

CHAPTER 9

**Spanish Loans and Evidence for Stratification in the Guarani Lexicon**

*Justin Pinta & Jennifer L. Smith*

1 INTRODUCTION

The contact between Spanish and Guarani<sup>1</sup> and the resulting linguistic borrowing has endowed Guarani with a wealth of Spanish loanwords. These loans are notable not only for their quantity but for the range of different phonological characteristics that they display. Not all loans show homogeneous adaptation strategies; moreover, implicational relationships hold between the different strategies applied to a given loan.

For example, loans that repair complex onsets always repair codas as well, but the reverse is not true. These implicational relationships reveal that the Guarani lexicon is stratified (Itô & Mester 1999), with a core-periphery structure. That is, lexical items are grouped into strata, but the strata are not entirely separate from one another. Instead, they are nested within one another such that they form a set-inclusion hierarchy.

This chapter begins with a discussion of the relevant points of both Guarani and Spanish native phonology, leading in turn to an analysis of the adaptation processes employed by Guarani when borrowing Spanish words. The subsequent section outlines what these repair strategies tell us about the nature of the Guarani lexicon. The lexicon is shown to be composed of various lexical strata arrayed in a core-periphery structure on the basis of the repairs they make to specific Spanish phonological structures. A brief discussion of morphological evidence for the synchronic relevance of Guarani strata followed by discussion of impossible nativization effects evident in Guarani concludes the chapter.

2 ADAPTATION STRATEGIES

2.1 NATIVE PHONOLOGIES<sup>2</sup>

2.1.1 GUARANI

Lexical stress in Guarani is most commonly word-final. 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 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 (1).<sup>3</sup>

---

<sup>1</sup> The term “Guarani” has been used to refer to a subgroup of the Tupi-Guarani language family, a dialect chain within that subgroup, and the specific language Paraguayan Guarani. In this work “Guarani” is used exclusively in reference to the single language of Paraguayan Guarani.

<sup>2</sup> Both the data and generalizations regarding native Guarani phonology in this section are taken from de Canese (1983) except where noted; generalizations regarding native Spanish phonology are taken from Hualde (2005).

<sup>3</sup> 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. Guarani has a regular nasal harmony process affecting both vowels and consonants (Gregores & Suárez 1967; Rivas 1975; Walker 1999), in which nasalization spreads in both directions from a stressed nasal vowel, and leftward from a medial prenasalized stop; spreading is blocked by a

(1)

		<b>Orthography</b>	<b>Transcription</b>	
(a)	Oxytone	ñandu	ɲa <sup>n</sup> dú	‘spider’
		ore	oré	‘we’
(b)	Monomorphemic non-oxytone	túva	túva	‘father’
		ára	ára	‘day’
(c)	Polymorphemic non-oxytone	ejumína <sup>4</sup>	e- ðʒu-mí-na	‘please come’
		ndéve	ndé-ve	‘to you’

Native Guarani syllable structure is (C)(G)V(G), where “V” represents a vowel acting as the syllabic nucleus and “G” (for “glide”) represents a non-nuclear vowel. Syllables may consist of a vowel only or may optionally contain an onset; however, complex onsets and codas of any kind are forbidden. Guarani makes extensive use of diphthongs. The Guarani phonemic inventory has 12 vowel phonemes consisting of oral and nasal versions of the six vowels /i ĩ u e o a/. All non-high vowels in Guarani (/e o a/) always constitute syllabic nuclei, while high vowels (/i ĩ u/) may or may not be nuclei depending on their surrounding segments. High vowels are less sonorous than other vowels (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 Guarani syllable types are found in (2).

(2)

V	ĩ	‘water’
CV	sɪ	‘mother’
CVG	mo.kóʃ	‘two’
CGVG	tu.ɣwáj	‘tail’

Guarani native phonology contains three prenasalized stops, /<sup>m</sup>b <sup>n</sup>d <sup>ŋ</sup>g/, which can be seen in both word-initial and word-medial position in (3) (Britton 2005).

(3)

(a)	Word-initial	<sup>m</sup> bo.vi.ví	‘to sew’
		<sup>n</sup> dɪ	‘saliva’
		<sup>ŋ</sup> gwa.ʔú	‘perhaps’
(b)	Word-	mo. <sup>m</sup> bo.ʔó	‘to enrage’

---

stressed oral vowel and its syllable onset. For the most part, transcriptions here abstract away from the effects of nasal harmony. However, following de Canese (1983) (and Guarani orthographic practice), nasal stops such as [m] are distinguished from prenasalized oral stops such as [<sup>m</sup>b], even though the distribution of these classes is determined by the nasal harmony process and is therefore predictable.

<sup>4</sup> Taken from Tase (2012: 136).

medial

re.no. <sup>n</sup> dé.pe	‘in front of’
no. <sup>n</sup> ga.tú	‘to keep’, ‘to preserve’

Guarani prenasalized stops are explicitly rendered in the orthography such that /<sup>m</sup>b/ is represented as orthographic *mb*, /<sup>n</sup>d/ as *nd* and /<sup>ŋ</sup>g/ as *ng*. When a Spanish loan containing one of the word-medial sequences /mb nd ŋg/ is borrowed (these sequences do not occur word-initially in Spanish) it is treated in Guarani in the same manner as the native Guarani /<sup>m</sup>b <sup>n</sup>d <sup>ŋ</sup>g/, i.e., it operates as one phonological unit. Instances in which the stop is voiceless (i.e., /nt/) are rarer in Guarani but do occur; Spanish /mp nt ŋk/ pattern with their voiced counterparts in being syllabified in Guarani as /<sup>m</sup>p <sup>n</sup>t <sup>ŋ</sup>k/.

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.<sup>m</sup>bV] and never as [Vm.bV]. Evidence for this comes from Guarani treatment of Spanish nasal codas, which are repaired by nasal coalescence as will be explicitly discussed in §2.2.4. 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. Particularly illustrative is the loan for *Finland* which is adapted as in (4).

(4)

<b>Spanish</b>	<b>Guarani</b>	
fin.lá.ndja	hĩ.la. <sup>n</sup> dja	‘Finland’

Here we see the first Spanish /n/ syllabified as a coda and therefore repaired through nasal coalescence. The second Spanish /n/ is syllabified as part of a prenasalized stop onset and consequently is not repaired.

### 2.1.2 SPANISH

Spanish lexical stress does not have one consistent surface location. Generally speaking Spanish primary lexical stress may fall on either the final, penultimate or antepenultimate syllable. This phonemic contrast leads to some triplet minimal sets such as the one seen in (5).

(5)

<b>Orthography</b>	<b>Transcription</b>	
término	tér.mi.no	‘term’
termino	ter.mí.no	‘I finish’
terminó	ter.mi.nó	‘he/she/it/you finished’

As in Guarani, things may become more complicated due to morphological factors such as the addition of a clitic to a verb (e.g., *cantando=nos=la*, ‘singing it to us’), and in such cases stress may fall even earlier than the antepenult. This surface diversity of Spanish stress becomes important when considering loans without word-final stress which are borrowed into Guarani.

Native Spanish syllable structure allows for more complex structures than that of Guarani. Spanish syllables may contain a vowel only or make optional 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 VCCCC strings in which syllabification results in a complex coda (i.e., VCC.CV). These scenarios produce the largest attested syllable in Spanish: CCVCC. Examples of Spanish syllable types can be found in (6).

