not the overtly prestigious language in the country. Most of the studies on the status of Guarani and Spanish have

agreed in that Spanish is perceived as the language of social mobility while Guarani is the language of national identity (Gynan, 1998; Choi, 2003). It is therefore not surprising that Guarani shows a large amount of structural and lexical borrowings from Spanish, even when it is not the language of a minority. In fact, many variationist studies of indigenous languages have shown “dramatic influences from socio-economically, politically and/or culturally powerful majority languages in direct, intrusive contact with the language of study” (Stanford & Preston, 2009:11). This dramatic influence can be seen in apparent-time studies, where older speakers show very different patterns of use than younger speakers (e.g. Dorian, 1981).

As mentioned before, scholars have claimed that Guarani is highly influenced by Spanish (Matras, 2007; Gómez Rendón, forthcoming). However, there are no corpus-based apparent time studies of Guarani. This gap in the literature prevents us from knowing whether the influence of Spanish is stable among generations or increases dramatically in younger speakers.

What is the motivation for the formation of bilingual verbs in Guarani? A variationist study might shed light on this question. On one hand, verb meanings that are exclusively expressed through a Spanish-origin stem across generations –that is, verb meanings that show null variation– might indicate that the motivation for the borrowing was need. On the other hand, verb meanings that are expressed through both a Spanish-origin stem and a Guarani-origin stem –that is, verb meanings that show variation– might indicate that the motivation for the borrowing is not need. If Guarani had a way of expressing the meaning with a native stem and yet, a Spanish-stem is borrowed, need might not be the reason of the borrowing. The motivation for the borrowing in this latter case would be harder to

identify. Based on Campbell (2009) and Muysken (2000), it could be proposed that the Spanish-origin verbs that are not borrowed to construe a new event are replacing Guarani-origin verbs with the same meaning. Because Spanish is overtly more prestigious than Guarani, speakers might choose the Spanish-origin verb form over the Guarani-origin form to index prestige. Another reason for variation among verb forms could be the lexical gap in the speakers' lexicon proposed by Silva-Corvalán (1994). This 'gap' would be wider in younger speakers because they are more exposed to both languages, Spanish and Guarani, than older speakers, who were more exposed to just one language, Guarani, in their youth. Thus, Guarani-forms of certain verb meanings may be more entrenched in the older speakers than in the younger speakers. Furthermore, the recent tendency towards bilingualism implies that older speakers would have had more interactions with monolingual speakers of either Guarani or Spanish than the younger generations, who are used to interact with bilingual speakers like themselves. Therefore, among the younger generations of bilinguals there is more room for borrowing, as the borrowed elements come from Spanish, a language that is spoken by most of the younger speakers' peers.

### **3. The study**

#### **3.1. Hypotheses**

The tendency towards bilingualism that has been described in urban areas of Paraguay presupposes that younger generations are more exposed to bilingual speakers of Guarani and Spanish than the older generations. In other words, older generations of urban

Guarani speakers' would have grown up in a community that was more divided into monolingual Guarani speakers and monolingual Spanish speakers. Thus, younger bilingual speakers whose peers are also bilingual may include more loanwords from Spanish in their Guarani. This increase of Spanish loanwords in younger Guarani speakers' speech would also be reinforced by the fact that Spanish is still the most spoken language in urban areas, Asunción in particular. Therefore, the first hypothesis of this research is that a larger percentage of Spanish-origin verbs will be found in the younger speakers than the older speakers. Secondly, as proposed by Tadmor, Haspelmath, & Taylor (2010) and Campbell (2009), it is hypothesized that the categorical Spanish-origin verbs will be "cultural borrowings" and will pertain to high-cultural-load semantic fields, such as *religion and belief*, and *clothing*. Finally, the third and main hypothesis of this research is that variable verbs, verb meanings that can be expressed with a Spanish-origin verb or a Guarani-origin verb, will be stratified by age. Thus, the use of the Guarani-origin form will be correlated with the older speakers, and the use of the Spanish-origin form will be correlated with the younger speakers. For instance, given the variable verb "to have a birthday", older speakers will tend to use the Guarani-origin form, *mboty*, and younger speakers will tend to use the Spanish-origin form, *cumpli*.

## **3.2. Methodology**

### *3.2.1 Data collection*

In June 2015, I conducted 35 sociolinguistic interviews in the capital of Paraguay, Asunción. According to the last census data, in 22,964 of the 117,209 households the language spoken is primarily Guarani (19.59%), while in 92,283 Spanish is the primarily



spoken language (78.73%) (Paraguay, 2003). As can be perceived, the preferred language in Asunción is Spanish. Nevertheless, this census did not include the option “both Guarani and Spanish” (as it had been done in the past, e.g. the 1982 census) as an answer. Thus, it cannot be known how many of these households are bilingual. There are some neighborhoods in Asunción and the metropolitan area, however, where Guarani is the primary language of communication. Unfortunately, there isn’t census data about the specific linguistic situation of each neighborhood. People in Asunción and in Paraguay do, however, have an impression of where each language is spoken. I, therefore, contacted people in Asunción and asked them for recommendations on where to conduct interviews in Guarani. One of these people was an acquaintance of mine, Israel Pedrozo, who suggested the under-privileged neighborhood San Cayetano (most known as Bañado Sur) of Asunción would be a good location for the interviews. As mentioned by the interviewees themselves, Guarani is the preferred language in the community, but Spanish is used as well. As Israel was not only born and raised in this neighborhood, but is also currently working there and is a well-known member of the community, I asked him if he would join me in the interviews. Once in Paraguay, I visited San Cayetano neighborhood almost everyday during the month of June 2015. Pedrozo accompanied me in most of the interviews there. There were some scarce cases where he wasn’t able to join me. In those cases, he would introduce me to the interviewees and I would conduct the interview by myself. In total, I recorded 23 interviews in Guarani in San Cayetano. In order to expand the study population, I also interviewed 8 more people in other neighborhoods of Asunción (only one of them with Pedrozo). Most of these people

outside San Cayetano work in Asunción but live in the surrounding cities: Ñemby, San Lorenzo, Mariano Roque Alonso, and Capiatá.

The 35 interviewees reported as bilinguals. Only two women said they were not able to speak Spanish but they were able to understand it. Regarding acquisition, most of them said they had learned Spanish at school, that is, they were sequential bilinguals. Only five participants reported they had learned both Guarani and Spanish simultaneously, that is, they self-reported as simultaneous bilinguals.

20 interviewees were women and 15 were men. The age range of the speakers was 18 to 79. Most participants were working class people. The most common job among men was drivers or gardeners, and most women were housewives. Only 2 interviewees had some college-level instruction, and one was finishing high school. The average years of formal education among all speakers was 6.6. Men's average of years of formal education was considerably higher than women's: 7.9 vs 5.3, respectively. Most participants who said why they dropped out of school claimed the reason was to start working.

The interviews were very spontaneous conversations. At the beginning of the interview, we would ask the participants to tell us about their life, where they were born, where they went to school, etc. Then, the interviewee would talk about the topic of their preference: their family, sports, the neighborhood, politics, etc. At the end of the interview, I asked participants to tell me something about their language use. The two basic questions about language use were what language they speak to whom, and what language their parents spoke to them as children.

### 3.3. Data analysis

#### 3.3.1. First hypothesis

In order to test the first hypothesis (*a larger percentage of Spanish-origin verbs will be found in the younger speakers than the older speakers*), I transcribed all the verb tokens of four interviews: the oldest male (age 75) and female (age 79) speakers and the youngest male (age 18) and female (age 24) speakers. Out of the four, only the oldest man switched to Spanish once in a while during the interview. I did not count those verbs, as the language of the study is Guarani. The other three interviewees did not code-switch during the interview.

The resulting list of verbs had 3553 tokens. I excluded the tokens produced by the interviewers and the ones produced by the interviewees that were clearly primed by the interviewers (as in “When did you *go* there?” “I *went* there yesterday”). The list reduced to 2930 verb tokens. These 2930 verb tokens corresponded to 331 verb meanings or types. I then separated the tokens into Spanish-origin and Guarani-origin and calculated the frequency of each category for each speaker. I also calculated the frequency for each verb meaning (type frequency). The frequencies per speakers can be observed in table 1 below:

Speaker (age)	Token frequency		Type frequency	
	Spanish-origin	Guarani-origin	Spanish-origin	Guarani-origin
Carlota (79)	12.54%	87.46%	32%	68%
Julián (75)	13.75%	86.25%	36%	64%
Paulina (24)	19.80%	80.20%	42%	58%
Aníbal (18)	22.81%	77.19%	45%	55%
total tokens: 2930			total types: 331	

*Table 2. Token and type frequency per verb category per speaker.*

As seen in the table, the younger the speaker, the higher the token and type frequencies of Spanish-origin verbs are. Thus, the first hypothesis is confirmed: younger speakers use more Spanish-origin verbs than older speakers. The average of Spanish-origin tokens among the four speakers is 17.23%, which is slightly lower than Gómez Rendón (forthcoming)'s results (18.3% of the total verbs were Spanish loanwords). It is worth noticing that the increase in frequency of Spanish-origin verbs from the oldest to the youngest speaker is only 10.27% (token frequency) and 13% (type frequency). Considering that the age difference between these two speakers is 61 years the increase in Spanish-origin tokens does not appear to be dramatic. However, cross-linguistic data is needed to judge the 'size' of the change.

