Opaque Nasalization in the Língua do Pê of Salvador, Brazil*

Maximiliano Guimarães (Universidade Federal do Paraná)

Andrew Nevins (Harvard University)

Abstract

We examine the pattern of opaque nasalization in two invented language games, as played by volunteers who were native speakers of the dialect of Brazilian Portuguese spoken in Salvador, Brasil. One group of speakers, on whom we focus here, 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 (e.g. Raimy 2000a,b) for reduplication and infixation; (ii) they demonstrate that nasalization is an active 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: Overapplication, underapplication, Multiprecedence, nasalization, invented language games, 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 **[bɐ̃nɐ̃nɐ]**, rather than **[banɐ̃nɐ]**, is a give-away that the talker is from the Northeast part of the country, because talkers from other parts of Brazil only nasalize vowels in open-syllables when they are stressed. In this section, we provide an overview of these nasalization processes.

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

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

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

Stress in nouns is generally penultimate in Brazilian Portuguese. More precisely, following the proposal of Arregi & Oltra-Massuet (2005) for Spanish,

* Conversations with Eric Raimy and Bill Idsardi were essential in the development of this paper. We would like to again thank our volunteers at UFBA for their participation in the experiment. Audience members at CUNY and MIT, too numerous to list here, contributed many important observations and questions. The first author was supported by a Research Grant by CAPES (Brazilian Government). The second author was partially supported by a Jorge Paulo Lemann Fund of the David Rockefeller Center for Latin American Studies at Harvard University.

one may say that stress in nouns is by default on the syllable immediately preceding the thematic vowel.

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’
b:	pa.no.ˈɾɐ̃.mɐ	‘panorama
c:	a.ˈɾɐ̃.ɲɐ	‘spider
d:	pow.ˈtɾõ.nɐ	‘easy chair’
e:	a.ˈɾõ.mɐ	‘aroma’
f:	ver.ˈgõ.ɲɐ	‘shame’
g:	ɐ̃n.ˈtẽ.nɐ	‘antenna’
h:	si.ˈnẽ.mɐ	‘movie theater’
i:	ˈsẽ.ɲɐ	‘password’
j:	foh.ˈtũ.nɐ	‘fortune’
k:	es.ˈpũ.mɐ	‘foam’
l:	ˈũ.ɲɐ	‘fingernail’
m:	bu.ˈzĩ.nɐ	‘horn-honk’
n:	ˈprĩ.mɐ	‘(female) cousin’
o:	ga.ˈlĩ.ɲɐ	‘hen’

All of the examples in (2) illustrate nasalization of a stressed vowel when the following consonant is heterosyllabic, as schematized in (1). Nasalization also occurs when the following consononant is tautosyllabic, as shown in (3).

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

a:	ˈsɐ̃m.bɐ	‘samba dance’
b:	ˈsɐ̃n.tɐ	‘(female) saint’
c:	ˈpõm.bɐ	‘pigeon’
d:	ˈpõn.tʊ	‘point’
e:	mɐ̃m.ˈbẽm.bɪ	‘amateur job’
f:	fa.ˈzẽn.dɐ	‘farm’
g:	ˈtũm.bɐ	‘grave’
h:	ˈmũn.dʊ	‘world’
i:	ta.ˈɾĩm.bɐ	‘know-how’
j:	ta.ma.ˈɾĩn.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[1]:

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

(5) /ˈsam.ba/ → [sɐ̃m.ba] (by (2)) → [sɐ̃.ba] (by (4))

Some evidence for the rule in (4) comes not only from the phonetic realization of word-internal coda consonants in BP, but from alternations involving word-final syllables[2]:

(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).[3]

[1] The rule in (4) is also motivated when considering that all sonorant codas lenite in Brazilian Portuguese. Lateral [l] in onset alternates with the glide [w] in codas: [ʒɔɦ.ˈnaw] / [ʒɔɦ.na.ˈlej.ɾʊ] 'newspaper/journalist'. The tap [ɾ] in onset alternates with the fricative [h] (and its homorganic [ɦ]) in codas: [floh]/[flo.ɾɪs] 'flower/flowers'. Other than nasals, laterals, and rhotics, the only other segment 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 [me(ɦ).mo]. Nobiling (1903) was perhaps the first to posit underlying nasal glide elements in coda position.

[2] Lipski (1975: 64) cites evidence for psychological reality of a coda from illiterate speakers who reduce /re.ˈzaN.do /'praying' to [re.ˈzã.nʊ], rather than [re.ˈza.dʊ] or [re.ˈzɐ̃ʊ]. This pattern of reduction is also attested in the speech of highly educated people, depending on the region. The point of Lipski's example is that even in the absence of orthographic representations, speakers posit a segmental nasal.