(6)

V	a	‘to’
CV	la	‘the’
CVC	tos	‘cough’
CCVC	tres	‘three’
CCVCC	trans.por.tár	‘to carry’

With regard to glides, Spanish patterns similarly to Guarani in the sense that high vowels (e.g., /i u/) may be nuclear or non-nuclear while non-high vowels (e.g., /e o a/) are (nearly) always nuclear.<sup>5</sup> Also similar to Guarani, any tautosyllabic combination of a high vowel and a non-high vowel will result in the non-high vowel as the nucleus. Adjacent non-high vowels form individual syllables while adjacent high vowels are tautosyllabic. Examples of Spanish syllables with glides can be found in (7).

(7)

VG	aj	‘there is/are’
CGV	pje	‘foot’
CGVG	bwej	‘ox’
CGVGC	lim.pjájs	‘you (plural) clean’

## 2.2 GUARANI TREATMENT OF SPANISH PHONOLOGICAL STRUCTURES

Having described the relevant characteristics of native Guarani and Spanish phonology, the discussion now turns to the Spanish loans in the Guarani lexicon and the relevant adaptations, or lack thereof, that the loans undergo in the process of nativization. All Guarani loan data used in this analysis and presented below was taken from a corpus of Spanish loanwords in Guarani consisting of 177 loans. The corpus was compiled by the first author, and 13 dictionaries, grammars, articles and other publications were used as sources for the corpus. The dictionaries and grammars, listed alphabetically by author's name, include Britton (2005), de Assis (2008), de Canese & Alcaraz (1997), Díaz (2006), Lustig (2005), Mayans (1980) and Morínigo (1931); academic works regarding Guarani include Gómez Rendón (2008), Tonhauser & Colijn (2010), Tonhauser et al. (2013) and Velázquez-Castillo (2013). The source Fritz (2004) is a Guarani-language Catholic missal obtained from central Paraguay. The final source was the Guarani version of the online encyclopedia Wikipedia. For further details concerning the corpus, in addition to the corpus itself in its entirety, see Pinta (2013).<sup>6</sup>

<sup>5</sup> Exceptions to /e o a/ being nuclear occur in some specific dialects of Spanish as well as in rapid speech.

<sup>6</sup> Gregores & Suárez (1967) and Zarratea (2013) also provide loanword corpora; the patterns seen there are largely congruent with the Pinta (2013) corpus. Smith & Pinta (2016) combines these three corpora and reaches very similar conclusions about evidence for strata as are presented here.

The phonological structures under discussion here are lexical stress, complex onsets, and codas, both non-nasal and nasal. Segmental adaptations are also attested but do not play an important a role in this analysis because they are found more sporadically in the corpus. For each phenomenon discussed there are some loans which undergo nativization processes and some which do not. Each pattern will be described in turn.

## 2.2.1 LEXICAL STRESS

### 2.2.1.1 ADAPTATION OF SPANISH STRESS

The loans of interest when considering lexical stress adaptation are those whose Spanish source form had non-final stress, as these are the only loans that present a conflict with native Guarani lexical stress patterns and thereby provide information regarding adaptation. Many loans entering with non-final stress undergo an adaptation which shifts the stress to word-final position. In (8)<sup>7</sup> we see some examples from the corpus in which this is observable.

(8)

<b>Spanish<sup>8</sup></b>	<b>Guarani</b>	
krís.to	ki.ri.tó	‘Christ’
bá.ka	va.ká	‘cow’
a.sú.kar	a.su.ká	‘sugar’
sa.pá.to	sa.pa.tú	‘shoe’
o.βé.xa	o.ve.ǰá	‘sheep’

### 2.2.1.2 TOLERANCE OF SPANISH STRESS

In contrast to the loans which repair non-native Guarani stress patterns, many loans show no stress repair whatsoever as seen in (9).

(9)

<b>Spanish</b>	<b>Guarani</b>	
es.pí.ri.tu	es.pí.ri.tu	‘spirit’
ka.tó.li.ka	ka.tó.li.ka	‘Catholic’
a.mé.ri.ka	a.mé.ri.ka	‘America’
bi.ná.ɣre	vi.ná.ɣre	‘vinegar’
bo.lí.tʃe	vo.lí.fo	‘store’

## 2.2.2 COMPLEX ONSETS

### 2.2.2.1 ADAPTATION OF SPANISH COMPLEX ONSETS

Spanish loans entering with 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

<sup>7</sup> For reasons of space examples here and henceforth are non-exhaustive (unless otherwise specified), but exhaustive lists can be found in Pinta (2013).

<sup>8</sup> With regard to the syllabification of the Spanish forms, here and in subsequent loan data tables (as well as in (4)) “Spanish” does not necessarily refer to the form of the original Spanish phonological grammar but rather to the perception by Guarani speakers of the syllabification of the Spanish form. The importance of this distinction will be further discussed in §2.3.

syllables. The epenthesis vowel is most commonly either /i/ or a copy of the vowel following the place of insertion. We see examples of this repair strategy in (10) with the complex onsets in the Spanish form and the epenthetic vowel in the Guarani form bolded.<sup>9</sup>

(10)

		<b>Spanish</b>	<b>Guarani</b>	
(a)	Word-initial	<b>gr</b> é.sja	<b>gi</b> .re.sjá	‘Greece’
		<b>kr</b> us	<b>ku</b> .ru.sú	‘cross’
		<b>kr</b> ís.to	<b>ki</b> ritó	‘Christ’
(b)	Word-medial	áws. <b>tr</b> ja	aw. <b>te</b> .rjá	‘Austria’
		e.ri. <b>tr</b> é.a	e.ri. <b>ti</b> .re.á	‘Eritrea’
		aws. <b>tr</b> á.lja	aw. <b>ta</b> .ra.ljá	‘Australia’

The second strategy seen in these cases is not epenthesis, but replacement of the second of the two segments by semivocalic /j/, as can be seen in (11). In these cases the second segment is always /l/ or /r/ given that Spanish only allows liquids as the second segment of a complex onset (Hualde 2005: 71).

(11)

		<b>Spanish</b>	<b>Guarani</b>	
(a)	Word-initial	<b>fr</b> án.sja	hjá.́.sja	‘France’
		<b>kr</b> o.á.sja	kjo.a.sjá	‘Croatia’
(b)	Word-medial	i. <b>ŋg</b> la.té.ra	i. <sup>h</sup> gja.te.rá	‘England’

#### 2.2.2.2 TOLERANCE OF SPANISH COMPLEX ONSETS

Some loans with complex onsets avoid making repairs and import the loan preserving the original Spanish sequence; in (12) we see examples of these loans.

(12)

		<b>Spanish</b>	<b>Guarani</b>	
(a)	Word-initial	<b>dro</b> .yár	dro.yá	‘to drug’
		<b>gra</b> .ná.ða	gra.na.dá	<i>kind of fruit</i>
		<b>tra</b> .tár	tra.tá	‘to treat’
(b)	Word-medial	ló. <b>nd</b> res	ló. <sup>n</sup> dre	‘England’
		bi.ná. <b>y</b> re	vi.ná.yre	‘vinegar’
		a.la. <b>kr</b> án	a.la.krán	‘scorpion’

<sup>9</sup> Justifications for the given syllabifications of word-medial CC or CCC strings such as those found in Spanish [krís.to] in (10a), for which more than one syllabification may seem plausible (e.g., for CC sequences: .CC, C.C, or CC.), are systematically offered in Pinta (2013: 31-33). See Smith & Pinta (2016) for additional discussion.

## 2.2.3 NON-NASAL CODAS

### 2.2.3.1 ADAPTATION OF SPANISH 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. In (13) we see loans with codas being repaired by deletion without discrimination on the basis of the segment in coda position.

