OPAQUE NASALIZATION IN LUDLINGS AND THE PRECEDENCE RELATIONS OF REDUPLICATION AND INFIXATION[1]

Maximiliano GUIMARÃES

Universidade Federal do Paraná. Ph.D. in Linguistics, University of Maryland at College Park
E-mail: maxguimaraes@ufpr.br

Andrew NEVINS

University College London. Ph.D. in Linguistics, Massachusetts Institute of Technology
E-mail: a.nevins@ucl.ac.uk

Abstract
We examine the pattern of opaque nasalization in two invented language games (*ludlings*), as played by volunteers who were native speakers of the dialect of Brazilian Portuguese spoken in the city of Salvador. One group of speakers showed an opaque pattern, in which nasalization over-applied in the reduplicative language game *Língua do Pê* but under-applied in the infixing language game *Língua do Ki*. These results have three consequences: (i) they provide further support for an abstract Multiprecedence-and-Linearization representation (Raimy, 2000a,b) for reduplication and infixation, such that infixation starts and ends between two segments that transitively precede each other, while reduplication does not; (ii) they demonstrate that, for a significant subgroup of our experimental participants, nasalization in Northeastern Brazilian Portuguese is an active

1 We thank Bernadete Abaurre, William Idsardi, Nicholas Neasom, Eric Raimy, Filomena Sandalo, and Thanasis Soultatis, for their insightful comments on an early draft. We also thank the audience of GLOW 2006, held in Barcelona, where we presented a previous version of this paper. We appreciate their input.

rule, and not simply a lexicalized pattern; (iii) they provide support for a novel condition on rule application: the iota-operator of Russell (1905), which imposes a uniqueness condition on structural descriptions.

Keywords
over/under-application; multiprecedence; nasalization; ludlings; Brazilian Portuguese

1. Nasalization in Brazilian Portuguese

Pretonic nasalization is a shibboleth of Northeastern dialects of Brazilian Portuguese. Thus, pronunciation of the name of the fruit *banana* as [bɐ̃ˈnɐ̃nɐ], rather than [baˈnɐ̃nɐ], is a give-away that the talker is from that region, because talkers from other regions only nasalize vowels in open syllables when they are stressed. In this section, we provide an overview of these nasalization processes. One of our goals in this paper is to use invented language games (*ludlings*) to investigate the empirical status of this process and subsequently present a formal model of its application.

1.1 Nasalization Patterns in All Dialects of BP ("Standard BP")

Nasalization of vowels in stressed syllables exists in all dialects of Brazilian Portuguese (BP), and can be schematized as in the following rule:

(1) $V_{[+stress]} \rightarrow [+nasal] \;/\; \_C_{[+nasal]}$

Importantly, the rule in (2) operates even when the following consonant is heterosyllabic, as can be seen in the following examples:

(2) *Tonic heterosyllabic regressive nasalization in BP*:

a:	ba.ˈnɐ̃.nɐ	'banana'	f:	ˈsẽ.ɲɐ	'password'
b:	a.ˈɾɐ̃.ɲɐ	'spider'	g:	is.ˈpũ.mɐ	'foam'
c:	pow.ˈtɾõ.nɐ	'easy chair'	h:	ˈũ.ɲɐ	'fingernail'
d:	a.ˈɾõ.mɐ	'aroma'	i:	bu.ˈzĩ.nɐ	'horn-honk'
e:	ɐ̃.ˈtẽ.nɐ	'antenna'	j:	ga.ˈlĩ.ɲɐ	'hen'

Nasalization also occurs if the following consononant is tautosyllabic, as in (3).

(3) *Tonic tautosyllabic regressive nasalization in BP*:

a:	ˈsɐ̃^{m}.bɐ	‘samba dance’	b:	ˈsɐ̃^{n}.tɐ	‘(female) saint’
c:	ˈpõ^{m}.bɐ	‘pigeon’	d:	ˈpõ^{n}.tʊ	‘point’
e:	ˈtẽ^{m}.pʊ	‘time’	f:	fa.ˈzẽ^{n}.dɐ	‘farm’
g:	ˈtũ^{m}.bɐ	‘grave’	h:	ˈmũ^{n}.dʊ	‘world’
i:	ta.ˈɾĩ^{m}.bɐ	‘know-how’	j:	ta.ma.ˈɾĩ^{n}. dʊ	‘tamarind’

There is some variation in whether the nasal consonant in the coda of the stressed syllable in (3) involves a distinct consonantal closure between the nasalized vowel and the following obstruent. Phonetic studies on Brazilian Portuguese, such as those of Shosted (2003), have revealed a brief and variable period of closure. The phonetic situation of fleeting/epenthetic coda nasals is similar to that of other languages with nasal vowels, such as Hindi, as studied by M. Ohala (1983) and Ohala & Ohala (1991), and Polish, as studied by Rubach (1977) and Bethin (1995). In terms of the traditional phonological understanding of BP, structuralist studies such as Mattoso Câmara Jr. (1970) have proposed that all nasal consonants are deleted in coda position[2]:

(4) $C_{[+nasal]} \rightarrow \varnothing \,/\, \_\_\,]_{\sigma}$

(5) /ˈsam.ba/ → [ˈsɐ̃^{m}.bɐ] (by (1)) → [ˈsɐ̃.bɐ] (by (4))

Evidence for the rule in (4) comes not only from the pronunciation of word-internal coda consonants in BP, but from alternations involving word-final syllables[3]:

(6) a: [lɐ̃] ‘wool’ b: [la.ˈnej.ɾʊ] ‘wool maker’

The deletion rule in (4) thus does not apply in cases of resyllabification, as in (6b). In this respect, the behavior of word-final nasal consonants in BP is reminiscent of Chomsky & Halle’s (1968) treatment of word-final nasal consonants in English alternations such as *damn* [dæm] ~ *damnation* [dæm.ˈnej.ʃən], which they analyzed as the result of a deletion rule applying to the coda nasal in *damn* but which fails to apply under resyllabification.

The distribution of nasalized vowels in BP extends beyond (1)-(6), however. Words with a coda nasal consonant in a non-stressed syllable still yield nasalization of the preceding vowel, as can be seen in the examples in (7).[4]

---

[2] Nobiling (1903) was perhaps the first to posit underlying nasal glide elements in coda position. The rule in (4) is motivated by the fact that all sonorant codas lenite in Brazilian Portuguese. Lateral [l] in onset position alternates with the glide [w] in codas: [ʒoɦ.ˈnaw] / [ʒoɦ.na.ˈlej.ɾʊ] ‘newspaper/newsboy’. The tap [ɾ] in onset alternates with the fricative [h] (and its homorganic [ɦ]) in codas: [floh]/[ˈflo.ɾɪs] ‘flower/flowers’. Other than nasals, laterals, and rhotics, the only other segments allowed in BP codas are the fricatives /s/ and /z/, and even these are subject to deletion/lenition, as can be found in high-frequency words such as *mesmo* ‘same’, often realized as [meɦ.mʊ].

[3] Lipski (1975: 64) cites evidence for the psychological reality of a coda from illiterate speakers who reduce /re.ˈzaN.do /‘praying’ to [re.ˈzɐ̃.nʊ], rather than [re.ˈzɐ.nʊ] or [re.ˈzɐ̃ʊ].

[4] Brazilian Portuguese has a fully productive and predictable rule of vowel reduction, raising post-tonic /o/ to [ʊ] and post-tonic /e/ to [ɪ].

(7) *Tautosyllabic regressive nasalization in BP*:

a:	/pandeiɾo/	[pɐ̃.ˈdej.ɾʊ]	'tambourine'
b:	/tɾombone/	[tɾõ.ˈbõ.nʲɪ]	'trombone'
c:	/dentista/	[dẽ.ˈtʃis.tɐ]	'dentist'
d:	/funduɾa/	[fũ.ˈdu.ɾɐ]	'depth'
e:	/sintuɾa/	[sĩ.ˈtu.ɾɐ]	'waist'

Unlike [ˈsɐ̃.bɐ] and the examples in (3), the nasalization of the pretonic vowels by a tautosyllabic coda nasal does not fall under the rule in (1). Thus, to maintain the analysis of all nasalized vowels as the result of a rule of nasalization, (1) needs to be split into two rules: a heterosyllabic rule of nasalization, applying only in stressed syllables, and a tautosyllabic rule of nasalization, applying everywhere:

(8) a: $V_{[+stress]} \rightarrow [+nasal] \,/\, \_\_\,]_{\sigma} \,.\, C_{[+nasal]}$

b: $V \rightarrow [+nasal] \,/\, \_\_\, C_{[+nasal]}\,]_{\sigma}$

We will note here that some researchers reject the treatment of vowel nasalization through two separate, formally similar rules, and prefer instead to analyze the vowels of [lɐ̃] and [ˈsɐ̃.bɐ] as underlyingly [+nasal], with the orthography simply reflecting two different conventions for indicating nasalization: <ã> for word-final [ɐ̃], and <am> for word-internal [ɐ̃]. When we turn to the pattern of nasalization in Northeastern BP (focusing in particular on the dialect spoken in Salvador, the capital of the state of Bahia, and the third largest city in Brazil), we will revisit the issue of the formal complexity of (8) and the question of underlying nasal vowels.

1.2 Nasalization patterns in Salvador BP

In addition to nasalization of stressed vowels by a heterosyllabic following nasal and nasalization of all vowels by a tautosyllabic coda consonant, Salvador BP has unstressed vowels undergo nasalization when preceding a heterosyllabic nasal. In (9) we compare the behavior of unstressed vowels in open syllables in Salvador BP with their pronunciations elsewhere in Brazil (i.e. "Standard BP", insofar as such a term represents an idealization based on the features of the language more commonly used in politics and mass-media entertainment). All of the following words have penultimate stress, and demonstrate pretonic nasalization:[5]

[5] At this point and various subsequent junctures throughout the paper we will mention that this allophonic rule does not apply 100% of the time, and that the pattern here represents an idealization of the actual pattern, which we estimate as highly regular. We venture that this is true of many allophonic rules in many languages, often described in categorical format, such as *s*-palatalization of codas in Rio de Janeiro BP (which Guy (1981) has shown to be variable), unstressed vowel reduction in BP, and flapping in English, all of which apply stochastically under the general model of variable rules proposed in Cedergren & Sankoff

(9)

	Orthography	*Salvador BP*	*Standard BP*	*Gloss*
	caneta	kɐ̃.ˈne.tɐ	ka.ˈne.tɐ	‘pen’
	familia	fɐ̃.ˈmi.ʎɐ	fa.ˈmi.ʎɐ	‘family’
	começo	kõ.ˈme.sʊ	ko.ˈme.sʊ	‘beginning’
	fonema	fõ.ˈnẽ.mɐ	fo.ˈnẽ.mɐ	‘phoneme’
	demora	dẽ.ˈmɔ.ɾɐ	de.ˈmɔ.ɾɐ	‘delay’
	cenoura	sẽ.ˈno.ɾɐ	se.ˈno.ɾɐ	‘carrot’
	fumaça	fũ.ˈma.sɐ	fu.ˈma.sɐ	‘smoke’
	luneta	lũ.ˈne.tɐ	lu.ˈne.tɐ	‘spyglass’
	cinema	sĩ.ˈnẽ.mɐ	si.ˈnẽ.mɐ	‘movie theater’
	limite	lĩ.ˈmi.tʃɪ	li.ˈmi.tʃɪ	‘limit’

Salvador BP thus includes a process which may schematized by the following rule, which may be contrasted with (8a) above, repeated as (11) below:

(10) *Salvador pre-tonic nasalization*:
$V \rightarrow [+\text{nasal}] \, / \, \_\_ \,]_{\sigma} . \, C_{[+\text{nasal}]}$

(11) *Standard BP tonic nasalization*:
$V_{[+\text{stress}]} \rightarrow [+\text{nasal}] \, / \, \_\_ \,]_{\sigma} . \, C_{[+\text{nasal}]}$

As Salvador BP also contains rules (8a) and (8b), these may all be collapsed:

(12) *Salvador nasalization*:
$V \rightarrow [+\text{nasal}] \, / \, \_\_ \, C_{[+\text{nasal}]}$

The arguments for underlying nasal vowels based on the redundancy of the structural change in (8a) and (8b) are thus no longer valid for Salvador BP, which only has the single nasalization rule in (12). However, a potential argument for underlying nasalized vowels based on the existence of minimal pairs such as (13) remains:

(13)	a:	[la]	<lá> ‘there’	b:	[sĩ]	<sim> ‘yes’
		[lɐ̃]	<lã> ‘wool’		[si]	<se> ‘if’

Notably, minimal pairs such as (13) only exist in word-final position, which led Mattoso Câmara to propose the deletion rule in (4), which was partially supported by alternations such as *<lã>/<laneiro>*. However, such pairs are admittedly few, and the possibility that some learners of BP may have lexically specified nasal vowels remains not only an open analytical option for a description of the language, but also an open option for the child learning which phonemes are contrastive. Importantly, if the explanation of (13) is based on underlying [+nasal] vowels, rather than nasalization by an underlying

(1974). As the stochastic and almost-but-not-quite fully consistent nature of the application of these rules is independent of their formal expression, we abstract away from this issue.

word-final nasal consonant, then the possibility of contrastive [+nasal] vowels becomes a potential option for other positions within the word.

