# Lexical Strata in Loanword Phonology: Spanish Loans in Guaraní

Justin Pinta

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Linguistics

Chapel Hill
2013

Approved by:

Jennifer L. Smith

Elliott Moreton

Bruno Estigarribia

# Abstract

JUSTIN PINTA: Lexical Strata in Loanword Phonology: Spanish Loans in Guaraní
(Under the direction of Jennifer L. Smith)

An analysis of a corpus of Spanish loanwords in Paraguayan Guaraní shows the stratified structure of the Guaraní lexicon evidenced by varying phonological repair strategies in the loans. Itô and Mester (1999 and earlier work) show that a language with a synchronically relevant stratified lexicon displays impossible nativization effects. The phonology and morphology of Guaraní provide evidence for the synchronic relevance of the stratification, and as expected the corpus shows specific nativization strategies which are unattested. A nonce-word experiment with native Guaraní speakers shows that in some cases, but not all, impossible nativizations are strongly avoided by native speakers. The Itô and Mester (1999) model handles the impossible nativizations within Optimality Theory through their proposed ranking consistency of faithfulness constraints across strata. Variable repair strategies of certain Spanish phonological structures in Guaraní in addition to the results of the experiment present a theoretical problem for ranking consistency.

To native speakers of the beautiful Guaraní language on which this thesis is based. Your consistent kindness and eagerness to lend a hand whenever possible not only facilitated this work but made its completion a thoroughly enjoyable and enriching experience.

# Acknowledgments

I am indebted to my brilliant committee for their dedication to this thesis and their patience with it as it evolved several times from its conception. Without their help it would be a different work and not for the better. Their subtle methods of teaching have taught me as much about what it is to be a good academic mentor as about the theoretical matters we were discussing. Thanks to Elliott for his help with the experiment; I regard myself as quite fortunate to have had such a wonderful and analytic mind with which to brainstorm experimental designs, and I learned a great deal by merely being around him and watching him reason. Thanks to Bruno not only for his key role as the sole member of the committee with both field experience and direct knowledge of Guaraní (not to mention Spanish!), but for his suggestions which much improved the clarity of several of the chapters. And special thanks to the wonderful chair of my committee Jen. I initially picked her as chair because I knew with her at the head of the project it would be much the better in the long run, and I was not mistaken. Her bottomless knowledge of all things phonology was a truly indispensable resource for me, and I was greatly honored and privileged to work with her.

Many thanks to the wonderful Guaraní-speaking informants who made much of this thesis possible by participating in the experiment. Thanks to Teresa Chapa at Davis Library at UNC-Chapel Hill for graciously going to such great lengths to get an obscure Paraguayan missal which ended up being an incredible resource. Special thanks as well to Chris Wiesen at UNC's Odum Institute whose assistance with the statistical portion of the experiment chapter was of tremendous help to me.

Thanks to my colleagues and friends Matt Fuller, Lúcia Fischer and Katherine Shaw for providing me with sound advice, wonderful ideas, and sometimes just an ear in which to vent when the going got rough. The three of them, the UNC Guaraní Working Group, my officemates Aziz Jabar and Mary Kohn, and the rest of the graduate students in the linguistics department at UNC-Chapel Hill made for a wonderfully supportive group which I couldn't imagine having gone through the program without. Special thanks as well to my good friend Jon Harter for showing me the light with LaTeX and for being available to deal with my questions at quite literally any hour of the day.

I also must thank my wonderful family and friends back home in Illinois who in spite of the geographic

distance showed genuine interest in my work and were quick to lend support at all times.

Last but not least, my wonderful partner Lizzie deserves special thanks for being there for me from the very beginning of this project until the very end. Her support was invaluable.

# Table of Contents

# Chapter 1

# Introduction

The contact between Spanish and Guaraní[1] and the resulting linguistic borrowing has endowed Guaraní with a wealth of loanwords. These loans are notable not just for their quantity but for their variation with respect to phonological characteristics. Not all loans show homogeneous adaptation strategies and the differing repairs made allow for clear insight into the structure of the Guaraní lexicon and consequently the grammar as a whole.

Examination of the repair strategies reveals a pattern of stratification of the Guaraní lexicon wherein can be seen that lexical strata are attested on the basis of the phonological characteristics of the loans. This stratification is most clearly seen in the core-periphery structure of the Guaraní lexicon. That is, strata are not entirely separate groupings of words but rather stacked one upon another such that they form a set-inclusion hierarchy. This hierarchy hints at the constraints on adaptation processes in the grammar and says much about the overall organizational principles observable in the lexicon.

One of the central claims of this thesis is that the lexical strata attested in Guaraní are synchronically relevant and not a mere relic of the history of the language. Evidence for this claim comes from the phonology and morphology of Guaraní, and it will argued that any account of the Guaraní lexicon must treat its strata as relevant in the grammar of modern-day native speakers.

Itô and Mester (1999 and earlier work) show through an analysis of the Japanese lexicon that synchronically relevant lexical strata in a language lead to impossible nativization effects where certain combinations of phonological repairs go unattested in a language's loanwords. An analysis of an original corpus of Spanish loans in Guaraní shows strong evidence for the existence of impossible nativizations in Guaraní; repair strategies, or in some cases the lack thereof, of Spanish phonological structures show systematic tendencies and are not combined at random.

Further evidence for the sychronic relevance of lexical strata in the Guaraní lexicon, in addition to

[1] Although the term "Guaraní" is used variously as a reference to a subgroup of the Tupí-Guaraní language family, a dialect chain within that subgroup, and the specific language Paraguayan Guaraní, in this thesis "Guaraní" is used exclusively in reference to the single language of Paraguayan Guaraní.

evidence supporting the existence of impossible nativizations, comes from an original nonce-word experiment in which Guaraní native speakers were asked for their judgments regarding possible nativizations of Spanish nonce forms. The experimental results, in which speakers showed an aversion to impossible nativizations in some cases but not all, simultaneously bolster claims made about stratal synchronic relevance and give more direct insight into the existence and nature of impossible nativizations in Guaraní than is attainable through a corpus study alone.

The theoretical account of impossible nativization effects as laid out in Itô and Mester (1999) handles these effects within Optimality Theory (Prince and Smolensky, 1993) by imposing limits on the cross-stratal rankings of faithfulness constraints. These limits, formulated in a condition named by the authors "ranking consistency" come into conflict with some of the facts presented here regarding Guaraní. Guaraní, seemingly unlike Japanese, shows variable repairs within some of its lexical strata and this variation presents serious theoretical problems for the model given its condition of ranking consistency. The results of the nonce experiment also call into question ranking consistency given the speakers' willingness to accept certain impossible nativizations which are predicted by the model (through its implementation of ranking consistency) to be categorically avoided.

This thesis is outlined as follows. Chapter 2 provides a discussion of the theoretical background of the subsequent analysis of the Guaraní lexicon, and serves to outline the central points of the analysis of the Japanese lexicon in Itô and Mester (1999) on which much of this thesis is based. Following, chapter 3 lays the groundwork for the Guaraní analysis by providing general descriptions of both Guaraní and Spanish phonology, as well as a systematic account of how Guaraní treats Spanish lexical stress and a few segmental structures in the process of loan nativization. Chapter 4 expands on the analysis of nativization by looking at the varied processes affecting loans with codas and complex onsets in the original Spanish forms, in addition to analyzing the Guaraní syllabification of Spanish loans. Chapter 5 provides a systematic optimality-theoretic analysis of the Guaraní grammar in which the stratification of the lexicon and the core-periphery structure which characterize it are discussed in detail, in addition to an in-depth analysis of the synchronic Guaraní grammar capable of handling the varied loan adaptation patterns attested in the lexicon. Chapter 6 outlines the details of an experiment carried out with the assistance of native Guaraní speakers, the purpose of which was to explicitly test speaker sensitivity to predicted possible and impossible nativizations in a forced choice nonce nativization test. Finally, chapter 7 concludes the thesis with a summary and further discussion of the facts presented regarding the structure of the Guaraní lexicon.

# Chapter 2

# Theoretical Background

## 2.1 Introduction

Itô and Mester (1999), following up on earlier work in Itô and Mester (1995a) and Itô and Mester (1995b), presents an analysis of the structure of the Japanese phonological lexicon outlining the organization of the lexicon into various lexical strata. Data from Japanese is shown as evidence of the stratified nature of the lexicon and the stratification can be seen through loanwords which Japanese has acquired from other languages. The lexicon is crucially shown to exhibit a core-periphery structure wherein the strata composing it are organized in a systematic and overlapping way as opposed to merely independent groupings of words. Their analysis aims to show how this core-periphery structure can be modeled within Optimality Theory to account for the fact that certain nativization strategies in loanword adaptation are unattested. Their investigation of the Japanese lexicon on this basis followed by a discussion of its relevance for both loanword nativization and modeling phonological lexica provides clear insight into crucially important issues in phonology.

The notion that the grammars of natural languages show stratification to one degree or another was not in and of itself original to this article, and previous work outlining the theoretical need to handle patterns of lexical stratification had been ongoing in the field of phonology for decades (see Chomsky and Halle (1968), Kiparsky (1968), Saciuk (1969) and Itô and Mester (1995a), among others). However, the discussion of this topic within the realm of loanword phonology, as well as the optimality-theoretic analysis provided (particularly the predictions made by the analysis) made this an important contribution to the literature. The model of lexical stratification provided has its roots in evidence from Japanese. The data provided from Japanese is both interesting and elucidating, yet in order for more broad generalizations to be made regarding the applicability of the model to natural languages in general (i.e. to cross-linguistic universals), its application to other languages follows as a natural means of bolstering or critiquing the theory.

This article and its modeling of lexical strata in the Japanese phonological lexicon, namely the modeling of its core-periphery structure, serves as the basis on which much of the analysis presented here on Spanish loans in the Guaraní lexicon is constructed. As a natural predecessor to the analysis of the Guaraní lexicon, this chapter serves to elucidate the model presented in Itô and Mester (1999) and summarize the core claims and predictions made therein. After an overview of the model and its most central ideas a discussion about the link between it and the phonological lexicon of Guaraní will conclude the chapter.

## 2.2 Structure of the Japanese lexicon

In their article, the authors begin their analysis with an overview of the structure of the Japanese lexicon. They outline four key strata[2] in Japanese which exhibit distinct phonological behavior. The first stratum is the native, or "Yamato" stratum which consists of the vocabulary representing the core of the Japanese lexicon. The next stratum, the "Sino-Japanese" stratum, is comprised of vocabulary which Japanese inherited from Chinese languages and is used in much of the technical vocabulary of the language. A comparison can be drawn between these two strata and the Germanic and Latinate sources of vocabulary in modern English, the Germanic representing native English vocabulary and the Latinate being used for much of the technical or academic vocabulary of the language.

Following the Yamato and Sino-Japanese strata comes the "foreign" stratum which includes words borrowed more recently than those in the Sino-Japanese stratum and having their origin in various languages (including English). Lastly is the "onomatopoetic/mimetic" stratum which includes words utilized for their onomatopoetic nature, and which are described by the authors as being of more importance to the Japanese phonological system than comparable words in English (e.g. *buzz*, *oink*, *tick-tock*, etc.). In (1) can be seen some examples of vocabulary coming from the various strata (Itô and Mester, 1999, 63).

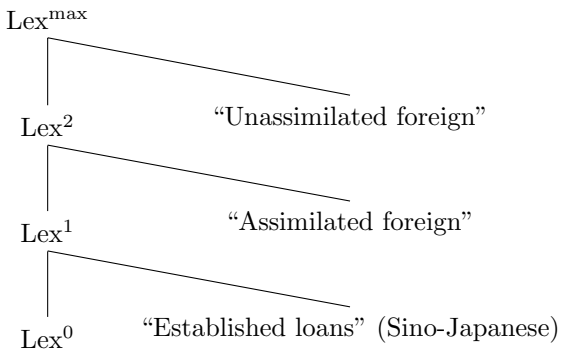
[2]Although the authors' way of describing the Japanese strata is not universally agreed upon, their stance is in accordance with the work of other scholars of Japanese phonology, see Martin (1952), McCawley (1968), Vance (1987), Shibatani (1990) and Kubozono (1995), among others.

(1)

| Stratum | Japanese | English gloss |
|---|---|---|
| Native (Yamato) | kotoba | 'word', 'language' |
| | oto | 'sound' |
| | kuruma | 'wheel', 'car' |
| Sino-Japanese | gen-go-gaku | 'linguistics' (speak-word-study) |
| | on-in-ron | 'phonology' (sound-rhyme-theory) |
| | den-wa | 'telephone' (electric-speak) |
| Foreign | san-tora | 'sound track' |
| | terefon-kādo | 'telephone card' |
| | pato-kā | 'patrol car', 'police car' |
| Onomatopoetic/Mimetic | kori-kori | 'crisply' |
| | sui-sui | 'lightly and quietly' |
| | mota-mota | 'slowly', 'inefficiently' |

Having described the differing lexical strata, the discussion appropriately turns to the relationship amongst the strata and what this tells us about the overall structure of the phonological grammar. Crucial regarding this relationship is the fact that stratification cannot be described as a mere partitioning of strata into parallel and non-overlapping sets but rather should be thought of as strata which overlap substantially and are characterized by their core-periphery structure. That is, the stratal hierarchy is organized on the notion of set inclusion, which can be depicted as in (2) (adapted from Itô and Mester (1999, 65)).

(2)




Of note here is the absence of the mimetic stratum in the remainder of the analysis and the breaking down of the "foreign" stratum into "assimilated foreign" and "unassimilated foreign". This is done due to the fact that foreign loans do not all enter at the same stage of assimilation, and the division between loans which are assimilated and those which are not (or at least much less so) is clearly visible on the

basis of the phonological adaptations, and as such they merit being placed into separate strata.

In (2) we see that $\text{Lex}^0$ represents the native (Yamato) stratum and is at the core of the hierarchy. All subsequent strata increasing in distance from the core stratum can be described in terms of set complementation such that "established loans" describes the set $\text{Lex}^1$ - $\text{Lex}^0$. In optimality-theoretic terms, lexical items in the core stratum ($\text{Lex}^0$) are those maximally subject to the wellformedness constraints which define the central area of the lexicon. That is, lexical items in $\text{Lex}^0$ are those exhibiting the most native (Yamato) phonological characteristics. Moving outwards from the core, each stratum violates increasingly more constraints until at the periphery only the most central constraints in the grammar which determine the most fundamental characteristics of the language are still exerting influence over lexical items.

Relationships such as these are in turn rooted in implicational relationships between lexical items across strata and the constraints they are subject to. In general terms, lexical items subject to a constraint $A$ are necessarily subject to constraint $B$, while those subject to $B$ are not necessarily subject to $A$. In this example constraints $A$ and $B$ could hold in, say the core stratum ($\text{Lex}^0$), while in the next stratum up the hierarchy, $\text{Lex}^1$, only constraint $B$ can exert influence over the forms. In this way any lexical item which is subject to the demands of constraint $A$ must also be subject to those of $B$ given that only forms in $\text{Lex}^0$ are subject to $A$ and they are by requirement also subject to $B$. Items in $\text{Lex}^1$ however are subject to $B$ but no longer to $A$ given their position further from the core. Such a relationship continues moving outward from the core such that the whole hierarchy is built upon this nested set-inclusion structure. It is in this sense that the lexicon is described as having a core-periphery structure, and this crucial attribute of the structure of the lexicon is the basis on which the overall analysis is constructed.

## 2.3 Optimality-theoretic analysis

### 2.3.1 The role of markedness constraints

The constraints at work in the grammar as presented by the authors are seen in (3).

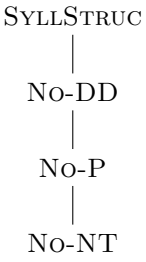
(3) **Constraints and definitions** - adapted from Itô and Mester (1999, 66)

- SyllStruc - Basic syllable structure constraints of Japanese (e.g. *Complex and CodaCond among others)
- NoVoicedGem (No-DD) - No voiced obstruent geminates (e.g. $*bb$, $*dd$, $*gg$, etc.)
- NoVoicelessLab (No-P) - No singleton-$p$: a constraint against nongeminate $p$

- NoNas⌒Voiceless (No-NT) - Post-nasal obstruents must be voiced (e.g. $*nt$, $*mp$, $*ŋk$, etc.)

These four markedness constraints form the basis of the phonological grammar in the sense that they are responsible for the phonological characteristics which define the most native Japanese lexical items, and it is through their interaction with lexical items that the stratified lexicon can be seen. As a result of the willingness of loans in a given stratum to adhere to (or ignore) the demands of the markedness constraints presented here, their rankings with respect to one another in the grammar can be obtained. Seen in (4) are their rankings in a Hasse diagram.

(4)




This ranking was arrived at through a systematic analysis of the loans and their violations of these constraints. For example, loans in the native stratum are subject to and obey the demands of all four markedness constraints presented here, as these are constraints which define the native Japanese phonological grammar. Moving up from the native stratum, the loans in the Sino-Japanese stratum obey all markedness constraints with the exception of No-NT (i.e. loans in this stratum are permissive of post-nasal voiceless obstruents). The next stratum moving away from the core is the assimilated foreign stratum, and here we see loans which neither obey No-NT nor No-P yet which are still subject to the demands of the other two wellformedness constraints, No-DD and SyllStruc. Lastly the most peripheral stratum, the unassimilated foreign, is only subject to the demands of SyllStruc and ignores all others.[3]

It is through this behavior of lexical items and the adherence to the markedness constraints that we can infer the ranking in (4). The willingness of various lexical items to disregard the demands of some constraints but not others gives crucial information about the rankings of the constraints with respect to one another. Depicted another way, in (5) we clearly see the relationship between the strata and the constraints modeled showing which constraints exert control over which strata (taken from Itô and Mester (1999, 69)).

[3]For the purpose of this summary, actual lexical items in Japanese justifying these claims about loans in certain strata obeying/disobeying markedness constraints are not included here but can be found in Itô and Mester (1999, 66-68).

(5)

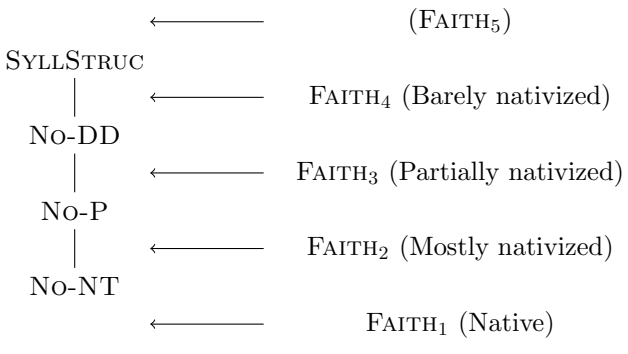
| | SyllStruc | No-DD | No-P | No-NT |
|---|---|---|---|---|
| a. Yamato | ✓ | ✓ | ✓ | ✓ |
| b. Sino-Japanese | ✓ | ✓ | ✓ | violated |
| c. Assimilated foreign | ✓ | ✓ | violated | violated |
| d. Unassimilated foreign | ✓ | violated | violated | violated |

This nesting of constraints clearly shows how, for example, every lexical item subject to No-P is also subject to No-DD while the reverse is not true. This is the central characteristic of this model; it is through this characteristic that important information regarding the structure of the Japanese lexicon is obtained.

### 2.3.2 The role of faithfulness constraints

Having established the markedness hierarchy at work in the grammar, the various behavior of the differing strata can be handled through the role of indexed faithfulness constraints. To clearly demonstrate the role of faithfulness in the system, all relevant faithfulness constraints are consolidated to the single constraint Faith, which incurs violations in any case where faithfulness to the original form is preferred over repair mandated by the markedness constraints discussed above. The constraint Faith has various indexed versions (e.g. $\text{Faith}_1$, $\text{Faith}_2$, etc.) which correspond to specific strata such that $\text{Faith}_1$ = Faith/Yamato, $\text{Faith}_2$ = Faith/Sino-Japanese, $\text{Faith}_3$ = Faith/Assimilated foreign, and $\text{Faith}_4$ = Faith/Unassimilated foreign. Including these faithfulness constraints in the wellformedness hierarchy produces a new hierarchy which is modeled in (6) where each indexed Faith constraint is included in the appropriate place to account for the patterns seen above in (5).

(6)




The position of the fifth indexed version of Faith is a hypothetical one where loans which show no tolerance to the markedness hierarchy whatsoever could be included.[4]

[4]As will be shown in subsequent chapters, this hypothetical position of an indexed Faith constraint is attested in the Guaraní grammar.

To better illustrate how these indexed FAITH constraints would work in the grammar, a series of tableaus (adapted from Itô and Mester (1999, 73-76)) can be seen in (7)-(11) modeling examples from the Japanese lexicon. Tableau inputs are indexed according to which stratum they belong to: UF (unassimilated foreign), AF (assimilated foreign), SJ (Sino-Japanese) or Y (Yamato). This indexing, shown in parentheses next to the input itself (e.g. /beddo/ (UF)), crucially decides which of the FAITH constraints is able to exert influence on the outcome.

#### 2.3.2.1 Faith$_4$ (Unassimilated foreign)

In the unassimilated foreign stratum the only constraint exerting influence over the forms is SYLLSTRUC and, as seen in (6), FAITH$_4$ appears ranked immediately below SYLLSTRUC such that all other markedness constraints are powerless with regard to faithfulness in this stratum.

Seen in (7) is the nativization for the loan coming from the English *bed*. In this tableau and subsequent tableaus in the section the FAITH constraints are abbreviated as F such that F$_4$=FAITH$_4$, F$_3$=FAITH$_3$, etc. The input candidate is given as the already adapted /beddo/ for the purpose of explicitly comparing it against other adaptations. This convention will be followed in the subsequent tableaus in this section as well (the authors' use of already-adapted input forms is inconsequential for the sake of demonstration here, but further discussion regarding the matter can be found in Itô and Mester (1999, 73-74)).

(7)

| /beddo/ (UF) | SYLLSTRUC | F$_4$ | NO-DD | F$_3$ | NO-P | F$_2$ | NO-NT | F$_1$ |
|---|---|---|---|---|---|---|---|---|
| a. ☞ beddo | | | * | | | | | |
| b. betto | | *! | | | | | | |

Here the winning candidate violates NO-DD yet still is picked as optimal due to candidate (b) violating the higher ranked FAITH$_4$ through the rendering of the geminate stop as voiceless instead of voiced.

#### 2.3.2.2 Faith$_3$ (Assimilated foreign)

Seen in (8) is the analysis of the nativization of the adapted /pabbu/, the adaptation resulting from English *pub*.

(8)

| /pabbu/ (AF) | SyllStruc | $F_4$ | No-DD | $F_3$ | No-P | $F_2$ | No-NT | $F_1$ |
|---|---|---|---|---|---|---|---|---|
| a. pabbu | | | *! | | * | | | |
| b. ☞ pabu | | | | * | * | | | |
| c. habu | | | | **! | | | | |
| d. habbu | | | *! | * | | | | |

Given the ranking of No-DD above $\text{Faith}_3$, the candidates (a) and (d) are eliminated due to their retention of the voiced geminate. Here the winning nativization avoids elimination by avoiding the geminate but incurs one violation of $\text{Faith}_3$ by reducing the geminate /bb/ to singleton [b]. With respect to candidate (c), it fares similarly to the winning (b), yet is knocked out by changing the word-initial consonant and thereby incurring an extra violation of $\text{Faith}_3$.

#### 2.3.2.3 Faith$_2$ (Sino-Japanese)

For the $\text{Faith}_2$ stratum we turn to the Sino-Japanese morpheme /paɴ/ meaning *group*.

(9)

| /paɴ/ (SJ) | SyllStruc | $F_4$ | No-DD | $F_3$ | No-P | $F_2$ | No-NT | $F_1$ |
|---|---|---|---|---|---|---|---|---|
| a. paɴ | | | | | *! | | | |
| b. ☞ haɴ | | | | | | * | | |

Here the losing (and faithful) candidate violates No-P which is crucially more highly ranked than the faithfulness constraint $\text{Faith}_2$. This eliminates candidate (a) leaving the winning candidate (b) which violates $\text{Faith}_2$ once by changing the word-initial consonant.

As an informative comparison, the form /paɴ/ also exists in the assimilated foreign stratum as a nativization of the Portuguese *pão*. As expected, the winning form changes given that it is subject to the more highly ranked version of Faith characteristic of this stratum, as seen in (10).

(10)

| /paɴ/ (AF) | SyllStruc | $F_4$ | No-DD | $F_3$ | No-P | $F_2$ | No-NT | $F_1$ |
|---|---|---|---|---|---|---|---|---|
| a. ☞ paɴ | | | | | * | | | |
| b. haɴ | | | | *! | | | | |

#### 2.3.2.4 Faith$_1$ (Yamato)

Lastly, the native stratum is subject to the demands of all four markedness constraints and $\text{Faith}_1$ is the most lowly ranked of all constraints in the grammar. This can be seen in the native Yamato [šiɴde] (gerund form of *die*) coming from the underlying /šiɴ-te/.

(11)

| /šin-te/ (Y) | SyllStruc | $F_4$ | No-DD | $F_2$ | No-P | $F_2$ | No-NT | $F_1$ |
|---|---|---|---|---|---|---|---|---|
| a. [šinte] | | | | | | | *! | |
| b. ☞ [šinde] | | | | | | | | * |

## 2.4 Possible and impossible nativizations

Central to this model of the Japanese lexicon is the way in which it handles the notion of impossible nativizations. Through its modeling of the crucial characteristic of the structure of the Japanese lexicon, its core-periphery structure, the model limits the ways in which native and non-native phonological characteristics can mix to arrive at nativizations. The optimality-theoretic analysis which constitutes the heart of the model is notable for capturing the fact that Japanese avoids specific mixing of phonological properties when nativizing loans.

To illustrate this more concretely, we turn to the same example put forth by the authors: the case of palatalization in plosives and fricatives in Japanese (data, constraints and tableaus adapted from Itô and Mester (1999, 77-80)). The wellformedness constraints *SI (where *S* represents fricatives) and *TI (where *T* represents plosives) result in the nativization of segments such as *si* and *ti* as *ši* and *či*, respectively. Interestingly however, in recent loans *SI is still enforced (e.g. English *sea* → Japanese *šii*, **sii*) while *TI is not (e.g. English *party* → Japanese *paatii*, **paačii*). This observation that the palatalization of fricatives is more important than that of plosives leads inevitably to the hierarchy seen in (12).

(12)

*SI
|
*TI

Taking for instance a loan coming from the English *city*, we see both a nonpalatalized fricative and plosive in the loan's original form. Given the hierarchy in (12), we are left with three possible nativizations depending on the stratum to which the loan pertains (used here are generic *X*, *Y* and *Z* for the purpose of demonstration). Tableaus illustrating the grammar producing the three nativizations are shown in (13)-(15).

(13)

| /siti/ (*stratum X*) | FAITH$_X$ | *SI | FAITH$_Y$ | *TI | FAITH$_Z$ |
|---|---|---|---|---|---|
| a. šiči | *!* | | | | |
| b. šiti | *! | | | * | |
| c. ☞ siti | | * | | * | |
| d. siči | *! | * | | | |

In (13) we see the fully faithful form emerge as the winner given the highly ranked position of the relevant version of FAITH.

(14)

| /siti/ (*stratum Y*) | FAITH$_X$ | *SI | FAITH$_Y$ | *TI | FAITH$_Z$ |
|---|---|---|---|---|---|
| a. šiči | | | **! | | |
| b. ☞ šiti | | | * | * | |
| c. siti | | *! | | * | |
| d. siči | | *! | * | | |

In (14) the partially nativized *šiti* wins given the relevant FAITH constraint outranking *TI but not *SI. This form is attested in the nativization of "Citibank" as *šitibaŋku*.

(15)

| /siti/ (*stratum Z*) | FAITH$_X$ | *SI | FAITH$_Y$ | *TI | FAITH$_Z$ |
|---|---|---|---|---|---|
| a. ☞ šiči | | | | | ** |
| b. šiti | | | | *! | * |
| c. siti | | *! | | * | |
| d. siči | | *! | | | * |

Lastly in (15) we see both markedness constraints exerting influence over the relevant FAITH constraint and as a result both the fricative and the plosive are palatalized. This form is also attested in the nativization of the brand "Citizen" as *šičizuɴ*.

As exemplified by the tableaus in (13)-(15), candidate (d) is a perpetual loser. Given the fixed markedness hierarchy *SI >> *TI, there is no possible ranking of any version of FAITH which will pick the form *siči* as optimal. Irrespective of the stratum to which they pertain, forms which palatalize a plosive must also palatalize the fricative (assuming there is one) and no scenario in which a plosive is palatalized without simultaneous palatalization of a fricative is possible. As such the model predicts *siči* as an impossible nativization of the input /siti/. This consequence of the OT analysis captures the implications of the core-periphery model and is of crucial importance to the theory.

## 2.5 Ranking consistency condition

The consolidating of all faithfulness constraints into the block constraint FAITH is a convenient and clear way of illustrating the role of faithfulness in the system; however, as recognized by the authors themselves, this is a problematic simplification in the sense that the unpacking of FAITH, necessary given the need to explicitly account for individual faithfulness constraints, complicates the model and raises issues which necessitate addressing. The complications arising from this threaten to deny the model of its central prediction of impossible nativizations by allowing individual faithfulness constraints to be freely ranked. Free ranking of faithfulness constraints allows for overall rankings which would allow for unattested nativizations, thereby denying the model its prediction.