[3] 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:	/pan.ˈdei.ɾo/	[pɐ̃.ˈdej.ɾʊ]	'tambourine'
b:	/tɾom.ˈbo.ne/	[tɾõ.ˈbõ.nɪ]	'trombone'
c:	/den.ˈtis.ta/	[dẽ.ˈʧis.tɐ]	'dentist'
d:	/fun.ˈdu.ɾa/	[fũ.ˈdu.ɾɐ]	'depth'
e:	/sin.ˈtu.ɾa/	[sĩ.ˈtu.ɾɐ]	'waist'

Unlike [ˈsɐ̃.bɐ] and the examples in (3), the nasalization of the pre-tonic 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 (as opposed to being underlying), (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: $V \rightarrow [+nasal] \ / \ \_\_ \ C_{[+nasal]} \ ]_{\sigma}$

Before proceeding, 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.ba] 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 for politics and mass-media entertainment). All of the following words have penultimate stress, and demonstrate pretonic nasalization:[4]

[4] 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 from the actual pattern, which we estimate as highly regular, and applying over 90% of the time. We venture that this is true of many, many allophonic rules in many languages, often described in the pages of this and other journals in categorical format, such as *s*-palatalization of codas in Rio de Janeiro BP (which is a variable rule; Guy 1981), unstressed vowel reduction in BP,

(9)

Orthography	Salvador BP	Standard BP	Gloss
banana	bɐ̃.ˈnɐ̃.nɐ	ba.ˈnɐ̃.nɐ	‘banana’
caneta	kɐ̃.ˈne.tɐ	ka.ˈne.tɐ	‘pen’
familia	fũ.ˈmi.ʎɐ	fa.ˈmi.ʎɐ	‘family’
começo	kõ.ˈme.sʊ	ko.ˈme.sʊ	‘beginning’
fonema	fõ.ˈne.mɐ	fo.ˈnẽ.mɐ	‘phoneme’
demora	dẽ.ˈmɔ.ɾɐ	de.ˈmɔ.rɐ	‘delay’
cenoura	sə̃e.ˈnoː.ɾɐ	se.ˈnoː.ɾɐ	‘carrot’
fumaça	fũ.ˈma.sɐ	fu.ˈma.sɐ	‘smoke’
luneta	lũ.ˈne.tɐ	lu.ˈne.tɐ	‘spyglass’
cinema	sĩ.ˈñe.mɐ	si.ˈñe.mɐ	‘movietheater’
limite	l̃i.ˈmi.ʧɪ	li.ˈmi.ʧɪ	‘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*:
$V \rightarrow [+nasal] \;/\; \_\_\ ]_{\sigma}\,.\ C_{[+nasal]}$

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

As Salvador BP also contains rule (8a) and (8b), these may all be collapsed into the following:

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

The arguments for underlying nasalized 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 process in (12). However, a potential argument for underlying nasalized vowels based on the existence of minimal pairs such as (13) remains:

---

and flapping in English, all of which apply stochastically under the general model of variable rules proposed in Cedergren & Sankoff (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 here. We recognize the importance of rule execution frequencies as a phenomenon of study, but place it outside the scope of this article, which focuses on the grammatical representation of the nasalization process itself, regardless of how it may interact with the many other factors that influence linguistic performance in a given time, place, and utterance situation.

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

Notably, minimal pairs such as (13) only exist in word-final position, which is what 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 limited, and the possibility that some learners of BP may have nasal vowels as part of their underlying inventory 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 word-final nasal consonant, then the possibility of contrastive [+nasal] vowels becomes a potential option for other positions within the word.

Thus, in considering the pronuciation of pretonic nasal vowels in open syllables, such as those in the left-hand column of (9), one might consider the possibility that these vowels are underlyingly nasal, and that indeed the locus of dialect variation between Salvador BP and Standard BP simply lies the lexical value of the [nasal] for pretonic vowels. Indeed, if you ask a native speaker of Salvador BP why they say [kɐ̃.ˈne.tɐ] and not [ka.ˈne.tɐ], their answer will likely be “I just learned the word that way”. In this paper, one of our goals is to demonstrate that, at least for some speakers of Salvador BP, the nasalization of pretonic vowels in open syllable such as [kɐ̃.ˈne.tɐ] is result of an active phonological rule, and not simply due to lexical listing.

In order to demonstrate that pretonic nasalization in BP is indeed an active process, and not simply a static generalization over the lexicon, we draw on two important sources of evidence that have often been used to support the existence of active phonological rules: demonstration that the rules apply in novel words, and demonstration that the rules show over- and under-application in certain contexts. Indeed, one of the best ways to show that the vowel in the initial syllable of Salvador BP [kɐ̃.ˈne.tɐ] is not underlyingly [+nasal] is to show that in certain contexts, that vowel shows up as [-nasal]. In the experiments to be described in which our volunteers played an invented language game called *Língua do Ki*, the vowel in [kɐ̃.ˈne.tɐ] was produced on the surface with [-nasal] value, due to certain conditions on rule application to which we will return. In short, one of the best ways of demonstrating the existence of a phonological rule turns out to be observing its failure to apply.

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 [kẽ.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.

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) 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 Spansh; 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.[5]

(15)	a:	ˈbɔ.lɐ	bɔ.<u>pɔ</u>.la.<u>pa</u>	‘ball’
	b:	mah	mah.<u>par</u>	‘sea’
	c:	ˈah.ʧɪ	ah.<u>par</u>.te.<u>pe</u>	‘art’
	d:	ˈtɾi.ʎʊ	tri.<u>pi</u>.ʎo.<u>po</u>	‘trail’
	e:	vɛɦ.ˈda.ʤɪ	vɛh.<u>peɦ</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’
	g:	a.ni.veh.ˈsa.ɾju	a.<u>pa</u>.ni.<u>pi</u>.veh.<u>peh</u>.sa.<u>pa</u>.rjo.<u>po</u>	‘birthday’