### 3.3.2. *Second and third hypotheses*

I separated the verb meanings into three categories to test the second and third hypotheses: 1) those that only had the Spanish-origin form, 2) those that only had a Guarani-origin form, and 3) those that had two forms: a Spanish-origin and a Guarani-origin form. The second hypothesis predicts that categorical Spanish-origin verbs will be cultural borrowings and will pertain to high-cultural-load semantic fields. Thus, if the second hypothesis were confirmed, the categorical Spanish-origin verbs will have a heavy cultural load, which will correspond to Tadmor, Haspelmath, & Taylor (2010)'s highly borrowable semantic domains. The third hypothesis predicts that verb meanings that can be expressed with a Spanish-origin verb or a Guarani-origin verb will be stratified by age. Thus, if the third hypothesis were confirmed, the variable verbs category would account for the token number difference in Spanish-origin verbs between the older and the younger speakers. That is, the older speakers would use the Guarani-origin form and the younger speakers would use the Spanish-origin form. Summarizing: the second hypothesis will be explored through the first verb meanings category: Spanish-origin only verbs; and the third hypothesis will be tested with the third verb meaning categories: the variable verbs. Categorical Guarani-origin verbs, the second category, will also be explored in order to contrast them with categorical Spanish-origin verbs. If Spanish-origin verbs are culturally-loaded, are Guarani-origin verbs not culturally-loaded? Because in later stages of this research, I will use verb tokens to run statistical analyses, I excluded those with a token frequency of less than 7.

### 3.3.2.1. The first category: Spanish-origin forms only

15 verb meanings with a token frequency higher than or equal to 7 for all speakers combined were expressed with the Spanish-origin form only. These verb meanings are presented in the table below.

	<b>Stem</b>	<b>Meaning</b>	<b>Tokens</b>
1.	<i>estudia</i>	study	21
2.	<i>agarra</i>	grab/catch	13
3.	<i>gusta</i>	like	13
4.	<i>falta</i>	lack	12
5.	<i>juga</i>	play	12
6.	<i>abri</i>	open	10
7.	<i>queja</i>	complain	9
8.	<i>ayuda</i>	help	9
9.	<i>lee</i>	read	9
10.	<i>topa</i>	find	8
11.	<i>gusto</i>	<i>be likeable</i>	8
12.	<i>paga</i>	pay	8
13.	<i>cocina</i>	cook	7
14.	<i>regala</i>	give a present	7
15.	<i>aprende</i>	learn	7

Table 3. High-frequency Spanish-origin verbs.

In the second hypothesis it was proposed that the action meanings that were expressed exclusively through Spanish-origin stems would be culturally loaded and that, therefore, the motivation to borrow them would be need. However, among the categorical Spanish-origin verbs, there are not many that are highly culturally loaded. In fact, most of these verbs belong to the second half of the Loanword Typology Project (2010) list of 22 semantic domains (number 1 being the most likely to be borrowed domain). Among the

Spanish-origin verbs that are categorical, that is, that do not have Guarani counterparts, there are verbs of cognition (11<sup>th</sup> in the borrowability scale): *study* and *learn*, and a verb of the food and drink domain (7<sup>th</sup> in the scale): *cook*. Also, *pay* and *open* (start a business, open a new store) are culturally loaded, and we have assigned them to the 5<sup>th</sup> domain: social and political relation. *Play* is most often used to talk about sports like soccer and volleyball, which are culture-specific activities (possible also belonging to the 5<sup>th</sup> domain: social and political relations). Surprisingly, we find somewhat culture-free meanings, like *like*, *lack*, *grab/catch*, and *find*. These belong into the emotions and values and the physical world semantic domains, which are near the bottom of the Loanword Typology (2010) scale. A table is provided to contrast these semantic domains with the Spanish loan verbs. It is worth noticing that the Loanword typology does not provide specific scales for specific word classes: their scale was based on loanwords in general and is probably highly biased toward concepts expressed as nouns, the most borrowable world class. The reason why categorical Spanish-origin loan verbs appear to go against the proposed hierarchy might be partially due to the fact that the hierarchy is not based exclusively on action words. In the table below the Spanish-origin verbs are placed to the right of the semantic fields they pertain.

<b>Semantic field In order of most to least borrowed</b>	<b>Categorical Spanish-origin verbs (# of tokens)</b>
1. Religion and belief	
2. Clothing and grooming	
3. The house	
4. Law	
5. Social and political relations	play [sports] (12) open [a store] (10) pay (8)

	give a present (7)
6. Agriculture and vegetation	
7. Food and drink	cook (7)
8. Warfare and hunting	
9. Possession	
10. Animals	
11. Cognition	study (21) learn (7)
12. Basic actions and technology	help (9)
13. Time	
14. Speech and language	read (9)
15. Quantity	
16. Emotions and values	like (13) lack (13) be likeable (17)
17. The physical world	grab/catch (13) complain (9) find (8)
18. Motion	
19. Kinship	
20. The body	
21. Spatial relations	
22. Sense perception	

Table 4. Loan verbs by semantic field according to the LWT.

Regarding the history of these fifteen Spanish-origin verbs, out of fifteen, five (*like*, *cook*, *lack*, *read*, *find*) were included in the 1931-book *Hispanismos en el Guaraní* by Marcos A. Morínigo (Morínigo, 1931). The fact that these verbs were present in the book indicates they have been part of the Guaraní lexicon for at least a century. However, more diachronic research has to be conducted to identify the motivation for the incorporation of these loanwords. This is difficult because Guaraní lexicographers have had extremely purist ideologies. For instance, one of the most complete Guaraní-Spanish dictionaries was written by Antonio Guasch, who is known to have had very negative attitudes towards Spanish loans in Guaraní (Penner, 2014). In a recently published book, Penner (2014) collects some of Guasch’s quotes on this matter. To cite a few: “Present day



Guarani has been impoverished a lot and corrupted [...], I wish we could completely ban Jopará...” (Guasch 1948:7, qtd in Penner 2014:28) and “The Guarani we speak is a degenerated language...” (Guasch 1956:10 qtd in Penner 2014:28). In the latest edition of his dictionary (2008), *cook*'s and *read*'s translations are, respectively, *japo tembi'u* (lit. *make food*), and *mo-ñe'ë* (lit. CAUS-*talk*) (Guasch & Ortiz, 2008). Given Guasch's negative attitudes towards loans, it is very likely that the Guarani equivalents of the Spanish forms *cocina* and *lee* had been already in disuse when Guasch first published the dictionary, in 1961; maybe these were never used at all and were proposed by him as 'alternatives' for the loans. The question of whether *japo tembi'u* and *mo-ñe'ë* were replaced by *cocina* and *lee* or not will remain unanswered for now. The same problem applies to the verb meaning 'find'. According to the dictionary, *johu* expresses that meaning, but there are no tokens of that form in our four-speaker corpus. They only use the Spanish-origin form *topa*. Can we say *topa* replaced *johu*? More diachronic research is needed to answer this question. If older recordings of Guarani speakers had tokens of *johu* where *topa* in present-day Guarani is expected, then a case could be made that *johu* has been replaced by *topa*.

*Gusta* ('like') and *falta* ('lack') are interesting loanwords because, unlike the ones previously mentioned, no Guarani equivalents can replace them without 'sacrificing' the grammatical construction of the sentence. That is, in order to express similar meanings with Guarani-origin words only, the whole sentence structure has to change. In fact, Guasch does not include *gustar* with the meaning 'like' in the dictionary and he translates *faltar* as *ndaipóri* (lit. *There isn't*) (Guasch & Ortiz, 2008). Then, it might be argued that these two verbs were borrowed for a need of grammatical convergence rather than a

lexical need. For instance, in sentence (3), *pelota* ('ball') is the object. If the form *falta* is used instead of *ndaipóri* ('there isn't'), then *pelota* becomes the subject. The sentence with the form *falta* would be *O-falta pelota kuri* (3SG-lack ball PST)

(3) Nda-ipó-ri pelota kuri

NEG-there.is-NEG ball PST

*There wasn't a ball.*

*Aníbal (18)*

In conclusion, it cannot be argued that categorical Spanish-origin verbs belong to semantic domains with a high cultural load, as proposed in the second hypothesis.