(13)

	Segment deleted	Spanish	Guarani	
(a)	r	a.te.ndér	a.te. <sup>n</sup> dé	‘to pay attention to’
		dok.tór	doɿ.tó	‘doctor’
(b)	s	a.rós	a.ró	‘rice’
		tí.fus	tí.fu	‘typhus’
(c)	l	ko.rál	ko.rá	‘corral’
		por.tu.yál	poɿ.tu.yá	‘Portugal’
(d)	ð	ko.mu.ni.ðáð	ko.mu.ni.ðá	‘community’

There is one notable exception in the corpus to the generalization that word-final non-nasal codas are repaired by deletion, seen in (14).

(14)

Spanish	Guarani	
krus	ku.ru.sú	‘cross’

In this case we see epenthesis in lieu of deletion used to satisfy the Guarani 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. It is however notable in being the only monosyllabic Spanish loan source in the corpus with a word-final non-nasal coda, and this may be a factor in its adaptation.

Seen in (15) are the loans whose source form contains a word-medial coda, and the process of repair in these cases is deletion as seen in the word-final cases.<sup>10</sup>

(15)

	Segment deleted	Spanish	Guarani	
(a)	r	mer.kú.rjo	me.ku.rjó	‘Mercury’

<sup>10</sup> The Spanish source forms in (15d–e) and also (18d–e) are syllabified, not according to the grammar of Spanish, but according to the Guarani perceived syllabification of these forms. See §2.3 below and Pinta (2013), Smith & Pinta (2016) for discussion.

		mor.sí.ʎa	<sup>m</sup> bu.sjá	‘blood sausage’
(b)	s	es.pi.ná.so	e.pi.na.só	‘spine’
		kris.to	ki.ri.tó	‘Christ’
(c)	l	ka.l.són	ka.só	‘pants’
		fa.l.tár	va.tá	‘to lack’
(d)	ð	sí.ð.ra	sí.ra	‘cider’
		pé.ð.ro	pe.rú	<i>proper name</i>
(e)	β	pwé.β.lo	pié.lo	‘town’
		pá.β.lo	pa.lí	<i>proper name</i>
(f)	p	ne.p.tú.no	ne.tu.nó	‘Neptune’
(g)	t	fút.bol	hú.vol	‘soccer’

There are two exceptions to the deletion pattern in (15), seen in (16).

(16)

Spanish	Guarani	
al.mi.ðón	a.ra.mi.ró	‘starch’
ká.β.ra	ka.va.rá	‘goat’

Similar to the case in (14), here we see epenthesis instead of deletion as a coda repair strategy. For example in the case of Spanish [al.mi.ðón] the first syllable's coda /l/ is retained while the epenthesis of /a/ causes resyllabification in which /l/ (adapted in Guarani as /r/) constitutes the onset of the following syllable. This kind of behavior with a coda in this position is only attested in these two forms in our corpus, though Gregores & Suárez (1967) give three other examples.

The second repair strategy for word-medial codas is, just as was seen for complex onsets in (11), replacement of the illicit segment by /j/. An exhaustive list of the word-medial codas in the corpus repaired by /j/ can be seen below in (17).

(17)

	Segment replaced	Spanish	Guarani	
(a)	r	sa.túr.no	sa.tu <sub>j</sub> .nó	‘Saturn’
		por.tu.γál	po <sub>j</sub> .tu.γá	‘Portugal’
		tur.kí.a	tu <sub>j</sub> .kjá	‘Turkey’
(b)	s	is.lá.ndja	i <sub>j</sub> .la. <sup>n</sup> dá	‘Iceland’
(c)	k	dok.tór	do <sub>j</sub> .tó	‘doctor’

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

### 2.2.3.2 TOLERANCE OF SPANISH NON-NASAL CODAS

In addition to loans which repair original Spanish codas, there are also those which are tolerant of codas. In (18) we see examples of loans from the corpus which faithfully preserve the codas from their Spanish source.



(18)

	<b>Coda segment</b>	<b>Spanish</b>	<b>Guarani</b>	
(a)	r	e.kwa.ðór már.te	e.kwa.tór már.te	‘Ecuador’ ‘Mars’
(b)	s	kris.tjá.no is.lám	kris.tjá.no is.lá	‘Christian’ ‘Islam’
(c)	l	mi.yél al.ma.sén	mi. <sup>h</sup> gél al.ma.sé	<i>proper name</i> ‘department store’
(d)	ð	péð.ro	ped.ro <sup>11</sup>	‘(St.) Peter’
(e)	β	páβ.lo	pav.lo	‘(St.) Paul’
(f)	t	at.lá.n <sup>h</sup> ti.ko	at.lá. <sup>n</sup> ti.ko	‘Atlantic’
(g)	k	ko.lek.tí.βo	ko.lek.tí.vo	‘bus’

#### 2.2.4 NASAL CODAS

##### 2.2.4.1 ADAPTATION OF SPANISH NASAL CODAS

Nasal codas are also avoided; however, they show a different repair strategy. As mentioned earlier, Guarani has a phonemic nasal/oral vowel contrast in which all vowels may appear as oral or nasal. When a Spanish loan with a nasal coda enters the language we see nasal coalescence as the repair strategy where the nasal coda coalesces with the previous vowel rendering it nasalized. Examples of such loans are seen in (19). It is worth mentioning that /n/ is by far the most frequently occurring nasal coda in Spanish (also possible are /m η/), and as such the vast majority of loans in the corpus with nasal codas in their original Spanish form have /n/.

(19)

	<b>Vowel nasalized</b>	<b>Spanish</b>	<b>Guarani</b>	
(a)	o	ka.mjón le.ón	ka.mjõ le.õ	‘truck’ ‘lion’
(b)	a	xwan o.mán	hwã o.mã	<i>proper name</i> ‘Oman’
(c)	i	pe.kín	pe.kí	‘Beijing’
(d)	e	al.ma.sén je.mén	al.ma.sé d̃ze.mé	‘department store’ ‘Yemen’

Word-medial nasal codas show identical behavior; examples are shown in (20).

(20)

<b>Spanish</b>	<b>Guarani</b>	
frán.sja	hiá.sja	‘France’
fin.lá.ndja	hi.la. <sup>n</sup> dja	‘Finland’

<sup>11</sup> Cases in which lexical stress is not represented are due to orthographic ambiguity in the original source for the loan. In the absence of absolute certainty regarding which syllable is stressed, that information is left out.

There is one exception to this pattern of coalescence, seen in (21). This loan is of interest due to the fact that here the nasal coda deletes without any coalescence; notably, it is a proper name. In the corpus there is a trend of proper names being subject to different phonological treatment than the other loans. Proper names seem conspicuously more susceptible to segment deletion, perhaps as the result of a prosodic template effect on the adaptation of names (most are disyllabic). Here this loan constitutes the only exception to the generalization exemplified by the data in (19) and (20).

(21)

<b>Spanish</b>	<b>Guarani</b>	
kons.tán.sja	ko.tá	<i>proper name</i>

#### 2.2.4.2 TOLERANCE OF SPANISH NASAL CODAS

Although numerically fewer, there are also instances in which nasal codas go unrepaired. The loans in which we see this are found in (22).