Língua do Pê is widespread as a children’s language game throughout Brazil, although many different variants of it exist. For instance, a second version of *Língua do Pê* does not copy the preceding vowel, but rather maintain *–pe-* as a fixed syllable, thus yielding *bɔ.<u>pe</u>.la.<u>pe</u>* for (15a). A third version of *Língua do Pê* uses *–pe-* as a fixed syllable, but places it before, rather than after each syllable, yielding *<u>pe</u>.bɔ.<u>pe</u>.la* for (15a). A fourth (and much more rare) version of Lingua do Pʧ uses *–pV(C)-* in a manner similar to (15), but places it before, rather than after each syllable, yielding *<u>pɔ</u>.bɔ.<u>pa</u>.la* for (15a). We have met speakers of all three of these other dialects, but our own experience is that the version of LdP in (14) is the most widespread, and indeed, is the version spoken by the first author

[5] The fact that the LdP form for a word like *castelo* [kaʃ.ˈtɛ.lo] is [kas.ˈpas.tɛ.ˈpɛ.lo.ˈpo], rather than [kas.ˈpas.tɛ.ˈpɛ.lʊ.ˈpʊ] may be a bit surprising at first, since it appears that the rule of post-tonic vowel raising (cf. footnote 3) is failing to apply. Notice, however, that LdP and LdK have a stress pattern of successive iambic feet, which makes the final syllable of *caspastepelopo* stressed, and hence immune from reduction. This explains why the last syllable is [po] rather than [pʊ]. It is still not so obvious why the syllable immediately preceding it is [lo] rather than [lʊ], since BP (including the dialect of Salvador) also has a rule of raising of pre-tonic mid-vowels. While post-tonic raising is uniform and well behaved, pre-tonic raising, on the other hand, is not uniform and has not been researched in any amount of detail. Pretonic mid vowels show a great deal of variation in Salvador BP, even for the same speaker (and, perhaps, even some lexical idiosyncrasies). They do not raise uniformly and sometimes even "exceptionally" lower, as shown in (i).

(i)	a:	/mo.ˈdɛr.no/ ‘modern’	[mo.ˈdɛɦ.nʊ]˜ [mɔ.ˈdɛɦ.nʊ]˜ [mu.ˈdɛɦ.nʊ]
	b:	/ʃo.ko.ˈla.te/ ‘chocolate’	[ʃɔ.kɔ.ˈla.ʧɪ] : *[ʃo.ko.ˈla.ʧɪ], *[ʃu.ku.ˈla.ʧɪ]
	c:	/ko.ˈʎɛr/ ‘spoon’	[ku.ˈʎɛh] : *[kɔ.ˈʎɛh], *[ko.ˈʎɛh]
	d:	/for.ˈna.ʎa/ ‘wood-stove’	[fɔɦ.ˈna.ʎɐ] : *[foɦ.ˈna.ʎɐ], *[fuɦ.ˈna.ʎɐ]
	e:	/kor.ne.ˈta/ ‘cornet’	[koɦ.ne.ˈtɐ] : *[kɔɦ.ˈne.tɐ], *[kuɦ.ˈne.tɐ]

Thus, the variation in the quality of the penultimate vowel in words like *caspastepelopo* is something we do not address here. Recall, though, that, if pre-tonic raising is a *bona fide* phonlogical rule that, by our hypothesis, should apply to pre-linearized struyctures (cf. section 3.4.1), then we expect the penultimate syllable of words like *caspastepelopo* not to show raising, because, at that level, there is a single token of it, which is associated with a stressed syllable (under the assumption that such rule also involves the iota operator (cf. section 3.4.4)).

of this paper. For this reason, we trained all volunteers using the pattern of LdP as characterized in (14) and exemplified in (15).

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-

We will delay a formal characterization of the process yielding Lingua do Ki (LdK) until Section 4.4. 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.ʧɪ	ah.<u>ki</u>.te.<u>ki</u>	'art'
	d:	ˈtɾi.ʎʊ	tri.<u>ki</u>.ʎo.<u>ki</u>	'trail'
	e:	vɛɦ.ˈda.ʤɪ	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'
	g:	a.ni.veh.ˈsa.ɾju	a.<u>ki</u>.ni.<u>ki</u>.veh.<u>ki</u>.sa.<u>ki</u>.rjo.<u>ki</u>	'birthday'

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.

2.3 *Training and Testing Methodology*

We recruited volunteers among undergraduate students from the Universidade Federal da Bahia, located in Salvador. Interviewing volunteers one at a time, we made sure that all volunteers were native speakers of the Salvador dialect through a pretest of questions designed to elicit words with pretonic nasal syllables, such as naming a picture of a pen and ensuring that the volunteer's response was [kɐ̃.ˈne.tɐ] and not [ka.ˈne.tɐ], or answering questions such as *Como se chama o lugar onde você vai asistir filmes*? ('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]. Volunteers who failed to apply pretonic nasalization at any point in this pretest were dismissed from further participation.

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 prospero 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 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 or LdK we presented nine test items to be transformed into the game, intermixed with nine filler items. The nine test items all contained vowels in open syllables that immediately preceded a heterosyllabic nasal consonant, such as [ka.ˈmi.za]. 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. All responses were recorded for further inspection and were checked post-hoc 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 results on the test items in LdP and LdK, the eleven volunteers who participated in our experiment can be divided into three groups[6]. We will call these groups the always-nasalizing, the never-nasalizing, and the Opaque-pattern groups, and discuss them in turn.