Turning to the example of the loan *pub* in which the input form is /pabbu/ (seen earlier in (8)), we can now unpack the block FAITH constraint into the individual faithfulness constraints IDENT-PLACE, prohibiting the changing of place features between corresponding segments, and IDENT-$\mu$, prohibiting the changing of moraic values between corresponding segments (constraints and figures adapted from Itô and Mester (1999, 81-84)). Given the fixed markedness hierarchy, the possible nativizations *pabu*, *pabbu* and *habu* can still be easily handled by the model, as shown in the tableaus (16)-(18).

(16)

| /pabbu/ | NO-DD | IDENT-PLACE | IDENT-$\mu$ | NO-P |
|---|---|---|---|---|
| a. pabbu | *! | | | * |
| b. ☞ pabu | | | * | * |
| c. habu | | *! | * | |
| d. habbu | *! | * | | |

(17)

| /pabbu/ | IDENT-PLACE | IDENT-$\mu$ | NO-DD | NO-P |
|---|---|---|---|---|
| a. ☞ pabbu | | | * | * |
| b. pabu | | *! | | * |
| c. habu | *! | * | | |
| d. habbu | *! | | * | |

(18)

| /pabbu/ | NO-DD | NO-P | IDENT-PLACE | IDENT-$\mu$ |
|---|---|---|---|---|
| a. pabbu | *! | * | | |
| b. pabu | | *! | | * |
| c. ☞ habu | | | * | * |
| d. habbu | *! | | * | |

Problematically however, the unpacking of FAITH into IDENT-$\mu$ and IDENT-PLACE now also potentially allows for the impossible nativization *habbu* to be chosen. This form is predicted as an impossible nativization by the model in the same way as the form *siči* was in the previous section; given that NO-DD is ranked higher that NO-P in the markedness hierarchy, no consolidated FAITH constraint can be ranked in such a way as to produce a form which adheres to NO-P but not to NO-DD (e.g. *habbu*). However through the free ranking of IDENT-$\mu$ and IDENT-PLACE, a grammar such as that modeled in (19) can now be produced whereby *habbu* is chosen as optimal.

(19)

| /pabbu/ | IDENT-$\mu$ | NO-DD | NO-P | IDENT-PLACE |
|---|---|---|---|---|
| a. pabbu | | * | *! | |
| b. pabu | *! | | * | |
| c. habu | *! | | | * |
| d. ☞ habbu | | * | | * |

This is problematic in the obvious sense that it robs the model of its central prediction of impossible nativizations, crucial to the model as being its chief empirical strength by theoretically accounting for the lack of certain specific nativization patterns. With the ability to freely rank the two faithfulness constraints in play here, the prediction of *habbu* as an impossible adaptation of *pabbu* is lost. To compensate for this, a condition on the ranking of faithfulness constraints is proposed by the authors and is outlined in (20).

(20)

**Ranking Consistency**, as presented in Itô and Mester (1999, 82):
Let F and G be two types of IO-faithfulness constraints (e.g. IDENT-PLACE and IDENT-$\mu$).
Then the relative rankings of the indexed versions of F and G are the same across all strata: $\forall AB\ (F/_A >> G/_A) \rightarrow (F/_B >> G/_B)$

Returning to the tableaus in (16)-(19), the ranking which separates (16)-(18) from (19) is IDENT-$\mu$ >> IDENT-PLACE. Given that the tableaus which predict possible nativizations can be arrived at through the ranking IDENT-PLACE >> IDENT-$\mu$, the condition of ranking consistency outlined here mandates that the rankings of IDENT-$\mu$ and IDENT-PLACE be the same across all strata, thereby outlawing the problematic ranking in (19) IDENT-$\mu$ >> IDENT-PLACE.

This has the desired effect of not permitting IDENT-$\mu$ to ever be able to exert influence on the candidates without IDENT-PLACE also being able to. As a result of this, forms such as *habbu* once again are predicted to be impossible nativizations due to the fact that they adhere to IDENT-$\mu$ but not IDENT-PLACE.

## 2.6 Relevance of the model to Spanish loans in Guaraní

The description of the structure of the Japanese lexicon in Itô and Mester (1999) serves as the groundwork on which will be built the analysis of the Guaraní lexicon presented in subsequent chapters. The parallels between the facts of the structure of the Japanese lexicon and that of the Guaraní lexicon are many, and the primary characteristics shown here for Japanese also hold for Guaraní.

The central characteristic, the core-periphery structure of the lexicon wherein strata are organized overlapping in such a way that constraints holding at the periphery also hold at the core but not vice versa, is also seen in Guaraní and will be explored in detail in subsequent chapters. Guaraní shows a well-defined markedness hierarchy similar to that of Japanese (seen in (4)) in which the fundamental characteristics of the language's phonology can be seen, and loan adaptations across strata show similar effects where the farther strata from the core violate an increasing number of the constraints in the hierarchy. It is in this similarity that the evidence from Guaraní does in many ways support the model presented in Itô and Mester (1999).

In spite of the similarities between the Japanese and Guaraní lexica which provide easy grounds for comparison of the two languages, there are differences as well. One difference is that while the analysis for Japanese included loans from various languages, the analysis here of Guaraní is solely concerned with the behavior of loans from Spanish. The Guaraní lexicon, while certainly containing loans from other languages than just Spanish, is inundated with Spanish loans in a way that allows for in-depth study to be done of the lexicon while only considering loans from Spanish. Another difference, and a much more important one, is found in the variable adaptation strategies which are attested in Guaraní that will ultimately prove difficult for the model presented here to handle given the restriction of ranking consistency. Further problematic issues which ranking consistency provides for this model are borne out in the experiment conducted on Guaraní native speakers, and this will be discussed in detail in later chapters.

# Chapter 3

# Loan Adaptation: Stress and Segmental Phenomena

## 3.1 Introduction

In this chapter we will look at the details of the native Spanish and Guaraní phonologies as they pertain to the ways in which Guaraní adapts loans from Spanish, in addition to some of the specifics of the Guaraní processes of adaptation. Although Guaraní utilizes an array of adaptation strategies to nativize loans, we are in this chapter only concerned with Guaraní treatment of Spanish non-final lexical stress and several segmental phenomena. These adaptations, constituting the less complicated details of Guaraní adaptation, will set the stage for the following chapter which will look at the more complex adaptations.

The chapter begins with the details of the corpus data and its collection, after which the discussion moves to the native phonologies of Guaraní and Spanish which are crucially important to understanding the patterns in loan adaptation. Following this will be a breakdown of some of the Guaraní adaptation strategies. Beginning with the treatment of Spanish stress, those loans which are adapted to Guaraní stress patterns will lead the discussion followed by those which tolerate Spanish stress even when it conflicts with native Guaraní patterns. Lastly a brief discussion of several segmental phenomena in the adaptations will close the chapter.

## 3.2 Data

All Guaraní loan data used in this analysis and presented hereafter was taken from a corpus of Spanish loan words in Guaraní consisting of 177 loans. The corpus was compiled by the author for the purpose of this research and 13 books, articles and other publications were used as sources for the corpus. The dictionaries/grammars, listed alphabetically by author's name, include Britton (2005), de Assis (2008), de Canese and Alcaraz (1997), Díaz (2006), Lustig (2005), Mayans (1980) and Morínigo (1931); academic works regarding Guaraní include Rendon (2008), Tonhauser and Colijn (2010), Tonhauser et al. (2013) and Velázquez-Castillo (2013). The source Fritz (2004) is a Guaraní-language Catholic missal obtained

from central Paraguay. The final source was the Guaraní version of the online encyclopedia Wikipedia.[5] All loans from this last source were collected from the title of the entry for a given term found in the online encyclopedia. The corpus is presented in its entirety in Appendix A.

## 3.3 Native phonologies

### 3.3.1 Guaraní

Lexical stress is Guaraní is most commonly word-final (both the data and information regarding native Guaraní phonology in this section is taken from de Canese (1983), except where noted). This is reflected in the orthography, in which oxytones are unmarked and words with lexical stress in any position except word-final must carry an acute accent mark to denote the position of primary stress. Exceptions to the generalization that stress is word-final include genuine monomorphemic exceptions as well as polymorphemic words in which stress may be non-final due to the combination of various morphemes. Relevant examples can be seen in (21).[6]

(21)

| Word type | Orthography | Transcription | English gloss |
|---|---|---|---|
| Oxytone | ñandu | ɲa$^{n}$dú | 'spider' |
| | ore | oɾé | 'we' |
| | mbyte | $^{m}$bɨté | 'center' |
| Monomorphemic non-oxytone | ta'ýra | taʔɨ́ɾa | 'son' |
| | túva | túva | 'father' |
| | ára | áɾa | 'day' |
| Polymorphemic non-oxytone | irũnguéra | iɾũ-$^{ŋ}$gwéɾa | 'friends' |
| | ndéve | $^{n}$dé-ve | 'to you' |
| | ajapóta[7] | a-d͡ʒapó-ta | 'I will make' |

Native Guaraní syllable structure is (C)(G)V(G).[8] Syllables may consist of a vowel only or may optionally contain an onset; complex onsets however, and codas of any kind, are forbidden. Guaraní makes extensive use of diphthongs. The Guaraní phonemic inventory has 12 vowel phonemes consisting

[5] http://gn.wikipedia.org/

[6] All transcriptions here and henceforth are standard IPA with the exception of the representation of primary lexical stress, which will be denoted with an acute accent over the stressed syllable nucleus. Morpheme boundaries are indicated in transcription by dashes when relevant.

[7] Taken from Tonhauser and Colijn (2010, 264).

[8] "V" is used to represent a vowel acting as the syllabic nucleus while "G" (for "glide") is used for non-nuclear vowels.

of oral and nasal versions of the six vowels /i ɨ u e o a/. All non-high vowels in Guaraní (e.g. /e o a/) always constitute syllabic nuclei, while their high counterparts (e.g. /i ɨ u/) may or may not depending on their surrounding segments. High vowels are less sonorous than their non-high counterparts (Zec, 2007) and consequently any combination of a non-high vowel followed by a high vowel will either result in two separate syllables (if the high vowel bears lexical stress) or in a diphthong with the non-high vowel as the syllabic nucleus (if the non-high vowel bears lexical stress). Two adjacent non-high vowels always constitute separate syllables, while two adjacent high vowels will be tautosyllabic. Examples of Guaraní syllable types in common words, with the relevant syllables bolded, are found in (22).

(22)

| **Syllable structure** | **Example** | **English gloss** |
|---|---|---|
| V | **ɨ** | 'water' |
| CV | **sɨ** | 'mother' |
| CVG | mo.**kő͡j** | 'two' |
| CGVG | tu.**ɣwáj** | 'tail' |

Guaraní native phonology contains three presnalized stops: /ᵐb ⁿd ᵑg/, which can be seen in both word-initial and word-medial position in (23)[9].

(23)

| **Position** | **Guaraní** | **English gloss** |
|---|---|---|
| *Word-initial* | **ᵐb**o.vɨ.vɨ́ | 'to sew' |
| | **ⁿd**ɨ | 'saliva' |
| | **ᵑg**wa.ʔú | 'perhaps' |
| *Word-medial* | mo.**ᵐb**o.ʔó | 'to enrage' |
| | ɾe.no.**ⁿd**é.pe | 'in front of' |
| | ɲo.**ᵑg**a.tú | 'to keep', 'to preserve' |

Guaraní prenasalized stops are explicitly rendered in the orthography such that /ᵐb/ is represented as orthographic *mb*, /ⁿd/ as *nd* and /ᵑg/ as *ng*. This is of importance when a Spanish loan containing one of the sequences /mb nd ŋg/ is borrowed, as will be seen in the following chapter.

### 3.3.2 Spanish

Spanish lexical stress is widely variable (both the data and information regarding native Spanish phonology in this section is taken from Hualde (2005)). Generally speaking Spanish words may place primary lexical stress on either the final, penultimate or antepenultimate syllable. This phonemic contrast leads to some triplet minimal sets such as the one seen below in (24).

[9]Data in figure (23) taken from Britton (2005).

(24)

| Orthography | Transcription | English gloss |
|---|---|---|
| número | número | 'number' |
| numero | numéro | 'I number' |
| numeró | numeɾó | 'he/she/it/you numbered' |

Similarly to Guaraní, things may become more complicated due to morphological reasons such as the addition of a clitic to a verb (e.g. *cantándonosla* = 'singing it to us') and in such cases stress may fall even earlier than the antepenult. This variability of Spanish stress becomes important when considering loans without word-final stress which make their way into Guaraní.

Native Spanish syllable structure allows for more complex structures than that of Guaraní. Spanish syllables may contain a vowel only, but may also make use of onsets and codas, both simple and complex. Simple and complex onsets are found commonly throughout the language, as are simple codas. Complex codas are rare, but occasionally result from word-medial VCCCV strings in which syllabification results in a complex codas (e.g. VCC.CV). These scenarios produce the largest attested syllable in Spanish: CCVCC. Examples of Spanish syllable types, with the relevant syllables bolded, can be found in (25).

(25)

| Syllable structure | Example | English gloss |
|---|---|---|
| V | **a** | 'to' |
| CV | **la** | 'the' |
| CVC | **tos** | 'cough' |
| CCVC | **tɾes** | 'three' |
| CCVCC | **tɾans**.poɾ.táɾ | 'to carry' |

With regard to glides, Spanish patterns similarly to Guaraní in the sense that its high vowels (e.g. /i u/) may be nuclear or non-nuclear while its non-high vowels (e.g. /e o a/) are always nuclear. Also similar to Guaraní, any tautosyllabic combination of a high vowel and a non-high vowel will result in the non-high vowel constituting the nucleus. Adjacent non-high vowels form individual syllables while adjacent high vowels will be tautosyllabic.. Examples of Spanish syllables with glides can be found in (26).

(26)

| Syllable structure | Example | English gloss |
|---|---|---|
| VG | **aj** | 'there is/are' |
| CGV | **pje** | 'foot' |
| CGVG | **bwej** | 'ox' |
| CGVGC | lim-**pjajs** | 'you (plural) clean' |

## 3.4 Guaraní treatment of Spanish stress

Having described the relevant characteristics of native Guaraní and Spanish phonology, the discussion now turns to the Spanish loans in the Guaraní lexicon and the relevant adaptations, or lack thereof, undergone by the loans in the process of nativization. Beginning this section are those loans which undergo nativization processes in the treatment of lexical stress followed by loans which show no nativization of stress. Completing the chapter is a short discussion of a few segmental adaptations encountered in the loans.

### 3.4.1 Stress adaptation

When considering lexical stress adaptation, the loans of interest are those whose original Spanish form had non-final stress, as these are the only loans who present a conflict with the native pattern of lexical stress in Guaraní and thereby provide information regarding adaptation. As would be expected, many loans entering with non-final stress undergo an adaptation which renders the stress word-final to pattern with native Guaraní words. In (27)[10] we see the examples from the corpus, all of which happen to have penultimate stress in their original Spanish forms, in which this is observable.

[10]Henceforth all tables with data from the corpus are exhaustive with regard to the number of examples they show and no loans have been left out (unless otherwise mentioned).

(27)

| Spanish[11] | Guaraní | English gloss |
|---|---|---|
| kɾís.to | kiɾitó | 'Christ' |
| bá.ka | vaká | 'cow' |
| a.sú.kaɾ | asuká | 'sugar' |
| es.pi.ná.so | epinasó | 'spine' |
| sa.pá.to | sapatú | 'shoe' |
| o.βé.xa | oveʃá | 'sheep' |
| gɾá.sja | gɾasjá | 'joke', 'grace' |
| ké.so | kesú | 'cheese' |
| ka.βá.ɾo | kavad͡ʒú | 'horse' |
| bo.rí.ka | $^{m}$buɾiká | 'donkey' |
| moɾ.sí.ɾa | $^{m}$busjá | 'blood sausage' |
| gɾa.ná.ða | gɾanadá | *kind of fruit* |
| ku.lán.tɾo | kuɾãtɰ̃ | 'coriander' |
| káɾ.men | kamé | *proper name* |
| kɾo.á.sja | kɨoasjá | 'Croatia' |
| káɾ.los | kaló | *proper name* |
| kaβ.ɾa | kavaɾá | 'goat' |
| péð.ɾo | peɾú | *proper name* |
| páβ.lo | palí | *proper name* |
| a.ntó.njo | toní | *proper name* |
| fɾan.sís.ka | ʃiká | *proper name* |
| ból.sa | vosá | 'bag' |
| o.βís.po | obispó | 'bishop' |
| es.tó.nja | etoŋá | 'Estonia' |
| a.le.má.nja | alemaŋá | 'Germany' |
| gɾé.sja | gɨɾesjá | 'Greece' |
| i.ŋgla.té.ra | i$^{ŋ}$gɨaterá | 'England |
| is.lá.ndja | iɨla$^{n}$dá | 'Iceland' |
| iɾ.lá.nda | ila$^{n}$djá | 'Ireland' |
| al.bá.nja | avaŋá | 'Albania' |
| áws.tɾja | awteɾjá | 'Austria' |
| tuɾ.kí.a | tuɨkjá | 'Turkey' |
| di.na.máɾ.ka | $^{n}$dinamaɨká | 'Denmark' |
| lí.βja | livjá | 'Libya' |
| aɾ.xé.lja | aɨheljá | 'Algeria' |
| e.ɾi.tɾé.a | eɾitɨɾeá | 'Eritrea' |
| meɾ.kú.ɾjo | mekuɾjó | 'Mercury' |
| sa.túɾ.no | satuɨnó | 'Saturn' |
| u.ɾá.no | uɾanó | 'Uranus' |
| aws.tɾá.lja | awtaɾaljá | 'Australia' |
| nep.tú.no | netunó | 'Neptune' |

### 3.4.2 Tolerance of original Spanish stress

Just as many loans from Spanish show repair strategies for avoiding phonological phenomena not found natively in Guaraní, many loans also show no repair whatsoever of the same phenomena. Starting with

[11]With specific regard to the syllabification of the Spanish forms, here and in subsequent data tables "Spanish" does not necessarily refer to the form of the original Spanish phonological grammar but rather to the perception of Guaraní speakers of the syllabification of the Spanish form, a distinction whose importance will be discussed with greater detail later in this chapter and following chapters.

the treatment of lexical stress, there is also a substantial amount of loans which tolerate the original Spanish stress as opposed to making an adaptation. Below we see the 56 loans in the corpus which pattern this way (out of a total of 97 unambiguous cases),[12] broken up into two tables in which (29) shows those cases where the original Spanish stress falls on the antepenult and (30) showing those cases where it falls on the penult.

(29)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| es.pí.ɾi.tu | espíritu | 'spirit' |
| ka.tó.li.ka | katólika | 'Catholic' |
| a.mé.ɾi.ka | améɾika | 'America' |
| áɾk.ti.ko | áɾktiko | 'arctic (ocean)' |
| at.lá.nti.ko | atlá$^{n}$tiko | 'atlantic (ocean)' |
| í.ndi.ko | í$^{n}$diko | 'indian (ocean)' |
| lí.βa.no | lívano | 'Lebanon' |
| xú.pi.teɾ | húpiteɾ | 'Jupiter' |

[12] In the corpus there are 120 loans which give us information about stress adaptation (i.e. whose original form has non-final stress). Of these, 41 repair the stress making it word-final (seen above in (27)), 56 tolerate non-final stress and the remaining 23 are ambiguous with regard to stress due to the Guaraní form deleting syllables. This is most often the case with borrowed versions of names (a more detailed discussion of this treatment of names is found in the following chapter), and several examples of the loans that pattern this way are seen in (28).

(28)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| do.ló.res | loló | *proper name* |
| kaɾ.ló.ta | kaló | *proper name* |
| kons.tán.sja | kotá | *proper name* |

(30)

| Spanish | Guaraní | English gloss |
|---|---|---|
| bi.ná.gɾe | vinágɾe | ‘vinegar’ |
| bo.lí.t͡ʃe | volíʃo | ‘store’ |
| bí.t͡ʃo | víʃo | ‘bug’ |
| kɾís.ma | kɾíhma | ‘sacrament of confirmation’ |
| en.sa.lá.ða | ensaláda | ‘salad’ |
| bwé.no | wéno | ‘good’ |
| ba.ké.ɾo | vakéɾo | ‘cowboy’ |
| bí.no | víno | ‘wine’ |
| bo.mbí.ɾa | go$^{m}$bíla | *straw used with maté* |
| e.nté.ɾo | e$^{n}$téɾo | ‘entire’ |
| ko.lek.tí.βo | kolektívo | ‘bus’ |
| e.lá.ða | eláda | ‘iced’ |
| a.ró.ɾo | aród͡ʒo | ‘stream’ |
| es.kwé.la | ekwéla | ‘school’ |
| pwéβ.lo | pɨélo | ‘town’ |
| pé.ɾes | péɾe | family name |
| es.kí.na | ekína | ‘corner’ |
| eɾ.má.na | eɾmána | ‘sister’ |
| eɾ.má.no | eɾmáno | ‘brother’ |
| bɾo.mís.ta | $^{m}$bɾomísta | ‘funny’ |
| t͡ʃí.ka | ʃíka | ‘girl’ |
| pjó.la | piʔóla | ‘cord’ |
| ko.mpu.ta.ðó.ɾa | ko$^{m}$butadóɾa | ‘computer’ |
| ko.ka.í.na | kokaína | ‘cocaine’ |
| xe.su.kɾís.to | hesukɾísto | ‘Jesus Christ’ |
| Ma.rí.a | María | ‘Mary’ |
| o.βís.po | ovíspo | ‘bishop’ |
| kɾis.tjá.na | kɾistjána | ‘Christian’ |
| kɾis.tjá.no | kɾistjáno | ‘Christian’ |
| a.ɾá.βja | aɾávja | ‘Arabian Peninsula’ |
| ó.t͡ʃo | óʃo | ‘eight’ |
| pí.ɲa | píɲa | ‘pineapple’ |
| tí.fus | tífu | ‘typhus’ |
| tí.fo | tífo | ‘typhus’ |
| tó.ɾo | tóɾo | ‘bull’ |
| ka.mí.sa | kamísa | ‘shirt’ |
| lá.ta | láta | ‘tin plating’, ‘can’ |
| lú.nes | lúne | ‘Monday’ |
| ew.ka.ɾis.tí.a | ewkaɾistía | ‘eucharist’ |
| ko.ló.mbja | kolómbja | ‘Colombia’ |
| bo.lí.βja | volívja | ‘Bolivia’ |
| fút.bol | húvol | ‘soccer’ |
| ló.mo | lómo | ‘loin’ |
| ka.né.la | kanéla | ‘cinnamon’ |
| ló.ndɾes | ló$^{n}$dɾe | ‘London’ |
| síð.ɾa | síɾa | ‘cider’ |
| e.ɾí.ða | eɾída | ‘wound’ |
| máɾ.te | máɾte | ‘Mars’ |

## 3.5 Segmental phenomena

Unsurprisingly, there are several consistent segmental phenomena encountered in the Guaraní adaptations of Spanish words. These phenomena do not play as important of a role in the large-scale analysis due to the fact that they are not seen as commonly in the corpus and consequently offer less information regarding patterns and the like. Nonetheless three particular phenomena which are seen enough that they merit mention are Guaraní's treatment of the Spanish phonemes /o/, word-initial /b/, and /l/.

### 3.5.1 Adaptation of /o/

Although there are scores of loans in the corpus which faithfully adapt the Spanish phoneme /o/, there are six which adapt it as Guaraní /u/. These loans are seen in (31). An explanation for this adaptation is not offered here, but it seems on the basis of the semantic and phonological properties of the words in which it occurs that these may have been among the first words Guaraní borrowed from Spanish.

(31)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| pé.ðɾ**o** | perú | *proper name* |
| sa.pá.t**o** | sapatú | 'shoe' |
| ké.s**o** | kesú | 'cheese' |
| ka.βá.ʎ**o** | kavad͡ʒú | 'horse' |
| b**o**.rí.ka | ᵐburiká | 'donkey' |
| ku.lán.tɾ**o** | kuɾãtu̇́ | 'coriander' |
| m**o**ɾ.sí.ʎa | ᵐbusjá | 'blood sausage' |

### 3.5.2 Adaptation of word-initial /b/

Also of interest with regard to segmental phenomena is the Guaraní adaptation of word-initial Spanish /b/. In the corpus there are 20 loans beginning with Spanish /b/ (orthographic *b* or *v*). As seen in (32), there are various adaptation strategies attested. In the majority of cases word-initial /b/ is adapted as Guaraní /v/;[13] however, there are also instances of it being adapted as /g/, as well as /ᵐb/. In one case it is adapted faithfully as /b/ (a notably rare occurrence in Guaraní, as /b/ is never found word-initially in native words and very rarely found in such a position in loans), and in the two cases in which /b/ is

[13]Possibly because /b/ undergoes intervocalic spirantization in Spanish rendering it [β], which is adapted into Guaraní as /v/. The realization of Spanish word-initial /b/ as [β] is actually quite common given the Spanish tendency for words to end in vowels, and for spirantization to affect the segment when not phrase-initial. Although perhaps less likely, orthographic interference may also play a role given the dual representation of /b/ as *b* and *v*.

followed by /w/ in the original Spanish form the /b/ is simply deleted.

(32)

| Segment replacing /b/ | Spanish | Guaraní | English gloss |
|---|---|---|---|
| v | ba.lér | valé | 'to be worth' |
| | bí.t͡ʃo | víʃo | 'bug' |
| | bá.ka | vaká | 'cow' |
| | bo.lí.t͡ʃe | volíʃo | 'store' |
| | ba.ké.ɾo | vakéɾo | 'cowboy' |
| | bí.no | víno | 'wine' |
| | bi.ná.gɾe | vinágɾe | 'vinegar' |
| | ból.sa | vosá | 'bag' |
| | beɾ.náɾ.ðo | vená | *proper name* |
| | baw.tís.mo | vawtismo | 'baptism' |
| | bɾa.síl | vɾasíl | 'Brasil' |
| | bo.lí.βja | volívja | 'Bolivia |
| | bu.tán | vutã́ | 'Bhutan' |
| g | bo.mi.tár | gomitá | 'to vomit' |
| | bo.mbí.ɾa | go$^{m}$bíla | *straw used with maté* |
| $^{m}$b | bɾo.mís.ta | $^{m}$bɾomísta | 'funny' |
| | bo.rí.ka | $^{m}$buɾiká | 'donkey' |
| b | baw.tís.ta | bawtista | 'baptist' |
| - | bwé.no | wéno | 'good' |
| | bwéj | wéj | 'ox' |

### 3.5.3 Adaptation of /l/

Similar to the case of /o/, the vast majority of loans in which Spanish /l/ appears see the resulting Guaraní lexical item adapting /l/ faithfully. Native Guaraní phonology had just one liquid phoneme, /ɾ/, but /l/ is seen in loans from other indigenous languages of the Americas as well as those from Spanish (Britton, 2005). In spite of this there are some loans in the corpus in which Spanish /l/ is adapted as native Guaraní /ɾ/, and these loans can be seen in (33).

(33)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| me.**l**ón | meɾő | ‘melon’ |
| ku.**l**án.tɾo | kuɾãtű | ‘coriander’ |
| pe.**l**ón | peɾő | ‘bald’ |
| a**l**.mo.á.ða | aɾmoxá | ‘pillow’ |
| a**l**.mi.ðón | aɾamiɾő | ‘bag’ |

## 3.6 Summary

To sum, an understanding of the native Spanish and Guaraní phonological grammars provides a necessary basis from which to analyze the Guaraní treatment of various Spanish phonological characteristics seen in loan adaptations. Here we have seen the behavior of Spanish lexical stress in loans in Guaraní in addition to that of several segmental phenomena. The behavior of the loans observed in the corpus with regard to their adaptation strategies (or in some cases the lack thereof) provides a good starting place to understand what the full picture of Guaraní loan adaptation tells us about native Guaraní phonology and its structure.

# Chapter 4

# Loan Adaptation: Codas and Complex Onsets

## 4.1 Introduction

Having discussed in the previous chapter the Spanish and Guaraní native phonologies in addition to the simplest adaptation characteristics of the loans seen in the corpus, we now focus our attention on the more complicated processes which will prove crucial to the subsequent analysis regarding the structure of the synchronic Guaraní grammar. This chapter will be concerned with Guaraní treatment of Spanish codas and complex onsets. Given that these phenomena are intimately concerned with syllable structure, the issue of syllabification, mainly how Guaraní syllabifies loans which it borrows from Spanish, will also be of crucial importance.

The chapter is structured beginning with those loans which show some form of adaptation as opposed to tolerance of Spanish codas and complex onsets. The discussion will begin with unambiguous codas and complex onsets (i.e. word-final and word-initial, respectively) before dealing with their word-medial counterparts. After a description of the various adaptation processes involved in Guaraní's handling of codas and complex onsets, those loans which are tolerant of these structures will be presented next. Closing the chapter is further discussion of syllabification and the crucial role of the perception of the syllabification of the Spanish form by Guaraní speakers.

## 4.2 Adaptation of codas and complex onsets

### 4.2.1 Loans with unambiguous codas/complex onsets

As seen in the previous chapter, neither codas nor complex onsets are allowed in native Guaraní words. Given the prevalence of both of these structures in Spanish, many loans entering Guaraní from Spanish possess one, the other or both, thereby forcing the Guaraní grammar to deal with them in some manner. To begin an analysis of how Guaraní handles such structures found in loans when nativizing, it is easiest

to begin with examples of unambiguous codas and complex onsets, i.e. word-final codas and word-initial complex onsets, and that is the concern of this immediate section.

#### 4.2.1.1 Non-nasal Codas

Spanish loans with word-final non-nasal codas which undergo repair show extremely consistent patterns of adaptation in which the codas are deleted to satisfy the Guaraní avoidance of codas. In (34) we see loans with codas being repaired by deletion without discrimination on the basis of the segment in coda position. The Spanish coda segment is bolded.