### 3.3.2.2. *More observations on semantic domains*

We have seen that the different semantic domains distinguished in Tadmor, Haspelmath & Taylor (2010)'s loanword typology can have both verbs of Spanish origin and verbs of Guarani origin. From a broad perspective, it appears that there isn't a semantic domain that is completely composed of either Guarani-origin or Spanish-origin verbs. However, a closer scrutiny of a larger corpus is needed to test this claim. From the corpus of four speakers, the 5<sup>th</sup> domain, social and political relations, in Tadmor, Haspelmath, & Taylor (2010)'s hierarchy has both Spanish-origin and Guarani-origin verbs. The same occurs with the religion domain (1<sup>st</sup> in the hierarchy), cognition (11<sup>th</sup> in the hierarchy), speech and language (14<sup>th</sup>). Here are some examples:

- Religion (1<sup>st</sup>): Guarani-origin: *ñembo'e* (pray); Spanish-origin: *ora* (pray), *adora* (adore), *venera* (worship).

- Social and political relations (5<sup>th</sup>): Guarani-origin: *ñemu* (trade, sell), *jogua* (buy). Spanish-origin: *cobra* (collect), *vende* (sell).
- Cognition (11<sup>th</sup>): *mbo'e* (teach), *aprende* (learn)
- Speech and language (14<sup>th</sup>) (and school-related): Guarani-origin *hasa* (pass [the test/course], possibly a calque from Spanish: *pasar* [la prueba]), *sẽ* (get [the scholarship], possibly a calque from Spanish: *salir* [la beca]); Spanish-origin: *lee* (read), *escribi* (write), *forma* (get trained), *exonera* (exonerate), *aplaza* (fail [the test/course]).

The presence of both Spanish-origin and Guarani-origin in the same semantic domains makes semantic fields a poor predictor of the form that will be the chosen to express a verb meaning. In other words, the semantic field a verb meaning belongs to would not predict whether the stem used to express it will be Spanish-origin or Guarani-origin.

### 3.3.2. *The second category: Guarani-origin forms only*

In the corpus of these four speakers, of the 71 meanings with a token frequency of 7 or higher, 46 are categorical in Guarani-origin form (cf. Spanish-origin categorical verbs: 15). In the Leipzig-Jakarta List of Basic Vocabulary, ‘go’ is the first action in the borrowability scale (third item overall). That is, it is the action meaning that is less likely to be borrowed from another language (Tadmor, Haspelmath, & Taylor, 2010). In Guarani, ‘go’ is not only categorical in its form, it is also the second most frequent verb. The most frequent verb is ‘tell/say’. Furthermore, as can be seen in table 4, all the Guarani-form categorical verbs are high in frequency. Tadmor, Haspelmath, & Taylor claim that “[i]n corpus linguistics, [the notion of basic vocabulary] may be equated with

the most frequent words” (2010:227). This study shows the correlation: the average token number of the 10 highest frequency categorical Spanish-origin verbs is 16; the average token number of the categorical Guarani-origin verbs is 147.4. Besides the verb *go*, out of the ten highest-frequency verbs, *come* and *give* are also in the Leipzig-Jakarta List of Basic Vocabulary (Tadmor, Haspelmath, & Taylor, 2010).

	<b>Stem</b>	<b>Meaning</b>	<b>Tokens</b>
1.	<i>e</i>	tell/say	208
2.	<i>ho</i>	go	199
3.	<i>japo</i>	do/make	108
4.	<i>reko</i>	have	105
5.	<i>ju</i>	come	102
6.	<i>me</i>	be (estar)	83
7.	<i>ĩ</i>	there is/are	75
8.	<i>me' ě</i>	give	73
9.	<i>ñe' ě</i>	speak	66
10.	<i>katu</i>	can	59

*Table 5. High-frequency Guarani-origin verbs.*

The fact that the highest frequency verbs are categorical Guarani-origin forms supports the claim that “frequently used words would also be highly resistant to borrowing, because more time and effort would be needed for the borrowing to become established” (Tadmor, 2009:74). According to Tadmor (2009), although logical, this relationship between borrowability and frequency hasn’t been extensively explored. Thus, apparent-time variationist studies may provide empirical data to study this correlation.

### 3.3.3. The third category: variable verbs

#### 3.3.3.1. The initial list: tokens from the 4 interviews

Out of the 331 verb meanings found in the list of verbs which excluded the tokens produced and primed by the interviewer (total tokens: 2930), only 10 that had a token frequency higher than 7 presented some degree of variability. However, because these verb meanings—the ones with variable forms—will determine the main linguistic variables of this study, I decided to use the list that included all the verb tokens produced by the interviewers and the interviewees (total tokens: 3553).

As a result, out of the 370 verb meanings found in the complete list (the one that included the tokens produced and primed by the interviewers), 13 with a token frequency higher than 7 had variable forms. The verbs and their number of tokens can be seen in table 5 below. (Notice that the verb *help* is in the list of Spanish categorical verbs and in the list of variable verbs. This is due to the fact that the list of variable verbs includes the tokens produced by the interviewers, who used the Guarani-form variant twice).

---

Meaning	Total tokens	Variants (Spanish-Guarani)	Spanish-bias
<b>know</b>	<b>110</b>	<i>conoce - kuaa</i>	0.018
<b>live</b>	<b>44</b>	<i>vivi - ko</i>	0.09
<b>leave (TR)</b>	<b>33</b>	<i>deja - heja</i>	0.242
<b>be born</b>	<b>29</b>	<i>nace - heñoi</i>	0.897
<b>sell</b>	<b>28</b>	<i>vende - ñemu</i>	0.857
<b>hit</b>	<b>23</b>	<i>golpea - nupã</i>	0.40
<b>finish</b>	<b>14</b>	<i>termina - pa</i>	0.714
<b>help</b>	<b>12</b>	<i>ayuda - pytyvõ</i>	0.833

---

<b>think</b>	<b>12</b>	<i>pensa - mo'ã</i>	0.25
<b>tell</b>	<b>12</b>	<i>relata - mongeta</i>	0.083
<b>pray</b>	<b>10</b>	<i>ora - ñembo'e</i>	0.40
<b>turn (age)</b>	<b>8</b>	<i>cumpli - mboty</i>	0.25
<b>yell</b>	<b>8</b>	<i>putea - jao</i>	0.25

Table 6. Variable verbs: Spanish-origin forms and Guarani-origin forms. The Spanish bias was calculated by dividing the amount of Spanish-origin tokens with the total number of tokens for the meaning.

Most of the verb meanings mentioned above lean towards one variant. For instance, on one hand, ‘know’ has by far more Guarani-form tokens than Spanish-form tokens; on the other, ‘sell’ has by far more Spanish-form tokens than Guarani-form tokens. Even though most of these verbs were almost categorical in their forms, I looked for more tokens of these 13 verb meanings in eight more interviews, which I selected randomly.

### 3.3.3.2. The new list of variable verbs: tokens from 12 interviews

After including the tokens of eight more interviews (the ones produced and primed by the interviewers were excluded), many verb meanings remained almost categorical in their forms, but a few verbs presented interesting patterns. These patterns will be described below:

- *End/finish*: Although there were almost equal numbers of Guarani-form tokens and Spanish-form tokens, 14 and 15 respectively, it was found that *termina* (the Spanish-form) was only used when the speaker was referring to academic studies. Thus, *pa* and *termina* are not interchangeable and cannot be considered variants of a variable verb.

(4) Las tres de la mañana-ma o-pa la quin[ce], nod-o-pá-i-katu ko.

At three in the morning-already 3SG.SUBJ-finish the fif[teen],

NEG-3SG.SUBJ-finish-NEG-well EMPH

*At three in the morning the fifteenth birthday party had already ended,*

*well it didn't really end.*

*Melisa (32)*

(5) O-termina la... kuri la i-sexto curso

3SG-finish the PST the his-sixth course

*He finished his sixth course [sixth secondary school year: 12<sup>th</sup> grade]*

*Melisa (32)*

It is worth noticing that in Spanish *terminar* 'to finish' appears in transitive constructions, e.g. 'el jefe terminó la fiesta' ('the boss ended the party'). However, *terminar* can also be used in passive constructions with the non-anaphoric *se*. In these cases, the theme becomes the subject of the clause, e.g. 'la fiesta se terminó' (the party INCH ended). Guarani only uses *termina* in the transitive construction. In the corpus of this study, inchoative uses of *terminar* are not found. The form *je-termina* (INCH-termina) is not present in the corpus.