[6] As mentioned in footnote 2, 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 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)). Importantly, 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 [kẽ.ˈmi.zɐ], the transformation into LdP yielded the result [kẽ.pẽ.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.

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

(19)	a:	kẽ.ˈmi.zɐ	kẽ.pẽ.mi.pi.za.pa	'shirt'
	b:	kẽ.ni.ˈbaw	kẽ.pẽ.ni.pi.baw.paw	'cannibal'
	c:	fẽ.ˈmi.ʎɐ	fẽ.pẽ.mi.pi.ʎa.pa	'family'
	d:	ʒẽ.ˈnɛ.lɐ	ʒẽ.pẽ.nɛ.pɛ.la.pa	'window'
	e:	ka.fũ.ˈnɛ	ka.pa.fũ.pũ.nɛ.pɛ	'caress'
	f:	kẽ.ˈme.lʊ	kẽ.pẽ.me.pe.lo.po	'camel'
	g:	pẽ.nɛ.ˈtõ.nʲɪ	pẽ.pẽ.nɛ.pɛ.tõ.põ.ne.pe	'pannetone bread'
	h:	ta.ˈksĩ.mɛ.tɾʊ	ta.pa.ˈksĩ.pĩ.mɛ.pɛ.tɾo.po	'taxi-meter'
	i:	sis.ˈtẽ.mɐ	sis.pis.tẽ.pẽ.ma.pa	'system'

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

(20) 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.

(21)	a:	kẽ.ˈmi.zɐ	kẽ.ki.mi.ki.za.ki	'shirt'
	b:	kẽ.ni.ˈbaw	kẽ.ki.ni.ki.baw.ki	'cannibal'
	c:	fẽ.ˈmi.ʎɐ	fẽ.ki.mi.ki.ʎa.ki	'family'
	d:	ʒẽ.ˈnɛ.lɐ	ʒẽ.ki.nɛ.ki.la.ki	'window'
	e:	ka.fũ.ˈnɛ	ka.ki.fu.ki.nɛ.ki	'caress'
	f:	kẽ.ˈme.lʊ	kẽ.ki.me.ki.lo.ki	'camel'
	g:	pẽ.nɛ.ˈtõ.nʲɪ	pẽ.ki.nɛ.ki.to.ki.ne.ki	'pannetone bread'
	h:	ta.ˈksĩ.mɛ.tɾʊ	ta.ki.ˈksĩ.ki.mɛ.ki.tɾo.ki	'taxi-meter'
	i:	sis.ˈtẽ.mɐ	sis.ki.tẽ.ki.ma.ki	'system'

Despite consistently producing a nasalized vowel in pretonic pre-nasal open syllables such as [kẽ.ˈmi.zɐ], these volunteers were clearly not applying an active nasalization rule, as they did not nasalize the infix *–ki-* in LdK, as can be seen in (21), despite the fact that the vowel of the infix meets the structural description for the putative rule of nasalization.

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 [ka.pa.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.

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

(23)	a:	kɐ̃.ˈmi.zɐ	kɐ.pɐ.mi.pi.za.pa	'shirt'
	b:	kɐ̃.ni.ˈbaw	kɐ.pɐ.ni.pi.baw.paw	'cannibal'
	c:	fɐ̃.ˈmi.ʎɐ	fɐ.pɐ.mi.pi.ʎa.pa	'family'
	d:	ʒɐ̃.ˈnɛ.lɐ	ʒɐ.pɐ.nɛ.pɛ.la.pa	'window'
	e:	ka.fũ.ˈnɛ	ka.pa.fu.pu.nɛ.pɛ	'caress'
	f:	kɐ̃.ˈme.lʊ	kɐ.pɐ.me.pe.lo.po	'camel'
	g:	pɐ̃.nɛ.ˈtõ.nʲɪ	pɐ.pɐ.nɛ.pɛ.to.po.ne.pe	'pannetone bread'
	h:	ta.ˈksĩ.mɛ.tɾʊ	ta.pa.ˈksi.pi.mɛ.pɛ.tɾo.po	'taxi-meter'
	i:	sis.ˈtẽ.mɐ	sis.pis.te.pe.ma.pa	'system'

For the volunteers who we group under the Never-Nasalizing pattern, they 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 shows nasalization, and nor does the infix which suddenly precedes a nasal consonant.

(25)	a:	kɐ̃.ˈmi.zɐ	kɐ.ki.mi.ki.za.ki	‘shirt’
	b:	kɐ̃.ni.ˈbaw	kɐ.ki.ni.ki.baw.ki	‘cannibal’
	c:	fɐ̃.ˈmi.ʎɐ	fɐ.ki.mi.ki.ʎa.ki	‘family’
	d:	ʒɐ̃.ˈnɛ.lɐ	ʒɐ.ki.nɛ.ki.la.ki	‘window’
	e:	ka.fũ.ˈnɛ	ka.ki.fu.ki.nɛ.ki	‘caress’
	f:	kɐ̃.ˈme.lʊ	kɐ.ki.me.ki.lo.ki	‘camel’
	g:	pɐ̃.nɛ.ˈtõ.nʲɪ	pɐ.ki.nɛ.ki.to.ki.ne.ki	‘pannetone bread’
	h:	ta.ˈksĩ.mɛ.tɾʊ	ta.ki.ˈksi.ki.mɛ.ki.tɾo.ki	‘taxi-meter’
	i:	sis.ˈtẽ.mɐ	sis.pis.te.pe.ma.pa	‘system’

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. As an example, let us 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.<u>pa</u>.dɾɐ̃w̃.<u>pɐ̃w̃</u>].