(34)

| Segment deleted | Spanish | Guaraní | English gloss |
|---|---|---|---|
| ɾ | ko.si.ná**ɾ** | ko.si.ná | 'to cook' |
| | se.ná**ɾ** | se.ná | 'to have dinner' |
| | me.ɾe.ndá**ɾ** | me.ɾe.$^{n}$dá | 'to have a snack' |
| | te.xé**ɾ** | te.xé | 'to weave' |
| | tɾa.tá**ɾ** | tɾa.tá | 'to treat' |
| | ma.ða.ɣas.ká**ɾ** | ma.da.ɣa.ká | 'Madagascar' |
| | dɾo.ɣá**ɾ** | dɾo.ɣá | 'to drug' |
| | bo.mi.tá**ɾ** | go.mi.tá | 'to vomit' |
| | ba.lé**ɾ** | va.lé | 'to be worth' |
| | a.sú.ka**ɾ** | a.su.ká | 'sugar' |
| | a.ko.pjá**ɾ** | a.ko.pjá | 'to stock up' |
| | es.kɾi.βí**ɾ** | kɾi.ví | 'to write |
| | o.ɣá**ɾ** | ó.ɣa | 'home', 'house' |
| | fal.tá**ɾ** | va.tá | 'to lack' |
| | a.te.ndé**ɾ** | a.te.$^{n}$dé | 'to pay attention to' |
| | dok.tó**ɾ** | doɨ.tó | 'doctor' |
| s | a.ní**s** | a.ní | 'anise' |
| | a.ró**s** | a.ɾó | 'rice' |
| | tí.fu**s** | tí.fu | 'typhus' |
| | lú.ne**s** | lú.ne | 'Monday' |
| | pé.ɾe**s** | pé.ɾe | *proper name* |
| | ka.pa.tá**s** | ka.pa.tá | 'overlord' |
| | káɾ.lo**s** | ka.ló | *proper name* |
| | ni.ko.lá**s** | ko.lá | *proper name* |
| | ló.ndɾe**s** | ló.$^{n}$dɾe | 'London' |
| ð | ko.mu.ni.ðá**ð** | ko.mu.ni.dá | 'community' |
| l | ko.rá**l** | ko.rá | 'corral' |
| | poɾ.tu.ɣá**l** | poɨ.tu.ɣá | 'Portugal' |
| | al.kó**:l** | al.kó | 'alcohol' |

There is one notable exception to the generalization that word-final codas are repaired by deletion found in the corpus, seen in (35).

(35)

| Spanish | Guaraní | English Gloss |
|---|---|---|
| kɾu**s** | ku.ɾu.sú | 'cross' |

In this case we see epenthesis in lieu of deletion used to satisfy the Guaraní ban on codas; word-final /s/ is retained and /u/ is epenthesized. Given the consistent behavior of the rest of the loans it seems that the assumption that this loan and its repair strategy are anomalous is a safe one. No explanation seems plausible on phonetic grounds given that there are no phonetic characteristics of this loan which distinguish it from the rest of the corpus (in (34) can be seen many other loans with word-final /s/ in the Spanish form which do not pattern this way), and as such it is difficult to know with certainty what caused its repair to differ from the others. It is notable in being the only monosyllabic form in Spanish with a word-final non-nasal coda and this may or may not be a relevant factor in its adaptation.

#### 4.2.1.2 Nasal Codas

As is expected, nasal codas are also avoided; however, they show a different repair strategy. As mentioned earlier, Guaraní has a fully developed phonemic nasal/oral vowel contrast in which all vowels may appear as oral or nasal. This becomes relevant when a Spanish loan with a nasal coda enters the language, and in these cases we see nasal coalescence as the repair strategy where the nasal coda coalesces with the previous vowel rendering it nasalized. Examples of such loans in which a word-final nasal consonant is repaired by coalescence are seen in (36). It is worth mentioning that /n/ is by far the most frequently occurring nasal coda in Spanish (also possible are /m ŋ/), and as such the vast majority of loans in the corpus with nasal codas in their original Spanish form have /n/. Occurring once however, /m/ patterns in the same manner; /ŋ/ appears in coda position but never word-finally.

(36)

| Vowel nasalized | Spanish | Guaraní | English gloss |
|---|---|---|---|
| o | ka.mjó**n** | ka.mjṍ | ‘truck’ |
| | le.ó**n** | le.ṍ | ‘lion’ |
| | me.ló**n** | me.ɾṍ | ‘melon’ |
| | pe.ló**n** | pe.ɾṍ | ‘bald’ |
| | a.βjó**n** | a.vjṍ | ‘airplane’ |
| | xa.pó**n** | ha.pṍ | ‘Japan’ |
| | kol.t͡ʃó**n** | ko.ʃṍ | ‘mattress’ |
| | xa.βó**n** | ha.vṍ | ‘soap’ |
| | al.mi.ðó**n** | a.ɾa.mi.ɾṍ | ‘bag’ |
| | plu.tó**n** | plu.tṍ | ‘Pluto’ |
| | t͡ʃi.t͡ʃa.ró**n** | ʃi.ʃa.ɾṍ | ‘pork rind’ |
| | kal.só**n** | ka.sṍ | ‘pants’ |
| a | xwa**n** | hwã | *proper name* |
| | o.má**n** | o.mã́ | ‘Oman’ |
| | bu.tá**n** | vu.tã́ | ‘Bhutan’ |
| | is.lá**m** | is.lã́ | ‘islam’ |
| i | pe.kí**n** | pe.kĩ́ | ‘Beijing’ |
| e | al.ma.sé**n** | al.ma.sẽ́ | ‘department store’ |
| | je.mé**n** | d͡ʒe.mẽ́ | ‘Yemen’ |

There is an exception to this pattern, seen in (37). This loan is of interest due to the fact that here the nasal codas deletes without any coalescence; notably, this exception is a proper name. Names at times seem to behave differently in the corpus with regard to their repairs, as they are more liberal with their deletion of segments (more on this later in the chapter). Here it constitutes the only exception to the generalization exemplified by the data in (36).

(37)

| Spanish | Guaraní | English gloss |
|---|---|---|
| káɾ.me**n** | ka.mé | *proper name* |

#### 4.2.1.3 Complex Onsets

Spanish loans entering with word-initial complex onsets show two strategies of repair. The first is vowel epenthesis, whereby the onset cluster is broken up by an epenthesized vowel creating two syllables.

The epenthesized vowel is most commonly either /ɨ/ or a duplicate of the vowel following the place of insertion. We see examples of this repair strategy in (38) with the epenthetic vowel bolded.

(38)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| gɾé.sja | g**ɨ**.ɾe.sjá | 'Greece' |
| kɾus | k**u**.ɾu.sú | 'cross' |
| kɾís.to | k**ɨ**.ɾi.tó | 'Christ' |

The second strategy seen in these cases is not epenthesis, but replacement of the second of the two segments by /ɨ/, as can be seen in (39).

(39)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| fɾán.sja | hɨã́.sja | 'France' |
| kɾo.á.sja | kɨo.a.sjá | 'Croatia' |

### 4.2.2 Loans with potentially ambiguous codas/complex onsets

It is simple to determine that a #CC sequence is indeed a complex onset or that a word-final consonant is a coda, but word-medial consonant strings can make things less clear. In potentially ambiguous situations, such as for example word-medial CC sequences which could be tautosyllabic if syllabified as .CC or not if syllabified as C.C, issues of syllabification become important. A brief discussion of some basic generalizations about sonority and how it can be used in Guaraní syllabification merits mention for the purpose of ruling out unlikely syllabifications for certain CC strings.

#### 4.2.2.1 Sonority

A logical place to begin an analysis of Guaraní syllabification patterns is with the sonority sequencing principle (SSP) (Selkirk (1984), Clements (1990)). The SSP captures the generalization that cross-linguistically segments constituting consonant clusters tend to increase in sonority as they approach the syllabic nucleus. This is to say that in a sequence such as $C_1C_2V$, $C_2$ tends to be more sonorous than $C_1$ given its position closer to the nucleus V. Similarly, in $VC_1C_2$ sequences $C_1$ tends to be the most sonorous for the same reason. Any two tautosyllabic consonants whose order is not increasing in sonority in the direction of the nucleus is in violation of the SSP. Although the SSP is not an absolute cross-linguistically, it seems a good basis for justifying that /pt/, /kt/ or other similar sequences are not tautosyllabic when syllabified by the Guaraní phonological grammar, especially given Guaraní's conservative natural syllable structure. As such it seems a reasonable assumption that in the case of two

consecutive consonants whose order is not increasing in sonority in the direction of the nucleus, the two consonants will be split such that they are not tautosyllabic. In these cases the first segment is treated as the coda of the first syllable and the second is treated as the onset of the syllable following it. A non-exhaustive list of examples of this in loans in Guaraní can be seen in (40). As a result of the first segment of a consonant cluster in these cases being syllabified as a coda, we see that segment undergoing repair by deletion as we would expect.

(40)

| CC sequence | Spanish | Guaraní | English gloss |
|---|---|---|---|
| pt | ne**p.t**ú.no | ne.tu.nó | 'Neptune' |
| ls | bó**l.s**a | vo.sá | 'bag' |
| lt | fa**l.t**áɾ | va.tá | 'to lack' |
| ɾl | ká**ɾ.l**os | ka.ló | *proper name* |
| ɾk | me**ɾ.k**ú.ɾjo | me.ku.ɾjó | 'Mercury' |
| ɾn | be**ɾ.n**áɾ.ðo | ve.ná | *proper name* |

Salient here is the fact that in these cases the first consonant of the CC sequence is deleted. As shown above, in unambiguous situations deletion is a repair strategy which is unique to codas and is not employed by complex onsets,[14] as demonstrated in (42).

(42)

| Spanish | Guaraní | English gloss |
|---|---|---|
| **gɾ**é.sja | gɨ.ɾe.sjá<br>*ge.sjá<br>*ɾe.sjá | 'Greece' |

This combined with the cross-linguistic generalizations captured by the SSP allow us to reasonably deduce the the first consonant in CC sequences such as those in (40) is indeed a coda, and that the sequence as a whole is not being treated as a complex onset.

In the loan in (43) we can see a clear example of these two differing repairs for codas and complex onsets at work in a loan with a CCC sequence.

(43)

| Spanish | Guaraní | English gloss |
|---|---|---|
| áw**s.tɾ**ja | aw.te.ɾjá | 'Austria' |

[14]There is a single anomalous exception to this, seen here in (41). In this case /ɾ/ is deleted from the original /tɾ/ cluster. The reason for this repair strategy is not conspicuous but may have to do with nasalization of the vowels surrounding the cluster in question. Explanation aside, this is the only loan in the corpus which patterns in this manner.

(41)

| Spanish | Guaraní | English gloss |
|---|---|---|
| ku.lán.tɾo | ku.ɾa̋.tũ | 'coriander' |

In the original Spanish form there is a CCC sequence (/stɾ/) which the Guaraní grammar must deal with in the process of adapting the loan. The resulting loan is [aw.te.ɾjá] where /s/ is deleted and an epenthetic vowel is inserted between /t/ and /ɾ/. Given the knowledge that deletion is a unique repair strategy to codas and vowel epenthesis unique to complex onsets, we can reason that the Guaraní grammar's syllabification of the original Spanish form must have been [áws.tɾja]. The /s/ is deleted as a result of its being syllabified as a coda and the vowel epenthesis is a result of /tɾ/ being syllabified as a complex onset.

#### 4.2.2.2 Deletion to repair word-medial codas

The diagnostic that in unambiguous cases the repair strategy of deletion is unique to codas allows us insight into word-medial CC sequences which could constitute perfectly viable complex onsets (i.e. either do not violate the SSP or violate it in a cross-linguistically common way such as /s/+stop clusters). Seen in (44) are the loans which delete the first consonant of a word-medial CC sequence which could plausibly be syllabified as a complex onset, thereby allowing us to assume the syllabification is C.C with the first consonant representing a coda.

(44)

| Segment deleted | Spanish | Guaraní | English gloss |
|---|---|---|---|
| s | es**s**.pi.ná.so | e.pi.na.só | 'spine' |
| | kɾí**s**.to | ki.ɾi.tó | 'Christ' |
| | es**s**.kwé.la | e.kwé.la | 'school' |
| | es**s**.kí.na | e.kí.na | 'corner' |
| | ma.ðа.ɣa**s**.káɾ | ma.da.ɣa.ká | 'Madagascar' |
| | fɾan.sí**s**.ka | ʃi.ká | *proper name* |
| | áw**s**.tɾja | aw.te.ɾjá | 'Austria' |
| | aw**s**.tɾá.lja | aw.ta.ɾa.ljá | 'Australia' |
| | es**s**.kɾi.βíɾ | kɾi.ví | 'to write' |
| | kon**s**.tán.sja | ko.tá | *proper name* |
| | es**s**.tó.nja | e.to.ɲá | 'Estonia' |
| | kɾi**s**.to.βál | ki.ɾi.tó | *proper name* |
| β | pwé**β**.lo | pɨé.lo | 'town' |
| | pá**β**.lo | pa.lí | *proper name* |
| ð | sí**ð**.ɾa | sí.ɾa | 'cider' |
| | pé**ð**.ɾo | pe.ɾú | *proper name* |

Seen in (45)[15] are similar cases differing in the sense that their word-medial CC sequences are not plausible (using the SSP as a diagnostic) complex onsets. Their repair is identical to that seen in (44).

(45)

| **Segment deleted** | **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|---|
| ɾ | ma**ɾ**.ga.ɾí.ta | ma.ᵑga.ɾí | *proper name* |
| | me**ɾ**.kú.ɾjo | me.ku.ɾjó | 'Mercury' |
| | mo**ɾ**.sí.ʎa | ᵐbu.sjá | 'blood sausage' |
| | ká**ɾ**.los | ka.ló | *proper name* |
| | be**ɾ**.ná**ɾ**.ðo | ve.ná | *proper name* |
| | ka**ɾ**.ló.ta | ka.ló | *proper name* |
| | ká**ɾ**.men | ka.mé | *proper name* |
| | i**ɾ**.lá.nda | i.la.ⁿdjá | 'Ireland' |
| l | a**l**.bá.nja | a.va.ɲá | 'Albania' |
| | ka**l**.són | ka.sṍ | 'pants' |
| | fa**l**.táɾ | va.tá | 'to lack' |
| | ko**l**.t͡ʃón | ko.ʃṍ | 'mattress' |
| | bó**l**.sa | vo.sá | 'bag' |
| p | ne**p**.tú.no | ne.tu.nó | 'Neptune' |
| t | fú**t**.bol | hú.vol | 'soccer' |

There is one notable exception to the deletion pattern in (44) and (45), seen in (46).

(46)

| **Spanish** | **Guaraní** | **English Gloss** |
|---|---|---|
| a**l**.mi.ðón | a.ɾa.mi.ɾṍ | 'bag' |

Similar to the previously mentioned case in (35), in this case we see epenthesis instead of deletion used to satisfy the Guaraní coda ban. The first syllable's coda /l/ is retained while the epenthesis of /a/ causes resyllabification in which /l/ (adapted in Guaraní as /ɾ/) constitutes the onset of the following syllable. As in (35), this behavior of codas in this position is only attested in this form and as such is assumed to be anomalous.

Worthy of note here is also the fact that word-medial nasal codas show no difference whatsoever when compared to their word-final counterparts. Word-final nasal codas are much more frequent in the corpus than word-medial ones, and as a result there are numerically fewer forms to show (of those which exhibit repair). The three forms which can be produced however are shown in (47) (with a further explanation

[15](45) represents an exhaustive list of the loans shown earlier in (40).

for the behavior of the last loan in the section immediately following).

(47)

| Spanish | Guaraní | English gloss |
|---|---|---|
| frá**n**.sja | hiã́.sja | 'France' |
| ku.lá**n**.tɾo | ku.ɾã.tṹ | 'coriander' |
| fi**n**.lá.ndja | hĩ́.la.ⁿdja | 'Finland' |

#### 4.2.2.3 Prenasalized Stops

Guaraní's native prenasalized stops also become relevant in the discussion of syllabification to account for the Guaraní treatment of loans entering with similar structures in their Spanish forms. Spanish words which contain word-medial sequences such as /mb/, /nd/ or /ŋg/ (these sequences do not occur word-initially in Spanish), are treated in Guaraní in the same manner as the native Guaraní /ᵐb ⁿd ᵑg/;[16] they operate as one phonological unit. For example a /VmbV/ sequence (where /m/ and /b/ could be any homorganic nasal/voiced plosive combination) in a Spanish loan will be syllabified as [V.ᵐbV] and never as [Vm.bV]. Evidence for this comes from Guaraní treatment of Spanish nasal codas, which are repaired by nasal coalescence as seen in (36). If the proper syllabification of a word-medial /VmbV/ sequence were [Vm.bV], we would expect nasal coalescence to affect the /Vm/ syllable, rendering it [Ṽ]. This is not observed however, as seen in the loans in (48).

(48)

| Spanish | Guaraní | English gloss |
|---|---|---|
| bo.mbí.ʎa | go.ᵐbí.la | *straw used with maté* |
| a.te.ndéɾ | a.te.ⁿdé | 'to pay attention to' |
| fin.lá.ndja | hĩ́.la.ⁿdja | 'Finland' |

Particularly illustrative here is the loan for *Finland* which becomes [hĩ́.la.ⁿdja] in Guaraní. Here we see the first Spanish /n/ syllabified as a coda and therefore repaired through nasal coalescence, while the second Spanish /n/ is syllabified as though it were part of a prenasalized stop and therefore is not repaired, as it is not in coda position.

#### 4.2.2.4 Epenthesis to repair word-medial complex onsets

Similarly, the diagnostic that in unambiguous cases the repair strategy of epenthesis is unique to complex onsets allows us to establish which word-medial consonant strings the Guaraní phonological grammar

[16] Versions of these native presnasalized stops in which the stop is voiceless (i.e. /nt/) are rare in Guaraní but do occur. As such Spanish /mp/, /nt/ and /ŋk/ pattern with their voiced counterparts in being syllabified in Guaraní as /ᵐp ⁿt ᵑk/.

syllabifies as containing a complex onset. The relevant examples are shown in (49).

(49)

| Spanish | Guaraní | English gloss |
|---|---|---|
| áws.**tɾ**ja | aw.te.ɾjá | 'Austria' |
| e.ɾi.**tɾ**é.a | e.ɾi.tɨ.ɾe.á | 'Eritrea' |
| aws.**tɾ**á.lja | aw.ta.ɾa.ljá | 'Australia' |

#### 4.2.2.5 Replacement with /ɨ/ to replace word-medial codas and complex onsets

Shown in (40) were those cases in which CC sequences in violation of the SSP are syllabified C.C with the result that the first consonant is deleted. In some cases however we see not deletion used to repair these sequences but replacement by /ɨ/, and these cases can be seen in (50).

(50)

| CC sequence | Spanish | Guaraní | English gloss |
|---|---|---|---|
| kt | do**k**.tóɾ | doɨ.tó | 'doctor' |
| ɾk | tu**ɾ**.kí.a | tuɨ.kjá | 'Turkey' |
| ɾn | sa.tú**ɾ**.no | sa.tuɨ.nó | 'Saturn' |
| ɾt | po**ɾ**.tu.ɣál | poɨ.tu.ɣá | 'Portugal' |

As outlined above, instances in which word-medial consonant strings are repaired by deletion or vowel epenthesis are not ambiguous. Here however the situation has grown mildly more complicated in the sense that while codas and complex onsets show repair patterns which differentiate them from one another (deletion and epenthesis, respectively), they also share the use of replacement by /ɨ/ as a repair strategy (seen above in (50) for codas and in (39), reproduced below in (51), for complex onsets). Crucially different in their use of this second strategy however is the fact that complex onsets which are repaired by replacement by /ɨ/ only replace the *second* of the two consonants and leave the first unchanged, as seen below in (51). However word-medial strings undergoing repair by /ɨ/ in which we know the first consonant to be a coda on the grounds of sonority see the *first* of the two consonants replaced (as seen above in (50)).

(51)

| Spanish | Guaraní | English gloss |
|---|---|---|
| **fɾ**án.sja | hɨã́.sja | 'France' |
| **kɾ**o.á.sja | kɨo.a.sjá | 'Croatia' |

Thus the repair strategy of replacement by /ɨ/, while shared by both word-medial codas and complex onsets, is not completely ambiguous given that the linear order of the consonants in an ambiguous CC sequence hints at the initial syllabification of the Guaraní phonological grammar. A CC sequence whose

first consonant is replaced by /ɨ/ is not tautosyllabic while a CC sequence whose second consonant is replaced by /ɨ/ is.

An exhaustive list of the word-medial codas in the corpus repaired by /ɨ/ can be seen below in (52).

(52)

| Segment replaced | Spanish | Guaraní | English gloss |
|---|---|---|---|
| s | i**s**.lá.ndja | iɨ.la.$^{n}$dá | 'Iceland' |
| ɾ | sa.tú**ɾ**.no | sa.tuɨ.nó | 'Saturn' |
| | a**ɾ**.xé.lja | aɨ.he.ljá | 'Algeria' |
| | po**ɾ**.tu.ɣál | poɨ.tu.ɣá | 'Portugal' |
| | tu**ɾ**.kí.a | tuɨ.kjá | 'Turkey' |
| | di.na.má**ɾ**.ka | $^{n}$di.na.maɨ.ká | 'Denmark' |
| k | do**k**.tóɾ | doɨ.tó | 'doctor' |

Evident is the fact that this repair strategy is not seen to correct codas which are word-final, and it may be that corresponding phonetic factors are at work causing the two differing repair strategies.

Word-medial complex onsets also show repair by replacement with /ɨ/, although just one example is found in the corpus. It can be seen in (53).

(53)

| Spanish | Guaraní | English gloss |
|---|---|---|
| i.ŋ**gl**a.té.ra | i.$^{ŋ}$gɨa.te.rá | 'England' |

### 4.2.3 VCCV Generalizations

Using the above diagnostics to disambiguate Guaraní syllabification, we see that the corpus is nearly entirely consistent in that a Spanish VCCV sequence is syllabified as VC.CV. The exceptions to this generalization are seen in (54).

(54)

| Spanish | Guaraní | English gloss |
|---|---|---|
| káβ.ɾa | ka.va.ɾá | 'goat' |
| e.ɾi.tɾé.a | e.ɾi.tɨ.ɾe.á | 'Eritrea' |

Here the first of the two exceptions is the more surprising. The syllabification seems to be [ká.βra] instead of the expected [káβ.ɾa] (which would be predicted to produce the form [ká.ɾa]). This syllabification is expected on the basis of loans in which word-medial fricative/liquid segments similar to this example are handled by deletion of the fricative as can be seen in (55).

(55)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| pá**β**.lo | pa.lí | *proper name* |
| pwé**β**.lo | pɨé.lo | 'town' |
| sí**ð**.ɾa | sí.ɾa | 'cider' |
| pé**ð**.ɾo | pe.ɾú | *proper name* |

The case of [káβ.ɾa] seen in (54) makes it difficult to tell whether it is syllabified different than every other such sequence in the corpus or whether it is syllabified as expected and has an unexpected repair strategy. For the sake of consistency however we will assume it is syllabified as the loans in (55), i.e. as [káβ.ɾa], and that its repair strategy is inconsistent. In any case however, it is a singular instance out of the entire corpus of such an example, and as such does not seem in a position to undermine the generalizations made regarding the patterns of the rest of the loans.

## 4.3 Tolerance of codas and complex onsets

### 4.3.1 Codas

Just as there are loans which go to lengths to repair original Spanish codas, there are also those which are tolerant of codas. Beginning with non-nasal codas, in (56) we see the loans from the corpus which preserve their original Spanish codas faithfully.

(56)

| Coda segment | Spanish | Guaraní | English gloss |
|---|---|---|---|
| s | e**s**.pí.ɾi.tu | es.pí.ɾi.tu | ‘spirit’ |
| | xe.su.kɾí**s**.to | he.su.kɾís.to | ‘Jesus Christ’ |
| | kɾi**s**.tjá.na | kɾis.tjá.na | ‘Christian’ |
| | kɾi**s**.tjá.no | kɾis.tjá.no | ‘Christian’ |
| | baw.tí**s**.ta | baw.tis.ta[17] | ‘baptist’ |
| | baw.tí**s**.mo | vaw.tis.mo | ‘baptism’ |
| | ew.ka.ɾi**s**.tí.a | ew.ka.ɾis.tí.a | ‘eucharist’ |
| | ka.te.kí**s**.ta | ka.te.kis.ta | ‘catechist’ |
| | i**s**.lám | is.lɐ́ | ‘islam’ |
| | o.βí**s**.po | o.bis.pó | ‘bishop’ |
| | o.βí**s**.po | o.vís.po | ‘bishop’ |
| | bɾo.mí**s**.ta | $^{m}$bɾo.mís.ta | ‘funny’ |
| ɾ | kon.fi**ɾ**.ma.sjón | kon.fiɾ.ma.sjón | ‘confirmation’ |
| | e.kwa.ðó**ɾ** | e.kwa.tóɾ | ‘Ecuador’ |
| | má**ɾ**.te | máɾ.te | ‘Mars’ |
| | xú.pi.te**ɾ** | hú.pi.teɾ | ‘Jupiter’ |
| | á**ɾk**.ti.ko | áɾk.ti.ko | ‘arctic’ (ocean) |
| | e**ɾ**.má.na | eɾ.má.na | ‘sister’ |
| | e**ɾ**.má.no | eɾ.má.no | ‘brother’ |
| l | a.pós.to**l** | a.pos.tol | ‘apostle’ |
| | bɾa.sí**l** | vɾa.síl | ‘Brazil’ |
| | mi.ɣé**l** | mi.$^{ŋ}$gél | *proper name* |
| | a**l**.ma.sén | al.ma.sɛ́ | ‘department store’ |
| | a**l**.mo.á.ða | aɾ.mo.xá | ‘pillow’ |
| k | ko.le**k**.tí.βo | ko.lek.tí.vo | ‘bus’ |
| | á**ɾk**.ti.ko | áɾk.ti.ko | ‘arctic’ (ocean) |
| t | a**t**.lá.nti.ko | at.lá.$^{n}$ti.ko | ‘atlantic’ (ocean) |
| ð | pé**ð**.ɾo | ped.ɾo | ‘(St.) Peter’ |
| β | pá**β**.lo | pav.lo | ‘(St.) Paul’ |

Although numerically fewer, there are also instances in which nasal codas also go unrepaired. The loans in which we see this are shown in (57).

(57)

| Spanish | Guaraní | English gloss |
|---|---|---|
| i.ɾá**n** | i.ɾán | ‘Iran’ |
| a.mé**n** | a.mén | ‘amen’ |
| sa**n** | san | ‘St.’ |
| xwa**n** | hwan | ‘(St.) John’ |
| ko**n**.fiɾ.ma.sjó**n** | kon.fiɾ.ma.sjón | ‘confirmation’ |
| a.la.kɾá**n** | a.la.kɾán | ‘scorpion’ |
| flo.ɾi.pó**n** | flo.ɾi.pón | *flower species* |
| e**n**.sa.lá.ða | en.sa.lá.da | ‘salad’ |

[17]Some loans were ambiguous in the original data source with regard to which syllable bears primary lexical stress, and those loans, the few that there are, are unmarked with regards to stress (even though inconsequential to the point being made here specifically).

### 4.3.2 Complex onsets

Similarly, a portion of the loans in the corpus with complex onsets also avoid making repairs and import the loan preserving the original Spanish sequences. In (58) we see a list of the loans in which tolerance of complex onsets is seen in lieu of repair.

(58)

| CC position | Spanish | Guaraní | English gloss |
|---|---|---|---|
| Word-initial | **bɾ**o.mís.ta | $^{m}$bɾo.mís.ta | 'funny' |
| | **kɾ**is.tjá.na | kɾis.tjá.na | 'Christian' |
| | **kɾ**is.tjá.no | kɾis.tjá.no | 'Christian' |
| | **bɾ**a.síl | vɾa.síl | 'Brazil' |
| | **pl**u.tón | plu.tṍ | 'Pluto' |
| | **dɾ**o.ɣár | dɾo.ɣá | 'to drug' |
| | **gɾ**á.sja | gɾa.sjá | 'joke', 'grace' |
| | **kɾ**ís.ma | kɾíh.ma | 'sacrament of confirmation' |
| | **fl**o.ɾi.pón | flo.ɾi.pón | *flower species* |
| | **gɾ**a.ná.ða | gɾa.na.dá | *kind of fruit* |
| | **tɾ**a.tár | tɾa.tá | 'to treat' |
| Word-medial | a.la.**kɾ**án | a.la.kɾán | 'scorpion' |
| | ló.n**dɾ**es | ló.$^{n}$dɾe | 'London' |
| | bi.ná.**gɾ**e | vi.ná.gɾe | 'vinegar' |
| | xe.su.**kɾ**ís.to | he.su.kɾís.to | 'Jesus Christ' |
| | es.**kɾ**i.βír | kɾi.ví | 'to write' |

## 4.4 Guaraní perceived syllabification of Spanish forms

While much has been said about syllable structure and syllabification in general of Spanish words, further discussion of its importance is merited. As has been shown, Guaraní differs in its treatment of codas and complex onsets with regard to repair strategies. This simple fact affects the resulting form in Guaraní due to the fact that the choice of syllabifying a given word-medial consonant affects what precisely happens to it. In the processing of loans, the Guaraní grammar must syllabify foreign structures before repairing them, given this fact that the nature of the syllabification can affect the corrections made. For example, a Guaraní speaker initially processing the Spanish word /áwstɾja/ must decide whether it is syllabified as in (a), (b) or (c) in (59). This decision affects the resulting loan, as can be seen in the last

column.