The structure of this paper is as follows. In Section 2, we will introduce the two invented language games that we asked our volunteers to play: *Língua do Pê* and *Língua do Ki*. In Section 3, we will discuss the results of *Língua do Pê*, in which many of our volunteers exhibited overapplication of nasalization in reduplicative structures based on [kɐ̃.ˈne.tɐ]. We will discuss four possible analyses of the overapplication facts: (a) underlying [+nasal] specification, (b) Base-Reduplicant Identity as implemented in Correspondence Theory, (c) two cycles of nasalization, and (d) a Multiprecedence-and-Linearization structure for reduplication. In Section 4, we discuss the results of *Língua do Ki*, in which many of our volunteers exhibited underapplication of nasalization in infixation structures based on [kɐ̃.ˈne.tɐ]. We will demonstrate that none of the analyses in (a)-(c) provide adequate models of the facts, and conclude that underapplication and overapplication of nasalization in Salvador BP are best handled by an active rule of nasalization similar to that in (12) above, but with the addition of a novel condition on rule application that becomes necessary for multiprecedence representations. Section 5 concludes the paper.

2. Língua do Pê and Língua do Ki

In drawing on evidence for phonological representations through observing the output of invented language games, we follow a rich tradition in the literature, including Sherzer (1970), Campbell (1980), Vago (1985), and Bagemihl (1995). Our specific interest here is in reduplicative and infixing language games, which require fundamental alteration of the basic precedence structure of words submitted to the game, and in so doing, may disrupt adjacency relations between segments and thereby destroy potential environments for rule application. We will describe each game in turn, and then describe our methodology for teaching volunteers the game and subsequently recording their productions for certain inputs of interest.

2.1. Língua do Pê

Língua do Pê (LdP) is widespread as a children's language game throughout Brazil, although many different variants of it exist. The variant we are dealing with (which is the most popular in Salvador) has an iterative infixing reduplicative pattern: it adds the syllable ***-pV(C)-*** after every syllable, where ***V(C)*** is a copy of the vowel and the coda consonant (if any) of the preceding syllable.

We present an informal characterization in (14):

(14) After each syllable Σ, insert a copy of Σ in which the onset consists of /p/

We will delay a formal characterization of the process yielding LdP until Section 3.4. At this point, it is worth noting that LdP bears a high resemblance to other forms of iterative infixation found in language games, such as ***-pV-*** reduplication in Jerigonza (played in Colombian Spanish (PIÑEROS, 1999)), ***-ub-*** infixation in the North American English game Ubbi Dubbi (popularized in the 1970s children's television show *Zoom*) and ***-Vv-*** reduplication (played in Hungarian (HARRISON; KAUN, 2000). The default stress pattern for LdP is binary iambic footing, with greater emphasis on the reduplicative syllable.

In (15), we provide sample inputs and outputs for LdP, indicating the added syllable by underlining. Note that /p/ replaces the entire onset, and not just the first member of a complex onset.[6]

(15)	a:	ˈbɔ.lɐ	bɔ.<u>pɔ</u>.la.<u>pah</u>	'ball'
	b:	mah	mah.<u>pah</u>	'sea'
	c:	ˈah.tʃɪ	ah.<u>pah</u>.te.<u>pe</u>	'art'
	d:	ˈtɾi.ʎʊ	tɾi.<u>pi</u>.ʎo.<u>po</u>	'trail'
	e:	vɛɦ.ˈda.dʒɪ	vɛh.<u>pɛɦ</u>.da.<u>pa</u>.de.<u>pe</u>	'truth'
	f:	ga.ˈɾo.tʊ	ga.<u>pa</u>.ɾo.<u>po</u>.to.<u>po</u>	'boy'

2.2 Língua do Ki

Língua do Ki (LdK) has an iterative infixing pattern: it adds the syllable –*ki*- after every syllable. We present an informal characterization in (16):

(16) After each syllable Σ, insert the syllable -ki-

The default stress pattern for LdK is binary iambic footing, with greater emphasis on the infixed syllable. In (17), we provide sample inputs and outputs for LdP, indicating the added syllable by underlining.

(17)	a:	ˈbɔ.lɐ	bɔ.<u>ki</u>.la.<u>ki</u>	'ball'
	b:	mah	mah.<u>ki</u>	'sea'
	c:	ˈah.tʃɪ	ah.<u>ki</u>.tʃɪ.<u>ki</u>	'art'
	d:	ˈtɾi.ʎʊ	tɾi.<u>ki</u>.ʎo.<u>ki</u>	'trail'
	e:	vɛɦ.ˈda.dʒɪ	vɛh.<u>ki</u>.da.<u>ki</u>.de.<u>ki</u>	'truth'
	f:	ga.ˈɾo.tʊ	ga.<u>ki</u>.ɾo.<u>ki</u>.to.<u>ki</u>	'boy'

Língua do Ki does not, to our knowledge, exist already as a language game among children in Brazil. We invented it for the purposes of contrasting iterative reduplication with iterative infixation, with the goal of keeping the fixed segmental material simple and distinct from that of LdP.

[6] LdP and LdK's stress pattern of successive iambic feet makes the final syllable stressed, hence immune from reduction. This explains why [o] and [e] can show up in the last syllable (rather than [ʊ] and [ɪ]).

2.3 Training and Testing Methodology

We recruited eleven volunteers among undergraduate students from the Federal University of Bahia, located in Salvador city. Interviewing volunteers one at a time, we made sure that all of them were native speakers of the local dialect through a pretest of questions designed to elicit words with pretonic nasal syllables, such as showing them a picture of a pen and asking what it was, and ensuring that the response was [kɐ̃.ˈne.tɐ] and not [kɐ.ˈne.tɐ]; or asking questions such as *"What do you call the place you go to watch movies?"* and ensuring that the answer was [sĩ.ˈnẽ.mɐ] and not [si.ˈnẽ.mɐ].

We taught each volunteer LdP and LdK in two different sessions, randomizing which game was taught first. For each game, we first gave the volunteer the explicit rule for playing the game as formulated in (14) or (16). Then we trained each volunteer on 25 words, increasing in complexity from monosyllabic words to bisyllabic to trisyllabic words. Finally, we assessed each volunteer's fluency by asking them to transform entire phrases into LdP or LdK, such as *Feliz natal e um próspero ano novo* ('Merry Christmas and a Happy New Year') into *Fe.pe.liz.piz na.pa.tal.pal e.pe ũ.pũ pros.pos.pe.pe.ro.po a.pa.no.po no.po.vo.po* or *Fe.ki.liz.ki na.ki.tal.ki e.ki ũ.ki. pros.ki.pe.ki.ro.ki. a.ki.no.ki.no.ki.vo.ki.* During this training period, we observed each volunteer's fluency in the game and ensured that by the end of the training period they produced each phrase with a minimal amount of hesitation between each syllable. Volunteers who were not fluent enough in LdP or LdK after this training period were dismissed from further participation.

After training each volunteer to the point of fluency in LdP and LdK, we presented nine test items to be transformed into the game (cf. (18)), intermixed with nine filler items. The nine test items all contained vowels in open syllables that immediately preceded a heterosyllabic nasal consonant, such as [kɐ̃.ˈmi.zɐ]. Eight of the items contained the vowel in question in a pretonic syllable, and one of the items contained the vowel in question in a stressed syllable. Seven of the nine test items had /a/ as the vowel of interest, because the effect of nasalization on /a/ also has a centralizing effect on the height of this vowel, thus making it extremely simple to perceive in real-time and with the naked ear whether nasalization occurred or not.

(18)	a:	kɐ̃.ˈmi.zɐ	'shirt'	f:	kɐ̃.ˈme.lʊ	'camel'
	b:	kɐ̃.nʲi.ˈbaw	'cannibal'	g:	pɐ̃.nɛ.ˈtõ.nʲɪ	'pannetone bread'
	c:	fɐ̃.ˈmi.ʎɐ	'family'	h:	ta.ˈksĩ.mɛ.trʊ	'taxi-meter'
	d:	ʒɐ̃.ˈnɛ.lɐ	'window'	i:[7]	siʃ.ˈtẽ.mɐ	'system'
	e:	ka.fũ.ˈnɛ	'caress'			

All responses were recorded for further inspection and were checked and

7 In Salvador BP, coronal sibilants are palatalized when they precede coronal segments (cf. footnote 9).

subsequently verified by the first author, a native speaker of the Salvador BP dialect. At the conclusion of each testing period, we asked our volunteers if they noticed a predominance of any particular type of word structure in the test items. None reported having noticed the existence of words with pretonic nasalization.

2.4 Classification of Results into Patterns

Based on their responses for the LdP and LdK versions of the test items, the informants can be divided into three groups,[8] which we call the always-nasalizing, the never-nasalizing, and the opaque-pattern groups, as shown in (19) and (20).