(22)

<b>Spanish</b>	<b>Guarani</b>	
i.rán	i.rán	‘Iran’
a.mén	a.mén	‘amen’
san	san	‘St.’
kon.fir.ma.sjón	kon.fir.ma.sjón	‘confirmation’
en.sa.lá.đa	en.sa.lá.da	‘salad’

### 2.3 GUARANI PERCEIVED SYLLABIFICATION OF SPANISH FORMS

As has been shown, Guarani differs in its treatment of codas and complex onsets with regard to repair strategies: codas are repaired by deletion (or /i/ replacement in  $\_C$  position or nasal coalescence), while complex onsets are repaired by epenthesis (or /i/ replacement in  $C\_$  position). This affects the resulting form in Guarani due to the fact that the syllabification of a given word-medial consonant affects what happens to it. In the adaptation of loans, the Guarani grammar must syllabify foreign structures before repairing them given that the syllable position of a consonant determines which repairs it undergoes. For example, a Guarani speaker initially processing the Spanish word /áwstrja/ must decide whether it is syllabified as in (23a), (23b) or (23c). This decision affects the resulting loan, as can be seen in the last column.

(23)

	<b>Perceived syllabification of Spanish form</b>	<b>Expected resulting Guarani lexical item</b>
(a)	áw.stɾja	aw.sV.tV.rjá
(b)	áws.tɾja	aw.tV.rjá
(c)	áwst.rja	aw.rjá

In (23a) we see a syllabification in which there are no codas; as a result no deletion takes place and all initial Spanish segments are present in the Guarani form. Given that Guarani most commonly repairs two tautosyllabic consonants by vowel insertion, we can posit that the result of such a syllabification would produce two epenthetic vowels (here represented generically as “V”). In (23b) /s/ is syllabified as a coda and /tr/ as a complex onset, and repairs are made accordingly. The /s/ is deleted and a vowel is inserted to break up the /tr/ sequence. The final possibility in (23c) would syllabify /st/ as a complex coda, leaving /r/ as a singleton onset of the following syllable. It is difficult to know how Guarani would handle such a form given that Spanish only rarely has complex codas, but one plausible scenario is that it would produce a form in which both segments are deleted given their coda position (see the last loan in (24) below).

As shown, the syllabification must be assigned first because the choice of syllabification has an impact on the final form of the word. Crucially, this syllabification is the *perceived* syllabification by the Guarani grammar of the Spanish source form.<sup>12</sup> This perceived syllabification has no direct connection to the actual Spanish syllabification of the word, and the two syllabifications of the same Spanish form could differ. Indeed this is attested, and the most salient examples are those Spanish forms with nasal/voiced plosive sequences (e.g., /mb nd ŋg/), where the nasal consonants in Spanish would be syllabified as codas and not as part of a prenasalized stop onset as in Guarani (e.g., the adaptation seen in (13a) resulting in Guarani [a.te.<sup>n</sup>dé] in which the Guarani perceived syllabification of the Spanish form is [a.te.ndér] and the actual Spanish syllabification is [a.ten.dér]).

The Guarani phonological grammar is consistent in its treatment of consonant strings found in the loans in the corpus, and generalizations can consequently be made. Spanish VCCV sequences are generally syllabified as VC.CV.<sup>13</sup> Spanish VCCCV sequences, while rarer, also show consistency in their syllabification as VC.CCV. Those loans containing original VCCCV sequences are shown in (24).

(24)

<b>Spanish</b>	<b>Guarani</b>	
áws. <b>tr</b> ja	aw.te.rjá	‘Austria’
aws. <b>trá</b> .lja	aw.ta.ra.ljá	‘Australia’
es. <b>kri</b> .βír	kri.ví	‘to write’
kons. <b>tán</b> .sja	ko.tá	<i>proper name</i>

Notable here is the last loan in the table, coming from the Spanish proper name *Constancia*. Here it is difficult to know whether the resulting form is representative of normal Guarani phonological processes of loan adaptation or is extraordinary due to either it being a proper name or, perhaps more importantly, it being the only loan whose original form contained a [nsC] string. One might expect nasal coalescence to occur here given the unambiguous nasal

<sup>12</sup> See Smith (2006, 2009) for an explicit model of the *perceived source-language representation*, a phonological representation of the borrowing-language speaker's knowledge of a loanword's source form.

<sup>13</sup> Excepting cases involving the Spanish sequences /mb mp nd nt ŋg ŋk/, as discussed above, as well as instances where the CC is a stop+liquid sequence, in which case it forms a complex onset. The preference for VC.CV syllabification includes Spanish source forms with lenited voiced obstruent + liquid clusters such as [ðr] or [βl], as seen in (15) and (18) above, even though these would be syllabified as onset clusters in Spanish.

coda in the first syllable, yet this is not observed. Setting this example aside, as well as those consonant sequences involving nasals interpreted as Guarani prenasalized stops, the standard syllabification of Spanish VCCCV sequences is VC.CCV.

### 3 STRUCTURE OF THE GUARANI LEXICON

Having established the nativization strategies of the Guarani phonological grammar when processing loans from Spanish, we now turn to what these strategies can tell us about the structure of the Guarani lexicon. The Guarani lexicon has clearly defined sublexica or strata, with distinct phonological characteristics; the existence of these strata can be inferred from the phonological behavior of the loans. Crucially, the core-periphery structure of the Guarani lexicon, outlined in the section immediately following, is the basis on which the case for clear strata can be made.

The majority of the lexical items comprising the Guarani lexicon are native to the language and are of Tupi-Guarani origin. Since the arrival of the Europeans in the Americas, a new source of words has presented itself via Guarani contact with Spanish. The synchronic situation in Guarani provides modern speakers with a range of loans varying in degree of phonological adaptation. 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 reconciling the fact that one synchronic Guarani grammar produces forms radically different from one another phonologically.

While the stratification of the lexicon is here argued to be synchronically relevant, the reason for the strata themselves is diachronic in nature. The new source of words from Spanish initially resulted in a series of borrowings which were repaired to the extent that they were phonologically indistinguishable from native Guarani words. As time went on and the number of Guarani-Spanish bilinguals rose, and with it a general familiarity in the Guarani-speaking community with Spanish phonology, loans began to enter which showed phonological characteristics disallowed in native Guarani (Morínigo 1931). This tolerance increased until the point where some Spanish loans began to enter unadapted from their original form. As a result the present-day Guarani lexicon has native Tupi-Guarani words, fully adapted loans from Spanish, partially adapted loans from Spanish, and loans from Spanish which are not adapted at all.<sup>14</sup>

Following is a discussion of the structure of the Guarani lexicon, beginning with a description of the core-periphery structure. After this is an outline and description of the specific strata followed by morphological evidence for the synchronic relevance of the strata. Finally a brief discussion of impossible nativization effects in Guarani will close the chapter.

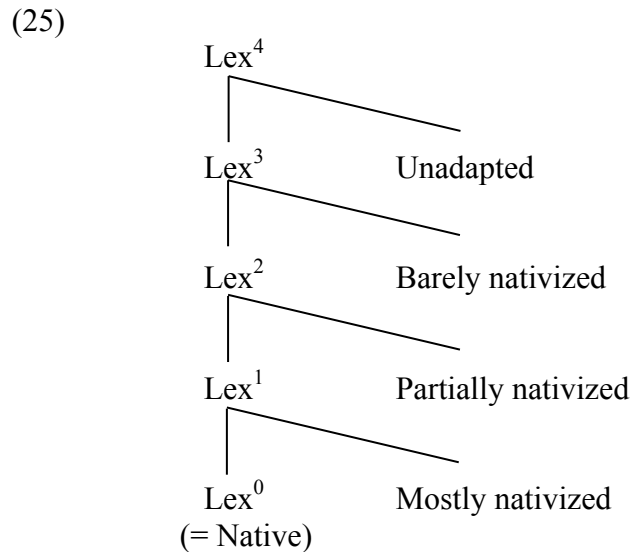
#### 3.1 CORE-PERIPHERY STRUCTURE

The lexical subclasses of Guarani cannot be described as a mere partitioning of the lexicon into parallel and non-overlapping sets with entirely distinct phonological grammars. Instead, the

---

<sup>14</sup>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 and English.

lexicon is better described in terms of strata which overlap substantially and are characterized by their core-periphery structure. That is, successively more peripheral strata tolerate successively greater numbers of phonological structures. The stratal hierarchy can be characterized by the notion of set inclusion, which can be depicted as in (25) where Lex<sup>n</sup> represents a sublexicon or stratum (adapted from Itô and Mester (1999: 65)).