(59)

| | Perceived syllabification of Spanish form | Expected resulting Guaraní lexical item |
|---|---|---|
| (a) | áw.stɾja | aw.sV.tV.ɾjá |
| (b) | áws.tɾja | aw.tV.ɾjá |
| (c) | áwst.ɾja | aw.ɾjá |

In (59a) we see a syllabification in which there are no codas, and as a result no deletion takes place and all initial Spanish segments are present in the resulting Guaraní form. Given that Guaraní commonly repairs two tautosyllabic consonants by vowel insertion,[18] we can posit that the result of such a syllabification would see two epenthetic vowels (whose quality cannot be known with certainty and is unimportant to the point, and are here represented by a generic 'V') to avoid two consecutive consonants. In (59b) we have /s/ syllabified as a coda and /tɾ/ as a complex onset, and repairs are made accordingly. The /s/ is deleted and a vowel is inserted to break up the /tɾ/ sequence, resulting in the form which we see attested in the corpus. The final possibility in (59c) would syllabify /st/ as a complex coda, leaving /ɾ/ as a singleton onset of the following syllable. It is difficult to know how Guaraní would handle such a form given that Spanish only rarely has complex codas, but regardless of whether it would have produced a form in which vowel epenthesis splits up /s/ and /t/ or whether both segments would be deleted given their coda position, the point that the word's initial syllabification has a clear influence on the final Guaraní form is unchanged.[19]

Most importantly, the syllabification must come first because the choice of syllabification has a dramatic impact on the final form of the word given that Guaraní does not show identical repair patterns for codas and complex onsets. Crucially, this syllabification is the *perceived* syllabification by the Guaraní grammar of the Spanish form.[20] Of note here is the fact that this perceived syllabification has nothing whatsoever to do with the actual Spanish syllabification of the word. Given that the Guaraní grammar is unconcerned with the Spanish grammar's treatment of the word, it would not be impossible for the two syllabifications of the same Spanish form to be different when assessed by the two different grammars. Indeed this is attested, and the most salient example of this are those Spanish forms with nasal/voiced plosive sequences as seen in (48), where the nasal consonants in Spanish would be syllabified as codas

[18]Although this is not the case for all CC strings, for the purpose of demonstration here epenthesis is chosen as the repair strategy with the understanding that if another repair strategy were used instead the point being advanced here would be unchanged.

[19]Similarly, it is noteworthy that although repairs by replacement with /i/ were not taken into account for these examples, they also do not undermine the point being made.

[20]See Smith (2006) for a discussion of the existence and importance of the mental representations of a borrowing language in loan adaptation as shown through Japanese loan doublets of English words.

and not as part of a prenasalized stop as in Guaraní. Further discussion of the importance of the Guaraní perception of syllabification will be seen in the following chapter.

This syllabification by the Guaraní phonological grammar is consistent in its treatment of consonant strings found in the loans in the corpus, and as a result generalizations can be made. As mentioned earlier, with the possible minor exception shown in (54), the corpus is consistent in that a Spanish VCCV sequence is syllabified VC.CV.[21] Spanish VCCCV sequences, while rarer, also seem to show consistency in their syllabification as VC.CCV. Those loans containing original VCCCV sequences are shown in (60).

(60)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| áw**s.tɾ**ja | aw.te.ɾjá | 'Austria' |
| aw**s.tɾ**á.lja | aw.ta.ɾa.ljá | 'Australia' |
| e**s.kɾ**i.βíɾ | kɾi.ví | 'to write' |
| ko**ns.t**án.sja | ko.tá | *proper name* |

Notable here is the last loan in the table, coming from the Spanish proper name *Constancia*. Salient in the corpus is the trend that those loans which are proper names seem to be subject to different phonological treatment during the process of loan adaptation. While a thorough discussion of this phenomenon will not be presented in this analysis, its mention is nevertheless of value. Proper names seem conspicuously more willing to delete segments or sequences of segments found in the original Spanish form than other loans, perhaps as the result of a prosodic template effect on the adaptation of names (i.e. most are disyllabic). To demonstrate this tendency, the proper names found in the corpus can be seen in (62), broken up by those which behave as expected given the rest of the corpus and those which show unexpected behavior.[22]

[21] In addition to the previously discussed exception of the Spanish sequences /mb mp nd nt ŋg ŋk/.

[22] One proper name in the corpus is left out of (62) given that it does not seem to pattern with either category. It preserves more of the original Spanish form than would be expected if it patterned with the rest of the corpus or if it patterned with most proper names:

(61)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| mi.ɣél | mi.$^{n}$gél | *proper name* |

(62)

| Repair pattern | Spanish | Guaraní | English gloss |
|---|---|---|---|
| Behavior consistent with rest of corpus | xwan | hwã | *proper name* |
| | káɾ.los | ka.ló | *proper name* |
| | péð.ɾo | pe.ɾú | *proper name* |
| | páβ.lo | pa.lí | *proper name* |
| Behavior unique to names (more deletion than expected) | xe.ɾáɾ.ðo | ki.ɾá | *proper name* |
| | kɾis.to.βál | ki.ɾi.tó | *proper name* |
| | ni.ko.lás | ko.lá | *proper name* |
| | a.ntó.njo | to.ní | *proper name* |
| | beɾ.náɾ.ðo | ve.ná | *proper name* |
| | fɾan.sís.ka | ʃi.ká | *proper name* |
| | si.mó.na | ʃi.mí | *proper name* |
| | do.ló.ɾes | lo.ló | *proper name* |
| | kaɾ.ló.ta | ka.ló | *proper name* |
| | káɾ.men | ka.mé | *proper name* |
| | kons.tán.sja | ko.tá | *proper name* |
| | li.βó.ɾja | li.vó | *proper name* |
| | xu.ljá.na | lu.lí | *proper name* |
| | maɾ.ga.ɾí.ta | ma.$^{n}$ga.ɾí | *proper name* |

Returning to the case of the Spanish *Constancia*, given that it is a proper name and as a result is more likely to delete segments, it is difficult to know whether the resulting form is representative of normal Guaraní phonological processes of loan adaptation or is extraordinary. For example one might expect nasal coalescence to occur here given the unambiguous nasal coda in the first syllable, yet this is not observed. Putting this example aside, the standard syllabification of Spanish VCCCV sequences is VC.CCV (again disregarding those consonant sequences involving nasals interpreted as Guaraní prenasalized stops).

## 4.5 Summary

Of the adaptation strategies under analysis in this thesis, those pertaining to Spanish codas and complex onsets are the most complex in that they are more variable than other adaptations. Similar to the stress and segmental adaptations seen in the previous chapter, some loans whose original forms have codas or

complex onsets are repaired in some way while others are not. Differentiating these loans from those discussed previously however is the fact that loans with codas and complex onsets which are repaired do not show uniform adaptation. Complex onsets may be repaired by epenthesis or by replacement of the second segment with /ɨ/ while codas are repaired by either deletion or, in word-medial cases, by replacement with /ɨ/. Nasal codas which are repaired show repair by coalescence. The crucial information which these treatment strategies provide us with regard to the structure of the Guaraní lexicon is discussed in the following chapter.

# Chapter 5

# Analysis

## 5.1 Introduction

Having established the nativization strategies of the Guaraní phonological grammar when processing loans from Spanish, we now turn to an analysis of such adaptations within the theoretical framework of Optimality Theory (henceforth OT). OT analyses will shed light on the internal processes of the grammar during the process of adaptation by modeling constraints on Guaraní forms whose rankings are responsible for the final form of the loan. As discussed in Itô and Mester (1999) OT analyses of loan adaptation are also important for their role in theoretically accounting for unattested nativizations, and this and how it pertains to Guaraní is one of the primary topics of discussion in this chapter.

The idea that a language's lexicon may show internal stratification with regard to the phonological characteristics of different lexical items is by no means a new topic of discussion within the literature. As shown in the previous chapter, the Guaraní lexicon seems to also pattern in this way as evidence of stratification is clear based on the phonological characteristics of the members of its lexicon. Morphological evidence in Guaraní, seen in detail in the following section, also points to evidence for the synchronic relevance of the strata. The sociolinguistic history of Guaraní has afforded linguists studying it a glimpse into how an influx of loanwords from another language (especially one distinct phonologically) can result in a lexicon with seemingly drastic differences in phonological characteristics from one word to another. These differences are not random however, and evidence for patterns within phonological stratification are abundant in Guaraní.

With the case of Guaraní the lexicon has clearly defined sublexica, the existence of which can be inferred from the phonological behavior of its members. Crucially, the core-periphery structure of the Guaraní lexicon is the basis on which the case for clear strata can be made. The non-random pattern of adaptation *across* strata says much about the lexicon's organization. Unsurprisingly the majority of the words comprising the Guaraní lexicon are native to the language and exist as an inheritance from

its linguistic predecessors as a member of the Tupí-Guaraní language family. Since the arrival of the Europeans in the Americas, a new source of words has presented itself via the contact of Guaraní with Spanish. The synchronic situation in Guaraní provides modern speakers with a range of loans varying in degree of phonological structure. These loans are grouped into strata whose differing phonological behavior must be accounted for by the grammar. In this sense a stratum is a group of words in the lexicon whose phonological characteristics are the product of one (stratal) grammar; the overall grammar of the language is then accounted for by the combination of the various stratal grammars. The use of stratal grammars to account for the vastly differing adaptation strategies is useful in allowing us to reconcile the fact that one synchronic Guaraní grammar produces forms radically different from one another phonologically.

While the value of the stratification of the lexicon within this analysis is in its synchronic relevance, the reason for the existence of strata is diachronic in nature. The new source of words from Spanish initially resulted in a series of borrowings from Spanish which were repaired to the extent that they were phonologically indistinguishable from native Guaraní words. As time went on and the number of Guaraní-Spanish bilinguals rose (and with it a general familiarity in the Guaraní-speaking community with Spanish phonology), loans began to enter which showed phonological characteristics disallowed in native Guaraní (Morínigo, 1931). This tolerance increased until the point was reached when some Spanish loans began to enter unadapted from their original form. As a result of this historical and sociolinguistic situation the present-day Guaraní lexicon has native Tupí-Guaraní words, fully adapted loans from Spanish, partially adapted loans from Spanish, and loans from Spanish which are not adapted at all.[23]

While the Guaraní lexicon has diachronic phenomena as the cause for its present-day situation, this analysis is concerned with the synchronic ramifications of Guaraní-Spanish contact. Although the synchronic situation was arrived at via diachronic means, the synchronic and the diachronic are distinct. For example those loans which entered from Spanish and underwent total repair are, for the purposes of the synchronic grammar, in no way distinct from those lexical items which came directly from Tupí-Guaraní predecessors of the language. That is to say that from the perspective of the present-day grammar it is not the etymological history of a given word that is important but rather how the word patterns in terms of its phonological characteristics.

The role of faithfulness constraints within OT in this analysis will be crucial. The increasing tolerance of phonological phenomena non-native to Guaraní can be handled within an OT framework by the

[23]This is in addition to loans from other languages such as neighboring (or formerly neighboring) indigenous South American languages as well as other European languages such as Portuguese, which will not be discussed here.

ranking of faithfulness constraints relative to the original markedness constraints responsible for giving native Guaraní lexical items the characteristics which they share. As such, the role and faithfulness and faithfulness constraints will play a central role in the modeling of the grammar.

This chapter is structured as follows. Beginning the chapter is a brief discussion of morphological evidence in the Guaraní grammar supporting synchronically relevant strata. Following is an outline of the core-periphery structure of the Guaraní lexicon and why this characteristic is highly important when discussing the lexicon's organization. Subsequently a slightly simplified (with regard to the role of faithfulness constraints) OT analysis of the stratal grammars is presented which is in turn followed by a discussion of the importance of impossible nativizations to the model. Following this is a more detailed breakdown of the role of faithfulness in the grammar and a discussion of problematic aspects of Guaraní adaptation strategies for the Itô and Mester (1999)-style modeling of faithfulness in the grammar. Concluding the chapter is a discussion of the the implications of Guaraní perceived syllabification on the modeling of Guaraní nativizations within OT.

## 5.2 Stratal evidence in Guaraní morphology

Some evidence for the synchronically relevant status of lexical strata in Guaraní comes from Guaraní morphology. One illustrative example comes from the Guaraní causative prefix *mbo-/mo-* which is used with intransitive verbs (Nordhoff, 2004). This native Guaraní morpheme, while used abundantly with native Guaraní verbs, does not attach to loan verbs (Bakker and Hekking, 2012). This discrimination suggests a distinction of some kind between loans and native verbs by the Guaraní morphology. An account of the Guaraní lexicon in which loans are described as entirely integrated with native lexical items such that no distinction is made between the two by the grammar would prove problematic when handling morphemes such as *mbo-/mo-*. Affixes showing sensitivity to whether the stem is native or borrowed bolster claims that strata in modern Guaraní are more than just diachronically relevant.

It is noteworthy that there is no irrefutable evidence that Guaraní uses borrowed Spanish morphemes productively, in spite of the rampant borrowing in vocabulary, and in many cases vocabulary with Spanish morphemes intact (e.g. plural markers, diminutive markers, agentive markers, adverbial markers, etc.) (Bakker and Hekking, 2012). In spite of this however there is at least one Spanish morpheme, the adverbial marker *-mente*, which remains controversial in this regard in that its behavior could be interpreted as productive, or at least on the path to productivity.

The common adverbial marker *-mente* is used productively in Spanish to form adverbs out of adjectives (Rendon, 2008). While Spanish adverbs are commonly marked by *-mente*, its use is not mandatory

and they may remain morphologically indistinct from their original adjectival forms while still functioning as adverbs (de Bruyne and Pountain, 1995). Spanish adjectives borrowed into Guaraní are attested with adverbial use both with the *-mente* suffix and with their bare adjectival forms, although the forms with *-mente* are more common; as a result there exist many Spanish loans in the Guaraní lexicon which exist as bare adjectives and as overtly marked adverbs Bakker and Hekking (2012). It is not readily clear whether the numerous loans in Guaraní which carry the *-mente* morpheme have therefore been borrowed from Spanish and subsequently lexicalized in that form or whether the attested adjectival forms have been productively turned in adverbs through the use of *-mente*. Relevant to the discussion here, if this latter explanation were the case it would be noteworthy due to the fact that there are no attested cases of native Guaraní adjectives being combined with *-mente* to be turned into adverbs.[24] The morpheme *-mente* therefore may constitute a second case in which the Guaraní grammar attaches specific affixes to specific stems on the basis of their status of native or loaned. Although this case can not be definitely made as the above case of *mbo-/mo-*, it would interestingly constitute the opposite of that case in being a morpheme which only attaches to loans and avoids native stems, providing evidence that the Guaraní grammar has at least some tendency to attach native affixes only to native stems and loan affixes only to loan stems.

## 5.3 Core-periphery structure

### 5.3.1 Implicational relationships

Evidence for the existence of lexical strata is also found in the core-periphery structure of the Guaraní lexicon. This structure shows that lexical groupings of words are not independent from one another but rather are stacked in a set-inclusion manner such that they overlap in predictable ways. This core-periphery structure can be clearly seen through several implicational relationships among the repair strategies of loans which outline the set-inclusion structure of the lexicon.

For example, those loans which make repairs to a Spanish complex onset also repair lexical stress, as can be seen in (63). Below in (63) and (64) are all loans in the corpus where a complex onset is repaired and the original stress is non-final.

[24] In fact, previous research has found that "Guaraní...does not seem to provide any examples of Spanish derivational or inflectional markers attached to native stems" (Bakker and Hekking, 2012, 199).

(63)

| Spanish | Guaraní | English gloss |
|---|---|---|
| **gɾ**é.sja | gɨ.ɾe.sjá | 'Greece' |
| **kɾ**us | ku.ɾu.sú | 'cross' |
| **kɾ**ís.to | ki.ɾi.tó | 'Christ' |
| **kɾ**o.á.sja | kɨo.a.sjá | 'Croatia' |
| áws.**tɾ**ja | aw.te.ɾjá | 'Austria' |
| e.ɾi.**tɾ**é.a | e.ɾi.tɨ.ɾe.á | 'Eritrea' |
| aws.**tɾ**á.lja | aw.ta.ɾa.ljá | 'Australia' |
| i.ŋ**gl**a.té.ra | i.ᵑgɨa.te.rá | 'England' |
| ku.lán.**tɾ**o | ku.ɾã.tű | 'coriander' |

As can be clearly seen, all of the above loans in (63) make repairs to lexical stress in addition to repairs to the complex onset. An additional case is seen in (64).

(64)

| Spanish | Guaraní | English gloss |
|---|---|---|
| **fɾ**án.sja | hɨã́.sja | 'France' |

With regard to this example, in native Guaraní phonology nasalized vowels attract stress (de Canese, 1983) and in the case of loanwords this is reflected in the fact that vowels which are nasalized in the process of repairing a nasal coda will bear lexical stress as they would in native Guaraní words.[25] Thus, putting aside the cases where vowel nasalization occurs, there are no cases in the corpus in which complex onsets are repaired without lexical stress also being repaired.

Another implicational relationship is in turn found between lexical stress and coda repair. In the loans which repair lexical stress, repairs of codas (both nasal and non-nasal) are also seen. All loans from the corpus which repair lexical stress and contain a coda (bolded for reference) are seen in (65).

[25]The notable case seen in (63) of Spanish [ku.lán.tɾo] → Guaraní [ku.ɾã.tű] patterns the way it does due to nasal spreading in Guaraní (Britton, 2005), a phenomenon not discussed here and of little importance to the analysis.

(65)

| Coda repaired | Spanish | Guaraní | English gloss |
|---|---|---|---|
| s | kɾí**s**.to | ki.ɾi.tó | 'Christ' |
| | e**s**.pi.ná.so | e.pi.na.só | 'spine' |
| | e**s**.tó.nja | e.to.ɲá | 'Estonia' |
| | i**s**.lá.ndja | iɨ.la.$^{n}$dá | 'Iceland' |
| | áw**s**.tɾja | aw.te.ɾjá | 'Austria' |
| | aw**s**.tɾá.lja | aw.ta.ɾa.ljá | 'Australia' |
| ɾ | a.sú.ka**ɾ** | a.su.ká | 'sugar' |
| | mo**ɾ**.sí.ʎa | $^{m}$bu.sjá | 'blood sausage' |
| | i**ɾ**.lá.nda | i.la.$^{n}$djá | 'Ireland' |
| | tu**ɾ**.kí.a | tuɨ.kjá | 'Turkey' |
| | di.na.má**ɾ**.ka | $^{n}$di.na.maɨ.ká | 'Denmark' |
| | a**ɾ**.xé.lja | aɨ.he.ljá | 'Algeria' |
| | me**ɾ**.kú.ɾjo | me.ku.ɾjó | 'Mercury' |
| | sa.tú**ɾ**.no | sa.tuɨ.nó | 'Saturn' |
| n | ku.lá**n**.tɾo | ku.ɾã.tṹ | 'coriander' |
| β | ka**β**.ɾa | ka.va.ɾá | 'goat' |
| | pá**β**.lo | pa.lí | *proper name* |
| ð | pé**ð**.ro | pe.ɾú | *proper name* |
| l | bó**l**.sa | vo.sá | 'bag' |
| | a**l**.bá.nja | a.va.ɲá | 'Albania' |
| p | ne**p**.tú.no | ne.tu.nó | 'Neptune' |
| Multiple codas | ká**ɾ**.me**n** | ka.mé | *proper name* |
| | ká**ɾ**.lo**s** | ka.ló | *proper name* |
| | fra**n**.sí**s**.ka | ʃi.ká | *proper name* |

In each of the cases of loans simultaneously containing codas and repairing lexical stress, which constitutes all of the loans seen in (65), we see the coda consistently repaired as well. The singular exception to this rule is seen in (66).

(66)

| Spanish | Guaraní | English gloss |
|---|---|---|
| o.βí**s**.po | o.bis.pó | 'bishop' |

With regard to this loan, its source was a Catholic missal (Fritz, 2004) only in print in rural Paraguay[26] and created for local use. The loan appears verbatim in the missal as *obispo*, and as such Guaraní orthography dictates the it be stressed word-finally. In the missal there are interestingly two loans coming from the original Spanish word *obispo*, and this is the less repaired of the two; it may be the case then that this word is not stressed word-finally but rather bears the stress on the penultimate syllable as in Spanish (in accordance with Spanish orthography which only marks stress if not penultimate). Indeed this would be expected given the pattern of the rest of the similar loans in the corpus. A simple error could in this case be to blame for the discrepancy and as such the loan may not constitute a genuine exception. Given the impossibility of knowing for sure however, the loan is included as an

[26]The only copy in existence outside of Paraguay is found in Walter Royal Davis Library at the University of North Carolina at Chapel Hill.

exception with the understanding that although it may not be, if it is a genuine exception it is the only one in the corpus.

Summing up the facts up to this point, we have established that in a loan the repair of complex onsets implies the repair of lexical stress, which in turn implies the repair of codas. Through this we have a hierarchical relationship between the repairs of the three phenomena which is simply modeled in (67) where "→" can be read "implies."

(67) *Complex onset repair* → *Lexical stress repair* → *Coda repair*

Of note is the fact that a set-inclusion relationship characterizes the relationships of the phenomena and their repairs. Repairing a complex onset implies repairing lexical stress but the reverse is not necessarily the case. The same can be said of the relationship between repairs of lexical stress and codas: the repair of lexical stress implies the repair of codas but not the other way around.

To illustrate the fact that the implications seen in (67) are unidirectional, in (68) we see all loans in the corpus which repair codas as well as contain non-final lexical stress. Out of a total of 31 loans which fit this description, 23 repair lexical stress but eight do not, showing that the repair of codas does not imply the repair of stress.

(68)

| Status of lexical stress | Spanish | Guaraní | English gloss |
|---|---|---|---|
| Repaired | kɾí**s**.to | ki.ɾi.tó | 'Christ' |
| | ká**β**.ɾa | ka.va.ɾá | 'goat' |
| | a.sú.ka**ɾ** | a.su.ká | 'sugar' |
| | e**s**.pi.ná.so | e.pi.na.só | 'spine' |
| | mo**ɾ**.sí.ʃa | $^{m}$bu.sjá | 'blood sausage' |
| | bó**l**.sa | vo.sá | 'bag' |
| | sa.tú**ɾ**.no | sa.tuɨ.nó | 'Saturn' |
| | ne**p**.tú.no | ne.tu.nó | 'Neptune' |
| | ká**ɾ**.lo**s** | ka.ló | *proper name* |
| | pé**ð**.ɾo | pe.ɾú | *proper name* |
| | a**ɾ**.xé.lja | aɨ.he.ljá | 'Algeria' |
| | aw**s**.tɾá.lja | aw.ta.ɾa.ljá | 'Australia' |
| | me**ɾ**.kú.ɾjo | me.ku.ɾjó | 'Mercury' |
| | pá**β**.lo | pa.lí | *proper name* |
| | fra**n**.sí**s**.ka | ʃi.ká | *proper name* |
| | ká**ɾ**.me**n** | ka.mé | *proper name* |
| | e**s**.tó.nja | e.to.ɲá | 'Estonia' |
| | i**ɾ**.lá.nda | i.la.$^{n}$djá | 'Ireland' |
| | i**s**.lá.ndja | iɨ.la.$^{n}$dá | 'Iceland' |
| | áw**s**.tɾja | aw.te.ɾjá | 'Austria' |
| | a**l**.bá.nja | a.va.ɲá | 'Albania' |
| | di.na.má**ɾ**.ka | $^{n}$di.na.maɨ.ká | 'Denmark' |
| | tu**ɾ**.kí.a | tuɨ.kjá | 'Turkey' |
| Unrepaired | tí.fu**s** | tí.fu | 'typhus' |
| | lú.ne**s** | lú.ne | 'Monday' |
| | e**s**.kí.na | e.kí.na | 'corner' |
| | e**s**.kwé.la | e.kwé.la | 'school' |
| | ló.ndɾe**s** | ló.$^{n}$dɾe | 'London' |
| | pwé**β**.lo | pɨé.lo | 'town' |
| | pé.ɾe**s** | pé.ɾe | *proper name* |
| | sí**ð**.ɾa | sí.ɾa | 'cider' |

Conspicuously absent above are nasal codas. This is again due to the fact that Guaraní nasal vowels, which appear when Spanish nasal codas are repaired, attract stress and thus give no information regarding the patterning of stress adaptation when nasal codas are repaired. Also, the majority of Spanish words in the corpus (and indeed in the language) containing nasal codas have them in word-final position and are nearly universally stressed word-finally, thus providing no information on Guaraní treatment of stress due to the lack of conflict they present with native Guaraní phonology. Nasal codas aside, (68) shows clearly that although repair of stress implies repair of codas, repair of codas does not necessarily imply repair of stress.

Similarly, repair of codas does not imply the repair of complex onsets. Shown in (69) are those loans which repair codas and also contain complex onsets. Out of the nine total loans which this encompasses, five repair the complex onset and four do not.

(69)

| Status of complex onset | Spanish | Guaraní | English gloss |
|---|---|---|---|
| Repaired | **kɾ**ús | ku.ɾu.sú | 'cross' |
| | **kɾ**ís.to | ki.ɾi.tó | 'Christ' |
| | **kɾ**is.to.βál | ki.ɾi.tó | *proper name* |
| | aw**s.tɾ**á.lja | aw.ta.ɾa.ljá | 'Australia' |
| | áw**s.tɾ**ja | aw.te.ɾjá | 'Austria' |
| Unrepaired | e**s.kɾ**i.βíɾ | kɾi.ví | 'to write' |
| | **tɾ**a.tá**ɾ** | tɾa.tá | 'to treat' |
| | ló.n**dɾ**e**s** | ló.$^{n}$dɾe | 'London' |
| | **dɾ**o.ɣá**ɾ** | dɾo.ɣá | 'to drug' |

Lastly, as was shown in (63), the repair of complex onsets implies the repair of lexical stress; however as can be seen in (70), the opposite is not the case. In these cases we see those loans which both repair stress and contain a complex onset. Out of these ten loans, eight do indeed repair the complex onset while two loans do not.

(70)

| Status of complex onset | Spanish | Guaraní | English gloss |
|---|---|---|---|
| Repaired | **kɾ**ís.to | ki.ɾi.tó | 'Christ' |
| | **gɾ**é.sja | gɨ.ɾe.sjá | 'Greece' |
| | i.ŋ**gl**a.té.ra | i.$^{ŋ}$gɨa.te.rá | 'England' |
| | áws.**tɾ**ja | aw.te.ɾjá | 'Austria' |
| | e.ɾi.**tɾ** é.a | e.ɾi.tɨ.ɾe.á | 'Eritrea' |
| | aws.**tɾ**á.lja | aw.ta.ɾa.ljá | 'Australia' |
| | **kɾ**o.á.sja | kɨo.a.sjá | 'Croatia' |
| | ku.lán.**tɾ**o | ku.ɾã.tű | 'coriander' |
| Unrepaired | **gɾ**a.ná.ða | gɾa.na.dá | *kind of fruit* |
| | **gɾ**á.sja | gɾa.sjá | 'joke', 'grace' |

There is a distinct hierarchical pattern in the repairs of these three phenomena when they appear in loans from Spanish. Some repairs imply others, and relationships of transitivity can be built making it predictable when some repairs will be made. Crucially however, these relationships are unidirectional. It is through these unidirectional implications that we see the clear core-periphery/set-inclusion structure of the lexicon. Repairs are not made at random but in a systematic way by the grammar and this says much about the synchronic organization of the Guaraní lexicon.

### 5.3.2 Strata

Having established the implicational relationships characterizing the lexicon, we now use these relationships as the basis for outlining the strata. Strata are delineated by the phonological characteristics, and now that these characteristics have been discussed we turn to the task of discussing the strata individually.

Before exploring the loans constituting the various strata however, a note about the frequency of occurrence of relevant forms is necessary. In several of the strata the loans which have the necessary characteristics to be unambiguously included as a member of that stratum are few in number. Important is the fact that this is a limitation of the corpus (and by extension the Guaraní lexicon) and not a limitation of the stratal interpretation. For example, the ideal situation by which to build a case regarding strata would be to have a large number of loans which show every possible structure as a way of discerning precisely how a given word treats any structure. Even given a corpus of over 175 loans, the actual situation falls short of the ideal. Not all Spanish words have complex onsets, codas and non-final stress simultaneously, and of the loans which do only some are repaired while others are not. Depending on whether repaired forms or non-repaired forms are of interest at a given time, only some of those forms will in turn have the other structures necessary to tell precisely where they fall along stratal lines. It is simply impossible to get large numbers of loans which show every possible repairable structure by which to get clear information regarding stratal patterns. In spite of this set back, that strata do exist and that they show predictable patterns seems supported by the evidence in the corpus given the implicational relationships between repairs seen previously.

#### 5.3.2.1 Stratum 1 (Native)

As previously discussed, the first Spanish loans which entered Guaraní and underwent total repair are indistinguishable from native Guaraní words phonologically speaking. The loans included in this "native" stratum are those which repair Spanish phonological characteristics not found in native Guaraní, and as discussed in the previous chapter these are most commonly seen via treatment of complex onsets, codas and lexical stress, as these are the primary distinguishing features which differentiate Spanish phonology from that of Guaraní.

Of note here is the fact that in order to gain a diagnostic with total certainty regarding which stratum a given loan belongs to, the loan would need to contain every possible phonological structure in question (e.g. a complex onset, a coda and non-final stress), as discussed above. Given the reality of the situation, loans which, for example, repair non-final stress but contain neither a complex onset nor a coda might be

possibly placed in any of the strata which repair non-final stress. In cases such like this loans which only contain one or two (but not all) repairable structures and repair them will be conservatively assumed to fall into stratum 1 with the understanding that this is neither empirically justifiable nor problematic for the analysis. The decision of where to place loans like this is essentially arbitrary and those loans which make repairs and which cannot be placed into a particular stratum on logical grounds are placed by default into stratum 1.[27]

Seen in (71) are examples of loans falling into stratum 1 which when faced when any of these non-native phonological characteristics or any combination thereof make repairs.