(19)	**Always-Nasalizing**	**Never-Nasalizing**	**Opaque-Pattern**
a:	kẽ.pẽ.mi.pi.za.pa	ka.pa.mi.pi.za.pa	kẽ.pẽ.mi.pi.za.pa
b:	kẽ.pẽ.nʲi.pi.baw.paw	ka.pa.nʲi.pi.baw.paw	kẽ.pẽ.nʲi.pi.baw.paw
c:	fẽ.pẽ.mi.pi.ʎa.pa	fa.pa.mi.pi.ʎa.pa	fẽ.pẽ.mi.pi.ʎa.pa
d:	ʒẽ.pẽ.nɛ.pɛ.la.pa	ʒa.pa.nɛ.pɛ.la.pa	ʒẽ.pẽ.nɛ.pɛ.la.pa
e:	ka.pa.fũ.pũ.nɛ.pɛ	ka.pa.fu.pu.nɛ.pɛ	ka.pa.fũ.pũ.nɛ.pɛ
f:	kẽ.pẽ.me.pe.lo.po	ka.pa.me.pe.lo.po	kẽ.pẽ.me.pe.lo.po
g:	pẽ.pẽ.nɛ.pɛ.tõ.põ.ne.pe	pa.pa.nɛ.pɛ.to.po.ne.pe	pẽ.pẽ.nɛ.pɛ.tõ.põ.ne.pe
h:	ta.pa.ksĩ.pĩ.mɛ.pɛ.tɾo.po	ta.pa.ksi.pi.mɛ.pɛ.tɾo.po	ta.pa.ksĩ.pĩ.mɛ.pɛ.tɾo.po
i:	sis.pis.tẽ.pẽ.ma.pa	sis.pis.te.pe.ma.pa	sis.pis.tẽ.pẽ.ma.pa

(20)	**Always-Nasalizing**	**Never-Nasalizing**	**Opaque-Pattern**
a:	kẽ.ki.mi.ki.za.ki	ka.ki.mi.ki.za.ki	ka.ki.mi.ki.za.ki
b:	kẽ.ki.nʲi.ki.baw.ki	ka.ki.nʲi.ki.baw.ki	ka.ki.nʲi.ki.baw.ki
c:	fẽ.ki.mi.ki.ʎa.ki	fa.ki.mi.ki.ʎa.ki	fa.ki.mi.ki.ʎa.ki
d:	ʒẽ.ki.nɛ.ki.la.ki	ʒa.ki.nɛ.ki.la.ki	ʒa.ki.nɛ.ki.la.ki
e:	ka.ki.fũ.ki.nɛ.ki	ka.ki.fu.ki.nɛ.ki	ka.ki.fu.ki.nɛ.ki
f:	kẽ.ki.me.ki.lo.ki	ka.ki.me.ki.lo.ki	ka.ki.me.ki.lo.ki
g:	pẽ.ki.nɛ.ki.tõ.ki.ne.ki	pa.ki.nɛ.ki.to.ki.ne.ki	pa.ki.nɛ.ki.to.ki.ne.ki
h:	ta.ki.ksĩ.ki.mɛ.ki.tɾo.ki	ta.ki.ksi.ki.mɛ.ki.tɾo.ki	ta.ki.ksi.ki.mɛ.ki.tɾo.ki
i:	sis.ki.tẽ.ki.ma.ki	sis.ki.te.ki.ma.ki	sis.ki.te.ki.ma.ki

[8] As mentioned in footnote 5 there is always variable performance in allophonic rules, especially when they are characteristic of a "dialect". Our grouping here represents careful analysis in which volunteers are only said to "nasalize" or "not nasalize" if they did so more than 77% of the time (e.g. 7 out of the 9 test items). Volunteers who nasalized anywhere above two but below seven of the experimental items for each condition were excluded from study in this article, due to uncertainty on our part on how to analyze their results. There were two such volunteers who we excluded on this basis. We judged these thresholds to be representative enough of a consistent pattern of rule application (which nevertheless may fire with a stochastic pattern of variable rule execution due to social, stylistic, and individual factors, following the model of Cedergren & Sankoff (1974)). A post-hoc item-analysis of the test stimuli revealed that there was little if any effect of lexical factors on the variable application of nasalization.

2.4.1: Always-Nasalizing Subjects

Three of our eleven volunteers consistently retained a nasalized vowel in words in both LdP and LdK. Thus, given an input word such as [kɐ̃.ˈmi.zɐ], the transformation into LdP yielded the result [kɐ̃.pɐ̃.mi.pi.za.pa]. This sort of output represents apparent "overapplication" because the vowel in the initial syllable is nasalized, even though it does not immediately precede a nasal consonant.

(21) Always-Nasalizers, LdP: a pretonic open syllable whose adjacency with a heterosyllabic following nasal consonant is disrupted nonetheless shows nasalization.

The volunteers who we group under the Always-Nasalizing pattern exhibited the following pattern in *Língua do Ki*: the lexical vowel remained nasalized, but the infix did not undergo nasalization:

(22) Always-Nasalizers (LdK): a pretonic open syllable whose adjacency with a heterosyllabic following nasal consonant is disrupted shows nasalization, but the infix which now precedes the nasal consonant does not.

Despite consistently producing a nasalized vowel in pretonic pre-nasal open syllables such as [kɐ̃.ˈmi.zɐ], these volunteers were clearly not applying an active nasalization rule, as they did not nasalize the infix *-ki-* in LdK (cf. (20)), even though the vowel of the infix meets the structural description for the nasalization rule.

The simplest explanation for the Always-Nasalizing pattern is that these subjects have indeed recorded a lexical specification of [+nasal] for pretonic prenasal vowels such as that in [kɐ̃.ˈmi.zɐ]. There is no evidence from the performance on these invented language games that these volunteers have generalized an active rule of nasalization. In the remainder of this article, we will not discuss this group's results any further, as they do not bear on the nature of the nasalization rule, nor on a dissociation in performance for reduplication versus infixation.

2.4.2: Never-Nasalizing Subjects

Two of our eleven volunteers never produced a nasalized vowel in words in either LdP or LdK. Thus, given an input word such as [kɐ̃.ˈmi.zɐ], the transformation into LdP yielded the result [kɐ̃.pɐ̃.mi.pi.za.pa]. This sort of output represents apparent "underapplication" because the vowel in the reduplicated syllable is not nasalized, even though it immediately precedes a nasal consonant.

(23) Never-Nasalizers, LdP: a pretonic open syllable whose adjacency with a heterosyllabic following nasal consonant is disrupted does not show nasalization, and nor does a pretonic open syllable which suddenly precedes a nasal consonant.

The volunteers who we group under the Never-Nasalizing pattern exhibited the following pattern in *Língua do Ki*: the lexical vowel did not become nor did the infix did not undergo nasalization:

(24) Never-Nasalizers, LdK: a pretonic open syllable whose adjacency with a heterosyllabic following nasal consonant is disrupted does not show nasalization, and nor does the infix which suddenly precedes a nasal consonant.

Given the fact that these volunteers consistenly applied pretonic nasalization for all of the words in question when presented in isolation, it was rather surprising to us that they did not apply nasalization at all when these same words underwent conversion to LdP and LdK. We offer a tentative explanation for this pattern of results.

The pretonic nasalization rule of Salvador BP has one consistent exception: it does not apply across morpheme boundaries that would traditionally be classified as Level 2 boundaries. For instance, consider the prefixes *ad*- and *re*- which behave as if they constitute a separate domain from stems to which they attach. To demonstrate this, let us take a brief detour into syllabification and syllable-position allophony.

In general, stop-liquid sequences in BP are syllabified as complex onsets; thus, *ladrão* 'thief' is syllabified as [la.ˈdɾɐ̃w̃]. There are two sources of evidence to support this syllabification. The first involves the realization of the rhotic. In Standard and Salvador BP, /r/ is realized as the fricative [h]~[ɦ] in the position of a simplex onset, and as a tap [ɾ] in the second position of a branching onset or in intervocalic position. In [la.ˈdɾɐ̃w̃] it is realized as a tap. The second source of evidence comes from *Língua do Pê* itself, in which *ladrão* is produced as [la.pa.ˈdɾɐ̃w̃.pɐ̃w̃].

In contrast, words formed by the prefix *ad*- and *re*- do not show evidence of resyllabification. Thus *ad*- + *rogar* [ɦo.ˈgah] results in [a.dʒi.ɦo.ˈgah] and not *[a.dɾo.ˈgah] 'to accept someone for adoption', and *re*- + *rasgar* [ɦaz.ˈgah] results in [ɦe.ɦaz.ˈgah], and not *[ɦe.ɾaz.ˈgah] 'to re-rip'.

Importantly, the prefix *re*- is ineligible for pretonic nasalization, even in Salvador BP. Thus *re+nomear* 'to re-name' does not surface as *[ɦẽ.nõ.me.ˈah], but as [ɦɛ.nõ.me.ˈah]. Similarly, *anormal* 'abnormal' does not surface as *[ɐ̃.nɔɦ.ˈmaw] but as [a.nɔɦ.ˈmaw]. We take this as evidence that Level 2 affixes are ineligible for pretonic nasalization, and for stem-level phonological processes quite generally.

Returning to a possible explanation for the Never-Nasalizing pattern exhibited by two of our volunteers, we suggest that these volunteers analyzed the reduplicative and fixed infixes -*pV*- and -*ki*- as Level 2 morphemes, ineligible for nasalization. Indeed, it seems likely that these volunteers may have understood iterative infixation with the metalinguistic awareness that it constitutes an "interruption" of the ordinary sequence of syllables within a word.

The simplest explanation for the Never-Nasalizing pattern is thus that these informants have induced a morphological structure for LdP and LdK that inhibits the application of pretonic nasalization, to the point where even words like [kẽ.ˈmi.za], which normally display nasalization, surface with oral vowels in the context of the ludling. In fact, the failure of the nasalization rule to apply under such conditions provides an interesting yet indirect way for diagnosing the fact that such a rule indeed exists, and that the vowel of [kẽ.ˈmi.za] is not underlyingly specified as [+nasal].[9]

Given that there is a potential unified explanation for the failure of nasalization to apply in either LdP or LdK, in the remainder of this article, we will not discuss this group's results any further. Our focus in this paper is on the nature of the nasalization rule, and on a dissociation in performance for reduplication versus infixation, and the Never-Nasalizers do not shed any light on this question.

2.4.3: Overapplication in LdP, Underapplication in LdK

Six of our eleven volunteers produced a distinct pattern of nasalization in LdP from that of LdK, which will occupy the focus of our discussion of the representation of the nasalization rule in subsequent sections of this article.

We cluster these informants under the label of Opaque-Pattern Group. In LdP, they consistenly produced an overapplication pattern of nasalization, in which both the original lexical pretonic vowel and its reduplicated copy underwent nasalization.

(25) Opaque-Pattern, LdP: a pretonic open syllable whose adjacency with a heterosyllabic following nasal consonant is disrupted nonetheless shows nasalization.

At first blush, these results recall the pattern of the Always-Nasalizers, for whom we concluded that the specification of the pretonic vowels was underlyingly [+nasal]. However, the behavior of the Opaque-Pattern group on LdK speaks against such an interpretation. A pre-theoretical statement of the observed pattern is that the nasalization rule is over-applying in the Opaque-Pattern group, as its structural description is not met by the vowel in the first

[9] We cannot altogether rule out the additional hypothesis that the Never-Nasalizers were relying on orthographic representations as an aid in producing LdP and LdK forms in real-time. However, one line of evidence speaks against this possibility. Many speakers of BP, particularly those from Rio de Janeiro, but, to a lesser extent, some from Salvador, palatalize all instances of /s/ in a coda position, producing [dojʃ] for the word *dois* 'two' [siʃ.tẽ.ˈmɐ] for the word *sistema* 'system'. This palatalization is an allophonic process that is not indicated by the orthography. If the Never-Nasalizers were relying on orthographic representations, and essentially "reading off" the pronunciation of LdP from a mental image of the written word, we would expect that coda-palatalization of /s/ would never take place for these speakers. However, one of our two Never-Nasalizers produced instances of *s*-palatalization.

syllable of *kã.pã.mi.pi.za.pa*, which stands two syllables away from a nasal consonant. This behavior of overapplication in LdP stands in contrast to the pattern observed for these same speakers in LdK.

(26) Opaque-Pattern, LdK: a pretonic open syllable whose adjacency with a heterosyllabic following nasal consonant is disrupted does not show nasalization, and nor does the infix which suddenly precedes a nasal consonant.

What is most interesting about this group of speakers is that they did not apply the nasalization rule to either the lexical syllable or to the infixed syllable, even though the infixed syllable stands in a position eligible for application of the nasalization rule. This behavior in LdK thus constitutes a case of "underapplication", as the nasalization rule does not apply to the infixes, even though they meet the structural description.

Speakers of the Opaque-Pattern thus show overapplication in LdP, while showing underapplication in LdK, as schematized below:

(27) *Overapplication in Reduplication (LdP)*: Nasalization applies to the first syllable of **kã**.pã.mi.pi.za.pa, even though it shouldn't.

(28) *Underapplication in Infixation (LdK)*: Nasalization fails to apply to the second syllable of ka.**ki**.mi.ki.za.ki, even though it should.