In (25) we see that Lex<sup>0</sup> represents the native 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 for example “mostly nativized” describes the set Lex<sup>1</sup> - Lex<sup>0</sup>. Lexical items in the core stratum (Lex<sup>0</sup>) are those with phonological characteristics which define the central area of the lexicon. That is, lexical items in Lex<sup>0</sup> are those exhibiting maximally native Guarani phonological characteristics. Moving outwards from the core, the members of each stratum begin to look increasingly less like native Guarani forms until at the periphery the words are unadapted from their original Spanish forms.<sup>15</sup>

Crucially, the ways in which a given stratum remains faithful to Guarani phonology hold in all strata “below” it in (25) as well. For example Lex<sup>3</sup> (barely nativized) remains faithful to Guarani phonology by avoiding nasal codas, and as such Lex<sup>2</sup>, Lex<sup>1</sup> and Lex<sup>0</sup> do as well (while Lex<sup>4</sup> does not). Similarly Lex<sup>2</sup> (partially nativized) is intolerant of non-nasal codas, as are Lex<sup>1</sup> and Lex<sup>0</sup> (while Lex<sup>3</sup> and Lex<sup>4</sup> are not).

### 3.1.1 IMPLICATIONAL RELATIONSHIPS

Set-inclusion relationships such as those seen in (25) are motivated by implicational relationships between particular repairs seen in lexical items across strata. For example, those loans which make repairs to a Spanish complex onset also repair lexical stress, as can be seen in (26).

(26)

---

<sup>15</sup> It is worth mentioning that the different sources for the loan corpus each contribute words that fall into different strata and the issue of differing adaptation strategies cannot be simply explained by disparate sources from which the data was obtained.

<b>Spanish</b>	<b>Guarani</b>	
gré.sja	gi.re.sjá	‘Greece’
krus	ku.ru.sú	‘cross’
krís.to	ki.ri.tó	‘Christ’
kro.á.sja	kɔ̃o.a.sjá	‘Croatia’
áws.trja	aw.te.rjá	‘Austria’
e.ri.tré.a	e.ri.ti.re.á	‘Eritrea’
aws.trá.lja	aw.ta.ra.ljá	‘Australia’
i.ŋgla.té.ra	i. <sup>h</sup> gia.te.rá	‘England’

The single exception to this generalization in our corpus is seen in (27).

(27)

<b>Spanish</b>	<b>Guarani</b>	
frán.sja	hiã.sja	‘France’

Here, additional factors are at work. In particular, the nasal coda has undergone coalescence to create a nasalized vowel in the Guarani form. Since only stressed vowels are contrastively nasalized in Guarani (Gregores & Suárez 1967; Rivas 1975), that additional factor may explain why stress does not become final in this example.

Another implicational relationship is in turn found between lexical stress and coda repair. In the loans which repair lexical stress, repairs of codas are also seen, as in (28).

(28)

	<b>Coda repaired</b>	<b>Spanish</b>	<b>Guarani</b>	
(a)	r	a.sú.kar	a.su.ká	‘sugar’
		mor.sí.ɫa	<sup>m</sup> bu.sjá	‘blood sausage’
		ir.lá.nda	i.la. <sup>n</sup> djá	‘Ireland’
		tur.kí.a	tuɿ.kjá	‘Turkey’
		di.na.már.ka	<sup>n</sup> di.na.maɿ.ka	‘Denmark’
		ar.xé.lja	aɿ.he.ljá	‘Algeria’
		mer.kú.rjo	me.ku.rjó	‘Mercury’
		sa.túr.no	sa.tuɿ.nó	‘Saturn’
(b)	s	krís.to	ki.ri.tó	‘Christ’
		es.pi.ná.so	e.pi.na.só	‘spine’
		es.tó.nja	e.to.ná	‘Estonia’
		is.lá.ndja	iɿ.la. <sup>n</sup> dá	‘Iceland’
		áws.trja	aw.te.rjá	‘Austria’
		aws.trá.lja	aw.ta.ra.ljá	‘Australia’
(c)	l	ból.sa	vo.sá	‘bag’
		al.bá.nja	a.va.ná	‘Albania’
(d)	ð	péð.ro	pe.rú	<i>proper name</i>
(e)	β	páβ.lo	pa.lí	<i>proper name</i>
(f)	p	nep.tú.no	ne.tu.nó	‘Neptune’

(g)	Multiple codas	kár.men	ka.mé	<i>proper name</i>
		kár.los	ka.ló	<i>proper name</i>
		fran.sís.ka	ʃi.ká	<i>proper name</i>

In each of the cases of loans simultaneously containing codas and repairing lexical stress, which constitutes all of the loans seen in (28), we see the coda consistently repaired as well. The single exception to this rule is seen in (29).

(29)

<b>Spanish</b>	<b>Guarani</b>	
o.βís.po	o.bis.pó	‘bishop’

With regard to this loan, its source was a Catholic missal (Fritz 2004) only in print in rural Paraguay and created for local use. The loan appears verbatim as *obispo*, so by the conventions of Guarani orthography it would seem to be stressed word-finally. Interestingly, there are two loans in the missal from the Spanish source word *obispo*, the other being [o.vís.po] (orthographic *ovísipo*); it may be the case then that this loan in (29) 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 orthographic error could in this case be to blame for the discrepancy and the loan may not constitute a genuine exception. However, given the impossibility of knowing with certainty, the loan is included as a single exception.

Summarizing the facts up to this point, we have established that 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 summarized in (30) where “→” is read “implies.”

(30)

*Complex onset repair* → *Lexical stress repair* → *Coda repair*

A set-inclusion relationship characterizes the relationships of the phenomena and their repairs. Repairing a complex onset implies repairing lexical stress but, crucially, the reverse is not the case. The same can be said of the relationship between 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 (30) are unidirectional, in (31) we see all non-finally stressed loans in the corpus which repair codas. 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.

(31)

	<b>Status of lexical stress</b>	<b>Spanish</b>	<b>Guarani</b>	
(a)	Repaired	kris.to	ki.ri.tó	‘Christ’

		káβ.ra	ka.va.rá	‘goat’
		a.sú.kar	a.su.ká	‘sugar’
		es.pi.ná.so	e.pi.na.só	‘spine’
		mor.sí.ʎa	<sup>m</sup> bu.sjá	‘blood sausage’
		ból.sa	vo.sá	‘bag’
		sa.túr.no	sa.tuᶞ.nó	‘Saturn’
		nep.tú.no	ne.tu.nó	‘Neptune’
		kár.los	ka.ló	<i>proper name</i>
		péð.ro	pe.rú	<i>proper name</i>
		ar.xé.lja	aᶞ.he.ljá	‘Algeria’
		aws.trá.lja	aw.ta.ra.ljá	‘Australia’
		mer.kú.rjo	me.ku.rjó	‘Mercury’
		páβ.lo	pa.lí	<i>proper name</i>
		fran.sís.ka	ʃi.ká	<i>proper name</i>
		kár.men	ka.mé	<i>proper name</i>
		es.tó.nja	e.to.ᶞá	‘Estonia’
		ir.lá.nda	i.la. <sup>n</sup> djá	‘Ireland’
		is.lá.ndja	iᶞ.la. <sup>n</sup> dá	‘Iceland’
		áws.trja	aw.te.rjá	‘Austria’
		al.bá.nja	a.va.ᶞá	‘Albania’
		di.na.már.ka	<sup>n</sup> di.na.maᶞ.ka	‘Denmark’
(b)	Tolerated	tur.kí.a	tuᶞ.kjá	‘Turkey’
		tí.fus	tí.fu	‘typhus’
		lú.nes	lú.ne	‘Monday’
		es.kí.na	e.kí.na	‘corner’
		es.kwé.la	e.kwé.la	‘school’
		ló.ndres	ló. <sup>n</sup> dre	‘London’
		pwéβ.lo	piᶞ.é.lo	‘town’
		pé.res	pé.re	<i>proper name</i>
		síð.ra	sí.ra	‘cider’