(71)

| **Phenomenon repaired** | **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|---|
| Lexical stress | lí.βja | li.vjá | 'Libya' |
| | b**á**.ka | va.ká | 'cow' |
| | sa.p**á**.to | sa.pa.tú | 'shoe' |
| Coda | fal.tá**ɾ** | va.tá | 'to lack' |
| | bo.mi.tá**ɾ** | go.mi.tá | 'to vomit' |
| | ba.lé**ɾ** | va.lé | 'to be worth' |
| Complex onset & coda | **kɾ**ús | ku.ɾu.sú | 'cross' |
| | **kɾ**is.to.βá**l** | ki.ɾi.tó | *proper name* |
| Complex onset & stress | e.ɾi.**tɾé**.a | e.ɾi.ti.ɾe.á | 'Eritrea' |
| | **gɾé**.sja | gi.ɾe.sjá | 'Greece' |
| | **kɾ**o.**á**.sja | kio.a.sjá | 'Croatia' |
| Coda & stress | k**áɾ**.lo**s** | ka.ló | *proper name* |
| | b**ól**.sa | vo.sá | 'bag' |
| | a**l**.b**á**.nja | a.va.ɲá | 'Albania' |
| All three | **á**w**s.tɾ**ja | aw.te.ɾjá | 'Austria' |
| | aw**s.tɾá**.lja | aw.ta.ɾa.ljá | 'Australia' |
| | **kɾís**.to | ki.ɾi.tó | 'Christ' |

#### 5.3.2.2 Stratum 2 (Mostly nativized)

In core-periphery terms stratum 1 represents the core of the lexicon due to the fact that loans in this stratum are indistinguishable from native Guaraní words. Moving one step out from the core, we see that the first non-native Guaraní structure to appear in the loans are complex onsets. Here appear the

[27]Later in the analysis the same will go for loans at the opposite end of the spectrum which tolerate given structures but do not contain others; they will by default be placed in stratum 5, the unadapted stratum.

patterns of a new stratum which is distinct phonologically from the native stratum. While there are no loans in the corpus which show tolerance of complex onsets as well as avoidance of both non-final stress and codas, we can see the combination of these features looking across various loans, showing that hypothetically a Spanish loan with the appropriate characteristics would pattern in this way. Shown in (72) are those loans in the corpus which permit complex onsets while still making repairs to codas or lexical stress.

(72)

| **Phenomenon repaired** | **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|---|
| Coda | tɾa.tá**ɾ** | tɾa.tá | 'to treat' |
| | plu.tó**n** | plu.tṍ | 'Pluto' |
| | e**s**.kɾi.βí**ɾ** | kɾi.ví | 'to write' |
| | dɾo.ɣár | dɾo.ɣá | 'to drug' |
| Stress | gɾa.n**á**.ða | gɾa.na.dá | *kind of fruit* |
| | gɾ**á**.sja | gɾa.sjá | 'joke', 'grace' |

#### 5.3.2.3 Stratum 3 (Partially nativized)

After those loans which show tolerance to complex onsets comes the next stratum of loans which are one step further removed from the core. These loans further increase their faithfulness to their original Spanish form while still making repairs of some sort. In this partially nativized stratum the next Spanish phonological structure to be tolerated is non-final lexical stress. Seen in (73) is the singular loan which tolerates both a complex onset and non-final stress while still repairing a Spanish coda.

(73)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| l**ó.ndɾe**s | ló.$^{n}$dɾe | 'London' |

Other loans included in this medial stratum due to their avoidance of codas but not non-final stress are seen in (74).

(74)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| t**í**.fu**s** | tí.fu | 'typhus' |
| l**ú**.ne**s** | lú.ne | 'Monday' |
| e**s**.k**í**.na | e.kí.na | 'corner' |
| e**s**.kw**é**.la | e.kwé.la | 'school' |
| pw**é**.**β**lo | pɨé.lo | 'town' |
| p**é**.ɾe**s** | pé.ɾe | *proper name* |
| s**íð**.ɾa | sí.ɾa | 'cider' |

#### 5.3.2.4 Stratum 4 (Barely nativized)

Next come the loans of another stratum, which tolerate not only complex onsets and non-final stress but codas as well. Here however the Guaraní grammar makes a distinction between nasal and non-nasal codas with regard to what is permissible and what is not. In the loans seen in (75) it is noteworthy that while non-nasal codas are being tolerated in this stratum, nasal codas are still avoided. Obviously the words in the corpus in which not only a nasal and non-nasal coda co-occur but in which they are treated in this manner are not as numerous as those with just one or the other; nonetheless the existence of some loans which pattern this way is suggestive of the existence of a stratum in which non-nasal codas are tolerated but nasal codas are not.

(75)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| i**s**.lá**m** | is.lã́ | 'islam' |
| a**l**.ma.sé**n** | al.ma.sẽ́ | 'department store' |

#### 5.3.2.5 Stratum 5 (Unadapted)

Nasal codas, while being the most consistently avoided of the Spanish phonological characteristics, are finally seen to also be tolerated in the loans in the last stratum. In this unadapted stratum Spanish loans are entering essentially unrepaired, and examples of this are seen in (76).

(76)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| ko**n**.fi**ɾ**.ma.sjó**n** | kon.fiɾ.ma.sjón | 'confirmation' |
| e**n**.sa.l**á**.ða | en.sa.lá.da | 'salad' |
| **fl**o.ɾi.pó**n** | flo.ɾi.pón | *flower species* |

Loans which might also be included in this stratum due to their tolerance of complex onsets, non-nasal codas and non-final stress, but which do not contain a nasal coda by which a positive diagnostic can be obtained are seen in (77).

(77)

| **Spanish** | **Guaraní** | **English gloss** |
|---|---|---|
| **bɾ**o.mí**s**.ta | $^{m}$bɾo.mís.ta | 'funny' |
| xe.su.**kɾí**s.to | he.su.kɾís.to | 'Jesus Christ' |
| **kɾ**i**s**.tj**á**.na | kɾis.tjá.na | 'Christian' |
| **kɾ**i**s**.tj**á**.no[28] | kɾis.tjá.no | 'Christian' |

[28]While it would be inappropriate to list [kɾis.tjá.na] and [kɾis.tjá.no] as separate words in Spanish due to its use of grammatical gender, these two words are listed are separate loans in Guaraní due to the lack of grammatical gender in that language.

### 5.3.3 Summary

As a means of summarizing the relationships amongst strata and the phonological characteristics which outline these relationships, the table in (78) demonstrates the treatment of the relevant phonological structures by the different strata ("N Codas" refers to nasal codas, #CC refers to complex onsets, and "Codas" refers to non-nasal codas).

(78)

| | N CODAS | #CC | NON-FINAL STRESS | CODAS |
|---|---|---|---|---|
| a. Native | Repaired | Repaired | Repaired | Repaired |
| b. Mostly nativized | Repaired | Repaired | Repaired | Tolerated |
| c. Partially nativized | Repaired | Repaired | Tolerated | Tolerated |
| d. Barely nativized | Repaired | Tolerated | Tolerated | Tolerated |
| e. Unadapted | Tolerated | Tolerated | Tolerated | Tolerated |

## 5.4 Modeling in OT

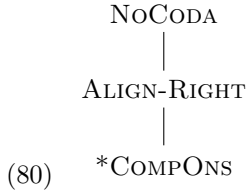
The discussion of complex onsets, codas and lexical stress points to key markedness constraints which are at work in the Guaraní grammar. These constraints and their definitions are seen in (79).

(79) **Markedness Constraints** - adapted from McCarthy (2008), except where noted

- NOCODA - Assign one violation mark for every coda consonant
- ALIGN-RIGHT - Assign one violation mark if the syllable receiving primary stress is not aligned to the right edge of the word (adapted from McCarthy and Prince (1993))
- *COMPLEXONSET - Assign one violation mark for every tautosyllabic cluster in the onset

The patterning of tolerance to Spanish phonological structures seen in the previous section gives a clear indication of the ranking of these three constraints. As seen in the implicational relationships earlier in the chapter, the first Spanish structure to be tolerated is complex onsets, followed by non-final lexical stress, in turn followed by codas. By this we can infer the ranking NOCODA >> ALIGN-RIGHT >> *COMPONS, as shown in the Hasse diagram in (80).

(80)



The more highly ranked constraints here are the most central to Guaraní phonology, as they are the last to be "let go" in the process of permitting non-native structures. In this way we see the strata moving up through the markedness constraints so to speak such that a given stratum which adheres to the demands of a given markedness constraint must also adhere to the demands of those markedness constraints which outrank it. For example a word in the stratum which only permits complex onsets (not non-final stress or codas) must adhere to demands of ALIGN-RIGHT, and as such is required to adhere to the demands of the markedness constraints ranked more highly than ALIGN-RIGHT, in this case NOCODA. Similarly, a given stratum which ignores the demands of a given markedness constraint must also ignore the demands of any markedness constraints which it outranks. Loans which only make repairs to codas are only subject to the requirements of NOCODA but not to ALIGN-RIGHT. This in turn requires these loans to also ignore the demands of *COMPONS, as they ignore the demands of ALIGN-RIGHT which outranks it.

Having established the core of Guaraní phonological demands through the ranking of the three central markedness constraints, we now turn to the important role of faithfulness constraints. The interaction of faithfulness constraints with the core markedness hierarchy seen in (80) is the means by which the grammar arrives at the different stratal patterns characterized by differing repair strategies.

As modeled in Itô and Mester (1999), it is for the purpose of demonstration easiest to begin with the active faithfulness constraints conflated to a single faithfulness constraint which will be referred to here, following the example of Itô and Mester (1999), as FAITH. The unpacking of this consolidated constraint and its theoretical implications will be discussed later in the chapter, but for the time being the position of the constraint FAITH will be used to evaluate the power faithfulness generally commands over the previously established wellformedness hierarchy.

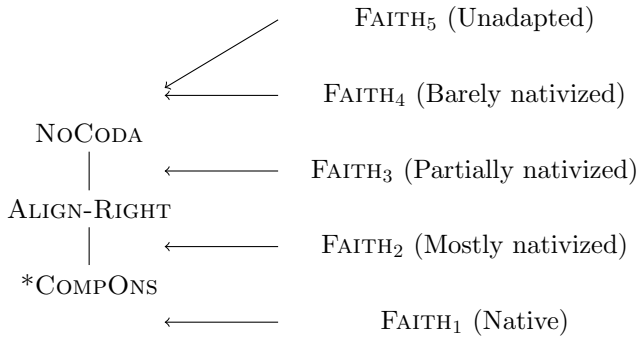
Given that the ultimate goal is to model the grammar synchronically, a synchronic re-ranking of faithfulness constraints to account for the differing adaptation strategies is theoretically unsound. That is, one synchronic grammar is composed of a fixed ranking of constraints. A Guaraní speaker born today would not inherit a grammar in which constraints were freely movable with regard to their rankings but rather one in which rankings are fixed in accordance with the nature of the language. For this reason, a synchronic grammar can be modeled using indexed faithfulness constraints, or in this case indexed versions of the conflated constraint FAITH.[29] That is to say that faithfulness in the grammar can be modeled such that each stratum has a particular FAITH constraint, which is an exact copy of every other FAITH constraint, to which it and only it is subject. This allows for differing phonological characteristics

[29]For a detailed discussion of the theoretical motivations for indexing faithfulness constraints and not markedness constraints see Itô and Mester (2008, 92-93).

of words while permitting all words to be ultimately produced by the same overall single grammar.

Using this representation, in (81) the indexed block FAITH constraints are added to the hierarchy in (80) thus accounting for the different behavioral patterns while maintaining a unified synchronic analysis. In accordance with the strata presented above, each stratum corresponds to a numbered version of the FAITH constraint, such that stratum 1 (the native stratum) is subject to the demands of $\text{FAITH}_1$, stratum 2 is subject to the demands of $\text{FAITH}_2$ and so on.

(81)




Building on the set-inclusion relationship of repair phenomena demonstrated earlier in the chapter, we can now see that this relationship is a reflection of the same relationship amongst constraints and their rankings with respect to one another. For example $\text{FAITH}_1$ is dominated by the same markedness constraints which exert their influence over $\text{FAITH}_2$, but not the other way around. The general idea is that the different strata created by this model are not completely independent of the other strata in the model, but rather all are linked by virtue of this structure of set-inclusion. It is in this sense that the lexicon has a core-periphery structure; the stratum subject to the demands of $\text{FAITH}_1$ is phonologically the same as the native Guaraní lexicon, and is thus seen as the most "core" of the strata. Moving out from this stratum, we find strata which are increasingly less "core", with their corresponding lexical items looking less and less like those found natively in Guaraní.

### 5.4.1 Tableaus - Conflated Faith

To demonstrate the implementation of the constraint rankings outlined in (81), tableau examples of these constraints in work in actual Guaraní lexical items can be seen in (82)-(86). In all loanword tableaus henceforth the inputs will be in pipes (e.g. |krís.to|) representing the Spanish form as perceived by the Guaraní grammar, which as was mentioned earlier and will be further discussed later, is syllabified (for another example of pipes used in loanword tableaus see Broselow (2009)).

In these tableaus any faithfulness violation with respect to complex onsets, codas and lexical stress will mean a violation of FAITH, and individual strata are only subject to the FAITH constraint they are

indexed to (and as such the non-relevant FAITH constraints are left out of the tableau with an ellipsis standing in their place). Here and in subsequent tableaus, the stratum to which a given loan is indexed is indicated in the tableau input with a superscript [F$x$] next to the loan, where $x$ is the number of the stratum to which it is indexed (e.g. |kɾís.to|$^{[F1]}$, meaning this particular loan is indexed to the stratum which is subject to the demands of FAITH$_1$). Each loan being shown can only be subject to one indexed version of FAITH, rendering other indexed versions powerless to exert influence on the loan. The absence of the non-relevant iterations of FAITH in a given tableau is therefore not meant to signify its absence in the grammar but rather its irrelevance in determining a winner for a given input.

#### 5.4.1.1 Faith$_1$ (Native)

(82)

| \|kɾís.to\|$^{[F1]}$ | … | NOCODA | … | ALIGN-RIGHT | … | *COMPONS | FAITH$_1$ |
|---|---|---|---|---|---|---|---|
| a. ☞ ki.ɾi.tó | | | | | | | *** |
| b. ki.ɾí.to | | | | *! | | | ** |
| c. ki.ɾis.tó | | *! | | | | | ** |
| d. ki.ɾís.to | | *! | | * | | | * |
| e. kɾi.tó | | | | | | *! | ** |
| f. kɾí.to | | | | *! | | * | * |
| g. kɾis.tó | | *! | | | | * | * |
| h. kɾís.to | | *! | | * | | * | |

In the native (i.e. fully nativized foreign) stratum, loans are adapted such that they become indistinguishable from native Guaraní lexical items. This is achieved by the ranking of the faithfulness constraints (again, here conflated to the single constraint, FAITH$_1$) below each of the three markedness constraints in question. The result is a loan which is fully compliant with the demands of native Guaraní phonology.

#### 5.4.1.2 Faith$_2$ (Mostly nativized)

(83)

| \|dɾo.ɣáɾ\|$^{[F2]}$ | … | NOCODA | … | ALIGN-RIGHT | FAITH$_2$ | *COMPONS | … |
|---|---|---|---|---|---|---|---|
| a. ☞ dɾo.ɣá | | | | | * | * | |
| b. dɾo.ɣáɾ | | *! | | | | * | |
| c. dɨ.ɾo.ɣá[30] | | | | | **! | | |
| d. dɨ.ɾo.ɣáɾ | | *! | | | * | | |

The mostly nativized stratum is characterized by the first loans which permit phonological structures disallowed in the native Guaraní lexicon, as can be seen in the Guaraní form [dɾo.ɣá]. In this case, FAITH$_2$ outranks *COMPONS but adheres to the demands of the other two markedness constraints, producing forms with are tolerant of complex onsets, but go to lengths to repair codas and lexical stress. Although the repair of lexical stress cannot be seen in (83) due to the loan having word-final stress in Spanish, it can be assumed on solid grounds due to the patterns seen in the rest of the corpus that if the Spanish stress had not been word-final, a repair would have been made. An example illustrative of this, in which ALIGN-RIGHT is satisfied but *COMPONS is not, can be seen in another loan from this stratum as shown in (84).

(84)

| /gɾá.sja/[F2] | … | NOCODA | … | ALIGN-RIGHT | FAITH$_2$ | *COMPONS | … |
|---|---|---|---|---|---|---|---|
| a. ☞ gɾa.sjá | | | | | * | * | |
| b. gɾá.sja | | | | *! | | * | |
| c. gɨ.ɾa.sjá | | | | | **! | | |
| d. gɨ.ɾá.sja | | | | *! | * | | |

#### 5.4.1.3 Faith$_3$ (Partially nativized)

(85)

| /ló.ndɾes/[F3] | … | NOCODA | FAITH$_3$ | ALIGN-RIGHT | … | *COMPONS | … |
|---|---|---|---|---|---|---|---|
| a. ☞ ló.$^{n}$dɾe | | | * | * | | * | |
| b. lo.$^{n}$dɾé | | | **! | | | * | |
| c. ló.$^{n}$dɾes | | *! | | * | | * | |
| d. lo.$^{n}$dɾés | | *! | * | | | * | |
| e. ló.$^{n}$dɨ.ɾes | | *! | * | * | | | |
| f. lo.$^{n}$dɨ.ɾés | | *! | ** | | | | |
| g. ló.$^{n}$dɨ.ɾe | | | **! | * | | | |
| h. lo.$^{n}$dɨ.ɾé | | | **!* | | | | |

In the partially nativized stratum, the faithfulness constraints are ranked such that only the markedness constraint NOCODA can exert influence on the resulting forms. ALIGN-RIGHT and *COMPONS are both outranked by FAITH$_3$, producing forms whose only adherence to Guaraní native phonology can be

[30] For the purpose of modeling a hypothetical complex onset repair, in this tableau as well as those following it, epenthesis of the vowel /ɨ/ was chosen for the sake of demonstration, bearing in mind that it or another vowel could be likely candidates (as in for example the attested repair of the Spanish [kɾís.to] → Guaraní [ki.ɾi.tó] in which /i/ is the epenthesized vowel, or in Spanish [krus] → Guaraní [ku.ru.sú] in which it is /u/).

seen in the repair of the Spanish form's coda.

#### 5.4.1.4 Faith$_5$ (Unadapted)

(86)

| /bɾo.mís.ta/[F5] | Faith$_5$ | NoCoda | ... | Align-Right | ... | *CompOns | ... |
|---|---|---|---|---|---|---|---|
| a. ☞ $^{m}$bɾo.mís.ta | | * | | * | | * | |
| b. $^{m}$bɾo.mis.tá | *! | * | | | | * | |
| c. $^{m}$bɾo.mí.ta | *! | | | * | | * | |
| d. $^{m}$bɾo.mi.tá | *!* | | | | | * | |
| e. $^{m}$bɨ.ɾo.mís.ta | *! | * | | * | | | |
| f. $^{m}$bɨ.ɾo.mis.tá | *!* | * | | | | | |
| g. $^{m}$bɨ.ɾo.mí.ta | *!* | | | * | | | |
| h. $^{m}$bɨ.ɾo.mi.tá | *!** | | | | | | |

Lastly,[31] in the unadapted stratum no influence from the three markedness constraints can be seen whatsoever. Faith$_5$ is subordinate to no markedness demands, and the resulting Guaraní forms tolerate complex onsets, codas and lexical stress which is not word-final.

## 5.5 Possible and impossible nativizations

As thoroughly discussed in Itô and Mester (1999), a central and crucially important attribute of this modeling of lexical strata in loanword phonology is the concept of impossible nativizations. Given that the model presents a fixed hierarchy of wellformedness constraints which interact with indexed faithfulness constraints to produce different lexical strata, there are certain forms which are not able to be picked as optimal by the model, and these forms represent nativizations which are not possible no matter what the ranking of the faithfulness constraints. This theoretically accounts for the fact that, independent of the formalism used to account for the grammar, there are unattested nativization patterns in the Guaraní loan corpus.

Taking for example the ranking in the Guaraní phonological grammar of Align-Right>> *CompOns, any loan subject to the demands of *CompOns is also necessarily subject to those of Align-Right, but not the other way around. There is no possible scenario in which a lexical item could

[31] At this stage in the analysis no tableau is given for the fourth stratum, the barely nativized stratum, given that the difference between stratum 4 and stratum 5 is modeled with specific faithfulness constraint interactions with NoCoda. That is to say that using the block Faith constraint does not allow for teasing the difference between the two strata apart and this must be done through unpacking the Faith constraint into its individual faithfulness constraints. This will be resolved later in the chapter.

adhere to the requirements of *CompOns while ignoring those of Align-Right given the fixed and superior ranking of Align-Right with respect to *CompOns. This is the central characteristic of the set-inclusion model and is well rooted in the implicational relationships seen in the data earlier in the chapter.

Looking at a specific example from the data in which these two constraints are at work, we return to the above loan in (83), Spanish [gɾá.sja] → Guaraní [gra.sjá]. For the sake of simplifying the example we will ignore the role of NoCoda due to its irrelevance in this loan which has no coda. This leads us with three possible nativizations allowed by the grammar depending on which stratum the loan is indexed to, the analyses of which can be seen in (87)-(89).[32]

(87)

| /gɾá.sja/[F1] | … | Align-Right | … | *CompOns | Faith$_1$ |
|---|---|---|---|---|---|
| a. gɾá.sja | | *! | | * | |
| b. gra.sjá | | | | *! | * |
| c. gɨ.ɾá.sja | | *! | | | * |
| d. ☞ gɨ.ra.sjá | | | | | ** |

(88)

| /gɾá.sja/[F3] | … | Align-Right | Faith$_3$ | *CompOns | … |
|---|---|---|---|---|---|
| a. gɾá.sja | | *! | | * | |
| b. ☞ gra.sjá | | | * | * | |
| c. gɨ.ɾá.sja | | *! | * | | |
| d. gɨ.ra.sjá | | | **! | | |

(89)

| /gɾá.sja/[F5] | Faith$_5$ | Align-Right | … | *CompOns | … |
|---|---|---|---|---|---|
| a. ☞ gɾá.sja | | * | | * | |
| b. gra.sjá | *! | | | * | |
| c. gɨ.ɾá.sja | *! | * | | | |
| d. gɨ.ra.sjá | *!* | | | | |

The important point here is that the form [gɨ.ɾá.sja] is a perpetual loser and cannot be picked as optimal no matter what the stratal affiliation of the loan. No matter what the ranking of Faith there is simply no way to achieve the effect of *CompOns being able to exert influence over the form without Align-Right also being able to do the same. Short of being able to freely rank these two markedness

[32] As shown earlier, the attested form is the winner picked by the tableau in (88), [gra.sjá], however the winners shown in (87) and (89) would be equally as possibly rendered by the grammar if the stratal affiliation of the lexical item were different.

constraints, this leaves any scenario in which the more lowly ranked of the two markedness constraints has power over the winning form while the more highly does not ruled out as an impossibility. We therefore see that through its repair of the complex onset but not of the coda, the candidate [gɨ.ɾá.sja] is harmonically bounded (Prince and Smolensky, 1993, 176-178) and consequently an impossible nativization of the Spanish form [gɾá.sja].

The model's prediction of impossible nativizations, while illustrated in this example using ALIGN-RIGHT>>*COMPONS, holds for all three constraint pair rankings and as such all loans have candidates which are rejected by the grammar as perpetual losers. This central idea to the model will be further discussed and tested in the following chapter through an experiment using Guaraní native speakers as participants.

## 5.6 Unpacking of the consolidated faithfulness constraint

Up until this point the use of the block faithfulness constraint FAITH as a means of showing the role of faithfulness in the stratification of the lexicon has served well for the purpose of demonstration. Conspicuous however is the fact that this simplification could be dangerous if not unraveled due to the wider range of outputs possibly chosen as optimal by free rankings of individual faithfulness constraints. In order to achieve a complete analysis the individual components of FAITH must be unpacked.

While the individual faithfulness constraints will be incrementally introduced throughout the analysis in this section, shown in (90) is a complete list of the constraints which will be discussed along with their definitions.

(90) **Faithfulness Constraints** - adapted from McCarthy and Prince (1995), except where noted

- DEP-IO (henceforth DEP) - Assign one violation mark for every segment in the output that lacks a correspondent in the input (≈ no epenthesis)
- MAX-IO (henceforth MAX) - Assign one violation mark for every segment in the input that lacks a correspondent in the output (≈ no deletion)
- $\text{MAX}_{\text{NASAL}}$-IO (henceforth $\text{MAX}_{\text{NASAL}}$) - Assign one violation mark for every [+nasal] segment in the input that lacks a correspondent in the output (≈ no deletion of nasals)
- $\text{MAX}_{\text{ONSET}}$-IO (henceforth $\text{MAX}_{\text{ONSET}}$) - Assign one violation mark for any element appearing in onset position in the input[33] which lacks a correspondent in the output (≈ no deletion of any member of the onset) (adapted from Beckman (1998))
- UNIFORMITY (henceforth UNIF) - Assign one violation mark for every output segment with multiple correspondents in the input (≈ no coalescence)

[33] "Input" here refers to the Guaraní grammar's perceived syllabification of the original Spanish form which, as discussed earlier and will be elaborated on later, is prosodified.

- IDENT$_{\text{CONS}}$ - Assign one violation mark for every segment in the input whose correspondent in the output differs in terms of the feature [consonantal]
- MATCHSTRESS - Assign one violation mark for every lexical stress-bearing vowel in the input whose correspondent in the output is not lexically stressed (adapted from Davidson and Noyer (1997))

### 5.6.1 Problems with unpacking the consolidated Faith

Some problematic issues arise when implementing the individual faithfulness constraints in the analysis and removing the consolidated FAITH. As shown up to now, the consolidated FAITH constraint allows the model to make accurate predictions about impossible nativizations. Deconsolidating FAITH however presents problems for these predictions by allowing individualized faithfulness constraints to be freely ranked thereby giving the model more possibilities for possible nativizations. In essence, unpacking FAITH threatens to rob the model of its predictions of impossible nativizations without extra adjustments being made.

Using one of the examples presented earlier (in (84)), Spanish [gɾá.sja] → Guaraní [gɾa.sjá], we now see the introduction of the faithfulness constraints DEP, MAX and MATCHSTRESS. These three faithfulness constraints interacting with the wellformedness constraints ALIGN-RIGHT and *COMPONS are all that is needed to model the specific nativization process for the Spanish [gɾá.sja].

Modeled below are the effects of these constraints being included in the analysis in a schematic representation in which the markedness constraints (M) and the faithfulness constraints (F) are shown on separate tiers for the purpose of demonstrating not only their interaction with one another (i.e. the interaction of M constraints with F constraints) but also within their respective groups (i.e. the interaction of M constraints with other M constraints as well as F constraints with other F constraints).

In (91) we see a ranking through which the fully nativized form [gi.ɾa.sjá] is arrived at. Here we see markedness constraints ranked most highly and faithfulness constraints ranked beneath all markedness constraints, with the result being a winner which tolerates no non-native phonological structures to Guaraní. This would correspond to stratum 1 (native).

**Input:** |gɾá.sja|

(91)

| | | | | | |
|---|---|---|---|---|---|
| M: | ALIGN-RIGHT >> | *COMPONS >> | | | |
| F: | | | DEP >> | MAX >> | MATCHSTRESS |

**Output:** gi.ɾa.sjá

By re-ranking the faithfulness constraints the slightly less nativized (and in this case the attested) form [gɾa.sjá] is produced in (92). This is a result of DEP and MAX outranking *COMPONS, the result

of which in turn forbids repair of complex onsets. This form would correspond to stratum 2 (mostly nativized).

**Input:** |gɾá.sja|

(92)

| | |
|---|---|
| M: | Align-Right >> *CompOns |
| F: | Dep >> Max >> MatchStress >> |

**Output:** gɾa.sjá

In (93) the fully faithful winner [gɾá.sja] could be simply produced by ranking all faithfulness constraints over all markedness constraints, in essence the opposite of the rankings seen in (91). This would correspond with stratum 5 (unadapted). Notably, thus far in the analysis there is no discernible difference between the unpacked faithfulness constraints and the ability of the consolidated Faith to produce the same forms.

**Input:** |gɾá.sja|

(93)

| | |
|---|---|
| M: | Align-Right >> *CompOns |
| F: | Dep >> Max >> MatchStress >> |

**Output:** gɾá.sja

Up until this point there has been no change in the rankings of the faithfulness constraints with respect to each other. Problematically however, swapping the rankings of Dep and MatchStress will allow for the choosing of the before-declared impossible nativization [gɨ.ɾá.sja] as optimal, as shown in (94) and for the sake of clarity in (95) in full tableau format.

**Input:** |gɾá.sja|

(94)

| | |
|---|---|
| M: | Align-Right >> *CompOns >> |
| F: | MatchStress >> Max >> Dep |

**Output:** gɨ.ɾá.sja

(95)

| |gɾá.sja| | MatchStress | Max | Align-Right | *CompOns | Dep |
|---|---|---|---|---|---|
| a. | gɾá.sja | | | * | *! | |
| b. | gɾa.sjá | *! | | | * | |
| c. ☞ | gɨ.ɾá.sja | | | * | | * |
| d. | gɨ.ɾa.sjá | *! | | | | * |
| e. | gá.sja | | *! | * | | |
| f. | ga.sjá | *! | * | | | |

In this case in (94) we see MatchStress most highly ranked, thereby demanding faithfulness to the original form with regard to lexical stress, while *CompOns is outranked by Max (preventing deletion as a repair strategy) yet outranking Dep (allowing for epenthesis as a repair strategy). This free ranking of faithfulness constraints is problematic for the model in that it threatens to rob it of the predictive power of impossible nativizations.