- *Hit and leave (tr)*: The Spanish counterparts of the Guarani-origin forms of these verbs were always produced with the Guarani morpheme *-je*, which can function as a passive or a reflexive (Palacios, 1999; Guasch, 1971). That is, *golpea* was produced as *je-golpea* and *deja* was almost always produced as *je-deja*. In the case of 'hit', *je-golpea* indicates

that the agent and patient participants of the action are the same entity (in Spanish, the equivalent would be *golpear-se*: hit-REFL). That is, all the cases of *-je* in the corpus functioned as a reflexive morpheme, not a passive. Although *nupã* was almost always used in the active voice, one speaker produced *nupã*, the Guaraní-origin form of *hit* with the morpheme *-ñe* (nasal allomorph of *-je*) three times. In all three, the morpheme *-ñe* had a passive function, not a reflexive function. So, although the stems of *a-ñe-nupã* and *a-je-golpea* have the same meaning and they both take the passive/reflexive morpheme, *-je* (*-ñe*), the data suggest that *golpea* is only used with the reflexive, and *nupã* is preferred with the passive and the active voice. This is the linguistic distribution found in the corpus: 1) *nupã* in an active sentence (example 6), 2) *nupã* with the passive morpheme (*-ñe*, the nasal allomorph of *je*) (example 7) and 3) *golpea* with the reflexive morpheme (*-je*) (example 8):

(6) Che ha'e-akue la a-pilla-ro, a-hecha-ro... a-troza-ta, ai-nupa-ta, a-juka-ta...

I say-PST the 1SG.SUBJ-catch-if, 1SG.SUBJ-see-if... 1SG.SUBJ-destroy-FUT,  
1SG.SUBJ-hit-FUT, 1SG.SUBJ-kill-FUT

*I used to say if I catch him [cheating], if I see him [cheating]... I will destroy him,  
I will hit him, I will kill him...*

*Melisa (32)*

(7) Arreador-pe a-ñe-nupã

Whip-with 1SG.OBJ-PASS-hit

*I was hit with a whip.*

*Julián (75)*



(8) Ha'a che... che jyva ári ha a-je-golpea-vai

1.SG.fall my... my arm on and 1.SG.OBJ-REFL-hit-bad

*I fell on my... my arm and I hit (hurt) myself badly.*

*Aníbal (18)*

The same complementation can be found with *leave*: *deja* almost always occurs with the morpheme *-je* as a reflexive (in Paraguayan Spanish, *dejarse* means ‘abandon’), while *heja* hardly ever occurs with that morpheme, and when it does (only one token in the corpus), it has a passive function. The following sentences illustrate the complementarity:

(9) Che ha'e "a-ha-ta a-mba'apo."

A-heja-pa la che memby-kuera.

I said "1SG.SUBJ-go-FUT 1SG.SUBJ-work-FUT."

1SG.SUBJ-leave-COMPL the my child-PL

*I said I will go to work.*

*I left all my children [to go to work].*

*Melisa (32)*

(10) O-je-heja-akue mita'i-pe plata amo-ite Argentina-pe

3SG.OBJ-leave-PST kid-to money there-AUG Argentina-in

*Money was left to the kid over there in Argentina.*

*Sara (36)*

(11) Mamá o-je-deja chugui

Mom 3SG.SUBJ-REFL-leave from.him

Mom abandoned him.

Paulina (24)

- *Yell*: it was found that *putea*, the Spanish-origin form, was used to mean the action of yelling was done by cursing angrily. The meaning is different from *jao*, which simply means yelling.

- *Sell, help and be born*: These verbs almost always occur with the Spanish-origin form. *Vende* (sell) and *nace* (be born) are both present in Morínigo (1931)'s inventory of Hispanic words in Guaraní. Unfortunately, Morínigo does not offer an explanation of the motivation of these loanwords in his inventory. However, in an article published in 1962, he states that *vende* arises from a need for specificity, given that *ñemu* means 'to trade' (and thus, it can be interpreted as either *sell* or *buy*) (Morínigo, 1962). The Guaraní-form tokens of 'be born' and 'sell' were produced by the two oldest speakers (79 and 75 years old). However, they also produced the Spanish-origin forms. In the case of 'help', only a 33-year-old produced the Spanish-forms. It can be presumed that the Spanish-origin forms took over the Guaraní-origin forms many years ago.

---

Meaning	Total tokens	Variants (Spanish – Guaraní)	Spanish bias
sell	66	<i>vende - ñemu</i>	0.045
help	37	<i>ayuda - pytyvõ</i>	0.054
be born	27	<i>nace- heñoi</i>	0.037

Table 7. 'sell', 'help' and 'be born', number of total tokens, stems, and Spanish bias

- *Pray*: it was claimed by a language consultant that *ora*, the Spanish-origin form of pray, is used by the Evangelic Protestants to differentiate themselves from the Catholics, who use *ñembo'e*. This information is corroborated by the fact that the person who produced the Spanish-origin tokens of 'pray' identified himself as an Evangelic Protestant.

- *turn (age), think, believe*: While looking for tokens of *think*, I noticed that *believe* (semantically very closely related to *think*) also presented a certain degree of variability. I included *believe* in the new list even though it was not part of the original list of variable verbs. Although there aren't many tokens of these verbs, it can be inferred from the data that these forms are interchangeable. However, more tokens are needed to have a stronger case. Here are two sentences where the stems could be interchanged without affecting the meaning of each utterance:

(12) Oi-mo'ã péa la hi-óga.

3SG.SUBJ-think that the her-house.

*She thought that was her house.*

*José (42)*

(13) En vez de o-pensa ha'e... la che país oĩ-ngo en extrema pobreza.

Instead of 3SG.SUBJ.think he... the my country is-EMPH in extreme poverty.

Instead of thinking that... my country is an extreme poverty situation.

*Ramiro (54)*

(14) Agosto a-mboty dieciocho ha a-sẽ.

August 1SG.1SUBJ-turn eighteen and 1SG.1SUBJ-left.

*In August I turned eighteen and I left.*

*Pipo (69)*

(15) Pe viernes a-cumpli setenta y cinco años.

This Friday 1SG.1SUBJ-turn seventy and five years.

*This Friday I turn seventy-five.*

*Julián (75)*

(16) Pediatra-pe che a-rovia.

*Pediatrician-OBJ I 1SG.SUBJ.believe.*

*I believe the pediatrician.*

*Jaime (42)*

(17) Nd-o-cree-i-akue chupe, la ñ-hermano.

NEG-3SG.SUBJ-believe-NEG-PST him, the his-brother.

*He didn't believe him, his brother.*

*Israel (35)*

Meaning	Total tokens	Variants (Spanish – Guaraní)	Spanish-bias
<b>think</b>	<b>40</b>	<i>pensa/malicia - mo'ã/hi'a</i>	0.75
<b>turn (age)</b>	<b>17</b>	<i>cumpli - mboty</i>	0.29
<b>believe</b>	<b>11</b>	<i>cree - (gue)rovia</i>	0.36

Table 8. 'think', 'turn' (age), and 'believe' according to degree of Spanish-bias.

- *Know* and *live*: Also interchangeable are the forms *kuaa* and *conoce* and *ko* and *vivi*.

Because these verb meanings have more tokens than the previous three (*think*, *turn* (age), and *believe*), I will use them to test which factors play a role in the production of the form.

The stems *kuaa* (know) and *ko* (live) have many meanings in Guaraní. That is, the equivalent stems in Spanish are polysemous. In the case of *kuaa*, the stem can be translated into Spanish as *conocer* (to know a person or a place) or *saber* (to know something, to have a skill). *Ko* (live) can be translated into Spanish as *vivir* (to live in a place), *funcionar* (to function/work), or *estar* (to be (e.g. ok)). It was found that the Guaraní-origin and the Spanish-origin forms were only interchangeable when the meanings were *know a person or place* (*kuaa* vs. *conoce*) and *live* (*ko* vs. *vivi*). However, even after excluding tokens where *kuaa* could not be translated as *conocer* and where *ko* could not be translated as *vivir*, there was clearly a preferred form: the Guaraní-origin form.

Meaning	Total tokens	Variants (Spanish-Guarani)	Spanish-bias
live	55	<i>vivi-ko</i>	0.15
know	50	<i>conoce-kuaa</i>	0.08

Table 9. 'live' and 'know', tokens from 12 interviews, according to Spanish bias.

Furthermore, the Spanish-origin tokens of 'live' were produced by the two oldest speakers (79 and 75 years old) and by two middle-aged speakers (54 and 42 years old). Similarly, the Spanish-origin tokens of *know* were produced by a 75-year-old and a 69-year-old man. In fact, Morínigo (1931) included the Hispanicism *conoce* in his inventory of Hispanicisms in Guarani. The long presence of the loan verb and the fact that it is produced by older generations more often than by younger generations provides evidence against the idea that *conoce* is replacing *kuaa*.

Even though there is evidence against the hypothesis that within variable verbs, older speakers prefer the Guarani-origin variant and younger speakers prefer the Spanish-origin variant (third hypothesis), it is worth looking at what social and linguistic factors predict one or the other. The verbs *know* and *live* are used because its two forms are interchangeable and they are present in almost every interview.