(27) and (28) constitute two types of opacity, where the surface representation does not contain the environments for application of the rule or lack thereof. The juxtaposition of overapplication in reduplication alongside underapplication in infixation constitutes the puzzle to be discussed and solved in the remainder of this paper. Section 3 discusses four possible approaches to overapplication in LdP. Section 4 examines the compatibility of these approaches to underapplication in LdK, concluding that a Multiprecedence-and-Linearization model best captures this difference between reduplication and infixation. Section 5 concludes the paper.

3. Overapplication in Língua do Pê

In the following four subsections, we discuss four possible analyses compatible with the pattern of overapplication of nasalization in LdP.

3.1 Lexical specification of [+nasal]

It is compatible with the facts of overapplication in LdP to simply say that, in the relevant dialect, words like [kɐ̃.ˈmi.zɐ] have a [+nasal] specification for the vowel in the initial syllable. On this view, there is no overapplication going on at all. There is simply reduplication of a nasal vowel.

(29) /kã.mi.za/ → [kɐ̃.pɐ̃.mi.pi.za.pa]

Under this approach, nasalization of the vowel in both the lexical and reduplicant syllable will occur, regardless of one's theory of reduplication. Given an underlying [+nasal] pretonic vowel, any theory of reduplication that makes an exact copy of the rime of each syllable for LdK would predict those vowels to be nasal.

In the following three subsections, we explore the consequences of models that do not include an underlying nasal specification for the pretonic vowel. All three of these models assume that an active process of nasalization is responsible for the pattern of overapplication in LdP, but they differ in their representations of reduplication and in their models of the morphology-phonology interface more generally.

3.2 Base-Reduplicant Correspondence

In the framework of Base-Reduplicant Correspondence (MCCARTHY; PRINCE, 1995), overapplication in LdP can be conceived as the result of a highly-ranked faithfulness constraint, demanding identity between the original lexical syllable and its copy in the reduplicant. For instance, in /ka.pa.mi.pi.za.pa/, the second syllable will become nasalized by virtue of preceding a nasal consonant, by the regular process of nasalization in Salvador BP. However, given that *ka* and *pa* are related by a Base-Reduplicant relation, they should strive to look as alike as possible in the output, even if this goes against what the regular phonology would predict. The following output-oriented constraints are crucial in understanding overapplication under this model:

(30) a: *Oral-V / __Nasal-C

A vowel immediately preceding a nasal consonant may not be oral

b: Ident-IO-[nasal]

The output form must not have a different value for [nasal] from the underlying representation

c: Ident-BR-[nasal]

The reduplicant and base must not have different values for [nasal] from each other

The interaction between these constraints that yields overapplication is depicted in the following tableau. Horizontal rows represent possible output candidates (where top-to-bottom order is irrelevant), and vertical columns represent constraint evaluation, where left-to-right order represents the extrinisic ordering of constraint evaluation. Each "*" in cell *x,y* in the tableau indicates that the output candidate in row *x* has incurred a single violation of the constraint in column *y*.

(31)

/ka.pa.mi.pi.za.pa/	*Oral-V/__N	Ident-BR-[nas]	Ident-IO-[nas]
a. ka.pẽ.mi.pi.za.pa		*	*
b. ☞kẽ.pẽ.mi.pi.za.pa			**
c. kẽ.pa.mi.pi.za.pa	*	*	*
d. ka.pa.mi.pi.za.pa	*		

The candidates (c) and (d) are excluded from surfacing, as they violate the most highly-ranked constraint. The candidate in (b) is preferred to the candidate in (a), as the latter violates the second-highest ranked constraint. The candidate in (b), which shows overapplication, is thus the optimal choice out of these four options.

The Base-Reduplicant Correspondence model thus provides an explanation for Overapplication in *LdP* under the ranking shown in (31). The crucial factor in the model which enables overapplication is the constraint Ident-BR-[nasal], which requires identity for [nasal] between the two syllables in a reduplication relation. We have abstracted away from the constraints that yield the actual iterative infixing reduplication pattern that constitutes the LdP game itself here.

3.3 Two Cycles of Nasalization

In the framework of Lexical Phonology (e.g., KIPARSKY, 1982), and in cyclic models of phonology more generally (e.g. HALLE; VERGNAUD, 1987), certain rules may apply more than once in the course of a phonological derivation. This is also called the "Persistent Serial Model" in McCarthy & Prince (1995), who discuss a version of Myers' (1991) proposal that certain rules that may apply any time their structural description is satisfied. For example, suppose that the rule of nasalization in Salvador BP in (32) – repeated from (12) – is able to apply both before and after the process of reduplication yielding LdP. This derivation is shown in (33).

(32) $V \rightarrow [+nasal] / \_\_ C_{[+nasal]}$

(33) *Derivation for /*kamiza*/ in LdP*:

a:	/ka.mi.za/	(Underlying Representation)
b:	kẽ.mi.za	(application of (32))
c:	kẽ.pa.mi.pi.za.pa	(application of LdP formation)
d:	kẽ.pẽ.mi.pi.za.pa	(application of (32) again)

Under this view, the apparent overapplication in [kẽ.pẽ.mi.pi.za.pa] is simply the result of the fact that the rule of nasalization had the chance to apply twice: once to the lexical syllable when it was immediately adjacent to the nasal consonant, and then again to the reduplicative infix when it in turn became adjacent to the nasal consonant.

3.4 A Multiprecedence-and-Linearization Representation

In this section we provide an analysis of overapplication in LdP in terms of the Multiprecedence-and-Linearization theory, pioneered by Raimy (2000a,b) for the study of reduplication, and subsequently developed in the works by Fitzpatrick & Nevins (2002, 2004), Iba & Nevins (2004), Nevins (2005a,b), Idsardi & Raimy (2005), inter alia. What follows is not meant to be a current overview of all aspects of the theory, but will provide a complete exposition of the analysis of the iterative reduplication of LdP.

3.4.1 Overview of Overapplication in Multiprecedence

Raimy (2000a,b) treats overapplication of nasalization in Malay as the result of a nasalization rule that applies to a structure which has multiple precedence relations between segments. Malay has a rule of nasalization that is progressive (rather than regressive, as in Salvador BP) in (34). The overapplication pattern can be seen in total reduplication, as shown in (35).

(34) $V \rightarrow [+nasal] / C_{[+nasal]}$ —

(35) /aŋen/ [aŋẽn] 'wind' [ãŋẽnãŋẽn] 'unconfirmed news'

In the Base-Reduplicant Correspondence analysis of McCarthy & Prince (1995), the overapplication of nasalization in the initial syllable is the result of constraint interaction that is essentially identical to that in (31), with the replacement of ***Oral-V/ __N** with ***Oral-V/ N__**.

Raimy (2000a,b), on the other hand, analyzes the overapplication in (35) as the result of the nasalization rule in (34) applying at a level of representation before there are two surface copies of the word /aŋen/, but where the immediate precedence relation between the final /n/ of the root and the initial /a/ has already been established (arrows denote immediate precedence, whereas |α| and |ω| denote initial and final word boundaries, respectively):

(36) |α| → a → ŋ → e → n → |ω|

If the nasalization rule in (34) is interpreted in terms of immediate precedence, then the immediate precedence relation between /n/ and /a/ meets its structural description, and the structural change of nasalization is applied. Subsequent linearization of (36) into a structure without multiple-precedence yields (37), where the nasalization on both occurrences of /a/ is the result of two linearized occurrences of a single nasalized token in the pre-linearization structure. For additional discussion of this case, see Raimy (2000a,b). Crucially, nasalization occurs prior to linearization.

(37) $|\alpha| \rightarrow \tilde{a}_1 \rightarrow \eta_1 \rightarrow \tilde{e}_1 \rightarrow n_1 \rightarrow \tilde{a}_2 \rightarrow \eta \rightarrow \tilde{e}_2 \rightarrow n_2 \rightarrow |\omega|$

We state that nasalization applies prior to linearization as the consequence of (38), which is inspired by the rule-ordering division between cyclic rules prior to tier conflation and post-cyclic rules after tier conflation in the parafixation theory of Mester (1988).

(38) All stem-level phonological rules apply to multiprecedence (i.e. pre-linearized) structures

We hold (38) as a working hypothesis, in need of further investigation, but eminently falsifiable. As the nasalization rule of Salvador BP does not apply across word boundaries (e.g. *a menina*, *[ẽ.mẽ.ˈnʲĩ.nɐ] 'the girl'), we will assume that it applies at the level of multiprecedence.

The next four sections are organized as follows. Section 3.4.1 discusses the introduction of multiple precedence into phonological representations and the distinction between immediate precedence and global precedence. Section 3.4.2 discusses the morphological operations that build a LdP structure. Section 3.4.3 discusses the linearization axioms that govern linearization of multiprecedence structures. Finally, Section 3.4.4 discusses the rule of Salvador BP nasalization in terms of immediate precedence and shows how its application to /kamiza/ yields overapplication in the output.

3.4.1. Immediate Precedence, Global Procedence, and Multiprecedence

Given that speech takes place over time, any phonological theory based on a combinatorial system that concatenates discrete units must assume that representations somehow encode precedence relations among those building blocks, so that the A(rticulatory)-P(erceptual) interpretive system can properly produce/recognize physical events in real time and relate them with grammatical representations through some mapping function. We take the standard view that such building blocks are segments associated with timing slots arranged on a skeleton representing the temporal axis.

All Multiprecedence-and-Linearization representations must incorporate the notions of initial and final word boundaries, which should be formally encoded in the input structure in a simple way. Defining initial and final boundaries as explicit symbols allows them to be trivially read off during linearization, given that the mapping procedure can be defined as a finite-state machine, which requires that a starting-point and a termination-point be deterministically defined.

There are, in principle, different ways of encoding the notions of beginning and end of a word in the representation. Following previous work in the Multiprecedence-and-Linearization framework, we assume that this is done simply through two formatives, which we denote here alpha $|\alpha|$ and omega $|\omega|$, that are inherently specified always to be the first and the last symbols in the string, respectively.[10]

Thus, the underlying form of the word ['vɛ.lɐ] 'candle' would be as in (39). For the sake of exposition alone, we omit timing slots from the notation throughout the paper, as if precedence relations held directly of segments themselves, like in (40).

(39) $$\begin{array}{ccccccccccc} |\alpha| & \rightarrow & X_1 & \rightarrow & X_2 & \rightarrow & X_3 & \rightarrow & X_4 & \rightarrow & |\omega| \\ & & | & & | & & | & & | & & \\ & & v & & \varepsilon & & l & & a & & \end{array}$$

(40) Π = { <|α|,v>, <|α|,ε>, <|α|,l>, <|α|,a>, <|α|,|ω|>, <v,ε>, <v,l>, <v,a>, <v,|ω|>, <ε,l>, <ε,a>, <ε,|ω|>, <l,a>, <l,|ω|>, <a,|ω|> }

Following Raimy (2000a,b), we assume that the grammar encodes only immediate precedence relations among segments (or, more precisely, timing slots) in phonological representations, and that global precedence relations (as in (40)) emerge during the mapping from 'deep' morpho-phonological structures (which constitute the input to cyclic rule application (cf. (37)) to 'surface' morpho-phonological structures (which constitute the input to phonetic implementation), coming for free from the transitivity inherent to real time. That is, global precedence is simply the transitive closure of the immediate precedence. From that perspective, the set of precedence relations in (40) actually reduces to the set of immediate precedence relations in (41).

(41) Π_1 = { <|α|,v>, <v,ε>, <ε,l>, <l,a>, <a,|ω|> }

The ordered-pair notation in (41) emphasizes the fact that immediate precedence is a binary and asymmetric relation. We can also represent immediate precedence by means of a set of conjoined boolean predicates that are true if the first argument immediately precedes the second argument, as in (42).

(42) Π_1 = { P(|α|,v) & P(v,ε) & P(ε,l) & P(l,a) & P(a,|ω|) }

Finally, we can also denote immediate precedence by means of the tail and head of an arrow, as in (43). This notation is perhaps the easiest one to read, which is why we will systematically adopt it throughout the paper.