### 5.6.2 Ranking consistency

The problem here, that a free ranking of individual faithfulness constraints will allow the grammar to produce forms which are predicted by the model to be impossible, is handled in Itô and Mester (1999) through a proposed consistency condition on the rankings of the individual faithfulness constraints across strata. The maintaining of the faithfulness constraints in a fixed order will prohibit the grammar from being able to rank them in such a way as to produce any winning form. A mildly adapted version of this condition, for the purpose of using constraint examples relevant to the immediate discussion about Guaraní, is stated in (96).

(96)

**Ranking Consistency**, adapted from (Itô and Mester, 1999, 82):
There are no strata where the relative rankings of the indexed unpacked faithfulness constraints are inconsistent:
$\forall AB\ (F/_A >> G/_A) \rightarrow (F/_B >> G/_B)$
i.e. if in a given stratum the indexed versions of, for example, Dep and Max are ranked Dep >> Max, there are no possible strata where the indexed versions of these constraints are ranked Max >> Dep.

This condition has the necessary effect of prohibiting the grammar from picking candidates which are predicted to be impossible nativizations as optimal. As seen above in (94), the only means by which the candidate [gɨ.ɾá.sja] can be chosen, given the fixed wellformedness hierarchy, is through MatchStress outranking Dep and Max, or in non-constraint terms, through it being more important to leave lexical stress unrepaired than complex onsets. Mandating that Dep and Max outrank MatchStress due to the fact that such a ranking is attested in other strata (i.e. freezing the ranking of the faithfulness constraints seen in (91)-(93)) has the desired effect of preventing any scenario in which MatchStress outranks Align-Right without Dep and Max also outranking *CompOns. This ensures that any case in which lexical stress repaired complex onsets are as well, thus restoring the predictions made the model regarding impossible nativizations.

### 5.6.3 Stratal analyses

Just as in Itô and Mester (1999), the condition of ranking consistency will be adopted here for the purpose of the stratal analyses featuring individual faithfulness constraints. This is done as a means of testing the imposed condition against the facts of the Guaraní lexicon and to see how its implementation affects the modeling of the structure of Guaraní. The rest of this section presents the constraint rankings for the five proposed lexical strata in the Guaraní lexicon. Important to remember is the notion that each faithfulness constraint is indexed to a particular stratum which it can solely affect and the grammar of the language would be formed by compiling all indexed versions of the individual faithfulness constraints.

Also, as a means of convention each markedness constraint when presented in a Hasse diagram will be marked using (M) (e.g. NoCoda = NoCoda(M)) as a means of more overtly showing the interaction between faithfulness and markedness constraints. Furthermore, faithfulness constraints will come indexed to the specific stratum they pertain to in the same way as the tableau inputs previously seen (e.g. Dep$^{[F1]}$ = indexed version of Dep only relevant for stratum 1).

#### 5.6.3.1 Stratum 1 (native)

Stratum 1, the stratum in which loans are completely nativized, is characterized by the ranking of markedness constraints over all relevant faithfulness constraints, such that in no scenario is faithfulness to the original Spanish form preferred over repair made by the Guaraní grammar. In this way we can construct the grammar producing this stratum by beginning with the well-established markedness hierarchy seen below in (97), a reproduction of that seen in (80).

(97)

NoCoda(M)
|
Align-Right(M)
|
*CompOns(M)

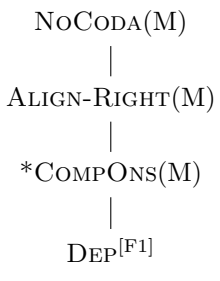
A series of simple ranking arguments using attested forms which fall into this stratum will provide information regarding the rankings of various faithfulness constraints for the stratum 1 grammar as well. As rankings are proven through comparative tableaus, the appropriate additions to the Hasse diagram will be made, thereby incrementally showing what we know for certain about the grammar. Beginning with (98) is the ranking argument *CompOns >> Dep$^{[F1]}$.

(98)

| \|gɾe.sja\|[F1] | *CompOns | Dep[F1] |
|---|---|---|
| a. ☞ gɨ.ɾe.sja | | * |
| b. gɾe.sja | * W | L |

Having established *CompOns >> Dep[F1] for the stratum 1 grammar, we can now add the position of Dep[F1] to the Hasse diagram as seen in (99).

(99)

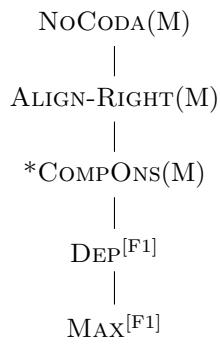



Established in (100) is the ranking of Dep[F1] over Max[F1],[34] which leads to the Hasse diagram in (101).

(100)

| \|a.ɾos\|[F1] | Dep[F1] | Max[F1] |
|---|---|---|
| a. ☞ a.ɾo | | * |
| b. a.ɾo.so | * W | L |

(101)




The ranking argument proving Max[F1] outranking Unif[F1], itself necessary to account for the differing repair strategies for nasal codas, is found in (102) which is followed by the updated Hasse diagram in (103).

[34] Although at this stage this ranking may seem problematic for the prior tableau in (98), the introduction of $\text{Max}_{\text{Onset}}$[F1] momentarily will resolve the matter.

(102)

| $\lvert x_1 w_2 a_3 n_4 \rvert^{[F1]}$ | MAX$^{[F1]}$ | UNIF$^{[F1]}$ |
|---|---|---|
| a. ☞ $h_1 w_2 \tilde{a}_{3/4}$ | | * |
| b. $h_1 w_2 a_3$ | * W | L |

(103)

NOCODA(M)
|
ALIGN-RIGHT(M)
|
*COMPONS(M)
|
DEP$^{[F1]}$
|
MAX$^{[F1]}$
|
UNIF$^{[F1]}$

Having established the hierarchy in (103), we now turn to those constraints whose location in the hierarchy cannot be unequivocally established. The three constraints in question here are MAX$_{\text{ONSET}}^{[F1]}$, MAX$_{\text{NASAL}}^{[F1]}$ and MATCHSTRESS$^{[F1]}$. The comparative tableaus in (104)-(106) show what we do know about the ranking of these constraints.

(104)

| $\lvert$kɾis.to$\rvert^{[F1]}$ | MAX$_{\text{ONSET}}^{[F1]}$ | DEP$^{[F1]}$ |
|---|---|---|
| a. ☞ ki.ɾi.to | | * |
| b. ki.to | * W | L |

(105)

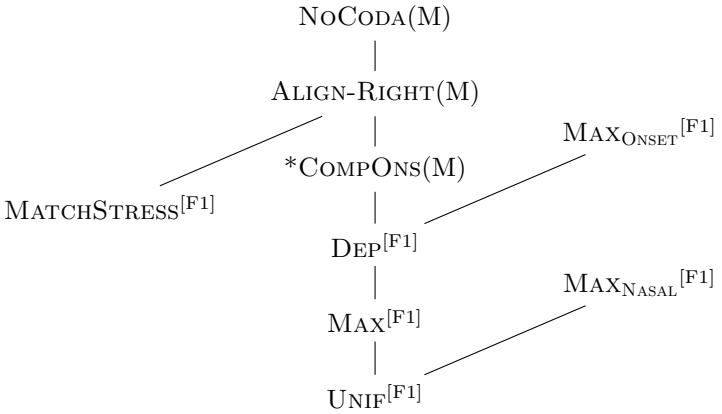
| $\lvert$xwan$\rvert^{[F1]}$ | MAX$_{\text{NASAL}}^{[F1]}$ | UNIF$^{[F1]}$ |
|---|---|---|
| a. ☞ hwã | | * |
| b. hwa | * W | L |

(106)

| $\lvert$sa.pá.to$\rvert^{[F1]}$ | ALIGN-RIGHT | MATCHSTRESS$^{[F1]}$ |
|---|---|---|
| a. ☞ sa.pa.tú | | * |
| b. sa.pá.tu | * W | L |

We know with certainty that MAX$_{\text{ONSET}}^{[F1]}$ outranks DEP$^{[F1]}$ but MAX$_{\text{ONSET}}^{[F1]}$ cannot be ranked with respect to other constraints due to the fact that they do not directly interact. Similarly, MAX$_{\text{NASAL}}^{[F1]}$ outranking UNIF$^{[F1]}$ is verifiable but other than MAX$_{\text{NASAL}}^{[F1]}$ being somewhere more highly ranked than UNIF$^{[F1]}$ its location cannot be found with certainty through the use of comparative tableau ranking arguments. Lastly, the constraints exerting influence on stress only interact with each other and as

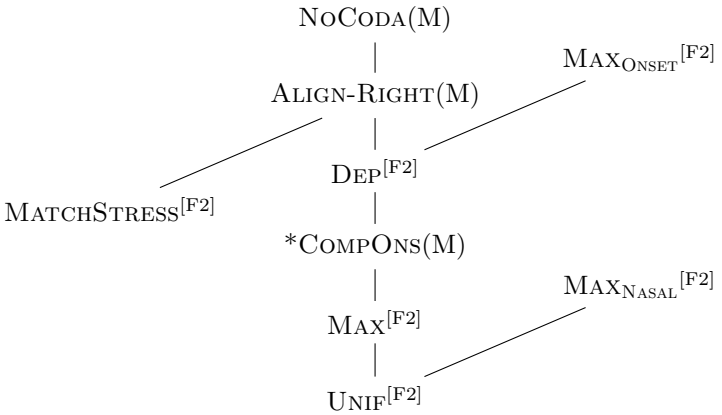
such are difficult to rank with respect to the other constraints. ALIGN-RIGHT must outrank MATCHSTRESS[F1] in this stratum due to the fact that non-final lexical stress is always repaired. This places MATCHSTRESS[F1] somewhere below ALIGN-RIGHT in the grammar but precisely where is not verifiable. Seen in (107) is the complete Hasse diagram for the stratum 1 grammar.




(107)

#### 5.6.3.2 Stratum 2 (Mostly nativized)

Through the condition of ranking consistency we can establish that the faithfulness hierarchy DEP[F1] >> MAX[F1] >> UNIF[F1] shown in (103) must hold for other strata as well. As shown earlier in the chapter, the defining characteristic of stratum 2 is the tolerance of complex onsets (and neither non-final stress nor codas). This can be simply handled by this stratal grammar by the ranking of DEP[F2] over *COMPONS, as seen in the Hasse diagram in (108).




(108)

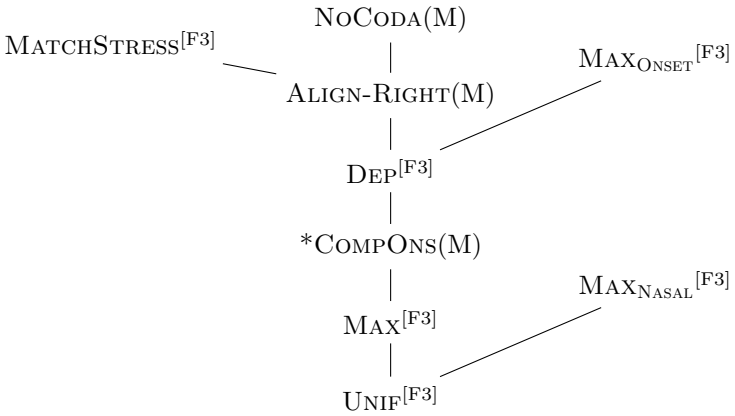
Crucially, the ranking of DEP[F2] and *COMPONS with respect to one another has not changed the rankings of the markedness hierarchy nor those of the faithfulness hierarchy as compared to the first stratum. That is to say that ranking consistency has not been compromised and this new stratal grammar is arrived at by a different interleaving of the markedness and faithfulness hierarchies. As

the strata get farther away from the native (i.e. core) stratum they are characterized by faithfulness constraints rising higher and higher while markedness constraints are moving down through the overall hierarchy.

#### 5.6.3.3 Stratum 3 (Partially nativized)

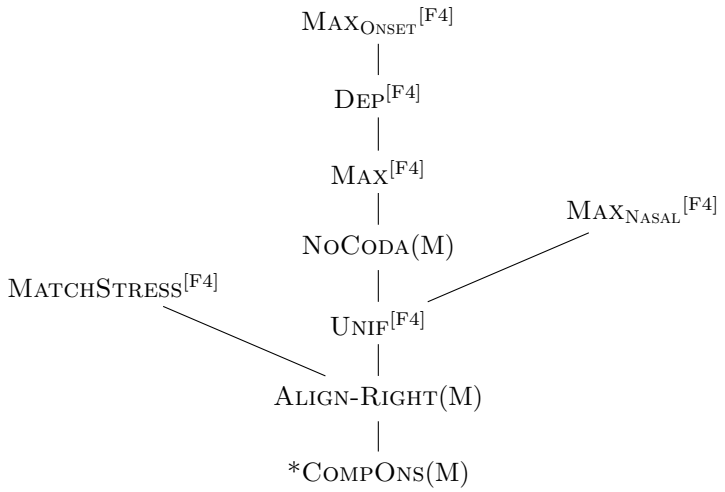
Stratum 3 is characterized by the tolerance of both complex onsets and non-final stress with the simultaneous repair of codas (both nasal and non-nasal). With respect to the constraints, this is handled by ranking MatchStress[F3] over Align-Right. Just as in the previous stratal grammars, the position of MatchStress[F3] with respect to other constraints is unknown due to a lack of conflict between it and the other constraints. As such the Hasse diagram for this stratum is nearly identical to that of stratum 2, with the only change being the understanding that the position of MatchStress[F3] is somewhere *above* Align-Right as opposed to being somewhere below it as with MatchStress[F2] and MatchStress[F1].

(109)




#### 5.6.3.4 Stratum 4 (Barely nativized)

The defining characteristic of stratum four, the last of the strata to make any corrections to Spanish loans, is the tolerance of everything with the exception of nasal codas. In this stratum, as has been shown, non-nasal codas are tolerated while nasal codas are still repaired by coalescence. This is handled by the grammar by ranking NoCoda between Max[F4] (which outranks it) and Unif[F4] (which it outranks) in order to prohibit outright deletion of codas but not coalescence as a repair strategy. Given the need to maintain fixed the markedness hierarchy as well as the faithfulness hierarchy, this means the moving up of all faithfulness constraints through the hierarchy (when compared to for example the stratum 3 grammar) until NoCoda is ranked between Max[F4] and Unif[F4] as seen in (110).

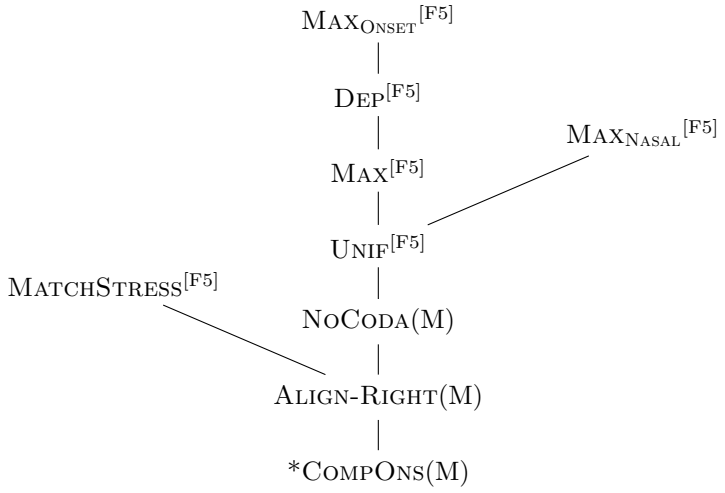



(110)

In this way the grammar can account for the repair strategies while maintaining the cross-stratum faithfulness hierarchy (DEP >> MAX >> UNIF) as well as the original markedness hierarchy (NOCODA >> ALIGN-RIGHT >> *COMPONS).

#### 5.6.3.5 Stratum 5 (Unadapted)

Lastly, the grammar of the unadapted stratum sees faithfulness being ranked over markedness in all scenarios. As such the faithfulness constraints have risen to the top of the overall hierarchy and the markedness constraints are found at the bottom. In this stratum even nasal codas are tolerated, and this means the ranking of UNIF[F5] over NOCODA which is the only difference between this stratal grammar and that of stratum 4. Shown in (111) is the Hasse diagram for the constraints in this final stratum (we know that MATCHSTRESS[F5] must outrank ALIGN-RIGHT but its precise location is not verifiable; MAX$_{ONSET}$[F5] and MAX$_{NASAL}$[F5] must be ranked above all markedness constraints given the high priority of faithfulness in this stratum).




(111)

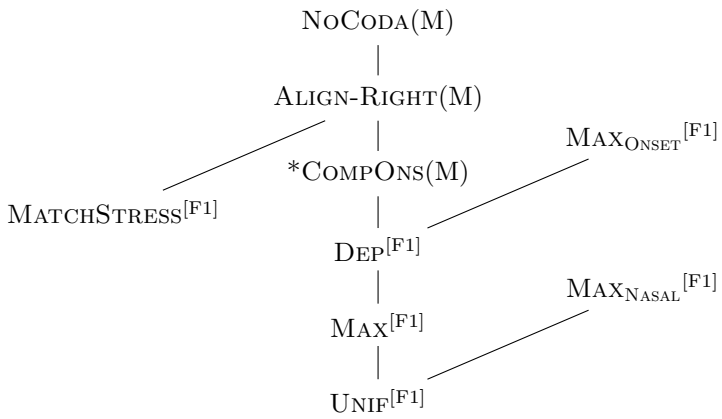
### 5.6.4 Illustration of combined stratal grammars

Having established the individual stratal grammars, we now turn to an illustrative example of how specific faithfulness constraints being ranked differently in different strata work in a single grammar. For the sake of example we will use the stratum 2 loan shown above in (72), Spanish [dɾo.ɣáɾ] → Guaraní [dɾo.ɣá].

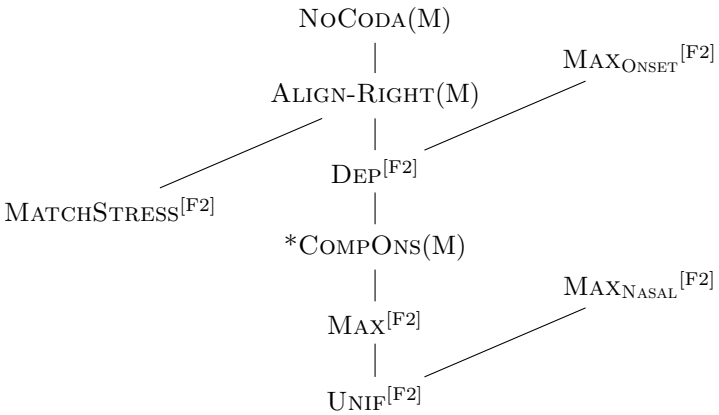
For the sake of clear illustration this example will be simplified in a few different ways. First, to limit the number of constraints, MatchStress and Align-Right will be set aside given that as this loan has word-final stress, neither of them will come directly into play. Secondly, we will see only the individualized faithfulness constraints from just two strata, again for the sake of keeping things clear. The faithfulness constraints from stratum 1 and stratum 2 (as the loan itself comes from stratum 2) will be shown with the understanding that the roles of the faithfulness constraints from strata 3 - 5 can be easily understood through this example without being explicitly represented.

Given that here will be modeled the constraints in the first and second strata, we begin with those individual stratal grammars, reproduced below in (112) and (113) respectively.

(112)



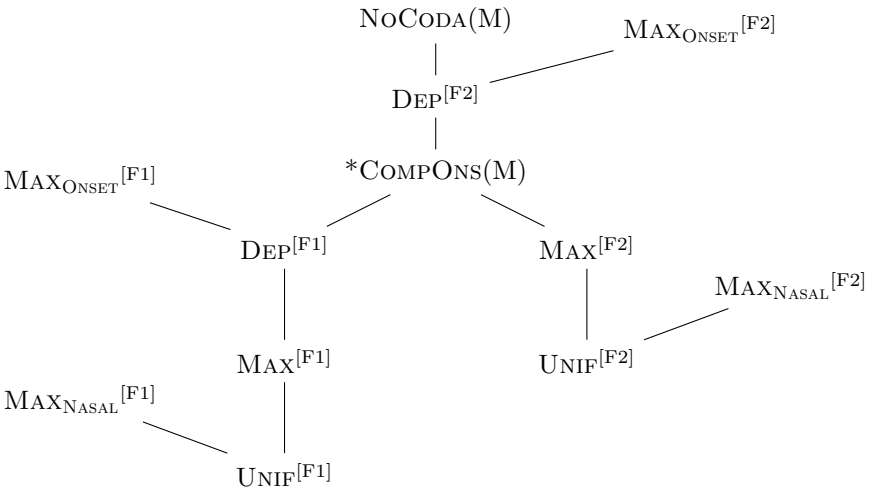

As originally discussed above, stratum 1 is characterized by the high ranking of markedness constraints in comparison to faithfulness constraints, thus producing forms which go to lengths to repair Spanish forms.




(113)

The distinguishing feature of stratum 2 when compared to stratum 1 is the ranking of Dep[F2] above *CompOns which produces the primary characteristic of loans in this stratum: the tolerance of complex onsets but neither non-final stress nor codas.

We now combine these individual stratal grammars into a single grammar, seen in (114), which would handle input forms according to what stratum they belong to.




(114)

Noteworthy here is that ranking the stratum-specific faithfulness constraints with respect to one another is generally speaking not possible given that a lack of conflict among them does not allow for ranking arguments to be made. The only exception to this is ranking through transitivity. As seen above, Dep[F2] for example can be ranked over Dep[F1] given the knowledge that Dep[F2] outranks *CompOns which in turn outranks Dep[F1].

We now turn to a tableau to show how such a grammar would actually work. One more simplification has been made in the tableau below; the Hasse diagrams above show all relevant constraints but not all constraints are needed to decide the winning output when comparing the stratum 1 and stratum 2 grammars. The more lowly ranked constraints which do not affect the winner, e.g. the Max, Max$_{Nasal}$,

and UNIF constraints, have been omitted. In addition to this, lexical stress is not overtly marked in the forms as a further simplification, given its irrelevance to this particular loan and to the point here being made.

Seen in (115) is a tableau showing the nativization of the Spanish [dɾo.ɣaɾ]. Crucially, this loan is indexed to stratum 2 and as such is only subject to the faithfulness constraints of that stratum, as previously discussed. We see the winning form here chosen as [dɾo.ɣa] ([dɾo.ɣá]), which is the attested nativization.

(115)

| /dɾo.ɣaɾ/[F2] | NoCoda | MaxOnset[F2] | Dep[F2] | *CompOns | MaxOnset[F1] | Dep[F1] |
|---|---|---|---|---|---|---|
| a. do.ɣa | | *! | | | | |
| b. do.ɣaɾ | *! | * | | | | |
| c. do.ɣa.ɾɨ | | *! | * | | | |
| d. ☞ dɾo.ɣa | | | | * | | |
| e. dɾo.ɣaɾ | *! | | | * | | |
| f. dɾo.ɣa.ɾɨ | | | *! | * | | |
| g. dɨ.ɾo.ɣa | | | *! | | | |
| h. dɨ.ɾo.ɣaɾ | *! | | * | | | |
| i. dɨ.ɾo.ɣa.ɾɨ | | | *!* | | | |

For the sake of illustration in (116) we see a tableau nearly identical to that in (115) with the exception of the hypothetical indexing of the loan to stratum 1. This clearly demonstrates how the same grammar is capable of producing different winning forms depending on which stratum a given loan is indexed to.

(116)

| |dɾo.ɣaɾ|[F1] | NoCoda | MaxOnset[F2] | Dep[F2] | *CompOns | MaxOnset[F1] | Dep[F1] |
|---|---|---|---|---|---|---|---|
| a. | do.ɣa | | | | | *! | |
| b. | do.ɣaɾ | *! | | | | * | |
| c. | do.ɣa.ɾɨ | | | | | *! | * |
| d. | dɾo.ɣa | | | | *! | | |
| e. | dɾo.ɣaɾ | *! | | | * | | |
| f. | dɾo.ɣa.ɾɨ | | | | *! | | * |
| g. ☞ | dɨ.ɾo.ɣa | | | | | | * |
| h. | dɨ.ɾo.ɣaɾ | *! | | | | | * |
| i. | dɨ.ɾo.ɣa.ɾɨ | | | | | | **! |

As expected, the winning form here is [dɨ.ɾo.ɣa]. If this loan patterned as the stratum 1 loans (i.e. if this loan were indexed to stratum 1) we would indeed expect the resulting form to repair the complex onset, and many such loans are attested in stratum 1.

The simplifying of this example to just using the relevant constraints at play in this loan, and to just using constraints from two strata, has allowed for a clear demonstration of what a simplified combined grammar could look like. Expanding on this it is not difficult to imagine the combination of all five stratal grammars, through which the entire Guaraní grammar would be produced. This combined grammar would be that which produces all loans seen in the corpus.

### 5.6.5 The problem of repair by replacement by /ɨ/

In the above stratal analyses the fact that Guaraní has two methods of repair for complex onsets (epenthesis and replacement of the second consonant by /ɨ/), as well as word-medial codas (deletion and replacement by /ɨ/) was ignored. Returning to this issue here, this variation in repair strategies is a salient way in which the facts of Guaraní seem to differ from the facts presented in Itô and Mester (1999) regarding Japanese. The presence of alternate repair strategies for a given phenomenon is something which is highly problematic for the model.

Beginning with the case of word-medial codas, the previous chapter's discussion of Guaraní repair strategies outlined how word-medial codas can either be deleted or replaced by /ɨ/. These repair strategies can be modeled through the interaction of the markedness constraint NoCoda and the faithfulness

constraints IDENT$_{\text{CONS}}$ and MAX (in this section unindexed for the purpose of demonstration). A grammar favoring replacement by /ɨ/ over deletion as a repair strategy for codas can be handled by a ranking such as NOCODA >> MAX >> IDENT$_{\text{CONS}}$ as seen in the tableau in (117); a grammar instead favoring deletion can be handled by swapping the rankings of MAX and IDENT$_{\text{CONS}}$ as seen in (118).

(117)

| \|sa.tuɾ.no\| | NOCODA | MAX | IDENT$_{\text{CONS}}$ |
|---|---|---|---|
| a. ☞ sa.tuɨ.no | | | * |
| b. sa.tu.no | | *! | |
| c. sa.tuɾ.no | *! | | |

(118)

| \|meɾ.ku.ɾjo\| | NOCODA | IDENT$_{\text{CONS}}$ | MAX |
|---|---|---|---|
| a. meɨ.ku.ɾjo | | *! | |
| b. ☞ me.ku.ɾjo | | | * |
| c. meɾ.ku.ɾjo | *! | | |

Given the condition of ranking consistency, the problem here for the model is obvious: there is simply no way to account for the variation in repair strategies for word-medial codas without being able to freely rank the two faithfulness constraints IDENT$_{\text{CONS}}$ and MAX. Even if the loans showing differing repair strategies were placed in different strata, ranking consistency does not allow for a free ranking of any faithfulness constraints.

This same problem applies in an identical way to the variation between repair of complex onsets through the interaction of the markedness constraint *COMPONS and the faithfulness constraints IDENT$_{\text{CONS}}$ and DEP, as shown in the tableaus (119) and (120).

(119)

| \|kɾo.a.sja\| | *COMPONS | DEP | IDENT$_{\text{CONS}}$ |
|---|---|---|---|
| a. ☞ kɨo.a.sja | | | * |
| b. kɨ.ɾo.a.sja | | *! | |
| c. kɾo.a.sja | *! | | |

(120)

| \|gɾe.sja\| | *COMPONS | IDENT$_{\text{CONS}}$ | DEP |
|---|---|---|---|
| a. gɨe.sja | | *! | |
| b. ☞ gɨ.ɾe.sja | | | * |
| c. gɾe.sja | *! | | |

Again the ability to freely rank faithfulness constraints, in this case IDENT$_{\text{CONS}}$ and DEP, is necessary in order to model the variation seen in the Guaraní grammar.

The notion of ranking consistency is an important aspect of the model of lexical strata as seen in Itô and Mester (1999) in that it is through ranking consistency that impossible nativizations are accounted for. While the facts of Guaraní seem to in large part line up with the model, the inability to model variation such as that shown here undermines ranking consistency which in turns undermines the prediction of impossible nativizations and by consequence the model itself.

As a side note, variation within OT has been modeled in various other ways, such as for example by using multiple grammars (Kiparsky, 1993) or Stochastic Optimality Theory (Boersma, 1998); while these theories may be better suited to account for the variation in Guaraní complex onset and coda repair, that specific issue is not the primary topic of discussion here and as such further investigation into the specifics of how these models and what they have to offer may shed light on the Itô and Mester (1999) model is a topic for future research. Exploiting the ability of these models to account for variation within OT might provide the Itô and Mester (1999) model with additional tools with which to handle grammars like that of Guaraní.

## 5.7 Pre-syllabified tableau inputs

Having analyzed the phenomena from an optimality-theoretic point of view and consequently introduced the relevant constraints, we now briefly turn back to the chapter 4 discussion of syllabification. As was discussed in detail there, Guaraní nativization processes of Spanish loans depend on the Guaraní grammar's syllabification of the Spanish form. This is illustrated by the fact that Guaraní does not treat codas and complex onsets the same with regard to repair strategies, and word-medial consonant strings may be syllabified such that a given consonant is in onset or coda position. This syllabification affects the loan's repair and ultimately changes the resulting Guaraní lexical item.