Absent in (31) are nasal codas. As noted above, the role of stressed nasal vowels in Guarani nasal harmony introduces an additional complication in these cases. 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, thereby not conflicting with Guarani phonology and providing no information on Guarani treatment of stress. Nasal codas aside, (31) demonstrates that although repair of stress implies repair of codas, repair of codas does not imply repair of stress.

Similarly, repair of codas does not imply the repair of complex onsets. Shown in (32) 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.

(32)	Status of complex onset	Spanish	Guarani
------	-------------------------	---------	---------



(a)	Repaired	<b>krus</b>	ku.ru.sú	‘cross’
		<b>kris.to</b>	ki.ri.tó	‘Christ’
		<b>kris.to.βál</b>	ki.ri.tó	<i>proper name</i>
		<b>aws.trá.lja</b>	aw.ta.ra.ljá	‘Australia’
		<b>áws.trja</b>	aw.te.rjá	‘Austria’
(b)	Tolerated	<b>es.kri.βír</b>	kri.ví	‘to write’
		<b>tra.tár</b>	tra.tá	‘to treat’
		<b>ló.ndres</b>	ló. <sup>n</sup> dre	‘London’
		<b>dro.yár</b>	dro.yá	‘to drug’

Lastly, as was shown in (26), the repair of complex onsets implies the repair of lexical stress; however, as seen in (33) the opposite is not the case. Here we see those loans which both repair stress and contain a complex onset. Out of these ten loans, eight repair the complex onset and two do not.

(33)				
	<b>Status of complex onset</b>	<b>Spanish</b>	<b>Guarani</b>	
(a)	Repaired	<b>kris.to</b>	ki.ri.tó	‘Christ’
		<b>gré.sja</b>	gi.re.sjá	‘Greece’
		<b>i.ŋgla.té.ra</b>	i. <sup>n</sup> gia.te.rá	‘England’
		<b>áws.trja</b>	aw.te.rjá	‘Austria’
		<b>e.ri.tré.a</b>	e.ri.ti.re.á	‘Eritrea’
		<b>aws.trá.lja</b>	aw.ta.ra.ljá	‘Australia’
		<b>kro.á.sja</b>	kɔ̃o.a.sjá	‘Croatia’
		<b>ku.lán.tro</b>	ku.ra.tú	‘coriander’
(b)	Tolerated	<b>gra.ná.ða</b>	gra.na.dá	<i>kind of fruit</i>
		<b>grá.sja</b>	gra.sjá	‘joke’, ‘grace’

There is a hierarchical pattern in the repairs of these three phenomena when they appear in loans from Spanish: certain repairs unidirectionally imply others. It is through these unidirectional implications that we see the core-periphery structure of the lexicon. Repairs are not made at random but in a systematic way by the grammar.

### 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 of their members, and here each stratum will be discussed in turn.

Before exploring the phonological characteristics of the various strata, however, a note about the frequency of the relevant phonological structures is necessary. In several strata the loans which have the necessary characteristics to be unambiguously included as a member of that stratum are few in number. This is a limitation of the corpus (and by extension the Guarani lexicon) and not a limitation of the stratal structure. 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; 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. Nevertheless, the implicational relationships between repairs discussed above support the view that there are strata in the Guarani lexicon that exhibit predictable patterns.

### 3.2.1 STRATUM 1 (NATIVE)

As previously discussed, the first Spanish loans which entered Guarani and underwent total repair are indistinguishable from native Guarani words phonologically speaking. The loans included in this “native” stratum are identified via their repair of complex onsets, codas and non-final lexical stress, as these are some of the core distinguishing features which differentiate Spanish phonology from that of Guarani.

As discussed above it is important to note that in order to have 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. 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 either of the strata which repair non-final stress. In cases such as 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 recognition that this is neither directly justifiable nor problematic for the analysis, as the decision of where to place loans like this is essentially arbitrary.<sup>16</sup>

Seen in (34) are examples of loans falling into stratum 1 which repair any and all of the relevant non-native phonological characteristics.

(34)	<b>Phenomenon repaired</b>	<b>Spanish</b>	<b>Guarani</b>	
(a)	Lexical stress	lí.βja bá.ka sa.pá.to	li.vjá va.ká sa.pa.tú	‘Libya’ ‘cow’ ‘shoe’
(b)	Coda	fal.tár bo.mi.tár ba.lér	va.tá go.mi.tá va.lé	‘to lack’ ‘to vomit’ ‘to be worth’
(c)	Complex onset & coda	<b>krus</b>	ku.ru.sú	‘cross’
(d)	Complex onset & stress	<b>kris.to.βál</b> e.ri.tré.a  gré.sja kro.á.sja	ki.ri.tó e.ri.ti.re.á  gi.re.sjá kjo.a.sjá	<i>proper name</i> ‘Eritrea’  ‘Greece’ ‘Croatia’

<sup>16</sup> 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.

(e)	Coda & stress	<b>kár.los</b> <b>ból.sa</b>	ka.ló vo.sá	<i>proper name</i> 'bag'
(f)	All three	<b>al.bá.nja</b> <b>krís.to</b> <b>áws.trja</b> <b>aws.trá.lja</b>	a.va.ná ki.ri.tó aw.te.rjá aw.ta.ra.ljá	'Albania' 'Christ' 'Austria' 'Australia'

### 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 is complex onsets. Here appear the 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 across various loans, showing that hypothetically a Spanish loan with all three of these phonological structures would pattern in this way. Shown in (35) are those loans in the corpus which permit complex onsets while still making repairs to codas or lexical stress.

(35)

	<b>Phenomenon repaired</b>	<b>Spanish</b>	<b>Guaraní</b>	
(a)	Coda	<b>tra.tár</b> <b>plu.tón</b> <b>es.kri.βír</b> <b>dro.yár</b>	tra.tá plu.tó kri.ví dro.yá	'to treat' 'Pluto' 'to write' 'to drug'
(b)	Stress	<b>gra.ná.ða</b> <b>grá.sja</b>	gra.na.dá gra.sjá	<i>kind of fruit</i> 'joke', 'grace'

### 3.2.3 STRATUM 3 (PARTIALLY NATIVIZED)

After those loans which show tolerance to complex onsets comes the next stratum of loans which further increase their faithfulness to their Spanish source 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 (36) is the single loan which tolerates both a complex onset and non-final stress while still repairing a Spanish coda.

(36)

<b>Spanish</b>	<b>Guaraní</b>	
<b>lón.dres</b>	lón. <sup>n</sup> dre	'London'

Other loans included in this medial stratum due to their avoidance of codas but not non-final stress are seen in (37).