In contrast, words formed by the prefix *ad-* and *re-* do not show evidence of resyllabification. Thus *ad-* + *rogar* [ɦo.gah] results in [a.ʤi.ɦo.ˈgah] and not [a.dɾo.ˈgah] ‘to accept someone for adoption’, and *re-* + *rasgar* [ɦazgah] 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 *[rẽ.nõ.me.ˈah], but as [rɛ.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 is 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.

We cannot altogether rule out an 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 [doiʃ] for the word *dois* 'two' and [siʃ.tẽ.mɐ] 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.

The simplest explanation for the Never-Nasalizing pattern is thus that these volunteers have induced a morphological structure for LdP and LdK that inhibits the application of pretonic nasalization, to the point where even words like [kɐ̃.ˈmi.zɐ] which these same volunteers nasalize in isolation surface with oral vowels in the environment of the game. In fact, the failure of the nasalization rule to apply under the conditions of the game provides an interesting indirect way for diagnosing the fact that such a rule indeed exists, and that the vowel of [kɐ̃.ˈmi.zɐ] is not underlyingly specified as [+nasal].

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 light on this question.

2.4.3: Overapplication in LdP, Underapplication in LdK

Six of our eleven volunteers produced 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 volunteers under the label of the "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.

(26) <u>Opaque-Pattern, LdP</u>: a pretonic open syllable whose adjacency with a heterosyllabic following nasal consonant is disrupted nonetheless shows nasalization

(27)	a:	kɐ̃.ˈmi.zɐ	kɐ̃.pɐ̃.mi.pi.za.pa	'shirt'
	b:	kɐ̃.ni.ˈbaw	kɐ̃.pɐ̃.ni.pi.baw.paw	'cannibal'
	c:	fɐ̃.ˈmi.ʎɐ	fɐ̃.pɐ̃.mi.pi.ʎa.pa	'family'
	d:	ʒɐ̃.ˈnɛ.lɐ	ʒɐ̃.pɐ̃.nɛ.pɛ.la.pa	'window'
	e:	ka.fũ.ˈnɛ	ka.pa.fũ.pu.nɛ.pɛ	'caress'
	f:	kɐ̃.ˈme.lʊ	kɐ̃.pɐ̃.me.pe.lo.po	'camel'
	g:	pɐ̃.nɛ.ˈtõ.nʲɪ	pɐ̃.pɐ̃.nɛ.pɛ.tõ.põ.ne.pe	'pannetone bread'
	h:	ta.ˈksĩ.mɛ.tɾʊ	ta.pa.ˈksĩ.pĩ.mɛ.pɛ.tɾo.po	'taxi-meter'
	i:	sis.ˈtẽ.mɐ	sis.pis.tẽ.pẽ.ma.pa	'system'

At first blush, these results recall the pattern of the Always-Nasalizers, for whom we concluded in Section 2.4.3 that the specification of the pretonic vowels was underlyingly [+nasal]. However, the behavior of the Interesting-Pattern group on LdK speaks against such an interpretation. A "pre-theoretical" statement of the pattern in shown above in (27) is that the rule of nasalization is over-applying in the Interesting-Pattern group, as its structural description is not met by the vowel in the first 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:

(28) 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 does not.

(29)	a:	kɐ̃.ˈmi.zɐ	kɐ.ki.mi.ki.za.ki	'shirt'
	b:	kɐ̃.ni.ˈbaw	kɐ.ki.ni.ki.baw.ki	'cannibal'
	c:	fɐ̃.ˈmi.ʎɐ	fɐ.ki.mi.ki.ʎa.ki	'family'
	d:	ʒɐ̃.ˈnɛ.lɐ	ʒɐ.ki.nɛ.ki.la.ki	'window'
	e:	ka.fũ.ˈnɛ	ka.ki.fu.ki.nɛ.ki	'caress'
	f:	kɐ̃.ˈme.lʊ	kɐ.ki.me.ki.lo.ki	'camel'
	g:	pɐ̃.nɛ.ˈtõ.nʲɪ	pɐ.ki.nɛ.ki.to.ki.ne.ki	'pannetone bread'
	h:	ta.ˈksĩ.mɛ.tɾʊ	ta.ki.ˈksi.ki.mɛ.ki.tɾo.ki	'taxi-meter'
	i:	sis.ˈtẽ.mɐ	sis.pis.te.pe.ma.pa	'system'

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 "under-application", 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 Lingua do Pe, while showing underapplication in Lingua do Ki, as schematized below:

(30) a: *Overapplication in LdP*:
Nasalization applies to the first syllable of ***kɐ̃****.pɐ̃.mi.pi.za.pa*, even though it shouldn't

b. *Underapplication in LdP*:
Nasalization fails to apply to the second syllable of *ka.****ki****.mi.ki.za.ki*, even though it should

(30a) and (30b) 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.

3. Overapplication in Língua do Pê

Based on their results on the test items in LdP, the eight of the volunteers showed overapplication in LdP, as demonstrated in (31) below, repeated from (27) above.