(43) $|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow l \rightarrow a \rightarrow |\omega|$

---

[10] The choice of this formalization with abstract symbols in immediate-precedence relations with the first and last segment of the word, instead of any other formalization, can be taken for now as just a notational convenience. In section 4.5, however, this formalism will reveal itself as crucial, and empirically motivated.

Since the role of the arrangement of timing slots in Π is to provide instructions for the A(rticulatory)-P(erceptual) system to sequence segments in real time, Π should in principle specify a linear order holding of all segments. That is, for any word W, given the set S of all segments of W, the relation R (which, in this case, is global precedence) holding of all segments of W must be a linear order on S. That is: R must have the properties in (44) (PARTEE; ter MEULEN; WALL, 1993: pp. 206-211).

(44) *Defining Properties of Linear Order*

a: *transitivity* $=_{def} \forall x, \forall y, \forall z\ [[[<x,y> \in R]\&[<y,z> \in R]] \rightarrow [<x,z> \in R]]$

b: *asymmetry* $=_{def} \forall x, \forall y\ [[<x,y> \in R] \rightarrow [<y,x> \notin R]]$

c: *totality/connectedness* $=_{def} \forall x, \forall y\ [[<x,y> \in R] \vee [<y,x> \in R]]$

However, as a matter of logic, nothing prevents the possibility that morpho-phonological representations far removed from the surface are not strictly linear, since it is not necessarily this structure which becomes phonetically implemented. As long as there is some mapping function that eventually converts it into a strictly linear representation fully interpretable by the A-P system, there is nothing wrong with structures like (45) – repeated from (36) – violating the asymmetry property in (44b), since, for instance, /a/ globaly precedes /e/ whereas /e/ globally precedes /a/.[11]

(45) $|\alpha| \rightarrow a \rightarrow ŋ \rightarrow e \rightarrow n \rightarrow |\omega|$

Once this structure is properly mapped into (46) – repeated from (37) –, the relevant relations hold of distinct occurrences of those tokens in (45), such that a strict linear order is achieved (i.e. $/a/_1$ precedes $/e/_1$ but not *vice versa*, while $/e/_1$ precedes $/a/_2$. but not *vice versa*, while $/a/_2$ precedes $/e/_2$ but not *vice versa*).

(46) $|\alpha| \rightarrow \tilde{a}_1 \rightarrow ŋ_1 \rightarrow \tilde{e}_1 \rightarrow n_1 \rightarrow \tilde{a}_2 \rightarrow ŋ \rightarrow \tilde{e}_2 \rightarrow n_2 \rightarrow |\omega|$

This set of mappings entails that morpho-phonology is conceived as a complex system involving two levels of representation. Both of them consist, roughly speaking, of immediate precedence relations holding of segments. The first one, which we can call *deep morpho-phonological structure* (or, alternatively, *multiprecedence structure*), does not need to exhibit a linear order for the transitive closure of its immediate precedence relations. The second one, which we can call *surface morpho-phonological structure* (or, alternatively, *linearized structure*), does.

11 Notice that *reflexivity* is also violated in (48). For instance, given the loop, /e/ ends up preceding itself (which is true of all segments in cases of total reduplication like this).

Our claim is that breaking down the morpho-phonological representation in two levels (one of which exhibits a linear order among segments while the other one does not) is indeed an adequate way of offering explanations on formal grounds for the speaker's linguistic intuitions when it comes to phenomena like reduplication and infixation, which show patterns of underapplication and overapplication of rules.

As for LdP, specifically, the deep and surface morpho-phonological structures will be like in (47) and (48), exemplified for the word [ˈvɛlɐ] 'candle'.

(47) $$\begin{array}{l} |\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow l \rightarrow a \rightarrow |\omega| \\ \qquad\qquad\;\; \downarrow\!\uparrow \qquad\; \downarrow\!\uparrow \\ \qquad\qquad\;\; p_1 \qquad\;\; p_2 \end{array}$$

(48) $|\alpha| \rightarrow v_1 \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l_1 \rightarrow a_1 \rightarrow p_2 \rightarrow a_2 \rightarrow |\omega|$

We pursue the strong hypothesis in (38) that all cyclic phonological rules apply only to deep morphophonological structures (which may contain multiprecedence patterns in some cases, like in (47)), and that the purpose of surface structures such as (48) is to feed the system of phonetic implementation. This makes the strong prediction that whenever a segment is involved in multiple precedence relations at deep morphophonological structure, all of its occurrences in the linearized surface structure will show up with the very same feature specification, since there was only one segment at the point where any rule may have applied. Whatever the feature specification of that segment was, it becomes multiplied into as many occurrences as necessary at the surface.[12]

3.4.2 How to Build a Língua do Pê Structure

We assume that the following morphological operations allow the building of a multiprecedence representation that will eventually yield LdP.

(49) For every syllable Σ:
- a: Add a new immediate precedence relation between the last segment of Σ and /p/;
- b: Add a new immediate precedence relation between /p/ and the nucleus of Σ

[12] Notice, however, that such a prediction is not incompatible with phenomena of apparent normal application of rules in reduplicative structures. Take, for instance, the phenomena of voicing of fricative codas in between two voiced segments, as in (i).

(i) /s/ → [z] / ___ [+voiced]

a:	*fisgar*	[fiz.gah]	'to bait'
b:	*fispisgarpar*	[fis.piz.gah.pah]	

Note that unlike regressive nasalization, voicing assimilation in Brazilian Portuguese occurs across word-boundaries as well (e.g. /as duas/ [az duas] ('the two-pl.')). Regular application of allophonic rules of this type follows from their post-lexical status.

In other words: $\forall\Sigma, \exists p \mid P(last(\Sigma), p) \ \& \ P(p,(nucleus(\Sigma))$, where last(__) and nucleus(__) are functions that return the last element and nucleus, respectively, of the syllable that is their argument. The use of last(__) is crucial as LdP copies the coda of each syllable. Note that for each syllable, there must be a token of the segment /p/ that is "available" for the application of rule (49). We assume that there is a new token of /p/ for each syllable in the word.

3.4.3 Linearization of Multiprecedence Structures

The following notation is necessary to understand the linearization algorithm:

(50) I = The input to the linearization algorithm, defined as a set of ordered pairs encoding immediate-precedence relations.
O = The output from the linearization algorithm, also defined as a set of ordered pairs encoding immediate-precedence relations.
I' = The transitive closure of I.
O' = The transitive closure of O.
P(x,y) = x immediately precedes y
P'(x,y) = x transitively precedes y

The linearization algorithm can be characterized as a mapping from I to O, in which ordered pairs may only be added to O in accordance with the Axioms below.

(51) *Starting Axiom*: The first ordered pair to be added to O must contain $|\alpha|$ as its first member.

(52) *Termination Axiom*: The last ordered pair to be added to O must contain $|\omega|$ as its last member.

(53) *Continuity Axiom*: If x is the second member of the ordered pair most recently added to O, then the next ordered pair to be added to O must contain a y such that $[\ P(x,y) \in I\ ]$.

(54) *Axiom of Choice of Path*: $\forall$x,y,z, whenever $[\ [\ P(x,y) \in I\ ] \ \& \ [\ P(x,z) \in I\ ]\ ]$ then $[\ P(x,y) \in O\ ]$ if and only if either (i) or (ii):
i: $[\ P(x,y) \notin O\ ] \ \& \ [\ [\ P'(y,z) \in I'\ ] \ \& \ [\neg\ P'(z,y) \in I'\ ]\ ]$
ii: $[\ P(x,z) \in O\ ]$

While Axioms (51-54) should be fairly clear, brief discussion is necessary on the properties of Axiom (54). This axiom determines what the linearization procedure must do whenever it reaches a state where there is a 'splitting point' in the path of immediate precedence relations in the input, i.e. right before a reduplicant or an infix.

Condition (i) of Axiom (54) is how the system formally encodes the general design property that every segment of the input must be represented onto the output (cf. Fitzpatrick & Nevins' (2002) *Completeness Condition*). Without this condition, whenever there is a choice between α and β, there would be no way to prevent the system randomly choosing α. If α were randomly chosen and there was no path that, from a graph-based representation, had 'a way back' to the original choice point, then there would be no way to map the other segment not chosen in the first pass though that bifurcation (i.e. β). Ultimately, this 'no segment left behind' condition is the way full interpretation is achieved in a strictly local fashion.

Condition (ii) of Axiom (54) is how the system formally encodes the general design property that segments from the input should not be unnecessarily overrepresented in the output (cf. Fitzpatrick & Nevins' (2002) *Economy Condition*). Segments may be represented more than once (yielding multiple occurrences out of the same token) only if necessary. Without this condition, the system could take a loop indefinitely many times, therefore overgenerating. Ultimately, this is an economy condition built into the system in a strictly local fashion.

These two conditions work together, so that the only situation where a segment is represented more than once are those where this is the only way to guarantee that (an)other segment(s) is also represented in the output.

An example derivation of the LdP for the input in (55) is provided in (56a-i).[13]

(55) $|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow l \rightarrow a \rightarrow |\omega|$

$$\begin{array}{ccccccccccc} |\alpha| & \rightarrow & v & \rightarrow & \varepsilon & \rightarrow & l & \rightarrow & a & \rightarrow & |\omega| \\ & & & & \downarrow\!\uparrow & & & & \downarrow\!\uparrow & & \\ & & & & p_1 & & & & p_2 & & \end{array}$$

[13] We adopt a graph-theoretic representation in the derivation rather than a set of ordered pairs simply as a means of making each step more perspicuous. Thus statements such as "concatenation of an occurrence of *x* at the end of a string" are equivalent to addition of an ordered pair whose second member is *x.*

(56) *Derivation of* [vε.pε.la.pa]

a:	By Axiom (51), the relation P(|α|,v) is identified and mapped into the output, through the concatenation of an occurrence of /v/ at the end of the string, as dictated by Axiom (53).	$\|\alpha\| \rightarrow v$
b:	By (53), P(v,ε) is mapped in the output.	$\|\alpha\| \rightarrow v \rightarrow \varepsilon$
c:	The relations P(ε,l) and P(ε,p_1) are both read off the input structure. At this point the system has to decide whether /l/ or /p_1/ will immediately follow /ε/ in the output. Axiom (54) – *Choice of Path* – demands that the next segment be the candidate /p_1/, because it transitively precedes /l/ (cf. the sub-path /$p_1 \rightarrow i_1 \rightarrow l$/) and not vice-versa.[14]	$\|\alpha\| \rightarrow v \rightarrow \varepsilon \rightarrow p_1$
d:	By (53), P(p_1,ε) is mapped in the output.	$\|\alpha\| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2$
e:	The system then comes back to the same bifurcation point of step (c). Axiom (54) demands that the next segment be /l/, because, at this point, concatenating another occurrence of /p_1/ at the end of the string would violate both conditions of Axiom (54), since P(ε,p_1) has already been mapped onto the output (contrary to what the first conjunct of condition (i) demands), and P(ε,l) has not yet been mapped onto the output (contrary to what condition (ii) demands). P(ε,l) is read off the input and it is represented in the output by the concatenation of an occurrence of /l/ at the end of the string being, as dictated by Axiom (53).	$\|\alpha\| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l$
f:	By (53), P(l,a_1) is mapped in the output.	$\|\alpha\| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l \rightarrow a_1$
g:	The relations P(a,|ω|) and P(a,p_2) are both read off the input structure. At this point the system has to decide whether |ω| or /p_2/ will immediately follow /a/ in the output. Axiom (54) demands that the next segment be the candidate /p_2/, because it transitively precedes |ω| (cf. the sub-path /$p_2 \rightarrow a \rightarrow \|\omega\|$/) and not vice-versa.	$\|\alpha\| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l \rightarrow a_1 \rightarrow p_2$
h:	By (53), P(p_2,a) is mapped in the output.	$\|\alpha\| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l \rightarrow a_1 \rightarrow p_2 \rightarrow a_2$
i:	Again, the system faces a bifurcation point. Since P(a,p_2) has already been mapped into the output and P(a,|ω|) has not, the next segment must be |ω|, by (54). By (53), |ω| is concatenated to the string. By (52), the derivation terminates. The resulting structure then feeds the interpretive component.	$\|\alpha\| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l \rightarrow a_1 \rightarrow p_2 \rightarrow a_2 \rightarrow \|\omega\|$

14 Since neither /ε→l/ nor /ε→p_1/ are in the output yet, all other requirements of (54) are trivially met.

The reader can verify that this linearization algorithm will succesfully yield the output structures of any LdP input created by the rule in (49).