### 3.3.3.3. Two variable verbs: 'know' and 'live'. Tokens from 35 interviews

I coded the meanings 'know' and 'live' in the corpus of 35 speakers. I obtained a total of 133 tokens for 'know' and 258 tokens for 'live'. The Spanish-bias of these meanings in the corpus of 35 speakers is very similar to the corpus of 12 speakers (cf. Table X and

Table X). For the total tokens of ‘live’, 11% is produced with the Spanish-origin form, *vivi*. For the total token of ‘know’, also 11% is produced with the Spanish-origin form, *conoce*. In order to determine what factors increase the likelihood of the production of the Spanish-origin form, I contrasted the tokens with two internal factors and three external factors. The internal factors were: presence or absence of other Spanish loanwords in the sentence, and person and number of the inflected verb. The external factors were age, gender, and level of formal education of the speaker.

<b>Meaning</b>	<b>Total tokens</b>	<b>Variants</b> (Spanish-Guarani)	<b>Spanish-bias</b>
live	258	<i>vivi-ko</i>	0.11
know	133	<i>conoce-kuaa</i>	0.11

*Table 10. ‘live’ and ‘know’, tokens from 35 interviews according to form.*

### 3.3.3.3.1. Know and live: Results

#### 3.3.3.3.1.1. Internal factors

- *Presence of Spanish in the sentence*: 63.2% of the sentences with tokens of live and know had other Spanish loanwords in the sentence. This percentage increases slightly in sentences where the Spanish-origin form is produced, from 62.4% where there isn't any other Spanish loanword in the sentence to 69.8% where there is another Spanish loanword in the sentence. If we analyze each verb separately, we find that 60% of the tokens of *conoce* have another Spanish loanword in the sentence, while 69% of the tokens of *kuaa* have a Spanish loanword in the sentence. With ‘live’, the percentage of

Spanish presence is higher with *vivi* (75%) than with *ko* (59%): there is a 16% difference. However, this difference does not provide enough evidence to claim that the presence of Spanish loanwords increases the likelihood of the production of the Spanish-form, especially keeping in mind that with ‘know’ the Spanish presence was higher with the Guarani-form. It might also be the case that the presence or absence of Spanish loanwords affects each verb meaning differently.

<b>meaning</b>	<b>form</b>	<b>% presence of other Spanish loanwords</b>
<b>know</b>	Spanish: <i>conoce</i>	60
	Guarani: <i>kuaa</i>	69.5
<b>live</b>	Spanish: <i>vivi</i>	75
	Guarani: <i>ko</i>	58.7

*Table 11. presence of Spanish loanwords in sentences where a form of ‘live’ and ‘know’ appear.*

- *Verb inflection:* The Spanish-bias for both meanings, ‘live’ and ‘know’, was 0.11. This bias is stable in almost all verb inflections. In ‘live’, it is stable with all person and number inflections. In ‘know’, the Spanish-bias increases with the first person plural exclusive and with the impersonal, 0.33 and 0.25, respectively. However, in raw numbers, this bias represents 4 and 2 tokens respectively and, thus, no conclusions can be drawn. However, what these 6 tokens combined have in common is that they all have the reciprocal or the passive/reflexive morpheme, *-jo/ño* and *-je/ñe*. The Guarani-form counterparts also have these morphemes. In the case of the impersonal inflections, all tokens have the passive morpheme *-je*. In the case of the first person plural exclusive only 2 out of 8 tokens have the reciprocal *-jo*. Because in other meanings (e.g. ‘hit’ and ‘leave’), the Spanish-form was preferred with certain morphemes, I will venture the



proposal that the reciprocal morpheme increases the likelihood of the Spanish-origin form of ‘know’.

### 3.3.3.3.1.2. External factors

- *age group*: The third hypothesis of this study proposed that when a verb meaning was variable, older speakers would produce the Guarani-origin form and younger speakers would produce the Spanish-origin form. However, the data does not show this correlation with the meanings ‘know’ and ‘live’. In the case of ‘know’, the bias is stable across age groups. In the case of ‘live’, the Spanish-bias increases from the older to the middle-age group generation but decreases again in the younger generations. The younger generation shows less Spanish-bias than the older generation: 0.02 vs 0.07, respectively.

meaning/age	younger (18-37)	mid (40-59)	older (60-79)
<b>know</b>	0.12	0.12	0.11
<b>live</b>	0.02	0.32	0.07

Table 12. Spanish bias according to age-group.

What makes the interpretation of the age distribution of ‘know’ and ‘live’ difficult is that the patterns are different for each verb. If the use of the Spanish-origin form of ‘know’ also decreased in younger generations, some explanations could be proposed. Instead, what the data seems to suggest is that each meaning changes differently and a general tendency cannot be proposed for Spanish-origin verbs in general. An explanation for why these two meanings in particular behave differently could be the fact that the forms of ‘live’, *ko* and *vivi* are not developing specialized uses as the forms of know may be

developing. As mentioned above, there is a tendency of the Spanish-origin form to pattern with the reciprocal and passive morphemes. As ‘live’ is an intransitive verb, it does not take the reciprocal, reflexive, or passive morphemes. We could say, then, that there is less room for specialization of each form. Then, the meaning ‘live’ is more vulnerable to social effects.

One external factor that might account for the decrease of Spanish-origin forms of know is the implementation of Guarani as a subject in schools through the 1994 Educational Reform. In Paraguay, people younger than 30 today, who attended school, had Guarani as a subject every school year. Although it is very likely that every school has a different approach to teaching Guarani, the Reform intended that the ‘variety’ taught at schools was ‘pure’, that is, it had to “exclude everything that was remotely similar to Spanish” (Penner 2014:75). Thus, middle-aged generations may have started the incorporation of more Spanish to index prestige. But younger generations, due to their exposure to purist ideology, may be less likely to find incorporation of Spanish loanwords as prestigious. However, earlier in the research we found that the two youngest speakers *did* use more Spanish-origin forms overall than the two oldest speakers.

- *gender*: the Spanish-origin form and the Guarani-origin form have the same distribution in both female and male speakers. In other words, no gender effect was found.

Meaning/Spanish-bias	female	male
<b>know</b>	0.08	0.17
<b>live</b>	0.10	0.12

Table 13. Spanish bias according to gender.

- *age and gender combined*: as seen in the age distribution of live, the middle-age generation is using the Spanish stem more than the other generations. When combined with gender, we can see both middle-aged women and men showing using vivi more than the younger and older generations. ‘Know’, however, shows an interesting distribution: the groups that have the highest percentage of the Spanish form are younger women and older men. At the same time, older women and younger men have 0 tokens of the Spanish form of ‘know’.

Meaning/age	younger	mid	older
<b>know</b>			
female	0.17	0.06	0
male	0	0.17	0.26
<b>live</b>			
female	0.03	0.34	0.06
male	0	0.29	0.07

Table 14. Spanish bias according to age and gender combined.

- *years of formal education*: I divided the speakers into two groups. Those who had 5 years of formal education, that is, they completed the 5<sup>th</sup> grade, and those who completed the 6<sup>th</sup> grade or more grades. The upper half group is mostly composed by men and some younger female speakers, and the lower half group is mostly composed by women and some older male speakers. That is why, to test an effect of the Educational Reform implemented in 1994 it is better to look at age groups rather than years in formal

education. The years in formal education division almost overlaps with the gender division. That is why we find no effect of years of formal education or gender.

<b>Meaning/Spanish-bias</b>	<b>lower</b>	<b>upper</b>
<b>know</b>	0.07	0.15
<b>live</b>	0.11	0.11

*Table 15. Spanish bias years of formal education (lower: up to 5 years, upper: 6 years or more).*

#### *3.3.3.4. Variable verbs and the third hypothesis*

On one hand, the third hypothesis proposed that the Spanish-form variant of the variable verbs would be present in the speech of younger speakers. However, this was not confirmed by the data. On the other hand, the first hypothesis stated that younger speakers would use more Spanish-origin verbs overall than older speakers. This hypothesis was confirmed with the corpus of four speakers, where an 8% increase of Spanish-origin verb token was found between the two oldest speakers and the two youngest speakers. The data confirmed that this 8% difference was not accounted by the variable verbs. Which verbs accounted for this difference?

The categorization of the all the verb tokens produced by the two youngest and two oldest speakers in the corpus put into evidence that categorical Spanish-origin verbs were 10 times more frequent than Guarani-origin verbs (the highest frequency for the Spanish-origin verbs was 208 and the highest frequency for the Guarani-origin verb was 21). This difference in frequency in the categories led us to believe that the 8% increase in the production of Spanish-origin verbs in the younger generations could be found in the low

frequency verbs. This claimed would also be supported by the fact that high frequency words are resistant to change (Bybee, 2015).

### 3.3.4. Frequency and Spanish-origin verbs

To test the effect that frequency would have on the production of verbs, the complete list of verbs of the 4-speaker corpus was separated into three categories: verbs with a token frequency of 35 in the four interviews combined (high frequency verbs), verbs with a token frequency of 7 to 34 (mid frequency verbs), and verbs with a token frequency of 7 or less.