(31)	a:	kɐ̃.ˈmi.zɐ	kɐ̃.pɐ̃.mi.pi.za.pa	'shirt'
	b:	kɐ̃.ni.ˈbaw	kɐ̃.pɐ̃.ni.pi.baw.paw	'cannibal'
	c:	fɐ̃.ˈmi.ʎɐ	fɐ̃.pɐ̃.mi.pi.ʎa.pa	'family'
	d:	ʒɐ̃.ˈnɛ.lɐ	ʒɐ̃.pɐ̃.nɛ.pɛ.la.pa	'window'
	e:	ka.fũ.ˈnɛ	ka.pa.fũ.pu.nɛ.pɛ	'caress'
	f:	kɐ̃.ˈme.lʊ	kɐ̃.pɐ̃.me.pe.lo.po	'camel'
	g:	pɐ̃.nɛ.ˈtõ.nʲɪ	pɐ̃.pɐ̃.nɛ.pɛ.tõ.põ.ne.pe	'pannetone bread'
	h:	ta.ˈksĩ.mɛ.tɾʊ	ta.pa.ˈksĩ.pĩ.mɛ.pɛ.tɾo.po	'taxi-meter'
	i:	sis.ˈtẽ.mɐ	sis.pis.tẽ.pẽ.ma.pa	'system'

In the following four subsections, we discuss four possible analyses compatible with the pattern in (31).

3.1 Lexical specification of [+nasal]

It is compatible with the facts in (31) to simply say that words like [kɐ̃mizɐ] have a [+nasal] specification for the vowel in the initial syllable. This hypothesis would hold that words like [kɐ̃mizɐ] in Salvador BP are indeed, simply the result of memorizing them as such. On this view, there is no overapplication going on at all in (31), there is simply reduplication of a nasal vowel.

(32) /kã.mi.za/ → [kẽ.pẽ.mi.pi.za.pa]

Under this possibility, 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 (31).

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 in (31), but they differ in their representation of reduplication and in their model of the morphology-phonology interface more generally.

3.2 Base-Reduplicant Correspondence

In the framework of Base-Reduplicant Correspondence, as developed by McCarthy & Prince (1995), the overapplication in (31) can be analyzed as the result of a highly-ranked faithfulness constraint, demanding identity between the original lexical syllable and its copy in the reduplicant. The idea is the following: 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:

(33) 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*.

(34)

/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 *Língua do Pê* under the ranking shown in (34). 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 (see, for example, 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 (35) – repeated from (12) – is able to apply both before and after the process of reduplication yielding LdP. This derivation is shown in (36).

(35) *Salvador BP nasalization*:
$V \rightarrow [+nasal] / \_\_ N$

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

a:	/ka.mi.za/	(Underlying Representation)
b:	kẽ.mi.za	(application of (35))
c:	kẽ.pa.mi.pi.za.pa	(application of LdP formation)
d:	kẽ.pẽ.mi.pi.za.pa	(application of (35) 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 theory of Multiprecedence-and-Linearization, pioneered by Raimy (2000a,b) for the study of reduplication, and subsequently developed in a number of papers, including Fitzpatrick & Nevins (2002, 2004), Nevins (2005a,b), and Idsardi & Raimy (2005). 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 infixing reduplication of LdP.

3.4.1 Overview of Overapplication in Multiprecedence

In Raimy (2000a,b), overapplication of nasalization in Malay is treated 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 (37). The overapplication pattern can be seen in total reduplication, as shown in (38).

(37) V → [+nasal] / N __

(38) /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 (34), with the replacement of *Oral-V/ __N with *Oral-V/ N__.

Raimy (2000a,b), on the other hand, analyzes the overapplication in (38) as the result of the nasalization rule in (37) 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):

(39) |α| → a → ŋ → e → n → |ω|

If the nasalization rule in (37) is interpreted in terms of immediate precedence, then the immediate precedence relation between /n/ and /a/ meets the structural description, and the structural change of nasalization is applied. Subsequent linearization of (39) into a structure without multiple-precedence yields (40), 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.

(40) $|\alpha| \rightarrow \tilde{a}_1 \rightarrow ŋ_1 \rightarrow e_1 \rightarrow n_1 \rightarrow \tilde{a}_2 \rightarrow ŋ_2 \rightarrow e_2 \rightarrow n_2 \rightarrow |\omega|$

We state the fact that nasalization applies prior to linearization as the consequence of (41), 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).

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

We hold (41) 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*, *[ɐ̃mẽnĩna] 'the girl'), we will assume that it applies at the level of multiprecedence. We discuss this in Section 3.4.4 below.

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 of concatenation of discrete units must necessarily 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.

Here we take the standard view that such building blocks are segments, which are 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 determined. Moreover, it is an undisputable fact that many phonological rules are sensitive to initial or final word boundaries. Therefore, the notions of begin and end of a word must be somehow formally encoded in morphophonological representations.

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, as in (42), in which P denotes the immediate precedence relation.

(42) a: $\exists|\alpha| \mid [\ \neg\exists x \mid P(x,|\alpha|)\ ]$ (start symbol)

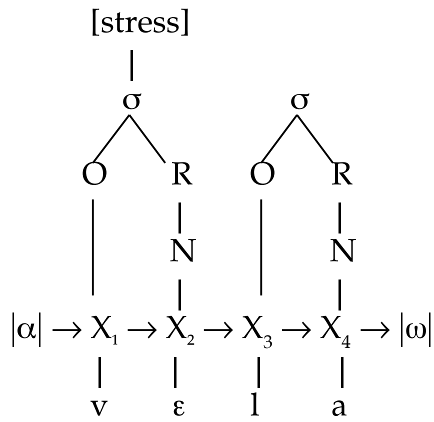
b: $\exists|\omega| \mid [\ \neg\exists y \mid P(|\omega|,y)\ ]$ (end symbol)

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.