3.4.4 Salvador BP Nasalization in terms of Immediate Precedence

The rule of nasalization thus far has been written in terms of an SPE-style structural description and structural change. Let us replace it with the following:

(57) Salvador BP Nasalization, *general definition*
Structural Description: $\forall y, \forall x \mid [P(x,y) \in I\,] \;\&\; V(x) \;\&\; C_{[+nasal]}(y)$,
Structural Change: x becomes [+nasal]

When applied to (58a), this rule yields (58b).

(58) a: $|\alpha| \rightarrow k \rightarrow a_1 \rightarrow m \rightarrow i \rightarrow z \rightarrow a_2 \rightarrow |\omega|$
b: $|\alpha| \rightarrow k \rightarrow \tilde{a}_1 \rightarrow m \rightarrow i \rightarrow z \rightarrow a_2 \rightarrow |\omega|$

Those simple cases are compatible with another formulation of the rule: (59).

(59) Salvador BP Nasalization, *uniqueness-based definition*
Structural Description: $\forall y, \iota x \mid [P(x,y) \in I\,] \;\&\; V(x) \;\&\; C_{[+nasal]}(y)$,
Structural Change: x becomes [+nasal]

The difference between (57) and (59) is that the segment affected by the rule is tied to the iota operator (RUSSELL, 1905), rather than the universal quantifier. The iota operator encodes uniqueness (i.e. "there exists a unique x"). In a nutshell, (57) states that "**AN** immediate preceder of the nasal consonant becomes nasalized" whereas (59) states that "**THE** immediate preceder of the nasal consonant becomes nasalized".

These different formalizations will not make any different predictions when it comes to simple cases like [kɐ̃.ˈmi.zɐ], since there is only one preceder of the nasal consonant anyway. In fact, for the majority of primary linguistic data that do not involve infixation or reduplication, the learner could in principle choose either. Our assumption here is that in accordance with the Subset Principle (BERWICK, 1985; MANZINI; WEXLER, 1987), the learner will by default choose the more restrictive rule, which is clearly the one with in (59) with iota, as it applies in a subset of the environments in with (57) applies.

Either (57) or (59) could generate forms like [ˈkɐ̃.mɐ] 'bed'. But when more complex structures involving multiprecedence arise, (57) and (59) are not equivalent, as the nasal consonant could have more than one preceder. In Section 4.4, we will show that the facts of underapplication in LdK additionally favor (59) over (57).

Adopting (63), the rule in (49) for producing LdP multiprecedence structures, and the linearization algorithm above, we now provide the derivation for /kamiza/ to the form [kẽ.pẽ.mi.pi.za.pa], which shows overapplication.

(60) *Derivation of* [kẽ.pẽ.mi.pi.za.pa]

a:	input to LdP rule (49)	$\|\alpha\| \rightarrow k \rightarrow a_1 \rightarrow m \rightarrow i \rightarrow z \rightarrow a_2 \rightarrow \|\omega\|$
b:	output from LdP rule (49)	$\|\alpha\| \rightarrow k \rightarrow a_1 \rightarrow m \rightarrow i \rightarrow z \rightarrow a_2 \rightarrow \|\omega\|$ ⇵ ⇵ ⇵ p_1 p_2 p_3
c:	Next, the nasalization rule in (59) applies. Notice that there is only one preceder of /m/, so the iota condition is met in this case.	$\|\alpha\| \rightarrow k \rightarrow ã_1 \rightarrow \mathbf{m} \rightarrow i \rightarrow z \rightarrow a_2 \rightarrow \|\omega\|$ ⇵ ⇵ ⇵ p_1 p_2 p_3
d:	Importantly, at the point of application of the nasalization rule to /a_1/, there is only a single token of /a_1/ that becomes nasalized.	$\|\alpha\| \rightarrow k \rightarrow ã_1 \rightarrow p_1 \rightarrow ã_2 \rightarrow m \rightarrow i_1 \rightarrow p_2$ $\rightarrow i_2 \rightarrow z \rightarrow a_4 \rightarrow p_3 \rightarrow a_4 \rightarrow \|\omega\|$

This account has successfully captured overapplication of nasalization in LdP. Both the vowel in *kã* and the vowel in *pã* are nasal because, at the time the rule applied, they were both the same vowel. Linearization yielded two distinct occurrences of /ã/. The opacity of [kẽ.pẽ.mi.pi.za.pa] is thus the result of the nasalization rule applying at a more abstract level of representation than the surface, namely one in which there is a single token of /a/ in two different immediate-precedence relations, one being P(a,m).

4. Underapplication in Língua do Ki

In this section we discuss analyses of the underapplication pattern in LdK as produced by the opaque group. Recall that the opaque group showed overapplication in LdP and underapplication in LdK. In Section 3, we discussed four possible analyses compatible with overapplication in LdP. In this section, we will examine whether they are compatible with Underapplication in LdK.

4.1 Lexical specification of [+nasal]

The analysis of overapplication in LdP as the result of lexical specification of vowels such as that of the initial syllable in *kãmiza* as underlyingly [+nasal] can not explain why that same vowel does not surface as [+nasal] in *ka.ki.mi.ki.za.ki*. If the vowel of *kãmiza* were lexically [+nasal], it should surface as such regardless of whether it is followed by an infix, a reduplicant, or its normal subsequent lexical syllable. This hypothesis is therefore untenable for explaining the behavior of the opaque group and as such insufficient in constituting a complete theory of the mental representation of pretonic nasal vowels in Salvador BP.

4.2 Base-Reduplicant Correspondence

The explanation for overapplication in LdP in terms of Base-Reduplicant Correspondence that was developed in Section 3 cannot be extended to the LdK results here, as infixation does not involve Base-Reduplicant Correspondence.

Since the opaque group is, by definition, composed of the same speakers applying overapplication in LdP and underapplication in LdK, we must assume that they are using the same constraint ranking for both games. Recall that the Correspondence model of LdP involved the following output-oriented constraints:

(61) a: *Oral-V / __Nasal-C
A vowel immediately preceding a nasal consonant may not be oral

b: Ident-IO-[nasal]
The output form must not have a different value for [nasal] from the underlying representation

c: Ident-BR-[nasal]
The reduplicant and base must not have different values for [nasal] from each other

The interaction between these constraints for the input /ka.ki.mi.ki.za.ki/ is depicted in the following tableau. Horizontal rows represent possible output candidates (where top-to-bottom order is irrelevant), and vertical columns represent constraint evaluation, where left-to-right order represents the extrinisic ordering of constraint evaluation. Each "*" in cell *x,y* in the tableau indicates that the output candidate in row *x* has incurred a single violation of the constraint in column *y*.

(62)

/ka.ki.mi.ki.za.ki/	*Oral-V/_N	Ident-BR-[nas]	Ident-IO-[nas]
a. ka.ki.mi.ki.za.ki	*		
b. kɐ̃.ki.mi.ki.za.ki	*		*
c. ☞ ka.kĩ.mi.ki.za.ki			*
d. kɐ̃.kĩ.mi.ki.za.ki			**

The candidates (a) and (b) are excluded from surfacing, as they violate the most highly-ranked constraint. The candidate in (c) is preferred to the candidate in (d), as the latter violates the third-highest ranked constraint twice, while the former violates it only once. This constraint ranking thus predicts that *ka.kĩ.mi.ki.za.ki* should be the optimal output among these four candidates, counter to fact. Recall that the actual output for the opaque group was *ka.ki.mi.ki.za.ki*, which is excluded by the high-ranked surface constraint *Oral-V / __N. The fact that this constraint wrongly excludes the actual output suggests that the process of nasalization is not well-modeled by a simple surface constraint of this form that applies only to an output representation.

The Base-Reduplicant Correspondence model, while successful for LdP overapplication, falls short as an explanation for underapplication in LdK under the ranking shown in (62). The real problem is not with Base-Reduplicant Correspondence, which is trivially satisfied and hence irrelevant in the case of infixation, but with the underapplication of nasalization even in a vowel contiguous to a nasal.

4.3 Two Cycles of Nasalization

In the Section 3.3 we discussed an implementation of a Persistent Serial Model in which the rule of nasalization in Salvador BP ((63), repeated from (12)) is able to apply both before and after the process of reduplication yielding LdP. The derivation for overapplication in LdP is repeated in (64).

(63) *Salvador BP nasalization*:
$V \rightarrow [+nasal] / \_\_ C_{[+nasal]}$

(64) *Derivation for /*kamiza*/ in LdP*:

a:		/ka.mi.za/	(Underlying Representation)
b:		kẽ.ˈmi.za	(application of (63))
c:		kẽ.pa.mi.pi.za.pa	(application of LdP formation)
d:		kẽ.pẽ.mi.pi.za.pa	(application of (63) again)

Under this view, the apparent overapplication in [kẽ.pẽ.mi.pi.za.pa] is simply the result of the fact that the rule of nasalization had the chance to apply twice: once to the lexical syllable when it was immediately adjacent to the nasal consonant, and then again to the reduplicative infix when it in turn became adjacent to the nasal consonant. This hypothesis predicts "overapplication" for the infixation pattern of LdK as well:

(65) *Derivation for /*kamiza*/ in LdK*:

a:		/ka.mi.za/	(Underlying Representation)
b:		kẽ.ˈmi.za	(application of (63))
c:		kẽ.ki.mi.ki.za.ki	(application of LdK formation)
d:	*	kẽ.kĩ.mi.ki.za.ki	(application of (63) again)

Recall that speakers who fall into the Opaque group produce *ka.ki.mi.ki.za.ki*, with no nasalization of either the infix or the lexical syllable in LdK. The persistent-application hypotheses is therefore untenable for explaining the behavior of the opaque group and thereby insufficient on its own in constituting a complete theory of the mental representation of pretonic nasal vowels in Salvador BP.

4.4 A Multiprecedence-and-Linearization Representation for LdK

Having provided a Multiprecedence account of overapplication in LdP, we demonstrate in this section how it extends to explain underapplication in LdK. In Section 4.4.1, we review the morphological operations of iterative infixation that yield LdK. In Section 4.4.2 we show how the linearization axioms from Section 3.4.3 apply to infixation. In Section 4.4.3 we show how the interpretation of the nasalization rule developed in Section 3.4.4 leads to underapplication in LdK. Section 4.4.4 provides a side demonstration of how the same iota operator leads to underapplication in a very different context, namely in the case of Javanese low-vowel rounding.

4.4.1 How to Build *Língua do Ki* Structures

Under the system we have adopted (as introduced in Section 3.4), the morphophonological operations constituting LdK are formally defined as follows:

(66) For every syllable Σ:
- a: Add a new immediate precedence relation between *x* (*x* = the last segment of Σ), and /k/ (which already immediately precedes /i/);
- b: Add a new immediate precedence relation between /i/ and the segment that *x* immediately precedes.

When applied to the input (67a), the rule (66) yields the structure in (67b).