It was found that both youngest and oldest speakers produced the high-frequency verbs with a Guarani-origin stem 100% of the times. The mid- and low-frequency verbs showed different percentages between the two groups. Out of all the mid-frequency, 70% were produced with a Guarani-origin stem by the oldest speakers, while the youngest speakers produced 65% of the mid-frequency verbs with a Guarani-origin stem. This difference widens in the low-frequency verbs. While the oldest speakers produce 54% of all the low-frequency with a Guarani-origin stem, the youngest speakers produce 44% of the low-frequency verbs with a Guarani-origin stem.

	Hfreq (35+)	Mfreq (7+)	Lfreq (<7)
OLDEST	100%	70%	54%
YOUNGEST	100%	65%	44%
Difference	0%	5%	10%

*Table 16. Percentages of Guarani-origin forms according to frequency for the two oldest speakers combines and the two youngest speakers combined.*

These results have two implications. On one hand, they imply that most meanings that will be borrowed from Spanish into Guarani, meanings that do not have Guarani-origin forms, will be low- or mid- frequency. On the other hand, these results show that the Guarani-origin forms that are likely to be replaced by Spanish-origin forms are the low-frequency verb meanings, primarily, and the mid-frequency, secondarily. In other words, if we are in the presence of a relexification process, where Guarani-origin forms are substituted by Spanish-origin form, the affected items will be the low-frequency ones. Because these items are low-frequency, a larger corpus is needed to study this phenomenon. This corpus should include recordings from last century and from areas of Paraguay where there is less exposure to Spanish. A higher change pace, that is more Spanish-origin words, would be expected in more recent recordings and in more urban places.

#### **4. Discussion and conclusion**

##### **4.1. First hypothesis**

The first hypothesis stated that younger speakers would use more Spanish-origin verbs overall than older speakers. The analysis of the speech of the two oldest and the two youngest speakers in the 35-speaker corpus confirmed this claim. An 8% increase of Spanish-origin verbs was found in the two youngest speakers. However, given that previous research conducted in Guarani concluded that present-day Guarani is a continuous code-mixing between Spanish and Guarani (e.g. Dietrich 2010), a higher

percentage difference was expected between the two groups. Also, the tendency towards bilingualism (Choi 2005) implies that younger Guarani speakers, especially in urban areas, are more exposed to both languages, Guarani and Spanish, than older speakers. Thus, younger speakers would have more opportunities to code-mix or to borrow more Spanish-origin lexical items. What the 8% difference between oldest and youngest speakers suggested is that the tendency to bilingualism does not presuppose a massive borrowing from one language into the other. On the contrary, borrowing from Spanish into Guarani –verbal borrowings, in particular– is rather systematic. All the verbs are borrowed following the direct insertion strategy proposed by Wichman and Wohlgemuth (2008). Also, when there seems to be Guarani form and a Spanish form for one verb meaning, one form is preferred with certain morphemes or with certain specific meanings.

#### **4.2. Second hypothesis**

The second hypothesis stated that the meanings that were only expressed with a Spanish-origin form would have a high-cultural load, as proposed by the Loanword Typology (Tadmor, Haspelmath, & Taylor, 2010). However, it was found that many Spanish-origin words that were categorical were in semantic domains that are less likely to be borrowed, for instance, ‘help’ (basic actions category), and ‘grab’ (the physical world category). Also, some domains, like *religion* and *social and political relations* that are very likely to be borrowed had both Spanish-origin stems and Guarani-origin stems. This is due to the fact that in some cases even when the meaning is new to the speech community, the form can either be a direct borrowing –and thus, have a foreign-origin stem– or be a calque – and thus, have a native-origin stem–. Therefore, the semantic domains proposed by the

Loanword Typology are not a good predictor of the type of stem –Spanish-origin or Guarani-origin– that a meaning will take. On the contrary, this study suggests that exploring the different meanings of each stem can account for borrowings in a more precise way than looking at semantic domains. That is, instead of finding patterns for all cognition verbs, we can look at a verb like ‘know’ and explore the different meanings that it has. As we saw it in Guarani, when ‘know’ means ‘to know or to meet a person’ it is more likely that the Spanish-origin stem will be used than when it means ‘to know something’. Had we tried to find a general pattern for all cognition verbs, we would have missed important patterns that emerge from the semantic nuances of each verb.

#### **4.3. Third hypothesis**

The third hypothesis stated that whenever a meaning could be expressed with a Spanish-origin verb or with a Guarani-origin verb, younger speakers would choose the Spanish-origin stem and older speakers would choose the Guarani-origin stem. However, among the very few variable verb meanings that were found in the corpus of 12 speakers, only a handful worked as true variable meanings. Even when the stems were semantically equal, one stem was favored in the presence of certain morphemes that add on the meaning of the stem. For instance, both *nupã* (Guarani-origin) and *golpea* (Spanish-origin) mean ‘hit’. However, *nupã* was preferred in the passive and active constructions and *golpea* was preferred in the reflexive constructions. Other variable verbs did not seem to have developed these linguistic preferences. Among these were ‘know’ and ‘live’, which were coded in the entire corpus of 35 speakers. As the other variable verbs, these had a Spanish-origin stem and a Guarani-origin stem. If our hypothesis were confirmed, there



would have been more tokens of the Guaraní-origin stem in the older generation and more tokens of the Spanish-origin stem in the younger generation. However, this pattern was not found. Instead, ‘live’ presented a stable distribution across generations (11-12% of the tokens were Spanish-origin in the three age groups) and the Spanish-origin stem for ‘know’ was more frequently used by the middle-aged generation than by the oldest and the youngest generation.

The results obtained with the analysis of the variable verbs disqualify the claim that younger generations prefer Spanish-origin stems in the presence of variable verbs. However, a deeper look at the first hypothesis showed that younger generations were using more Spanish-origin verbs when the meanings were low- or mid-frequency. Therefore, younger generations might indeed have a preference for Spanish-origin variant when the verb meaning is variable. But, because this preference is triggered by low-frequency meanings, it is extremely difficult to quantify it. Thus, a larger corpus will be needed to re-test this hypothesis. Also, the fact that ‘know’ and ‘live’ showed very different distributions across generations points at the importance of exploring these verbs not only from a broad perspective but also from a specific one.

#### **4.4. Conclusion**

This study was based on a corpus of 35 interviews conducted in Guaraní, in Asunción and its metropolitan area in June of 2015. An initial count of all the verbs produced by the two oldest and the two youngest speakers of the corpus showed younger speakers use more Spanish-origin verbs overall than older speakers. All the tokens of these verb meanings were categorized in three groups: those that were only expressed through a

Spanish-origin form; those that were only expressed through a Guarani-origin form; and those that were variable, that is, they had a Spanish-origin form and a Guarani-origin form. It was expected that the categorical Spanish-origin verb meanings would have a high-cultural load (Tadmor, Haspelmath, & Taylor 2010). However, these belonged to very diverse semantic domains, some of which, like *the physical world*, are unlikely to be borrowed, according to the Loanword Typology Project (Tadmor, Haspelmath, & Taylor 2010). Also, out of the 331 meanings found in the initial count of the four speakers, twelve showed some degree of variability. These were then coded in eight more interviews of the corpus. However, it was found that although the Guarani stem and the Spanish stem of these verbs were semantically equal, the inflected verbs patterned with different constructions. For instance, the Spanish-origin form would be preferred in a reflexive construction while the Guarani-origin form would be preferred in a passive construction. Thus, after excluding the verbs that were showing a complementary distribution, only five variable verbs remained: ‘think’, ‘believe’, ‘turn (age)’, ‘know’ and ‘live’. Because the latter two were the most frequent of the five, they were coded in the remaining 23 interviews of the corpus. The tokens of these two variable verbs were divided into three age groups to test whether the youngest group had produced more tokens of the Spanish-origin form, as proposed in one of the hypotheses of this study. However, it was found that neither the Spanish-origin form or ‘live’ nor ‘know’ were preferred over the Guarani counterpart by the younger generations. This result conflicted with a previous result in the study: the finding that younger speakers were using more Spanish-origin verbs than older speakers. However, a re-examination of the initial verb count showed that the difference in preference for Spanish-origin forms between the

oldest and the youngest speakers was in the low-frequency verbs. This result not only confirms the idea that frequent lexical items are resistant to change (Bybee, 2015), even in cases of language contact, but also that if younger speakers show preference for the Spanish-origin form only when the meaning is low- or mid-frequency. Because low-frequency verbs are intrinsically very unlikely to appear in spontaneous conversations, a much larger corpus of interviews will be needed to test the hypothesis Spanish-origin forms are replacing that low-frequency Guarani-origin forms.