Thus, the underlying form of the Portuguese word [ˈvɛ.la] 'candle' would be as in (43).

(43)
$$\begin{array}{ccccccccccc} |\alpha| & \rightarrow & X_1 & \rightarrow & X_2 & \rightarrow & X_3 & \rightarrow & X_4 & \rightarrow & |\omega| \\ & & | & & | & & | & & | & & \\ & & \text{v} & & \varepsilon & & \text{l} & & \text{a} & & \end{array}$$

Syllabic (and additional prosodic/metric) structure would be built from this skeleton of timing slots, ultimately yielding something like (44).

(44)


What is implicit in the simplified notation in (43) above is that each phonological representation is formally defined as in (45), as a pair <Π,Φ>, where Π is a set of ordered pairs encoding precedence relations among timing slots, whereas Φ is a set of ordered pairs encoding associations between each timing slot and a segment.

(45) $<\Pi,\Phi>$, such that: $\Pi = \{ <|\alpha|,X_1>, <|\alpha|,X_2>, <|\alpha|,X_3>, <|\alpha|,X_4>, <|\alpha|,|\omega|>, <X_1,X_2>, <X_1,X_3>, <X_1,X_4>, <X_1,|\omega|>, <X_2,X_3>, <X_2,X_4>, <X_2,|\omega|>, <X_3,X_4>, <X_3,|\omega|>, <X_4,|\omega|> \}$

$$\Phi = \{ <X_1,v>, <X_2,\varepsilon>, <X_3,l>, <X_4,a>\}$$

For the sake of exposition, we omit timing slots from the notation throughout the paper, as if precedence relations held directly of segments themselves, like in (46). The reader should keep in mind, though, that examples like (46) are mere shortcuts standing for richer structures like (45),

(46) $\Pi = \{ <|\alpha|,v>, <|\alpha|,\varepsilon>, <|\alpha|,l>, <|\alpha|,a>, <|\alpha|,|\omega|>, <v,\varepsilon>, <v,l>, <v,a>, <v,|\omega|>, <\varepsilon,l>, <\varepsilon,a>, <\varepsilon,|\omega|>, <l,a>, <l,|\omega|>, <a,|\omega|> \}$

In plain English, (46) reads as in (47).

(47) |α| precedes /v/, |α| precedes /ε/, |α| precedes /l/, |α| precedes /a/, |α| precedes |ω|
/v/ precedes /ε/, /v/ precedes /l/, /v/ precedes /a/, /v/ precedes |ω|
/ε/ precedes /l/, /ε/ precedes /a/, /ε/ precedes |ω|
/l/ precedes /a/, /l/ precedes |ω|
/a/ precedes |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 (45) and (46)) emerge during the mapping from 'deep' morpho-phonological structures (which constitute the input to cyclic rule application (cf. (41)) 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 (46) actually reduces to the set of immediate precedence relations in (48).

(48) $\Pi_I = \{ <|\alpha|,v>, <v,\varepsilon>, <\varepsilon,l>, <l,a>, <a,|\omega|> \}$

Actually, immediate precedence can be represented in three distinct ways. The ordered-pair notation in (48) 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 (49).

(49) $\Pi_I = \{ P(|\alpha|,v) \ \& \ P(v,\varepsilon) \ \& \ P(\varepsilon,l) \ \& \ P(l,a) \ \& \ P(a,|\omega|) \}$

Finally, we can also denote immediate precedence by means of the tail and head of an arrow, as in (50). This notation is perhaps easiest one to read, which is why we will systematically adopt it along the paper.

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

It is important to emphasize that all of these notations are formally equivalent, though of course some make more salient certain aspects of the representation (thus better satisfying "depictive adequacy").

Since the role of the arrangement of timing slots in Π is to provide instructions for the A(rticulatory)-P(erceptual) system to string segments in real time, Π should in principle specify a linear order among 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. In other words, R must have the three properties in (51) (cf. Partee, ter Meulen & Wall 1993: 206-211).[7]

(51) *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 linearizes it, converting it into a strictly linear representation fully interpretable by the A-P system, there is nothing wrong, for example, with structures like (39) – repeated below as (52) – violating the asymmetry property in (51b), since, for instance, /a/ globaly precedes /e/ whereas /e/ globally precedes /a/.[8]

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

[7] The fourth property inherent to linear order is *irreflexivity* (called aliorelativity by Russel 1919: chapter IV), defined as follows: *irreflexivity* $=_{def} \forall x, \forall y \ [[<x,y> \in R] \rightarrow [x \neq y]]$ (i.e. $\forall x, [<x,x> \notin R]$). Actually, it is not necessary to make this fourth property explicit, since every symmetric relation is, by definition, an irreflexive one. Alternatively, we can do the other way around, making irreflexivity explicit and omitting (presupposing) asymmetry, since, among transitive relations, all irreflexive ones are also asymmetric, and vice versa (see Russel 1919: chapter IV).
[8] Notice that *reflexivity* (cf. previous footnote) is also violated in (52). For instance, given the loop, /e/ ends up preceding itself (which is true of all segments in cases of total reduplication like this).