With regard to the OT tableaus, it is of theoretic importance to bring to light the necessity of the inputs to come presyllabified. For the aforementioned reasons, the syllabification and its ability to alter the resulting form are crucial to the OT analysis. While OT inputs are traditionally represented unsyllabified (Prince and Smolensky, 1993), it seems here that with the specific case of loanword tableaus, the case of Guaraní demonstrates the need for prior syllabification.

As a Guaraní speaker processes a Spanish form for nativization, the decision must be made with regard to the form's syllabification. Once syllabified however, it is appropriate to ask where this specific syllabification comes from. As discussed in the previous chapter, in most cases Spanish VCCV sequences are syllabified as VC.CV by the Guaraní grammar. For example, Spanish [kɾís.to] → Guaraní [ki.ɾi.tó], in which the /ís.to/ sequence is syllabified VC.CV accounting for the deletion of /s/ as it is processed

as a coda. Returning to the question however, where does this VC.CV syllabification come from?

One place it does not seem to come from is the well-attested markedness hierarchy in the Guaraní grammar. The markedness hierarchy (NoCoda >> Align-Right >> *CompOns) has NoCoda as the most highly ranked of wellformedness constraints and it notably outranks *CompOns. If these constraints were to influence in some way the Guaraní grammar's choice of syllabification one would indeed expect V.CCV to be the preferred syllabification of a VCCV sequence, in accordance with the constraint ranking whose highest priority is to ban codas.

If not coming from the constraint rankings, another logical possibility would be from the Spanish grammar's syllabification. Again however, as discussed in the previous chapter, there are cases in which the Guaraní syllabification of the Spanish form differs from the actual Spanish syllabification. This is the case with /mb nd ŋg/ segments which are tautosyllabic in Guaraní and not in Spanish. The illustrative loan coming from the Spanish for *Finland* shows this well given its original Spanish grammar syllabification of [fin.lán.dja] and yet the Guaraní rendering [hĺ.la.$^{n}$dja], suggesting a Guaraní grammar syllabification of the Spanish form [fin.lá.ndja], explainable by the differing adaptation behavior of the two nasal consonants.

In the end, the Guaraní perceived syllabification is not coming *solely* from the Guaraní grammar nor from the Spanish grammar, as neither seems capable by itself to account for the syllabification patterns. The patterns do interestingly show cross-linguistic principles of syllabification (VCCV → VC.CV), but the full story of where this prosodified representation comes from is a topic for future research.

# Chapter 6

# Experiment

## 6.1 Introduction

As has been discussed, a crucial aspect of the structure of the Guaraní lexicon is its core-periphery structure. This structure in conjunction with demands on the rankings of individualized faithfulness constraints produces one of the central predictions of the Itô and Mester (1999) model: the existence of nativizations which the grammar does not produce. These impossible nativizations are accounted for theoretically with the notion of ranking consistency which locks the faithfulness hierarchy such that the constraints may not be freely ranked across strata. Discussed in this chapter is an experiment whose goal was to test the sensitivity of native speakers to impossible nativizations.

Having observed the markedness constraint ranking NoCoda >> Align-Right >> *CompOns in the data, it is prudent that the experiment test each individual ranking where one constraint dominates another for impossible and possible nativizations of words which are subject to their influence. That is to say that NoCoda >> Align-Right, Align-Right >> *CompOns and NoCoda >> *CompOns must all be individually evaluated, and in the experiment nonce words were presented to the speakers which were specifically designed to do this. The model predicts that in all possible cases the impossible nativizations should be rejected as ungrammatical by the native speakers. The goal of the experiment was to investigate the effect that the constraints active in given nonce forms would have on the likelihood that impossible nativizations would be chosen.

This chapter begins with the details of the experimental design and methodology. Following are the details of the results of the experiment in turn followed by a discussion of the linguistic implications of the results in light of the model of lexical strata under discussion in previous chapters. The experiment is shown to prove problematic for the model, as the condition of ranking consistency does not allow for an interpretation of the results consistent with the model. The experimental results show that speakers are sensitive to impossible nativizations under some contexts but not all, a finding which simultaneously

calls into question ranking consistency while providing direct evidence for the synchronic relevance of lexical strata in Guaraní.

## 6.2 Methods

### 6.2.1 Participants

The participants were either Guaraní speakers which the author knew or speakers found online from either the social media site Facebook[35] or the online language forum WordReference.[36] Members contacted via Facebook were found on the basis of their membership in Guaraní language groups on the site. WordReference is an online community of over a half a million language learners from all over the world; participants there were found using member searches, filtered by references to "Guaraní" or "Paraguay" in user profiles. On both websites users were contacted via private message. After an initial message verifying that participants were both willing to participate and that their native language (or one of them) was Guaraní, the questionnaire was delivered and then returned upon completion. In addition to judgments about nonce words, participants were asked to provide their age, country of origin and native language(s).

A total of eight participants participated in the experiment. Each of the participants was a native speaker of Guaraní and each was fluent in Spanish as well. Seven of the eight participants were of Paraguayan nationality while the eighth was from Argentina.[37]

### 6.2.2 Materials

In the electronic questionnaire participants were presented with 16 nonce Spanish words. The nonce forms were created with specific attention to native Spanish phonotactics. They are based on real Spanish words, with changes to one or two phonemes being made to create words which do not exist in the Spanish lexicon. The words were presented to a native Spanish speaker who assisted with their forms until all seemed plausible as native Spanish forms.

Each nonce word presented to participants came accompanied by two possible Guaraní adaptations of the word, labeled "A" and "B". Next to the two options was a blank in which participants typed in "A" or "B" in order to make their choice of which seemed to be the more natural of the two adaptations.

[35] http://www.facebook.com/

[36] http://forum.wordreference.com/

[37] Presumably the northern part of the country where Guaraní is spoken natively.

Of the 16 nonce words there were four individual groupings, each consisting of four forms to produce a total of 16. Three of the four groupings, representing 12 of the 16 forms, were to test each constraint ranking referenced in the above section. The first grouping is concerned with evaluating the ranking NoCoda>>Align-Right, and as such its Spanish forms contain both a coda and non-final lexical stress but not a complex onset. The second group evaluates the ranking Align-Right>>*CompOns and its forms therefore contain non-final lexical stress and a complex onset but no coda. The third group was designed to evaluate the ranking NoCoda>>*CompOns and therefore contains nonce Spanish forms with both a coda and a complex onset, as well as word-final lexical stress to eliminate interference from Align-Right. The remaining four forms were fillers in which vowels were changed to help occlude the purpose of the experiment from participants and were of no relevance to the testing of the constraint rankings.

Found in (121) is a summary of the forms of theoretical interest to the experiment, organized by grouping; in (122) are found the fillers. In both examples the respective orthographic representations are given in parentheses.[38]

(121)

| **Relevant constraint ranking** | **Spanish nonce form** | **Option 1 (predicted possible)** | **Option 2 (predicted impossible)** |
|---|---|---|---|
| NoCoda>> Align-Right | pes.té.sa (*pesteza*) | pe.té.sa (*petésa*) | pes.te.sá (*pestesá*) |
| | xís.to (*gisto*) | hí.to (*híto*) | his.tó (*histó*) |
| | twáɾ.to (*tuarto*) | twá.to (*tuáto*) | twaɾ.tó (*tuartó*) |
| | gól.de (*golde*) | gó.de (*góde*) | gol.dé (*goldé*) |
| Align-Right>> *CompOns | plá.βo (*plavo*) | pla.vó (*plavó*) | pa.lá.vo (*palávo*) |
| | tɾá.sja (*tracia*) | tɾa.sjá (*trasiá*) | ta.ɾá.sja (*tarásia*) |
| | kɾí.sjo (*cricio*) | kɾi.sjó (*krisió*) | ki.ɾí.sjo (*kirísio*) |
| | glá.βo (*glabo*) | gla.vó (*glavó*) | ga.lá.vo (*galávo*) |
| NoCoda>> *CompOns | plo.mél (*plomel*) | plo.mé (*plomé*) | pɨ.lo.mél (*pylomél*) |
| | tɾe.ðás (*tredaz*) | tɾe.dá (*tredá*) | tɨ.ɾe.dás (*tyredás*) |
| | gɾu.βás (*grubaz*) | gɾu.vá (*gruvá*) | gu.ɾu.vás (*guruvás*) |
| | bla.sál (*blazal*) | bla.sá (*blasá*) | ba.la.sál (*balasál*) |

[38]In Guaraní orthography lexical stress is only marked when it is non-final; for the purpose of this experiment however stress in the two options for Guaraní adaptations is always overtly marked even when unnecessary to avoid any Spanish orthographic interference with lexical stress position.

(122)

| Spanish nonce form | Option 1 | Option 2 |
|---|---|---|
| bɾí.se (*brice*) | bɾi.sé (*bricé*) | bɾú.se (*brúce*) |
| ka.βá.lo (*cavalo*) | ka.va.ló (*kavaló*) | ka.vá.lu (*kaválu*) |
| pár.so (*parso*) | pá.so (*páso*) | pá.su (*pásu*) |
| u.ɲál (*uñal*) | u.ɲá (*uñá*) | u.ɲa.lá (*uñalá*) |

For the purposes of randomization there were two different versions of the form, *a* and *b*, each being administered in half of the cases. The versions are identical in all ways, with the exception of the order of the 16 tasks. Both the orders for version *a* and those for version *b* were randomized using a simple script written in the programming language Perl.

With regard to internal randomization, the questionnaire (both *a* and *b*) was set up such that if participants performed as expected, that is if they always chose the predicted possible nativization as opposed to the predicted impossible nativization, they would end up picking half "A" responses and half "B". This was also done on the level of the four groupings within the 16 forms, such that in each respective grouping to pick the predicted possible nativizations would be to pick two "A" responses and two "B" responses.

The questionnaire itself can be found in appendix B in both the original Spanish form as it was delivered to speakers as well as an English translation for reference. Version *a* is that found in Appendix B; version *b* was not included given that save the order of presentation of the nonce forms it is identical to version *a*. Also included are the "expected results" for the questionnaire in which the predicted possible nativizations have been chosen.

### 6.2.3 Experimental procedure

The experiment was administered as an electronic questionnaire (.doc file) entirely in Spanish. It was sent to the participants who filled it out electronically and then sent it back. In the questionnaire speakers were given a Spanish nonce word along with two corresponding Guaraní forms of the Spanish word. Participants were informed that it was a nativization test, and then were asked to pick which of the two given Guaraní forms seemed the most natural as an adaptation of the given Spanish word. They were urged to pick the best of the two options even in the event that both seemed strange or some unlisted option was their pick for the nativization of the word given. Time limits on the questionnaire were not imposed, and the only mention of decision time was to encourage participants not to overthink the decision but rather go with their instinct.

## 6.3 Results

As mentioned before, the goal of the experiment was to investigate the effect that the constraints active in given nonce forms would have on the likelihood that impossible nativizations would be chosen. Given that in the questionnaire each grouping had four tasks, and that there were eight participants, each grouping produced a total of 32 judgments. Presented in (123) are the results of the possible and impossible nativizations chosen for each of the three groupings, as well as a summary of all three groupings together. "Possible nativization rate" refers to the percentage of the time that possible nativizations, as predicted by the model, were chosen over impossible nativizations.

(123)

| **Constraint ranking being tested (i.e. grouping)** | **Possible nativizations chosen** | **Impossible nativizations chosen** | **Possible nativization rate** |
|---|---|---|---|
| NoCoda >> *CompOns | **31** | **1** | **96.88%** |
| NoCoda >> Align-Right | **19** | **13** | **59.38%** |
| Align-Right>> *CompOns | **16** | **16** | **50%** |
| Overall | **66** | **30** | **68.75%** |

Statistical analysis of the data bears out the conspicuous: the behavior of the group concerned with NoCoda>>*CompOns is largely different than the other two groups. To statistically verify this, a Wald Chi-Square test was performed in the context of logistic regression accounting for multiple observations within subjects. Seen in (124) is a reference used for the grouping keys used in (125), in which the results of the statistical analysis are presented.

(124)

| **Constraints** | **Grouping** |
|---|---|
| NoCoda >> *CompOns | 1 |
| NoCoda >> Align-Right | 2 |
| Align-Right>> *CompOns | 3 |

(125)

| Grouping comparison | Degrees of freedom | Wald Chi-Square value | $p$-value |
|---|---|---|---|
| 1 vs. 2 & 3 | 2 | 11.4774 | **0.0032** |
| 1 vs. 3 | 1 | 10.4476 | **0.0012** |
| 1 vs. 2 | 1 | 5.6362 | **0.0176** |
| 2 vs. 3 | 1 | 0.0827 | **0.7737** |

In the first grouping comparison, we compare the first grouping with the other two as a means of determining whether or not its unique behavior is statistically significant; the $p$-value for this comparison is 0.0032 thus allowing us to say that grouping 1 is statistically different from the others. Subsequent comparisons seen in (125) show that when comparing grouping 1 individually to grouping 2 as well as 3, statistically significant results are also achieved. The final comparison, that of groupings 2 and 3, bears out the expected result that their difference is far from statistical significance.

For the purpose of discussing the results, a reminder of the markedness hierarchy in the Guaraní phonological grammar motivated in previous chapters is reproduced in (126).

(126)

NoCoda

|

Align-Right

|

*CompOns

In light of the markedness hierarchy, an initial interpretation of the data shows two highly interesting trends. The first is the high level of predicted possible nativizations chosen by native speakers in the group of words testing the interaction of NoCoda and *CompOns. As seen in the hierarchy, these two constraints are the farthest separated with respect to each other, as NoCoda occupies the highest rank with *CompOns occupying the lowest rank. In only one single word did one of the speakers choose an impossible nativization in this category, and the numbers largely speak for themselves with regard to native speaker perception of the perceived acceptability of impossible nativizations in this group.

In the other two groups however, those designed to test NoCoda >> Align-Right and Align-Right >> *CompOns, this trend toward perceiving the predicted impossible nativizations as illegitimate was not borne out. The nonce forms testing the interaction between NoCoda and Align-Right showed participants choosing the possible nativization at a rate of 58.38%, a near 50/50 split, while the forms testing Align-Right and *CompOns indeed did show a clean 50/50 split.

The overall results showed that participants picked predicted possible nativizations 68.75% of the time, or 66 times out of 96 tasks. Given that the model predicts that no impossible nativizations should be chosen and that there was such a discrepancy between the NoCoda >> *CompOns group and the other two, it is clear that while the prediction of the model was not found to be upheld by the results of the experiment some effect was clearly demonstrated.

The clear delineating factor between the groups comparing the interaction of NoCoda and *CompOns and the other two is the distance in the hierarchy that separates NoCoda and *CompOns. They are the most distant with respect to each other and it appears that this distance is related to the increase in the likelihood of native speakers to perceive impossible nativizations in this category as unacceptable. Comparing NoCoda >> Align-Right and Align-Right >> *CompOns, in both sets of constraints the individual constraints are adjacent to one another in the hierarchy and this has seemingly had the consequence of speakers being less sensitive to predicted impossible nativizations with respect to these comparisons.

## 6.4 Discussion

### 6.4.1 Relevance of results for the model

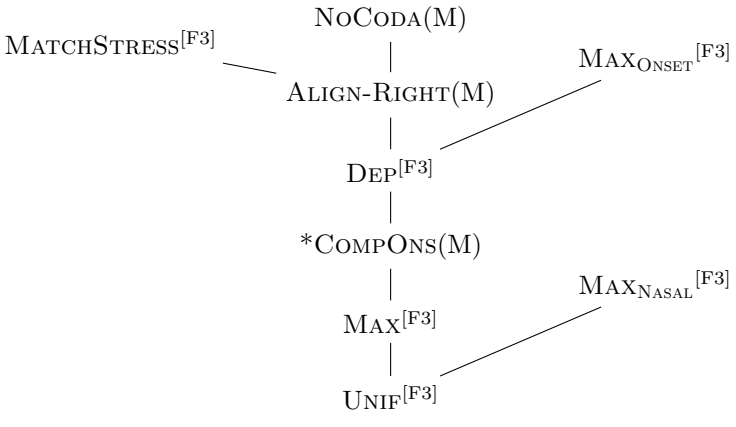
The results of the experiment indicate that in some cases speakers are sensitive to impossible nativizations and in other cases they are not. This is significant due to the fact that the model predicts that impossible nativizations should be categorically avoided by native speakers. As shown in (123), in the overall results impossible nativizations were picked in 30 of the 96 tasks (at a rate of 31.25%). Impossible nativizations being picked nearly one third of the time shines light on the ranking consistency condition described in the previous chapter. Ranking consistency is necessary when individualizing the faithfulness constraints in order to maintain the model's predictions of impossible nativizations, yet the results of the experiment suggest that speakers are not uniformly adverse to grammars which yield impossible nativizations as a result. This, in conjunction with the before-raised problems of ranking consistency with regard to variable adaptation strategies in Guaraní, further raises questions about the plausibility of ranking consistency with regard to its application to Guaraní. A speaker who actively prefers a predicted impossible nativization over one which is predicted to be possible is in effect advocating a grammar in which the relevant rankings of the faithfulness constraint are reversed.

Using an example from the experiment, the nonce Spanish loan [xís.to] was one of the four nonce words utilized to compare the interaction of NoCoda and Align-Right. The two options given as

nativizations were [hí.to] (predicted possible) and [his.tó] (predicted impossible). In a grammar operating under the condition of ranking consistency, [his.tó] is predicted as impossible for the following reason.

In the previous chapter, the grammar producing stratum 3 was shown to necessitate the ranking of MatchStress[F3] above Align-Right in order to account for the behavior of loans in that stratum with regard to tolerating non-native (i.e. non-final) stress. The grammar producing this stratum is reproduced below in (127).

(127)




Through transitivity, the ranking MatchStress[F3] >> Max[F3] is uncontroversially established. Given the fixed markedness hierarchy we also know the ranking NoCoda >> Align-Right. Given these two rankings, the condition of ranking consistency will not allow for the form [his.tó] to be picked as optimal no matter what the grammar. The impossible form [his.tó] requires a grammar in which Max is ranked above NoCoda to account for the presence of the coda, yet in order for final stress to be observed Align-Right must dominate that grammar's MatchStress. The conflict here lies in the fact that given the fixed markedness ranking NoCoda >> Align-Right, there is no possible way for a grammar to rank Max high enough to allow for the coda without it simultaneously outranking MatchStress, a ranking which is forbidden by ranking consistency given the grammar for stratum 3.

Returning to the issue of the nonce Spanish loan [xís.to] in the experiment, it is notable due to the fact that amongst the eight participants, four chose [hí.to] (predicted possible) and four chose [his.tó] (predicted impossible). The four which chose the impossible [his.tó] are by means of their rejecting [hí.to] for [his.tó] advocating a grammar in which Max must outrank MatchStress. The notion of ranking consistency explicitly forbids this and it is in this sense that the experimental results place the notion of ranking consistency under scrutiny.

A potential modification to the model which would let it more satisfactorily account for the experimental results would be the removal of ranking consistency and the addition of some form of recognition of distance effects on the predicted impossible nativizations. The model must account for the fact that

the farther apart the constraints on the hierarchy, the more averse speakers are to impossible nativizations concerning those two constraints. OT-derived models such as Stratal OT (Boersma and Hayes, 2001) do take into account constraint distance and provide at least one example of how representing it is possible using OT. The tendency for speakers to *not* avoid impossible nativizations in situations comparing adjacent constraints is reason to rethink ranking consistency and to look for more appropriate methods of accounting for gaps in nativization patterns.

### 6.4.2 Further discussion

One alternative interpretation of the experimental results might be that participants chose "possible" versus "impossible" nativizations on the basis of which of the two forms presented contained structures that are more frequently attested. This explanation however does not seem consistent with the distribution of the results across the three groups as seen in (123). As shown, the group with the outstanding results is group 1 in which words presented have both codas and complex onsets. Groups 2 and 3 which by and large show similar results compared to one another are those whose words are meant to test sensitivity to codas/stress and stress/complex onsets, respectively. If frequency sensitivity were a primary motivator for the choices of the participants then one would expect results skewed in such a way as to avoid whichever of the structures was that being avoided on account of its infrequency. For example, if in a hypothetical scenario speakers were avoiding codas on the basis that they were less statistically frequent, then the results of group 2 (the group whose forms had codas and non-final stress) would be expected to show skewing in favor of the forms without codas, and yet in this group no such significant trend was attested. A hypothetical example involving complex onsets and group 3 shows the same. The very experimental design involving the three groups testing each pairing of relevant structures severely mitigates the plausibility of this interpretation given that the results showed no specific structure consistently avoided in the groups.

Another potential explanation for the discrepancy might be that group 1 is different from groups 2 and 3 as a result of participants outright ignoring stress and consequentially not seeing any impossible nativizations in groups 2 and 3, on account of the stress information being disregarded. This alternative explanation seems inadequate however given that the role of stress in the experiment was if anything drawn attention to (unintendedly yet unavoidably) by the fact that the stress of a word was always overtly marked even in cases when the Guaraní orthographic rules do not permit this. This was necessary to block possible interference from Spanish orthography, which has much different rules for representing lexical stress. That stress was seen by the participants as relevant in some way was further evidenced by

the fact that several of those who participated in the experiment (and indeed all of those who had any speculation or comment on what the purpose of the experiment might have been) replied after having completed the form with a brief discussion or mention of stress in Guaraní, and the tendency of Guaraní to have words, including loans, which are stressed word-finally. Based on the replies of the participants it seems that the role of lexical stress was perhaps more conspicuously important to the overall experiment than was that of codas or complex onsets.

A potential interpretation for the results in the group which compared complex onsets and codas, in which the predicted possible nativizations were picked over 96% of the time, might be that speakers were not paying attention to impossible nativizations but rather to the fixed markedness hierarchy, and were making their choices on the basis of which word violated the highest ranked markedness constraint. For example, one of the nonce words in the category was Spanish [gɾu.βás], for which the two given nativization options were the predicted possible nativization [gɾu.vá] and the predicted impossible [gu.ɾu.vás]. The argument might be made that speakers were making their choice not on the basis of "possible" versus "impossible" but rather on the fact that the most highly ranked markedness constraint is NoCoda, and this form violates that constraint while the other does not. This explanation however does not account for the fact that the other two groupings showed drastically different effects in which speakers picked impossible nativizations over possible ones nearly half of the time. If speakers were making their decision on the basis of which of the forms violated the most highly ranked markedness constraint, that pattern would be expected to hold across all constraint rankings but in this case is found only in the ranking of NoCoda >> *CompOns.

The position that speakers were not paying specific attention to which of the violated markedness constraints is most highly ranked but rather the distance between the violated markedness constraint and the satisfied markedness constraint is less easily dismissed by the results of the experiment. While the results were enlightening in showing that a distance effect between markedness constraints is attested, further expansion of the experiment would shed further light on this matter. Returning to the example of [gɾu.βás] and its two choices (the predicted possible nativization [gɾu.vá] and the predicted impossible [gu.ɾu.vás]), a subsequent experiment might also include the other two possibilities of nativization: [gu.ɾu.vá] and the faithful [gɾu.vás], both of which are predicted to be possible. If in such an experiment speakers were asked to order the four forms from most acceptable to least acceptable, stronger evidence regarding the effect of constraint rankings (and constraint distance) would be readily available. This would be possible through a compiling of all responses of all speakers in order to get an average ranking of the four forms by native speakers, thus allowing for insight into which forms are perceived as more acceptable than others (and consequently insight into the roles of the interaction of the constraints in

question in a given form). In such a scenario the tendency of speakers to pick the same rankings, or perhaps the tendency to pick the same "best" or "worst" forms but vary with regard to the in-between forms, would help get at the specific nature of the distance effects seen here. For example, one might expect that speakers show less response variation for forms testing constraints more distant from each other on the hierarchy, while forms testing adjacent constraints may cause more variation indicating weaker judgments.

Problematic for this experiment design might be the difficulty for speakers of the task and the consequential variability amongst speakers in their rankings of the forms. This could be overcome however by substantially increasing the number of participants in order to get more reliable results. The scope of this project did not allow for such an experiment, yet the results attested here shine light on the fact that further exploration of the sensitivity of speakers to the constraint rankings could lead to highly interesting and statistically persuasive results.

The decision to present speakers with just two options to choose from as nativizations was made with several factors in mind. Ease of the task for the participant was a large factor, especially given that it was designed to be sent as an electronic document. While more information could have been gleaned from a task in which four possible nativizations were given as choices, narrowing the choices down to two still allowed the task to shine light on whether or not speakers would under any circumstance choose impossible nativizations, while making the decision easier on the participants. Yet another reason for this was for the anticipated low participant number for the experiment, which ended up being borne out in the fact that just eight speakers were found.[39] Only allowing two choices allowed for more robust statistical generalizations to be made in light of the paucity of participants.

The experiment carried out and described here is informative by showing the willingness of native speakers of Guaraní to choose nativizations predicted by the model to be impossible over nativizations predicted to be possible. In one environment however, in which the two most distant markedness constraints are compared, speakers nearly completely avoided impossible nativizations. This interesting trend lays the foundation for further experiments of this type in order to reach conclusions which are more conclusive in explaining precisely why the experimental results turned out the way they did; the role of distance effects in the Guaraní markedness hierarchy and the effects they have on loan adaptation and in the Guaraní grammar as a whole provides an intriguing avenue of further research.

[39]Scores of messages were sent out on various social websites and to anyone who was identifiable as a potential Guaraní speaker. The majority of messages went unanswered, while some speakers did reply saying that they spoke Guaraní but not natively. These factors in conjunction only allowed for eight willing native speakers of Guaraní to be tested.

### 6.4.3 Evidence for synchronic relevance of lexical strata

In a broader sense, the experimental results provide direct evidence for the synchronic relevance of strata. Some recent work in loanword phonology has called into question the stratal interpretation of loans and favored an interpretation saying that loans in effect expand that native grammar itself and that individual strata are unnecessary (see Rice (2006) for such a proposal with the case of stress adaptation in Norwegian). Under such an interpretation of the facts in Guaraní one would argue that Guaraní has in essence become stratum 5 and is now fully tolerant of any and all phonological characteristics found in the wealth of loans from Spanish.

Assuming this to be the case however, one would expect to see no impossible nativization effects whatsoever in the results of the experiment. This leads from the fact that if Guaraní is now tolerant to codas, complex onsets and non-final stress, then any combination of these elements should be perceived as perfectly grammatical by speakers, as the grammar now allows them. The notion of impossible nativizations regarding repair strategies of these three phenomena should be completely unattested if the strata were without synchronic relevance, and yet the group in which impossible nativizations were avoided over 96% of the time makes clear the fact that impossible nativizations are not entirely fictitious in modern Guaraní. This clearly leads to the necessity of addressing strata as something of synchronic relevance to the Guaraní grammar and not merely as a relic of previous periods in the language's history.

# Chapter 7

# Conclusion

Presented here was an analysis of the phonological adaptation strategies of Spanish loans in Guaraní and a discussion of what the patterns of adaptation tell us about the structure of the Guaraní lexicon. Loans imported into Guaraní from Spanish show a wide variety of adaptation patterns, ranging from loans repaired to be fully compliant with native Guaraní phonology, loans which show only partial adaptation and loans which show no adaptation whatsoever and are imported as is from Spanish.

An analysis of the corpus of Spanish loans in Guaraní presented in this thesis reveals the stratal nature of the Guaraní lexicon. Loans form groupings on the basis of their phonological characteristics, and these strata are grouped in a set-inclusion pattern in which their phonological characteristics are seen to overlap and stack upon each other. In this sense the lexicon is characterized by its core-periphery structure, and this crucial aspect of the lexicon is what provides evidence for the stratification patterns seen in Guaraní.

The model in Itô and Mester (1999) applied to Guaraní is shown to account for much of the characteristics of the structure of the Guaraní phonological lexicon. As discussed in this work, a language with synchronically relevant strata shows impossible nativization effects in its adaptation of loanwords, and these impossible nativizations are strongly evidenced in Guaraní by both its lexicon and the results of the nonce experiment discussed in the previous chapter.

The lexical strata in Guaraní are shown to be more than mere historical relics and are indeed synchronically relevant. This is evidenced not only by features of Guaraní phonology and morphology but by the experimental results as well. An account of synchronically irrelevant lexical strata would expect speakers to show no sensitivity to impossible nativizations given the acceptance of the natively-forbidden phonological structures. Such an account is incapable of handling the results of the experiment given that speakers nearly totally avoided impossible nativizations of those Spanish nonce words containing codas and complex onsets.

In the experiment speakers were shown to prefer predicted impossible nativizations over predicted

possible ones nearly one third of the time. Impossible nativizations were avoided in the grouping of words which involved the two most distant markedness constraints from one another on the markedness hierarchy of the grammar. In addition to showing possible distance effects with regard to markedness constraints, this has provocative theoretical implications for theories of loanword phonology predicting total speaker sensitivity to impossible nativizations (Itô and Mester, 1999) as well those predicting no speaker sensitivity whatsoever (Rice, 2006). That speakers showed any sensitivity at all supports the notion that stratal patterns in Guaraní are still productive and that the grammar overall has not merely become openly and indiscriminately permissive of Spanish phonological characteristics.

The condition of ranking consistency proposed in Itô and Mester (1999) whereby faithfulness constraints are held consistent in their ranking to one another across strata cannot completely account for the facts shown here regarding Guaraní. The experimental results show speaker willingness to choose impossible nativizations; this advocating for grammars which could only be produced by constraint rankings forbidden by ranking consistency proves problematic for the model. In addition to this, problems arise for ranking consistency in the variation of Guaraní adaptation strategies for some Spanish structures. Accounting for these variations within the model necessitates free ranking at least to some extent of the individual faithfulness constraints, something forbidden by the model.