Overall, the results of this study suggest that verbal borrowings from Spanish into Guarani are systematic: Guarani speakers do not borrow items randomly, as previously proposed (e.g. Palacios, 1999; Kallfell, 2006). If they did, we would have seen a lot more variability among speakers. Thus, the low variability in the expression of verb meanings helps discredit the idea that spoken Guarani is a random, on-the-spot mixture of Guarani and Spanish. If every speaker produced Present Day Guarani differently, each speaker would borrow different words and, thus, different verbs as well. This study showed that Guarani speakers do not borrow verbs from Spanish randomly. First, all the borrowings are incorporated to the Guarani lexicon through the direct insertion strategy proposed by Wichman and Wohlgemuth (2008). Second, borrowed verbs fulfill specific communicative purposes (e.g. with the form *golpea* of the meaning ‘hit’ the reflexive meaning is reinforced). Third, if Spanish forms are replacing Guarani forms, the relexification is occurring in the low frequency verb meanings, not in *all* verbs. Most importantly, the present study shows that the systematic exploration of contact phenomena in Guarani is not only plausible but can also contribute meaningfully to the language contact theory.

## 5. References

- Bybee, J. (2015). *Language change*. Cambridge University Press.
- Campbell, L. (2009). *Historical linguistics: an introduction*. Edinburgh: Edinburgh University Press.
- Choi, J. K. (2003). Language attitudes and the future of bilingualism: the case of Paraguay. *International Journal of Bilingual Education and Bilingualism*, 6(2), 81-94.
- Choi, J. K. (2005). The linguistic situation in urban Paraguay: a tendency toward Spanish-Guaraní bilingualism? *Spanish In Context*, 2(2), 175-201.
- Dorian, N. C. (1981). *Language death: The life cycle of a Scottish Gaelic dialect*. University of Pennsylvania Press.
- Dietrich, W. (2010). Lexical evidence for a redefinition of Paraguayan “jopara”. *Language Typology and Universals*, 63(1), 39-51.
- Garvin, P. L., & Mathiot, M. (1968). The urbanization of the Guarani language: a problem in language and culture. In J. Fishman (Ed.). *Readings in the sociology of language*. De Gruyter, Berlin, Boston.
- Gómez Rendón, J. A. (forthcoming). Language contact in Paraguayan Guarani. To appear in: Anthony Grant (Ed.). *The Oxford Handbook of Language Contact*. Oxford: Oxford University Press.
- Gregores, E., & Suárez, J. A. (1967). *A description of colloquial Guarani*. The Hague: Mouton & Co.
- Guasch, A. (1948). *El idioma guaraní. Gramática y antología de prosa y verso*. Segunda edición mejorada y acrecentada. Buenos Aires: Sarandi.

- Guasch, A (1956): *El idioma guaraní. Gramática y antología de prosa y verso. Tercera edición refundida y acrecentada*. Asunción: Casa América – Moreno Hermanos.
- Guasch, A., & Ortiz, D. (2008). *Diccionario castellano-guaraní, guaraní-castellano: sintáctico, fraseológico, ideológico*. Asunción: Centro de Estudios Paraguayos "Antonio Guasch".
- Gynan, S. N. (1998). Attitudinal dimensions of Guarani-Spanish bilingualism in Paraguay. *Southwest Journal of Linguistics*, 17(2), 35-59.
- Haspelmath, M. (2009). Lexical borrowing: concepts and issues. In M. Haspelmath & U. Tadmor (Eds.). *Loanwords in the world's languages: a comparative handbook*. Berlin: Walter de Gruyter.
- Kallfell G. (2006). Uso de las voces verbales del yopará, en comparación con las del guaraní. In W. Dietrich & H. Symeonidis (Eds.). *Guaraní y Mawetí-Tupí-Guaraní. Estudios históricos y descriptivos sobre una familia lingüística de América del Sur*. Berlin: Lit.
- Matras, Y. (2007). The borrowability of structural categories. In *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton de Gruyter.
- Morínigo, M. A. (1931). *Hispanismos en el guaraní: Estudio sobre le penetración de la cultura española en la guaraní, según se refleja en la lengua*. Buenos Aires: Casa J. Peuser.
- Morínigo, M. A. (1962). Influencia del español sobre el léxico del guaraní. *Filología*, 8, 213-220.

- Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge, UK: Cambridge University Press.
- Palacios Alcaine, A. (1999). *Introducción a la lengua y cultura guaraníes*. Valencia: Universitat de València, Departament de Teoria dels Llenguatges.
- Paraguay. Dirección General de Estadística y Censos. (1983). Censo nacional de población y viviendas del Paraguay, 1982. *Revista Paraguaya de Sociología*, 20(56), 161-174.
- Paraguay. Dirección General de Estadística y Censos. (2003). *Principales resultados del Censo 2002: Vivienda y población*. Asunción, Paraguay: DGEEC.
- Penner, H. (2014). *Guaraní aquí. Jopara allá: reflexiones sobre la (socio)lingüística paraguaya*. Bern, Switzerland: Peter Lang.
- Rubin, J. (1968). *National bilingualism in Paraguay*. The Hague: Mouton.
- Sankoff, D., Poplack, S. & Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Language Variation and Change*, 2, 71-101.
- Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. New York: Oxford University Press.
- Stanford, J. N., & Preston, D. R. (2009). *Variation in indigenous minority languages*. Amsterdam: John Benjamins.
- Tadmor, U. (2009). Loanwords in the world's languages: findings and results. In M. Haspelmath & U. Tadmor (Eds.). *Loanwords in the world's languages: a comparative handbook*. Berlin: Walter de Gruyter.
- Tadmor, U., Haspelmath, M., & Taylor, B. (2010). Borrowability and the notion of basic vocabulary. *Diachronica: International Journal For Historical*

*Linguistics/Revue Internationale Pour La Linguistique*

*Historique/Internationale Zeitschrift Für Historische Linguistik*, 27(2), 226-

246.

Wichmann, S., & Wohlgemuth, J. (2008). Loan verbs in a typological perspective. *Empirical approaches to language typology*, 35, 89.

Wilson, D. V., & Dumont, J. (2015). The emergent grammar of bilinguals: The Spanish verb hacer 'do' with a bare English infinitive. *International Journal of Bilingualism*, 19(4), 444-458.

## Appendix 1

*List of verb meanings present in the interviews of the two oldest and the two youngest speakers, from most to least frequent*

Verb meanings	Spanish-origin tokens	Guarani-origin tokens	Total tokens
1. tell/say		208	208
2. go		199	199
3. do/make		108	108
4. have		105	105
5. come		102	102
6. be (estar)		83	83
7. there is/are		75	75
8. give		73	73
9. speak		66	66
10. can		59	59
11. know		55	55
12. work		51	51
13. be (ser)		44	44
14. see		37	37
15. take (llevar)		37	37
16. live	4	30	34
17. enter		32	32
18. leave		30	30
19. leave (TR)	7	22	29
20. name		28	28
21. want		26	26
22. die		24	24
23. remember		23	23
24. bring		22	22
25. stay		22	22
26. study	21		21
27. put		19	19
28. teach		19	19
29. hit	4	12	16
30. sell	11	4	15
31. arrive		14	14
32. be born	12	1	13
33. grab/catch	13		13
34. like	13		13
35. finish	10	2	12



36. lack (faltar)	12		12
37. marry		12	12
38. play	12		12
39. release (soltar)		12	12
40. start/begin		12	12
41. wash		11	11
42. ask for		10	10
43. buy		10	10
44. climb		10	10
45. cry		10	10
46. open	10		10
47. be happy		9	9
48. complain	9		9
49. help	9		9
50. hurt	1	8	9
51. look for		9	9
52. need		9	9
53. read	9		9
54. eat		8	8
55. eat a meal		8	8
56. find	8		8
57. likeable (to be)	8		8
58. pay	8		8
59. remove		8	8
60. sit		8	8
61. think	2	6	8
62. walk		8	8
63. yell	2	6	8
64. cook	7		7
65. give a present	7		7
66. hear/listen to		7	7
67. learn	7		7
68. love		7	7
69. sleep		7	7
70. wait		7	7
71. work (funcionar)		7	7
72. call		6	6
73. kick out		6	6
74. pass		6	6
75. save	6		6
76. tell	1	5	6
77. turn out		6	6
78. know (conocer)		6	6