(37)

<b>Spanish</b>	<b>Guarani</b>	
tí.fus	tí.fu	‘typhus’
lú.nes	lú.ne	‘Monday’
es.kí.na	e.kí.na	‘corner’
es.kwé.la	e.kwé.la	‘school’
pwéβ.lo	pié.lo	‘town’
pé.res	pé.re	<i>proper name</i>
síð.ra	sí.ra	‘cider’

#### 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 Guarani grammar distinguishes nasal and non-nasal codas with regard to what is permissible. In the loans seen below in (38) non-nasal codas are tolerated while nasal codas are still avoided. 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 few; nonetheless the existence of loans which pattern this way motivates the existence of this stratum.

(38)

<b>Spanish</b>	<b>Guarani</b>	
is.lám	is.lá	‘Islam’
al.ma.sén	al.ma.sé	‘department store’

#### 3.2.5 STRATUM 5 (UNADAPTED)

Nasal codas, while being the most consistently avoided of the Spanish phonological characteristics, are finally seen to be tolerated in the loans in the last stratum. In this unadapted stratum Spanish loans are entering essentially unrepaired. Examples of this are seen in (39).

(39)

<b>Spanish</b>	<b>Guarani</b>	
kon.fir.ma.sjón	kon.fir.ma.sjón	‘confirmation’
en.sa.lá.ða	en.sa.lá.da	‘salad’
flo.ri.pón	flo.ri.pón	<i>flower species</i>
a.la.krán	a.la.krán	‘scorpion’

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 (40).

(40)

<b>Spanish</b>	<b>Guarani</b>	
bro.mís.ta	<sup>m</sup> bro.mís.ta	‘funny’
xe.su.krís.to	he.su.krís.to	‘Jesus Christ’

kris.tjá.na kris.tjá.na ‘Christian’  
 kris.tjá.no<sup>17</sup> kris.tjá.no ‘Christian’

### 3.2.6 SUMMARY OF STRATA

As a means of summarizing the relationships among strata and the phonological characteristics which determine these relationships, the table in (41) demonstrates the treatment of the relevant phonological structures by the different strata (“N Codas” refers to nasal codas, “Codas” refers to non-nasal codas, and “#CC” refers to complex onsets).

(41)

	N CODAS	CODAS	NON-FINAL STRESS	#CC
1. Native	<i>Repaired</i>	<i>Repaired</i>	<i>Repaired</i>	<i>Repaired</i>
2. Mostly nativized	<i>Repaired</i>	<i>Repaired</i>	<i>Repaired</i>	<i>Tolerated</i>
3. Partially nativized	<i>Repaired</i>	<i>Repaired</i>	<i>Tolerated</i>	<i>Tolerated</i>
4. Barely nativized	<i>Repaired</i>	<i>Tolerated</i>	<i>Tolerated</i>	<i>Tolerated</i>
5. Unadapted	<i>Tolerated</i>	<i>Tolerated</i>	<i>Tolerated</i>	<i>Tolerated</i>

### 3.3 EVIDENCE FOR STRATAL SYNCHRONIC RELEVANCE IN GUARANI MORPHOLOGY

Some recent work in loanword phonology has called into question the stratal interpretation of loan adaptation and favored an interpretation that loans in effect expand the native grammar and individual strata are unnecessary (see Rice (2006) for such a proposal concerning stress adaptation in Norwegian). Under such an interpretation of Guarani one would argue that Guarani has in essence become stratum 5 and is now fully tolerant of any and all phonological characteristics found in loans from Spanish. There is evidence that this is not a viable account of the facts in Guarani (see Pinta (2013) and Smith & Pinta (2016) for the results of a nonce-loan nativization experiment in which native speakers show synchronic knowledge of the kinds of implicational relationships among loan repairs that motivate lexical strata). Additional evidence for the synchronically relevant status of lexical strata in Guarani comes from the language’s morphology.

For example, Guarani has a causative prefix, *mbo-/mo-*, which is used with intransitive verbs (Nordhoff 2004). This native Guarani morpheme, while used abundantly with native verbs, does not attach to loan verbs (Bakker & Hekking 2012). An account of the Guarani 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 in accounting for morphemes such as *mbo-/mo-*. Affixes showing sensitivity to whether the stem is native or borrowed bolster claims that strata in modern Guarani are more than just diachronically relevant.

There is no irrefutable evidence that Guarani uses borrowed Spanish morphemes

<sup>17</sup> While it would be inappropriate to list [kris.tjá.na] and [kris.tjá.no] as separate words in Spanish due to its use of grammatical gender, these two words are listed as separate loans in Guarani due to the lack of grammatical gender in that language.

productively, in spite of the rampant borrowing in vocabulary with Spanish morphemes intact (e.g., plural markers, diminutive markers, agentive markers, adverbial markers, etc.) (Bakker & Hekking 2012). However, there is at least one Spanish morpheme, the adverbial marker *-mente*, which could be interpreted as productive or at least on the path to productivity in Guarani.

The common adverbial marker *-mente* is used productively in Spanish to form adverbs out of adjectives (Gómez Rendón 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 & Pountain 1995). Spanish adjectives borrowed into Guarani 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 Guarani lexicon which exist as bare adjectives and as overtly marked adverbs (Bakker & Hekking 2012). It is not readily clear whether the numerous loans in Guarani 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 into adverbs through the use of *-mente*. If this latter explanation were the case it would be noteworthy due to the fact that there are no attested cases of native Guarani adjectives being combined with *-mente* to be turned into adverbs.<sup>18</sup> The morpheme *-mente* therefore may constitute a second case in which the Guarani grammar attaches specific affixes to specific stems on the basis of their status of native or loaned. Although this case cannot be made as definitively as the case of *mbo-/mo-*, it would constitute the opposite of that case in being a morpheme which only attaches to loans and avoids native stems, providing evidence that the Guarani grammar has at least some tendency to attach native affixes only to native stems and loan affixes only to loan stems.

The morphological evidence thus suggests that the grammar of Guarani makes a distinction between native and loan morphemes; such a native/loan split does not independently provide evidence for five distinct strata in the Guarani lexicon, but it does bolster the claim that the Guarani grammar does not universally treat loans identically to native forms.

#### 4 POSSIBLE AND IMPOSSIBLE NATIVIZATIONS

As extensively discussed in Itô & Mester (1999), languages that have synchronically relevant lexical strata show “impossible-nativization effects,” that is, logically possible adaptation forms that are ungrammatical. Guarani is no exception to this. The core-periphery structure of the lexicon allows for a given loan to be adapted in a number of possible ways, but not all logically feasible adaptations are attested.

Looking at a specific example, we turn to the loan for *grace*, Spanish [grá.sja] → Guarani [gra.sjá]. There are three possible nativizations allowed by the grammar depending on the stratal affiliation of the loan. An adaptation of Spanish [grá.sja] falling into stratum 1, the native stratum, would see repairs made to make the form indistinguishable from native Guarani lexical items. This process would repair the complex onset as well as the lexical stress, yielding a Guarani adaptation of [gì.ra.sjá] (where /ì/ is chosen as the epenthetic vowel for the sake of illustration). If this loan were affiliated with stratum 2 (mostly nativized) we would expect tolerance of the complex onset but not of non-final stress, yielding [gra.sjá], which in this case is the attested nativization of the form. Affiliation with stratum 3 (partially nativized) would require

---

<sup>18</sup> In fact, previous research has found that “Guarani...does not seem to provide any examples of Spanish derivational or inflectional markers attached to native stems.” (Bakker & Hekking 2012: 199).

tolerance of both the complex onset and non-final stress; this would produce the unrepaired form [grá.sja]. In fact, if the adaptation patterned with loans from strata 3, 4 or 5 then we would in all cases expect the unrepaired form [grá.sja] given this particular loan's lack of a coda – nasal or non-nasal – the treatments of which are the distinguishing features of strata 3-5.