Future studies proposed include investigation of the source of the Guaraní syllabification of Spanish forms, which was shown to not be able to be accounted for on the basis of the Guaraní or Spanish grammars alone. While not the central topic of investigation of this thesis, the necessary prosodified inputs of Guaraní loanword tableaus and their theoretical ramifications also merit further investigation. In addition to this, modifications to the experiment to more clearly investigate the role of distance effects and sensitivity to impossible nativizations would also benefit from subsequent research. Both additions to and subtractions from the model presented in Itô and Mester (1999) are also shown to be necessary in order to account for the facts in Guaraní, and ways of accounting for variable repair strategies as well as tolerance of predicted impossible nativizations make for sensible avenues of future research of both the Guaraní lexicon and the lexica of other languages as a means of finding more out about general patterns of lexical organization and what they mean for the organization of grammars in a larger sense.

# Appendix A

# Corpus

In this appendix are the 177 words comprising the corpus of Spanish loans in Guaraní used for and discussed in this thesis. The loans are listed in alphabetical order (by the transcription of the original Spanish word) and come with the references for their source(s). Loans listed with two sources can be found in identical forms in either. The source reference key is included in a table before the corpus itself. Sources which include page numbers are those sources which are not organized in such a way that the loan is easily retrievable (i.e. are not organized alphabetically such as in the case of dictionaries).

All transcriptions are standard IPA with the exception of the use of an acute accent mark to indicate the nucleus of the syllable bearing lexical stress. Some multisyllabic loans have no accent mark and this is due to source ambiguity; in the absence of reliable information about stress it was left out.

Syllabification of the Spanish forms does not represent the Spanish grammar syllabification but rather the perceived syllabification by the Guaraní grammar, which is relevant in the determination of phonological repairs by this grammar and as such included as a tool by which to compare the Guaraní repaired lexical item to the original Spanish lexical item.

The two loans coming from Velázquez-Castillo (2013) are taken from a handout of a talk given by the author at UNC-Chapel Hill on April 12, 2013. Loans used in the handout have their ultimate origin in Zarratea (1981).

Those loans coming from the Guaraní Wikipedia were chosen only if they were the title of an article.

## A.1 Source Key

| | |
|---|---|
| 1 | Britton (2005) |
| 2 | Morínigo (1931) |
| 3 | de Canese and Alcaraz (1997) |
| 4 | Mayans (1980) |
| 5 | Díaz (2006) |
| 6 | de Assis (2008) |
| 7 | Rendon (2008) |
| 8 | Lustig (2005) |
| 9 | Fritz (2004) |
| 10 | Tonhauser and Colijn (2010) |
| 11 | Velázquez-Castillo (2013) |
| 12 | Tonhauser et al. (2013) |
| 13 | Wikipedia (2012) |

## A.2 Corpus

| **Spanish** | **Guaraní** | **English gloss** | **Source(s)** |
|---|---|---|---|
| a.ko.pjáɾ | akopjá | 'to stock up' | 3, 6 |
| a.la.kɾán | alakɾán | 'scorpion' | 2 |
| al.bá.nja | avaɲá | 'Albania' | 13 |
| a.le.má.nja | alemaɲá | 'Germany' | 13 |
| al.kóːl | alkó | 'alcohol' | 2, 6 |
| al.ma.sén | almasẽ́ | 'department store' | 6, 2 |
| al.mi.ðón | aɾamiɾṍ | 'bag' | 4, 7 (p. 279) |
| al.mo.á.ða | aɾmoxá | 'pillow' | 7 (p. 278) |
| a.mén | amén | 'amen' | 9 (p. 12) |
| a.mé.ɾi.ka | amérika | 'America' | 13 |
| a.nís | aní | 'anise' | 2, 6 |
| a.ntó.njo | toní | *proper name* | 8 (p. 94) |
| a.pós.tol | apostol | 'apostle' | 9 (p. 54) |
| a.ɾá.βja | aɾávja | 'Arabian peninsula' | 13 |
| áɾk.ti.ko | áɾktiko | 'arctic (ocean)' | 13 |
| a.rós | aɾó | 'rice' | 3, 6 |
| a.ró.ʎo | aród͡ʒo | 'stream' | 6 |
| aɾ.xé.lja | aɨheljá | 'Algeria' | 13 |
| a.sú.kaɾ | asuká | 'sugar' | 1, 6 |
| a.te.ndéɾ | ate$^{n}$dé | 'to pay attention to' | 7 (p. 279) |
| at.lá.nti.ko | atlá$^{n}$tiko | 'atlantic (ocean)' | 13 |
| aws.tɾá.lja | awtaɾaljá | 'Australia' | 13 |

*Continued on next page*

*Continued from previous page*

| Spanish | Guaraní | English gloss | Source(s) |
|---|---|---|---|
| áws.tɾja | awteɾjá | 'Austria' | 13 |
| a.xé.no | ahe᷄no | 'another's' | 7 (p. 278) |
| a.βjón | avjo᷄ | 'airplane' | 6, 2 |
| bá.ka | vaká | 'cow' | 3, 6 |
| ba.ké.ɾo | vakéro | 'cowboy' | 2 |
| ba.lér | valé | 'to be worth' | 2, 6 |
| baw.tís.mo | vawtismo | 'baptism' | 9 (p. 42) |
| baw.tís.ta | bawtista | 'baptist' | 9 (p. 32) |
| beɾ.nár.ðo | vená | *proper name* | 8 (p. 94) |
| bi.ná.gɾe | vinágre | 'vinegar' | 2 |
| bí.no | víno | 'wine' | 2 |
| bí.t͡ʃo | víʃo | 'bug' | 1, 6 |
| bo.lí.t͡ʃe | volíʃo | 'store' | 1, 6 |
| bo.lí.βja | volívja | 'Bolivia' | 13 |
| ból.sa | vosá | 'bag' | 6, 7 (p. 281) |
| bo.mbí.ʎa | gombíla | *straw used with maté* | 2 |
| bo.mi.tár | gomitá | 'to vomit' | 1, 3 |
| bo.rí.ka | ᵐburiká | 'donkey' | 2, 3 |
| bɾa.síl | vɾasíl | 'Brazil' | 13 |
| bɾo.mís.ta | ᵐbɾomísta | 'funny' | 7 (p. 279) |
| bu.tán | vuta᷄ | 'Bhutan' | 13 |
| bwéj | wéj | 'ox' | 1, 6 |
| bwé.no | wéno | 'good' | 2, 6 |
| di.na.már.ka | ⁿdinamaɨká | 'Denmark' | 13 |
| dok.tór | doɨtó | 'doctor' | 3, 7 (p. 279) |
| do.ló.ɾes | loló | *proper name* | 8 (p. 94) |
| dɾo.ɣár | dɾo.ɣá | 'to drug' | 12 (p. 80) |
| e.kwa.ðór | ekwatóɾ | 'Ecuador' | 13 |
| e.lá.ða | eláda | 'iced' | 3 |
| en.sa.lá.ða | ensaláda | 'salad' | 2 |
| e.nté.ɾo | eⁿtéro | 'entire' | 3 |
| e.ɾí.ða | eɾída | 'wound' | 10 (p. 259) |
| e.ɾi.tɾé.a | eɾitɨɾeá | 'Eritrea' | 13 |
| eɾ.má.na | eɾmána | 'sister' | 12 (p. 97) |
| eɾ.má.no | eɾmáno | 'brother' | 6 |
| es.kí.na | ekína | 'corner' | 2, 6 |
| es.kɾi.βíɾ | kɾiví | 'to write' | 3, 6 |
| es.kwé.la | ekwéla | 'school' | 2, 6 |
| es.pi.ná.so | epinasó | 'spine' | 2, 6 |
| es.pí.ɾi.tu | espíɾitu | 'spirit' | 9 (p. 12) |

*Continued on next page*

*Continued from previous page*

| Spanish | Guaraní | English gloss | Source(s) |
|---|---|---|---|
| es.tó.nja | etoɲá | 'Estonia' | 13 |
| ew.ka.ɾis.tí.a | ewkaɾistía | 'eucharist' | 9 (p. 42) |
| fal.táɾ | vatá | 'to lack' | 6, 7 (p. 279) |
| fin.lá.ndja | hḯla$^{n}$dja | 'Finland' | 13 |
| flo.ɾi.pón | floɾipón | *flower species* | 2 |
| fɾan.sís.ka | ʃiká | *proper name* | 8 (p. 94) |
| frán.sja | hiã́sja | 'France' | 13 |
| fút.bol | húvol | 'soccer' | 13 |
| gɾa.ná.ða | gɾanadá | *kind of fruit* | 6 |
| gɾá.sja | gɾasjá | 'joke', 'grace' | 1, 4 |
| gɾé.sja | giɾesjá | 'Greece' | 13 |
| í.ndi.ko | í$^{n}$diko | 'indian (ocean)' | 13 |
| i.ŋgla.té.ɾa | i$^{ŋ}$giaterá | 'England' | 13 |
| i.ɾán | iɾán | 'Iran' | 13 |
| iɾ.lá.nda | ila$^{n}$djá | 'Ireland' | 13 |
| is.lám | islã́ | 'islam' | 13 |
| is.lá.ndja | iɨla$^{n}$dá | 'Iceland' | 13 |
| je.mén | d͡ʒemẽ́ | 'Yemen' | 13 |
| kal.són | kasṍ | 'pants' | 2 |
| ka.mí.sa | kamísa | 'shirt' | 3 |
| ka.mjón | kamjṍ | 'truck' | 6, 2 |
| ka.na.ðá | kanatã́ | 'Canada' | 13 |
| ka.né.la | kanéla | 'cinnamon' | 2 |
| ka.pa.tás | kapatá | 'overlord' | 6, 7 (p. 280) |
| káɾ.los | kaló | *proper name* | 8 (p. 94) |
| kaɾ.ló.ta | kaló | *proper name* | 8 (p. 94) |
| káɾ.men | kamé | *proper name* | 8 (p. 94) |
| ka.te.kís.ta | katekista | 'catechist' | 9 (p. 42) |
| ka.tó.li.ka | katólika | 'Catholic' | 9 (p. 12) |
| ka.βá.ɾo | kavad͡ʒú | 'horse' | 1, 6 |
| káβ.ɾa | kavaɾá | 'goat' | 1, 5 |
| ké.so | kesú | 'cheese' | 1, 6 |
| ko.ka.í.na | kokaína | 'cocaine' | 6 |
| ko.lek.tí.βo | kolektívo | 'bus' | 12 (p. 85) |
| ko.ló.mbja | kolóᵐbja | 'Colombia' | 13 |
| kol.t͡ʃón | koʃṍ | 'mattress' | 2 |
| ko.mpu.ta.ðó.ɾa | koᵐbutadóɾa | 'computer' | 3, 6 |
| ko.mu.ni.ðáð | komunidá | 'community' | 9 (p. 30) |
| kon.fiɾ.ma.sjón | konfiɾmasjón | 'confirmation' | 9 (p. 42) |
| kons.tán.sja | kotá | *proper name* | 8 (p. 94) |

*Continued on next page*

*Continued from previous page*

| Spanish | Guaraní | English gloss | Source(s) |
|---|---|---|---|
| ko.rál | korá | ‘corral’ | 2, 6 |
| ko.si.náɾ | kosiná | ‘to cook’ | 2 |
| kɾís.ma | kɾíhma | ‘confirmation’ | 2 |
| kɾis.tjá.na | kɾistjána | ‘Christian’ | 9 (p. 30) |
| kɾis.tjá.no | kɾistjáno | ‘Christian’ | 9 (p. 48) |
| kɾís.to | kiɾitó | ‘Christ’ | 2, 6 |
| kɾis.to.βál | kiɾitó | *proper name* | 8 (p. 94) |
| kɾo.á.sja | kɨoasjá | ‘Croatia’ | 13 |
| kɾús | kuɾusú | ‘cross’ | 2, 5 |
| ku.lán.tɾo | kuɾãtṹ | ‘coriander’ | 2 |
| lá.ta | láta | ‘tin plating’, ‘can’ | 3, 6 |
| le.ón | leṍ | ‘lion’ | 6, 2 |
| lí.βa.no | lívano | ‘Lebanon’ | 13 |
| lí.βja | livjá | ‘Libya’ | 13 |
| li.βó.ɾja | livó | *proper name* | 8 (p. 94) |
| ló.mo | lómo | ‘loin’ | 2 |
| ló.ndɾes | ló$^{n}$dɾe | ‘London’ | 12 (p. 82) |
| lú.nes | lúne | ‘Monday’ | 1, 3 |
| ma.ða.ɣas.káɾ | madaɣaká | ‘Madagascar’ | 13 |
| maɾ.ga.ɾí.ta | mangaɾí | *proper name* | 8 (p. 94) |
| ma.rí.a | maría | ‘Mary’ | 9 (p. 12) |
| máɾ.te | márte | ‘Mars’ | 13 |
| me.lón | meɾṍ | ‘melon’ | 6, 2 |
| me.ɾe.ndáɾ | meɾe$^{n}$dá | ‘to have a snack’ | 2 |
| meɾ.kú.ɾjo | mekuɾjó | ‘Mercury’ | 13 |
| mi.ɣél | mingél | *proper name* | 8 (p. 94) |
| moɾ.sí.ʔa | $^{m}$busjá | ‘blood sausage’ | 3, 5 |
| nep.tú.no | netunó | ‘Neptune’ | 13 |
| ni.ko.lás | kolá | *proper name* | 8 (p. 94) |
| o.ɣár | óɣa | ‘home’, ‘house’ | 5, 6 |
| o.mán | omã́ | ‘Oman’ | 13 |
| ó.t͡ʃo | óʃo | ‘eight’ | 1, 3 |
| o.βé.xa | oveʃá | ‘sheep’ | 3, 4 |
| o.βís.po | obispó | ‘bishop’ | 9 (p. 20) |
| o.βís.po | ovíspo | ‘bishop’ | 9 (p. 42) |
| pa.pá | papá | ‘Pope’ | 9 (p. 20) |
| pa.ró.kja | parokja | ‘parish’ | 9 (p. 44) |
| páβ.lo | palí | *proper name* | 8 (p. 94) |
| páβ.lo | pavlo | ‘(St.) Paul’ | 9 (p. 32) |
| péð.ɾo | peɾú | *proper name* | 8 (p. 94) |

*Continued on next page*

*Continued from previous page*

| Spanish | Guaraní | English gloss | Source(s) |
|---|---|---|---|
| péð.ɾo | pedɾo | ‘(St.) Peter’ | 9 (p. 32) |
| pe.kín | pekĩ́ | ‘Beijing’ | 13 |
| pe.lón | peɾṍ | ‘bald’ | 7 (p. 278) |
| pé.ɾes | pére | *proper name* | 11 (p. 4) |
| pí.ɲa | píɲa | ‘pineapple’ | 1 |
| pjó.la | piʔóla | ‘cord’ | 7 (p. 281) |
| plu.tón | plutṍ | ‘Pluto’ | 13 |
| poɾ.tu.ɣál | poɨtuɣá | ‘Portugal’ | 13 |
| pwéβ.lo | pɨélo | ‘town’ | 11 (p. 5) |
| san | san | ‘St.’ | 9 (p. 32) |
| sa.pá.to | sapatú | ‘shoe’ | 1, 4 |
| sa.túɾ.no | satuɨnó | ‘Saturn’ | 13 |
| se.náɾ | sená | ‘to have dinner’ | 2 |
| se.βó.ɾa | sevój | ‘onion’ | 5, 7 (p. 280) |
| síð.ɾa | síɾa | ‘cider’ | 2 |
| si.mó.na | ʃimí | *proper name* | 8 (p. 94) |
| te.xéɾ | te.xé | ‘to weave’ | 2 |
| tí.fo | tífo | ‘typhus’ | 1 |
| tí.fus | tífu | ‘typhus’ | 3 |
| tó.ɾo | tóɾo | ‘bull’ | 1, 4 |
| tɾa.táɾ | tɾatá | ‘to treat’ | 10 (p. 258) |
| t͡ʃí.ka | ʃíka | ‘girl’ | 6, 7 (p. 279) |
| t͡ʃi.t͡ʃa.rón | ʃiʃarṍ | ‘pork rind’ | 13 |
| tuɾ.kí.a | tuɨkjá | ‘Turkey’ | 13 |
| u.ɾá.no | uɾanó | ‘Uranus’ | 13 |
| xa.pón | hapṍ | ‘Japan’ | 13 |
| xa.wáj | haváj | ‘Hawaii’ | 13 |
| xa.βón | havṍ | ‘soap’ | 6, 2 |
| xe.ɾár.ðo | kiɾá | *proper name* | 8 (p. 94) |
| xe.su.kɾís.to | hesukɾísto | ‘Jesus Christ’ | 9 (p. 12) |
| xo.sé | hosé | ‘(St.) Joesph’ | 9 (p. 32) |
| xu.ljá.na | lulí | *proper name* | 8 (p. 94) |
| xú.pi.teɾ | húpiteɾ | ‘Jupiter’ | 13 |
| xwan | hwã | *proper name* | 8 (p. 94) |
| xwan | hwan | ‘(St.) John’ | 9 (p. 32) |

# Appendix B

# Experiment Questionnaire

## B.1 Given version of questionnaire (Spanish)

¡Mil gracias por su ayuda! Por favor rellene la caja con la información básica antes de terminar el resto del cuestionario.

**Información básica sobre usted:**

| | |
|---|---|
| Edad: | |
| País de origen: | |
| Lengua(s) materna(s): | |

**Instrucciones:**

Abajo usted encontrará 16 palabras simuladas que parecen palabras del español. Imagine usted que las palabras han entrado en el guaraní como préstamos a través del español. Todas las palabras serían sustantivos.
Para cada palabra "española" hay dos posibles palabras "guaraníes" que podrían ser la forma de la palabra adaptada al guaraní.

¿Cuál de las dos posibilidades es la **más** natural? Es decir, ¿cuál **más** parece una verdadera palabra guaraní?

Aunque haya una mejor opción que no sea una de las dos (o aunque las dos le resulten malas), usted sólo tiene que elegir la mejor opción de las dos que se dan.
Por favor elija solamente una sola opción. Si es difícil determinar cuál es mejor, elija la que le pareció mejor al leer las dos por primera vez. ¡Pensándolo demasiado puede hacerlo más difícil!

Los acentos en las palabras, aunque no son siempre necesarios, sirven para aclarar cuál de las sílabas es la con acento.

Por ejemplo: galávo = ga-LA-vo, galavó = ga-la-VO

Para elegir, por favor ponga **a** o **b** en la columna "*Adaptación más natural*".

| *Palabra en español* | *Adaptaciones en guaraní* | | *Adaptación más natural* |
|---|---|---|---|
| cricio | (a) krisió | (b) kirísio | |
| gisto | (a) híto | (b) histó | |
| tracia | (a) tarásia | (b) trasiá | |
| plomel | (a) pylomél | (b) plomé | |
| uñal | (a) uñá | (b) uñalá | |
| cavalo | (a) kavaló | (b) kaválu | |
| pesteza | (a) pestesá | (b) petésa | |
| tredaz | (a) tredá | (b) tyredás | |
| tuarto | (a) tuartó | (b) tuáto | |
| brice | (a) brisé | (b) brúse | |
| glabo | (a) galávo | (b) glavó | |
| golde | (a) góde | (b) goldé | |
| parso | (a) páso | (b) pásu | |
| grubaz | (a) gruvá | (b) guruvás | |
| plavo | (a) plavó | (b) palávo | |
| blazal | (a) balasál | (b) blasá | |

## B.2 Translated version of questionnaire (English)

Thanks so much for your help! Please fill out the box with the basic information before completing the rest of the form.

**Basic information about yourself:**

| | |
|---|---|
| Age: | |
| Home country: | |
| Native language(s): | |

**Instructions:**

Below you will find 16 simulated words that look like Spanish words. Imagine that these words have entered Guaraní as loans from Spanish. All of the words would be nouns.
For each "Spanish" word there are two possible "Guaraní" words which could be the form of the adapted word in Guaraní.

Which of the two possibilities is the **most** natural? That is to say, which seems **most** like an actual Guaraní word?

Even if there is a better option that isn't one of the two listed (or if both seem bad), you only have to pick the best option of those given.
Please pick **only one** option. If it is difficult to determine which is best, pick the option which seemed best to you when you first read the two choices. Overthinking it can make choosing more difficult!

The accent marks, although not always necessary, are there to clarify which of the syllables is emphasized.
For example: galávo = ga-LA-vo, galavó = ga-la-VO

Please pick your answer by putting **a** or **b** in the column "*Most natural adaptation*".

| *Spanish word* | *Guaraní adaptations* | | *Most natural adaptation* |
|---|---|---|---|
| cricio | (a) krisió | (b) kirísio | |
| gisto | (a) híto | (b) histó | |
| tracia | (a) tarásia | (b) trasiá | |
| plomel | (a) pylomél | (b) plomé | |
| uñal | (a) uñá | (b) uñalá | |
| cavalo | (a) kavaló | (b) kaválu | |
| pesteza | (a) pestesá | (b) petésa | |
| tredaz | (a) tredá | (b) tyredás | |
| tuarto | (a) tuartó | (b) tuáto | |
| brice | (a) brisé | (b) brúse | |
| glabo | (a) galávo | (b) glavó | |
| golde | (a) góde | (b) goldé | |
| parso | (a) páso | (b) pásu | |
| grubaz | (a) gruvá | (b) guruvás | |
| plavo | (a) plavó | (b) palávo | |
| blazal | (a) balasál | (b) blasá | |

## B.3 Expected results of questionnaire

"Expected" here refers to the options which are predicted by the model to be possible nativizations. Also, the response " - " is given for those tasks whose forms were fillers and are therefore irrelevant.

| *Palabra en español* | *Adaptaciones en guaraní* | | *Adaptación más natural* |
|---|---|---|---|
| cricio | (a) krisió | (b) kirísio | A |
| gisto | (a) híto | (b) histó | A |
| tracia | (a) tarásia | (b) trasiá | B |
| plomel | (a) pylomél | (b) plomé | B |
| uñal | (a) uñá | (b) uñalá | - |
| cavalo | (a) kavaló | (b) kaválu | - |
| pesteza | (a) pestesá | (b) petésa | B |
| tredaz | (a) tredá | (b) tyredás | A |
| tuarto | (a) tuartó | (b) tuáto | B |
| brice | (a) brisé | (b) brúse | - |
| glabo | (a) galávo | (b) glavó | B |
| golde | (a) góde | (b) goldé | A |
| parso | (a) páso | (b) pásu | - |
| grubaz | (a) gruvá | (b) guruvás | A |
| plavo | (a) plavó | (b) palávo | A |
| blazal | (a) balasál | (b) blasá | B |

# References


Bakker, D. and E. Hekking (2012). Constraints on morphological borrowing: evidence from latin america. In L. Johanson and M. Robbeets (Eds.), *Copies versus cognates in bound morphology*, pp. 187220. Leiden, Boston: Brill.

Beckman, J. (1998). *Positional Faithfulness*. Ph. D. thesis, University of Massachusetts Amherst.

Boersma, P. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Ph. D. thesis, University of Amsterdam.

Boersma, P. and B. Hayes (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry 32*(1), 45–86.

Britton, A. S. (2005). *Guaraní-English/English-Guaraní Concise Dictionary*. New York: Hippocrene Books Inc.

Broselow, E. (2009). Stress adaptation in loanword phonology: perception and learnability. In P. Boersma and S. Hamann (Eds.), *Phonology in Perception*, pp. 191–229. Berlin: Mouton de Gruyter.

Chomsky, N. and M. Halle (1968). *The Sound Pattern of English*. New York: Harper & Row.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston and M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and the physics of speech*, pp. 283–333. Cambridge: Cambridge Universtiy Press.

Davidson, L. and R. Noyer (1997). Loan phonology in huave: nativization and the ranking of faithfulness constraints. In B. Agbayani and S.-W. Tang (Eds.), *Proceedings of the West Coast Conference on Formal Linguistics*, Volume 15, pp. 65–80. Stanford: CSLI.

de Assis, C. F. (2008). *Dicionário Guarani-Português/Português-Guarani*. São Paulo: Cecy Fernandes de Assis.

de Bruyne, J. and C. J. Pountain (1995). *A comprehensive Spanish grammar*. Oxford: Blackwell Publishers.

de Canese, N. K. (1983). *Gramática de la lengua Guaraní*. Asunción: Colección Ñemitỹ.

de Canese, N. K. and F. A. Alcaraz (1997). *Diccionario Guaraní-Español/Español-Guaraní*. Asunción: Colección Ñemitỹ.

Díaz, M. R. D. (2006). *Gran Diccionario de la Lengua Guaraní*. Corrientes: Moglia Ediciones.

Fritz, F. M. (Ed.) (2004). *Misal Romano: Edición bilingüe Guaraní - Nivaclé*. Asunción: Vicariato Apostólico del Pilcomayo.

Hualde, J. I. (2005). *The Sounds of Spanish*. Cambridge: Cambridge University Press.

Itô, J. and A. Mester (1995a). The core-periphery structure of the lexicon and constraints on reranking. In J. Beckman, S. Urbanczyk, and L. Walsh (Eds.), *Papers in Optimality Theory*, pp. 181–210. Amherst: GLSA.

Itô, J. and A. Mester (1995b). Japanese phonology. In J. Goldsmith (Ed.), *Handbook of Phonological Theory*, pp. 817–838. Cambridge: Blackwell.

Itô, J. and A. Mester (1999). The phonological lexicon. In N. Tsujimura (Ed.), *A handbook of Japanese linguistics*, pp. 62–100. Oxford: Blackwell Publishers.

Itô, J. and A. Mester (2008). Lexical classes in phonology. In S. Miyagawa and M. Saito (Eds.), *The Oxford Handbook of Japanese Linguistics*, pp. 84–106. Oxford: Oxford University Press.

Kiparsky, P. (1968). How abstract is phonology? In O. Fujimura (Ed.), *Three Dimensions of Linguistic Theory*, pp. 5–56. Tokyo: TEC.

Kiparsky, P. (1993). An ot perspective on phonological variation. Handout from Rutgers Optimality Workshop 1993, also presented at NWAVE 1994. Available at http://www.stanford.edu/ kiparsky/Papers/nwave94.pdf.

Kubozono, H. (1995). *Gokeisei to On-in-kouzou [Word Formation and Phonological Structure]*. Tokyo: Kurosio.

Lustig, W. (2005). *Guarani für Paraguay: Wort für Wort*. Bielefeld: Reise Know-How.

Martin, S. E. (1952). *Morphophonemics of Standard Colloquial Japanese. Language Dissertation No 47*. Baltimore: Linguistic Society of America.

Mayans, A. O. (1980). *Nuevo Diccionario Español-Guaraní/Guaraní-Español*. Buenos Aires: Editorial Universitaria de Buenos Aires.

McCarthy, J. J. (2008). *Doing Optimality Theory: Applying Theory to Data*. Oxford: Blackwell Publishers.

McCarthy, J. J. and A. Prince (1993). Generalized alignment. In G. Booij and J. van Marle (Eds.), *Yearbook of Morphology*, pp. 79–153. Dordrecht: Kluwer.

McCarthy, J. J. and A. Prince (1995). Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers in Linguistics 18*, 249–384.

McCawley, J. D. (1968). *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.

Morínigo, M. A. (1931). *Hispanismos en el Guaraní*. Buenos Aires: Facultad de Filosofía y Letras de la Universidad de Buenos Aires.

Nordhoff, S. (2004). Nomen/verb-distinktion im guaraní. Master's thesis, University of Cologne, Germany.

Prince, A. and P. Smolensky (1993). Optimality theory: constraint interaction in generative grammar. Ms, Rutgers University & University of Colorado, Boulder.

Rendon, J. G. (2008). *Typological and social constraints on language contact : Amerindian languages in contact with Spanish*. Ph. D. thesis, University of Amsterdam.

Rice, C. (2006). Norwegian stress and quantity: The implications of loanwords. *Lingua 116*(7), 1171–1194.

Saciuk, B. (1969). The stratal division of the lexicon. *Papers in Linguistics 1*, 464–532.

Selkirk, E. (1984). On the major class features and syllable theory. In J. Kingston and M. E. Beckman (Eds.), *Language Sound Structure: Studies in Phonology*, pp. 107–136. Cambridge: MIT Press.

Shibatani, M. (1990). *The Languages of Japan*. Cambridge: Cambridge University Press.

Smith, J. L. (2006). Loan phonology is not all perception: Evidence from japanese loan doublets. In T. J. Vance and K. A. Jones (Eds.), *Japanese/Korean Linguistics, Volume 14*, pp. 63–74. Stanford: CSLI.

Tonhauser, J., D. Beaver, C. Roberts, and M. Simons (2013, March). Toward a taxonomy of projective content. *Language 89*(1), 66–106.

Tonhauser, J. and E. Colijn (2010). Word order in paraguayan guaraní. *International Journal of American Linguistics 76*(2), 255–288.

Vance, T. (1987). *An Introduction to Japanese Phonology*. New York: SUNY Press.

Velázquez-Castillo, M. (2013, 12 April). The deictic nature of guaraní evidentials: Distance, directionality and perspective. Handout from talk given at UNC-Chapel Hill.

Wikipedia (2012). Vikipetã, opaite tembĩkuaa hekosãsóva renda. Online: `http://gn.wikipedia.org/`; accessed October 2012.

Zarratea, T. (1981). *Kalaíto Pombéro*. Asunción, Paraguay: Asunción: Ediciones NAPA.

Zec, D. (2007). The syllable. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology*, Chapter 8, pp. 161–194. Cambridge: Cambridge University Press.