79. close		5	5
80. listen to (hacer caso)	5		5
81. pray	4	1	5
82. prepare	5		5
83. put in		5	5
84. put together (things)		5	5
85. spoil	1	4	5
86. stand		5	5
87. accept	4		4
88. be afraid		4	4
89. be late	4		4
90. dress		4	4
91. exonerate (academic)	4		4
92. fight		4	4
93. flood (inundar)		4	4
94. get close		4	4
95. get up		4	4
96. grow		4	4
97. make someone go through		4	4
98. pass (academic)		4	4
99. pass by (the bus)		4	4
100. pick up	2	2	4
101. rest	1	3	4
102. speak (conversar)		4	4
103. take an exam	4		4
104. treat	4		4
105. turn (age)	2	2	4
106. abandon		3	3
107. ask		3	3
108. be brothers/sisters	3		3
109. break		3	3
110. clean		3	3
111. date (romantic)	3		3
112. descend	2	1	3
113. fail	3		3
114. feel		3	3
115. get dirty		3	3
116. have children		3	3
117. have dinner	3		3
118. hurt (acertar)		3	3
119. invite	3		3
120. leave		3	3
121. lie down		3	3

122.	look		3	3
123.	make sth pretty		3	3
124.	mean		3	3
125.	meet	3		3
126.	pay attention	2	1	3
127.	practice	3		3
128.	reach	3		3
129.	respect	3		3
130.	run	3		3
131.	show	1	2	3
132.	steal		3	3
133.	throw away		3	3
134.	turn		3	3
135.	understand	3		3
136.	use		3	3
137.	apologize	2		2
138.	apply	2		2
139.	be in a course (cursar)	2		2
140.	be left over (sobrar)		2	2
141.	be quiet		2	2
142.	become desgraciado	2		2
143.	become true (cumplirse la profesía)	2		2
144.	been long		2	2
145.	believe	2		2
146.	breath	2		2
147.	build a house		2	2
148.	calculate/think	2		2
149.	change	2		2
150.	consider	2		2
151.	dance		2	2
152.	deliver (entregar)	2		2
153.	descomponer		2	2
154.	expose		2	2
155.	fall		2	2
156.	feed		2	2
157.	feel sorry		2	2
158.	file (desfilar)	2		2
159.	fill (cargar)	2		2
160.	fix (arreglar)	2		2
161.	get (conseguir)	2		2
162.	get (the scholarship)		2	2
163.	get along	2		2

164.	get educated (formarse)	2		2
165.	get ruined (se fundió)	2		2
166.	get together	2		2
167.	go for a walk (pasear)	2		2
168.	graduate	2		2
169.	greet	2		2
170.	have surgery (operarse)	2		2
171.	introduce	2		2
172.	jog	2		2
173.	join together	2		2
174.	let know	2		2
175.	look for/deserve		2	2
176.	lose	1	1	2
177.	make sth stop	2		2
178.	matter		2	2
179.	miss	2		2
180.	move		2	2
181.	practice (a profession)	2		2
182.	pursue	2		2
183.	rain		2	2
184.	raise		2	2
185.	recognize	1	1	2
186.	remember/celebrate	2		2
187.	request	1	1	2
188.	sacar (dinero)		2	2
189.	separate		2	2
190.	shout		2	2
191.	succeed		2	2
192.	take a shower		2	2
193.	take care		2	2
194.	taste good		2	2
195.	touch		2	2
196.	tremble		2	2
197.	try	1	1	2
198.	visit	2		2
199.	waste		2	2
200.	write	2		2
201.	act	1		1
202.	adore/worship	1		1
203.	advance		1	1
204.	alcanzar		1	1
205.	appreciate	1		1
206.	attract attention	1		1

207.	baptize	1	1
208.	be dedicated	1	1
209.	be full of		1 1
210.	be happy (cause to)		1 1
211.	be in charge (encargarse)	1	1
212.	be interested	1	1
213.	be possible		1 1
214.	be someone's turn (me toca)	1	1
215.	be thankful	1	1
216.	be wrong	1	1
217.	become an adult		1 1
218.	behave	1	1
219.	belong	1	1
220.	bite		1 1
221.	blend	1	1
222.	blow		1 1
223.	boil		1 1
224.	borrow		1 1
225.	break (their word)		1 1
226.	bring down		1 1
227.	bury		1 1
228.	clean (arreglar)	1	1
229.	clip the nails		1 1
230.	collect (cobrar)	1	1
231.	condemn	1	1
232.	congratulate (felicitar)	1	1
233.	continue	1	1
234.	cough		1 1
235.	count	1	1
236.	cover (cubrir, ser cómplice)	1	1
237.	cover (tapar)	1	1
238.	criticize	1	1
239.	cut		1 1
240.	demand	1	1
241.	deserve	1	1
242.	draw	1	1
243.	drink mate		1 1
244.	drink terere		1 1
245.	dry		1 1
246.	earn	1	1
247.	envy	1	1
248.	escape	1	1
249.	experience		1 1

250.	experience (experimentar)	1	1
251.	fail (a test, class or course)	1	1
252.	fail (fracasar)	1	1
253.	fish	1	1
254.	fly		1 1
255.	forget		1 1
256.	get (alcanzar)	1	1
257.	get angry		1 1
258.	get drunk		1 1
259.	get together (congregarse)	1	1
260.	get used to	1	1
261.	give milk		1 1
262.	guard/protect	1	1
263.	harm (perjudicar)	1	1
264.	have a girl		1 1
265.	have an accident	1	1
266.	have breakfast		1 1
267.	have time	1	1
268.	heal		1 1
269.	hide		1 1
270.	hold		1 1
271.	include	1	1
272.	inflate		1 1
273.	insult		1 1
274.	integrate (integrar)	1	1
275.	it is cool		1 1
276.	it will be 4 years	1	1
277.	it will be 5 years	1	1
278.	laugh		1 1
279.	let	1	1
280.	lick		1 1
281.	lie		1 1
282.	look upwards		1 1
283.	misbehave	1	1
284.	obey (cumplir la orden)	1	1
285.	offer	1	1
286.	order		1 1
287.	overflow (rebosar)	1	1
288.	party	1	1
289.	play (an instrument)		1 1
290.	please	1	1
291.	post (pegar)	1	1
292.	praise	1	1

293.	pull		1	1
294.	put oil on sth	1		1
295.	quit		1	1
296.	record	1		1
297.	recover	1		1
298.	reject/repudiate	1		1
299.	rent	1		1
300.	resist (aguantar)	1		1
301.	result		1	1
302.	rub (frotar)		1	1
303.	search for carefully (rebuscarse)		1	1
304.	see-be happy		1	1
305.	see/look		1	1
306.	send		1	1
307.	serve (god)	1		1
308.	sew	1		1
309.	share	1		1
310.	show up	1		1
311.	soak (mojar)		1	1
312.	solve (resolver)	1		1
313.	stand out	1		1
314.	start (desatar la revolución)	1		1
315.	suffer	1		1
316.	survive	1		1
317.	sweep		1	1
318.	take (llevar) age difference		1	1
319.	take advantage	1		1
320.	taste	1		1
321.	have mercy		1	1
322.	thank	1		1
323.	throw rocks	1		1
324.	turn around		1	1
325.	turn one year old	1		1
326.	use	1		1
327.	wait (expect)		1	1
328.	wash hands		1	1
329.	wear		1	1
330.	worry	1		1
331.	worship	1		1

## Appendix 2

*Corpus of Guaraní speakers, from youngest to oldest*

<b>Name (Gender)</b>	<b>Age</b>	<b>Last school year completed</b>
Aníbal (M)	18	11 <sup>th</sup>
Paulina (F)	24	2 <sup>nd</sup> year of college
Hugo (M)	30	1 <sup>st</sup> year of college
Lila (F)	31	7 <sup>th</sup>
Melisa (F)	32	9 <sup>th</sup>
Augusto (M)	32	8 <sup>th</sup>
Rodolfo (M)	32	some high school
Rita (F)	33	9 <sup>th</sup>
Raquel (F)	33	3 <sup>rd</sup>
Camilo (M)	33	8 <sup>th</sup>
Sara (F)	36	some elementary
Basilía (F)	37	3 <sup>rd</sup>
Mirta (F)	37	6 <sup>th</sup>
Jaime (M)	42	Some elementary
Leticia (F)	43	6 <sup>th</sup>
Susana (F)	45	4 <sup>th</sup>
Eusebio (M)	47	6 <sup>th</sup>
Fabio (M)	47	9 <sup>th</sup>
Pablino (M)	50	6 <sup>th</sup>
Ángeles (F)	54	some elementary
Ramiro (M)	54	some elementary
Mónica (F)	55	5 <sup>th</sup>
Marta (F)	56	6 <sup>th</sup>
Mariana (F)	59	6 <sup>th</sup>
Petunia (F)	60	3 months
Graciana (F)	60	4 <sup>th</sup>
Antonio (M)	60	9 <sup>th</sup>
Donato (M)	62	6 <sup>th</sup>
Leoncia (F)	68	3 <sup>rd</sup>
Rumilda (F)	69	some elementary
Pipo (M)	69	9 <sup>th</sup>
Felipe (M)	72	7 <sup>th</sup>
Ñata (F)	74	2 <sup>nd</sup>
Julián (M)	75	2 <sup>nd</sup>
Carlota (F)	79	2 <sup>nd</sup>