(67) a: $|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow l \rightarrow a \rightarrow |\omega|$

b:

$$\begin{array}{ccccccccccc} |\alpha| & \rightarrow & v & \rightarrow & \varepsilon & \rightarrow & l & \rightarrow & a & \rightarrow & |\omega| \\ & & & & \uparrow & & \downarrow & & \uparrow & & \downarrow \\ & & & & k_1 & \rightarrow & i_1 & & k_2 & \rightarrow & i_2 \end{array}$$

This is a case where it becomes clear why the edge symbols (|α| and |ω|) are crucial. At first blush, it seems that word boundaries can be defined without such primitives. For instance, the first segment can be defined in relational terms, as the only segment x for which there is no segment y such that y immediately precedes x. Similarly, the last segment can be defined as the only segment x for which there is no segment y such that x immediately precedes y. However, if word boundaries are defined in that way, iterative-infixation rules like the one in (66) above cannot be defined in a trivial fashion, as the infixation of the last instance of /ki/ in a structure like (68) would require a different (and most likely disjunctive) rule than the one above in (66), given that there is no “segment that /a/ immediately precedes” in the structure.

(68) a: $v \rightarrow \varepsilon \rightarrow l \rightarrow a$

b:
$$\begin{array}{l} v \rightarrow \varepsilon \rightarrow l \rightarrow a \\ \quad\downarrow \quad \uparrow \quad \downarrow \\ k_1 \rightarrow i_1 \quad k_2 \rightarrow i_2 \end{array}$$

The same reasoning applies to the starting symbol |α|. For instance, consider another dialect of LdP (cf. Section 2.1) which involves the insertion of the fixed syllable /pe/ right before each syllable, so that [ˈvɛ.lɐ] becomes [pe.vɛ.pe.la]. In a system with boundary symbols, the rule can be trivially formalized as follows:

(69) For every syllable Σ:

a: Add a new immediate precedence relation between /e/ (which is already immediately preceded by /p/) and *x* (*x* = the first segment of Σ)

b: Add a new immediate precedence relation between /p/ and the segment that immediately precedes *x*

(70) a: $|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow l \rightarrow a \rightarrow |\omega|$

b:
$$\begin{array}{l} |\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow l \rightarrow a \rightarrow |\omega| \\ \downarrow \quad \uparrow \quad \downarrow \quad \uparrow \\ p_1 \rightarrow e_1 \quad p_2 \rightarrow e_2 \end{array}$$

Without |α| and |ω|, the infixation of the first instance of /pe/ would require a different rule, since there is no "segment that immediately precedes /v/" in the input.

(71) a: $v \rightarrow \varepsilon \rightarrow l \rightarrow a$

b:
$$\begin{array}{l} v \rightarrow \varepsilon \rightarrow l \rightarrow a \\ \downarrow \quad \uparrow \quad \downarrow \\ p_1 \rightarrow e_1 \quad p_2 \rightarrow e_2 \end{array}$$

We thus adopt the formalization of LdK in (66), which explicitly includes boundary symbols as participants in immediate-precedence relations.

4.4.2 The Linearization of Infixes

A sample derivation of the *LdK* output in (67b) from (67a) is provided below.

(72) *Derivation of* [vɛ.ki.la.ki]

a:	By Axiom (51), the relation P(|α|,v) is identified. By (53), P(|α|,v) is read off the input and represented in the output.	|α| → v
b:	By (53), P(v,ε) is mapped in the output.	|α| → v → ε
c:	P(ε,l) and P(ε,k_1) are both read off the input structure. At this point the system has to decide whether /l/ or /k_1/ will immediately follow /ε/ in the output. By Axiom (54), the path ε→k_1 is chosen, because /k_1/ transitively precedes /l/ (cf. the sub-path /k_1→i_1→l/) and not vice-versa.	|α| → v → ε → k_1
d:	By (53), P(k_1,i_1) is mapped in the output.	|α| → v → ε → k_1 → i_1
e:	By (53), P(i_1,l) is mapped in the output.	|α| → v → ε → k_1 → i_1 → l
f:	By (53), P(l,a) is mapped in the output.	|α| → v → ε → k_1 → i_1 → l → a
g:	P(a,|ω|) and P(a,k_2) are both read off the input. Axiom (54) demands that the next segment be /k_2/, because it transitively precedes |ω| (cf. the sub-path /k_2→i_2→|ω|/) and not vice-versa.	|α| → v → ε → k_1 → i_1 → l → a → k_2
h:	By (53), P(k_2,i_2) is mapped in the output.	|α| → v → ε → k_1 → i_1 → l → a → k_2 → i_2
i	By (53), P(i_2,|ω|) is mapped in the output. By (52), the derivation terminates. The resulting structure then feeds the interpretive component.	|α| → v → ε → k_1 → i_1 → l → a → k_2 → i_2 → |ω|

When the mapping is complete, there have been two immediate precedence relations left unmapped, namely P(ε,l) and P(a,|ω|). This is not "harmful" at all since (i) all segments have been mapped, and (ii) the correct predictions are made, as far as both order of segments and rule application are concerned. Infixation thus provides a crucial case in which the axioms of linearization yield unmapped immediate precedence relations from a multiprecedence input.

The reader can verify that this Linearization algorithm will succesfully yield the output structures of any LdK input created by the rule in (66).

4.4.3 The Failure of Nasalization

Given the nasalization rule in (59) above, defined in terms of a *unique preceder*, it now becomes clear why underapplication occurs in LdK structures, as shown below.

(73) *Derivation of* [kɐ̃.ki.mi.ki.za.ki]

a:	input	$\lvert\alpha\rvert \rightarrow k \rightarrow a_1 \rightarrow m \rightarrow i_4 \rightarrow z \rightarrow a_2 \rightarrow \lvert\omega\rvert$
b:	infixation	$\lvert\alpha\rvert \rightarrow k \rightarrow a_1 \rightarrow m \rightarrow i_4 \rightarrow z \rightarrow a_2 \rightarrow \lvert\omega\rvert$ $\downarrow \; \uparrow \; \downarrow \; \uparrow \; \downarrow \; \uparrow$ $k_1 \rightarrow i_1 \quad k_2 \rightarrow i_2 \quad k_3 \rightarrow i_3$
c:	nasalization fails to apply, since both /a_1/ and /i_1/ precede /m/	$\lvert\alpha\rvert \rightarrow k \rightarrow a_1 \rightarrow m \rightarrow i_4 \rightarrow z \rightarrow a_2 \rightarrow \lvert\omega\rvert$ $\downarrow \; \uparrow \; \downarrow \; \uparrow \; \downarrow \; \uparrow$ $k_1 \rightarrow i_1 \quad k_2 \rightarrow i_2 \quad k_3 \rightarrow i_3$
d:	linearization	$\lvert\alpha\rvert \rightarrow k \rightarrow a_1 \rightarrow k_1 \rightarrow i_1 \rightarrow m \rightarrow i_4 \rightarrow k_2 \rightarrow i_2 \rightarrow z \rightarrow a_2 \rightarrow k_3 \rightarrow i_3 \rightarrow \lvert\omega\rvert$

The crucial step is step (73c): the multiprecedence representation of infixes is such that *both* the lexical segment /a/ (which undergoes ordinarily nasalization in [kɐ̃.ˈmi.zɐ] and the infix vowel /i_1/ (which should undergo nasalization based on looking at the surface) immediately precede the nasal consonant /m/. As the structural description of the rule requires a unique preceder, it fails to apply and neither segment is nasalized.

4.4.4 Extensions of the Iota Operator

At this point we deem it useful to demonstrate that the iota operator is not simply an artifact designed for underapplication in LdK, and that it plays a role in the interpretation of structural descriptions of rules in situations with no infixation at all.

4.4.1 Underapplication of Low Vowel Rounding in Javanese

A concrete case can be found in Javanese[15], which exhibits a rule of low-vowel rounding in word-final position (DUDAS, 1976, p. 206).

(74)	a:	medʒɔ	'table'	medʒa-ku	'my table'	medʒa-ne	'his table'
	b:	dʒiwɔ	'soul'	dʒiwa-ku	'my soul'	dʒiwa-ne	'his soul'
	c:	dɔngɔ	'color'	dɔnga-ku	'my color'	dɔnga-ne	'his color'

Dudas (1976) provides the rule in (75), which we formulate in precedence-based terms in (76).

(75) $a \rightarrow ɔ \,/\, \_\_ \, \lvert\omega\rvert$

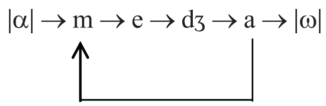
[15] We heartily thank Eric Raimy and William Idsardi for bringing the relevance of this case to our attention.

(76) Structural Description: y such that |α| (y), ιx | [P(x,y) ∈ I] & low(x)
Structural Change: x becomes [+round]

In reduplication, the rule in (76) overapplies, as shown in (77). This is a consequence of the graph in (78)

(77)	a:	medʒɔ	'table'
	b:	medʒɔ-medʒɔ	'tables'
	c:	dɔngo	'prayer'
	d:	dɔngo-dɔngo	'prayers'

(78)

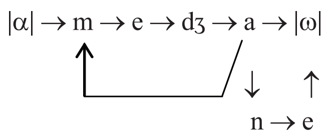


Since /a/ remains the unique preceder of |ω| in (78), it is expected that the low-vowel rounding rule of (76) should apply. Interestingly, when reduplication and suffixation interact, the rule of low-vowel rounding *underapplies*, as shown in (79):

(79)	a:	medʒa-medʒa-ne	'his tables'
	b:	dɔnga-dɔnga-ne	'his prayers'

When one examines the multiprecedence representation of (80), which contains both reduplication and suffixation, the underapplication of low-vowel rounding is in fact entirely to be expected, given the rule in (76).

(80)



The rule in (76) cannot apply to (80), as there is no *unique* preceder of |ω|: both /e/ and /a/ are in immediate precedence relations with |ω|. Just like nasalization in LdK, underapplication here too results from the interaction of the iota operator with a multiprecedence representation. The parallel between overapplication in reduplication and underapplication in affixation in these two unrelated languages is striking confirmation for the iota operator as a condition on rules in multiprecedence structures.

Before concluding, we will note that the cases of total reduplication such as (77b,d) provide additional evidence for defining word boundaries in terms of abstract symbols ('first' and 'last'), that are in bona-fide immediate-precedence relations. Consider again the structure in (80) above, but without the boundary symbols.

(81) a: *input to total reduplication*
m → e → dʒ → a
b: *output from total reduplication*
m → e → dʒ → a

Let us briefly consider the consequences of assuming that the first and the last segment can be defined in relational terms, respectively, as (i) the only segment which is not preceded by anything and (ii) the only segment which does not precede anything.

Given this relational definition, it is true that in (81a), /m/ and /a/ can be identified as the first and the last segments, respectively. However, the same doesn't hold for (81b). Once the morphology of total reduplication demands that an immediate-precedence loop from /m/ to /a/ is created, there is simply no first segment and no last segment in the structure. An immediate problem thus arises with respect to linearization. There is no way to identify a segment not preceded by anything.

Also, even if the first segment could be identified, a problem arises with respect to the loop. After the linearization procedure reaches the state /a/, Axiom (53) requires that the next state be /m/ again, since the relation P(m,a) still needs to be mapped. But after the second occurrence of /m/ is written in the output, the linearization algorithm would have no way to stop: since there is no 'last segment' in the path anymore. The linearization procedure would fall into an infinite loop. One could argue that the infinite loop can be avoided if the system incorporates some principle requiring that segments/relations from the input be mapped into the output only when necessary. Going through all the segments of the loop indefinitely many times would violate this general economy principle. However, this is not enough to derive the facts. If that were true, then we would expect the linearized form of (81b) to be (82b) rather than (82a), since (82b) realizes all the segments and relations from the input.