Crucially, no matter what the stratal affiliation of the loan, the form [gi.rá.sja] is not a possible outcome; the core-periphery structure shown in (41) will not produce a form in which the complex onset is repaired without lexical stress also being repaired. With the previously discussed exception of the single loan in (27), no loan in the corpus whose original form contains a complex onset and non-final lexical stress sees this type of adaptation, and the hypothetically possible [gi.rá.sja] is thus deemed an impossible nativization. These effects hold for all other combinations of phonological characteristics of which Guarani is intolerant as well, such that for any loan whose original Spanish form has more than one characteristic disallowed in native Guarani there will be a hypothetically possible nativization of that loan whose form is unattested.

The existence of these impossible nativization effects is not a new topic of discussion among loanword phonologists. Recent works have modeled them within Optimality Theory (Prince & Smolensky 1993) through the interaction of a fixed hierarchy of faithfulness constraints with freely ranked markedness constraints.<sup>19</sup> Crucially, Guarani native speakers have also been shown to exhibit strong avoidance of impossible nativizations in forced choice nonce nativization tests (Pinta 2013), although not all speakers seem to have exactly the same set of synchronically relevant strata (Smith & Pinta 2016).

## 5 CONCLUSION

This chapter has presented an analysis of the phonological adaptation of Spanish loans in Guarani and a discussion of what the patterns of adaptation tell us about the structure of the Guarani lexicon. An analysis of a corpus of Spanish loans in Guarani reveals the systematic and stratal nature of the Guarani lexicon evidenced by the differing adaptation strategies of the loans. Loans form groupings on the basis of their phonological characteristics, with some loans repaired to be fully compliant with native Guarani phonology, some which show only partial adaptation, and some which show no adaptation whatsoever. These patterns determine strata that are grouped in a set-inclusion pattern in which their phonological characteristics are seen to overlap so that progressively more phonological structures become possible in progressively more peripheral strata. In this sense the lexicon is characterized by its core-periphery structure.

The sociolinguistic history of Guarani has afforded 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. As has been shown, however, these differences are not random, and evidence for patterns within loan adaptation is abundant in Guarani.

The structure of the lexicon in conjunction with the behavior of some Guarani morphemes provides evidence for the synchronic relevance of the lexicon's stratification. Moreover, evidence for impossible nativization effects (Itô & Mester 1999) is confirmed in Guarani on the basis of experimental results (see Pinta (2013) and Smith & Pinta (2016)). Formal optimality-theoretic analyses have been able to account for many aspects of the impossible nativization effects observed in various languages, but the precise details of the nature of

---

<sup>19</sup> For in depth discussion of the strengths and weaknesses of optimality theoretic analyses of this kind see Itô & Mester (1999, and earlier work) for Japanese, Karvonen (1998) for Finnish, and Pinta (2013) for Guarani.

faithfulness and markedness constraint interaction are still debated and are a topic of further research. Guarani provides a useful case study of a language with extensive borrowing, showing lexical stratification and impossible nativization effects as well as other consequences of prolonged language contact.

#### REFERENCES

- Bakker, D., & Hekking, E. (2012). Constraints on morphological borrowing: Evidence from Latin America. In L. Johanson & M. Robbeets (Eds.), *Copies versus cognates in bound morphology* (pp. 187-220). Leiden: Brill.
- Britton, A. S. (2005). *Guarani-English/English-Guarani concise dictionary*. New York: Hippocrene Books Inc.
- de Assis, C. F. (2008). *Dicionário guarani-português/português-guarani*. São Paulo: Cecy Fernandes de Assis.
- de Bruyne, J., & Pountain, C.J. (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. & Alcaraz, F.A. (1997). *Diccionario guaraní-español/español-guaraní*. Asunción: Colección Ñemitỹ.
- Díaz, M. R. D. (2006). *Gran diccionario de la lengua guaraní*. Corrientes: Moglia Ediciones.
- Fritz, M., (Ed.). (2004). *Misal romano: Edición bilingüe guaraní-nivaclé*. Asunción: Vicariato Apostólico del Pilcomayo.
- Gómez Rendón, J.A. (2008). *Typological and social constraints on language contact: Amerindian languages in contact with Spanish*. Utrecht: LOT
- Gregores, E., & Suárez, J.A. (1967). *A description of colloquial Guarani*. The Hague-Paris: Mouton.
- Hualde, J. I. (2005). *The sounds of Spanish*. New York: Cambridge University Press.
- Itô, J. & Mester, A. (1999). The phonological lexicon. In N. Tsujimura (Ed.), *A handbook of Japanese linguistics* (pp. 62-100). Oxford: Blackwell Publishers.
- Karvonen, D. (1998). *Finnish loanword phonology and the core-periphery structure of the lexicon*. Unpublished manuscript, UC Santa Cruz.
- Lustig, W. (2005). *Guarani für Paraguay: Wort für Wort*. Bielefeld: Reise Know-How.
- Mayans, A. O. (1980). *Nuevo diccionario español-guaraní/guaraní-español*. Buenos Aires: Editorial Universitaria de Buenos Aires.
- 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 Guarani* (Master's thesis). University of Cologne, Germany.
- Pinta, J. (2013). *Lexical strata in loanword phonology: Spanish loans in Guarani* (Master's thesis). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 1550078)



- Prince, A., & Smolensky, P. (1993). *Optimality theory: Constraint interaction in generative grammar*. Unpublished manuscript, Rutgers University & University of Colorado, Boulder.
- Rice, C. (2006). Norwegian stress and quantity: The implications of loanwords. *Lingua*, 116(7), 1171-1194.
- Rivas, A. (1975). Nasalization in Guarani. In E. Kaisse & J. Hankamer (Eds.), *Proceedings of NELS*, Volume 5. 134-143.
- Smith, J. L. (2006). Loan phonology is not all perception: Evidence from Japanese loan doublets. In T. J. Vance & K. A. Jones (Eds.), *Japanese/Korean linguistics, volume 14* (pp. 63-74). Stanford: CSLI.
- Smith, J. L. (2009). Source similarity in loanword adaptation: Correspondence theory and the posited source-language representation. In S. Parker (Ed.), *Phonological argumentation: Essays on evidence and motivation* (pp. 155-177). London, Equinox.
- Smith, J. L., & Pinta, J. (2016). *Aggressive core-periphery phonology in Guarani: Implications for stratal faithfulness*. Manuscript in preparation. The University of North Carolina at Chapel Hill & The Ohio State University.
- Tase, P. (2012). *Discursos y mensajes: Excelentísimo Presidente Don Fernando Lugo*. Lulu Press.
- Tonhauser, J., Beaver, D., Roberts, C., & Simons, M. (2013). Toward a taxonomy of projective content. *Language*, 89(1), 66-106.
- Tonhauser, J. & Colijn E. (2010). Word order in Paraguayan Guarani. *International Journal of American Linguistics*, 76(2), 255-288.
- Walker, R. (1999). Guarani voiceless stops in oral versus nasal contexts: An acoustical study. *Journal of the International Phonetic Association*, 29(1), 63-94.
- Velázquez-Castillo, M. (2013). The deictic nature of Guarani evidentials: Distance, directionality and perspective. [Talk handout]. Department of Romance Studies, University of North Carolina, Chapel Hill, NC.
- Zarratea, T. (2013). *Gramática elemental del guaraní paraguayo*. Asunción: ServiLibro.
- Zec, D. (2007). The syllable. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 161-194). Cambridge: Cambridge University Press.