(82) *potential outputs from the rule of low-vowel rounding, given input* (81b)
a: m → e → dʒ → a → m → e → dʒ → a (actually attested form)
b: m → e → dʒ → a → m (wrongly predicted form)

In summary, Javanese total reduplication constitutes further evidence for the necessity of abstract boundary symbols as crucial to a coherent linearization algorithm.

4.5 Summary of Approaches to Underapplication in LdK

We conclude this section by considering again the essential fact of underapplication in LdK, as repeated in (83):

(83) kẽ.ˈmi.zɐ ka.<u>ki</u>.mi.<u>ki</u>.za.<u>ki</u>

While it may not be surprising that nasalization does not apply to the original lexical syllable, as its adjacency to the nasal consonant has been disrupted, it is extremely surprising from the surface representation that the infix –*ki*– should not be nasalized in a dialect of Portuguese in which all pre-nasal vowels are nasalized, and moreover for speakers for whom infixes that are reduplicants do undergo nasalization. It is difficult to imagine any output constraint that would block nasalization in this case. Indeed, we hope to have shown that only through appealing to a more abstract level of representation can one find a natural explanation for underapplication: the triggering segment (the nasal) is in two immediate-precedence relations at once, and the rule of nasalization cannot apply under those conditions.

5. Conclusion

While there have been a few recent proposals as to the proper representation of infixation (e.g. Halle (2001) proposes that infixation is the result of metathesis; and Yu (2003) proposes that an infix "subcategorizes" for a particular position in the word), our proposal here is the first to address the treatment of *iterative* infixation. We have treated iterative infixation, of the type found in LdK, as requiring universal quantification within the structural description and structural change of the rule that adds the infix and the corresponding immediate precedence relations. Up until now, there has been little reason, aside from theory-internal grounds, to suppose that a multiprecedence representation for infixation is correct. In particular, when the word /kamiza/ comes out as [kakimikizaki], it is not obvious that the immediate-precedence link between /a/ and the following /m/ still exists once the prefix *ki* is inserted. However, we have discovered a way to diagnose the presence of this link: the fact that it blocks application of a rule requiring a unique preceder. In particular, we know that the immediate precedence relation between /a/ and the following /m/ still exists at the level of multiprecedence in [kakimikizaki] because its presence leads to failure of the /m/ to have a unique preceder, and hence underapplication of nasalization.

More generally, we have pointed to a difference between infixation and reduplication. They do not differ in overapplying or underapplying due to any differences in morphological status. Both *Língua do Pê* and *Língua do Ki* involve fixed segmental material inserted into the precedence structure of a word. The difference is that infixation starts and ends between two segments that transitively precede each other, while reduplication does not. This is a basic consequence of the precedence structures and is independent of any diacritic facts or privileged constraints about the morphological status of these items.

We would like to point out that, having taught our volunteers these two invented language games, and observing overapplication in LdP and underapplication in LdK, there is the pretheoretically imaginable possibility that some speaker might have yielded the opposite pattern, as schematized in (84):

(84) *Unattested Pattern across the Two Games*
Overapplication of LdK output: kẽ.<u>kĩ</u>.ˈmi.<u>ki</u>.za.<u>ki</u>
Underapplication of LdP output: ka.<u>pa</u>.mi.<u>pi</u>.za.<u>pa</u>

The unnatested response pattern is ruled out by our representations. (84a) would be a case in which there was no iota operator on the nasalization rule, and both immediate preceders of the nasal /m/ should undergo nasalization, yielding surface application of nasalization to both the infix and the lexically-preceding /a/ of the first syllable. While it is logically possible that speakers could entertain a rule of nasalization without the iota operator (and hence apply nasalization to any or all vowels satisfying the structural description of immediately preceding a nasal consonant in (84a)), if this were the case, these same speakers would not be able to underapply in LdP, as the single vowel /a/ satisfies the condition on immediate precedence as well in (84b). This pattern is thus impossible under any conditions on the rule of nasalization.

However, the attested opposite pattern, of overapplication in LdP and underapplication in LdK, finds a natural explanation in the theory here, given that infixation always involves a situation with two immediate precedence relations pointing to the same endpoint (the lexical immediate-preceder, and the newly-introduced immediate precedence from the last segment of the infix), while reduplication need not involve two immediate precedence relations pointing to the same endpoint. Hence, if a uniqueness condition, imposed by the iota operator, will apply, it must necessarily cause rule-blocking for the case of infixation.

In terms of our contribution to the understanding of dialect-specific phonotactics, we hope to have shown that pretonic nasalization in Salvador BP is, at least for some speakers, the consequence of a rule-governed process, and not simply the result of memorizing static patterns in the lexicon. The fact that two invented language games reveal very different results for whether nasalization is kept or not demonstrate that (at least some) speakers cannot simply have recorded nasalization in the underlying representation, but must be applying nasalization by a rule which can be diagnosed by the very fact that it fails to apply under certain conditions. Thus, while the word [bẽ.ˈnẽ.nɐ] happens to have a pretonic nasal vowel in its un-infixed and un-reduplicated form, it ain't necessarily so. ◪

GUIMARÃES, M.; NEVINS, A. NASALIZAÇÃO OPACA EM JOGOS DE LINGUAGEM E AS RELAÇÕES DE PRECEDÊNCIA NA REDUPLICAÇÃO E NA INFIXAÇÃO

Resumo
Investigamos o padrão de nasalização opaca obtido quando solicitamos informantes voluntários (falantes nativos de Português Brasileiro (PB), dialeto de Salvador-BA) a traduzirem estímulos de PB para versões codificadas em dois jogos de codificação de linguagem: Língua do Pê (reduplicativo) e Língua do Ki (infixativo). Um grupo de informantes exibiu um padrão opaco, no qual houve sobre-aplicação de nasalização com reduplicação, e sub-aplicação com infixação. Tais resultados têm três consequências: (i) eles corroboram a tese de que há uma representação abstrata em termos de Multiprecedência-&-Linearização (RAIMY, 2000a,b) para reduplicação e infixação, em que a infixação (mas não a reduplicação) começa e termina entre dois segmentos que precedem um ao outro transitivamente; (ii) eles evidenciam que, para um subgrupo significativo dos participantes em nosso experimento, a nasalização é uma regra ativa dos dialetos nordestinos do PB, e não simplesmente um padrão lexicalizado; (iii) eles dão suporte adicional para uma 'nova' condição sobre aplicação de regras: o operador iota, de Russell (1905), que impõe uma condição de unicidade às descrições estruturais.

Palavras-chave
sobre/sub-aplicação; multiprecedência; nasalização; jogos lingüísticos; português brasileiro

References

BAGEMIHL, B. Language games and related areas. In J. Goldsmith (ed.), *The Handbook of Phonological Theory*. Blackwell, 1995.

BERWICK, R. *The Acquisition of Syntactic Knowledge*. Cambridge, MA: MIT Press, 1985.

BETHIN, C. "Polish nasal vowels." *International Journal of Slavic Linguistics and Poetics, volume xxxviii*: 33-71, 1995.

CAMPBELL, L. The Psychological and Sociological Reality of Finnish Vowel Harmony. In *Issues in Vowel Harmony* pp. 245-270. Amsterdam: John Benjamins, 1980.

CEDERGREN, H.; SANKOFF, D. Variable rules: Performance as a Statistical Reflection of Competence. *Language* 50.2: 333-355, 1974.

CHOMSKY, N.; HALLE, M. *The Sound Pattern of English*. New York: Harper and Row, 1968.

DUDAS, K. *The Phonology and Morphology of Modern Javanese*. Doctoral dissertation, University of Illinois, Urbana-Champaign, 1976.

FITZPATRICK, J.; NEVINS, A. Phonological Occurrences: relations and copying. Cambridge: MIT, ms, 2002. <http://ling.auf.net/lingBuzz/000051>

FITZPATRICK, J.; NEVINS, A. Linearizing Nested and Overlapping Precedence in Multiple Reduplication. In Sudha Arunachalam and Tatjana Scheffler (eds.), *The Proceedings of the 27th Penn Linguistics Colloquium.* Penn Working Papers in Linguistics 10.1: 75-88, 2004.

GUY, G. Linguistic Variation in Brazilian Portuguese: Aspects of the Phonology, Syntax and Language History. Ph.D. dissertation. University of Pennsylvania, 1981.

HALLE, M.; VERGNAUD, J-R. *An Essay on Stress*. Cambridge: MIT Press, 1987.

HALLE, M. Infixation vs. Onset Metathesis in Tagalog, Chamorro, and Toba Batak. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language,* pp.153-168. Cambridge: MIT Press, 2001.

HARRISON, D.; KAUN, A. Pattern-Responsive Lexicon Optimization. In NELS 30. GLSA, University of Massachusetts, Amherst, 2000.

IDSARDI, W.; RAIMY, E. Remarks on Language Play. Ms., University of Maryland and Swarthmore College, 2006.

KIPARSKY, P. Lexical phonology and morphology. In I.S. Yang, ed., *Linguistics in the Morning Calm*. Seoul: Hanshin, 1982.

LIPSKI, J. Brazilian Portuguese vowel nasalization: secondary aspects. *Canadian Journal of Linguistics* 20: 59-77, 1975.

MANZINI, R.; WEXLER, K. Parameters, Binding Theory, and Learnability. *Linguistic Inquiry* 18.3: 413-444, 1987.

MATTOSO CÂMARA Jr., J. *Estrutura da Língua Portuguesa*. Petrópolis: Vozes, 1970.

MCCARTHY, J.; PRINCE, A. Faithfulness and Reduplicative Identity. In *University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory,* 249–384. GLSA, University of Massachusetts, Amherst, 1995.

MESTER, A. *Studies in Tier Structure*. Outstanding Dissertation in Linguistics Series. Garland: New York, 1988.

MYERS, S. Persistent Rules. *Linguistic Inquiry* 22: 315-344, 1991.

NEVINS, A. Overwriting does not Optimize in Nonconcatenative Word Formation. *Linguistic Inquiry* 36.2, 2005.

______. What UG Can and Can't Do to Help the Reduplication Learner. In A. Csirmaz, A. Gualmini, and A. Nevins (eds.) *MIT Working Papers in Linguistics 48: Plato's Problems*. Cambridge, MA: MITWPL, 2005.

NOBILING, O. Die Nasalvokale in Portugiesischen. *Die Neuren Sprachen* 11: 129-153, 1903.

OHALA, M. *Aspects of Hindi Phonology*. Delhi: Motilal Banarsidass, 1983.

OHALA, M.; OHALA, J. Nasal epenthesis in Hindi. *Phonetica* 1991: 207-220, 1991.

PARTEE, B.; ter MEULEN, A.; WALL, R. *Mathematical Methods in Linguistics*. Dordrecht: Kluwer Academic Publishers, 1993.

PIÑEROS, C.-E. Head-dependence in Jerigonza, a Spanish language game. In *Advances in Hispanic Linguistics, Vol. 1*. pp. 265-277. Somerville: Cascadilla Press, 1999.

RAIMY, E. Remarks on Backcopying. *Linguistic Inquiry* 31.3:541-552, 2000a.

______. *The Phonology and Morphology of Reduplication*. Berlin: Mouton de Gruyter, 2000b.

RUBACH, J. Nasalization in Polish. *Journal of Phonetics*, 5: 17-25, 1977.

RUSSELL, B. On Denoting. *Mind* 14: 479-493, 1905.

SHERZER, J. Talking backwards in Cuna: the sociological reality of phonological descriptions. *Southwestern Journal of Anthropology* 26: 343-353, 1970.

SHOSTED, R. Nasal coda restoration in Brazilian Portuguese. In M.J. Solé, D. Recasens, and J. Romero (eds.). *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Universitat Autonoma, 2003.

VAGO, R. The treatment of long vowels in word games. *Phonology* 2: 329-342, 1985.

YU, A. The phonology and morphology of infixation. Ph.D. dissertation, University of California, Berkeley, 2003.