Once this structure is properly mapped into (53) – repeated from (40) –, the relevant relations hold of distinct occurences of those tokens in (52), 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*).

(53) $|\alpha| \rightarrow \tilde{a}_1 \rightarrow ŋ_1 \rightarrow e_1 \rightarrow n_1 \rightarrow \tilde{a}_2 \rightarrow ŋ_2 \rightarrow e_2 \rightarrow n_2 \rightarrow |\omega|$

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 exhibit a linear order for the transitive clorure of its immediate precedence relations. The second one, which we can call *surface morpho-phonological structure* (or, alternatively, *Linearized Structure*), does.

This formalism, in which morpho-phonology is factored out into two levels, is obviouly not the only logical possibility for the architecture of grammar. One could easily conceive a single level of morphophonology where linear order holds, and where phenomena like overapplication and underapplication in reduplicative structures are handled differently, by correspondence rules and the like. However, we will demonstrate in this paper that such approaches are empirically inadequate to handle the full range of phenomena into consideration. The main point is ultimately empirical, and the multiprecedence alternative is formally coherent.

This reasoning goes back to Chomsky's (1975 [1955]) considerations on the advantage (and/or the necessity) of postulating multiple distinct levels of representation in linguistic theory.

> A language is an enormously complex system. Linguistic theory attempts to reduce this imense complexity to menageable proportions by the construction of a system of linguistic levels, each of which makes a certain descriptive apparatus available for the characterization of linguistic structure. A grammar reconstructs the total complexity of a language stepwise, separating out the contribution of each linguistic level. The adequacy of a linguistic theory containing a given set of abstractly formulated levels can be tested by determining whether the grammars resulting from rigorous application of this theory (i) meet certain formal conditions of simplicity; [and] (ii) lead to intuitively satisfactory analyses – i.e. offer explanations on formal grounds for the linguistic intuition of the native speaker.
>
> (Chomsky 1975 [1955]: 63)

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 a descriptively and explanatorily adequate way of offering explanations on formal grounds for the linguistic intuition of the native speaker, when it comes to phenomena like reduplication and infixation, which show patterns of underapplication and overapplication of rules.

As for *Lingua do Pê*, specifically, the deep and surface morpho-phonological structures will be like in (54) and (55), exemplified for the word [vɛla] 'candle'.

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

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

We pursue the strong hypothesis in (41) that all cyclic phonological rules apply only to deep morphophonological structures (which may contain multiprecedence patterns in some cases, like in (54)), and that the only purpose of surface structures such as (55) 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 occurences 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 occurences as necessary at the surface.[9]

[9] 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.')). Whether or not facts like normal application of postlexical allophony may be taken as counter-examples to our hypothesis, depends on the phonetic patterns shown in the allophony, which can only be properly observed through systematic experimental measurements. The prediction is that only allophonic alternations originating from strictly cyclic phonological manipulation of symbol,s in abstract representations should not exhibit normal application. Nothing, however, prevents that there be some allophonies originating at the very level of postcyclic rules, and through dynamic mechanisms of implementation. The ultimate prediction is that cases of normal application should be found across morphemes, words, phrases, and so on, in a gradient fashion; whereas cases of overapplication and underapplication should be restricted to the immediate morphophonological domain defined by Multiprecedence 'loops', and that the latter cases should not exhibit gradient behavior.

Our choice for not having post-linearization rules in the system is not just the urge to have a more restrictive/prohibitive (hence more falsifiable) theory. Actually, there are conceptual reasons motivating this move, which bear on the type/token distinction. Consider (56).

(56) *Inventory of phonemes* *(level-0)*

{ k, p, t, d, k, b, v, m, n, ɛ, o, i, u, a, e, ɛ, ... }

The /ɛ/ in (56) is a *type* (of phoneme): a unique abstract entity that is represented at other levels of representation less removed from the surface by instantiations of it. Let us first consider the simple case in (57).

(57) *Pre-linearized structure* *(level-1)*

$|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow l \rightarrow a \rightarrow |\omega|$ <vela> 'candle'

The /ɛ/ in (57) is a representation of that type /ɛ/ from (56) at a lower level of abstraction. Technically speaking, the /ɛ/ in (57) it is a *token* of that type /ɛ/ from (56). It is still an abstract entity quite removed from the surface (physical) level, but it is not that very same abstract unique type /ɛ/ from the inventory. Crucially, at this level, there can be many entities with the very same formal properties (i.e. [+back], [–round], [–high], etc), as opposed to the unique /ɛ/ in (56). There can even be many distinct tokens of the same type inside a single morpheme, as shown in (58).

(58) *Pre-linearized structure* *(level-1)*

$|\alpha| \rightarrow p \rightarrow \varepsilon_1 \rightarrow l \rightarrow \varepsilon_2 \rightarrow |\omega|$ 'Pelé' (proper name)

Neither (57) nor (58) is yet a structure that is interpreted/implemented by the Articulatory-Perceptual interface component. They are still higher-level formal objects that have to be further mapped, by means of a linearization mechanism, onto something directly readable by the A-P system. This becomes clear when we consider morphologically complex cases like the one in (59), which is mapped onto (60).

(59) *Pre-linearized structure* *(level-1)*

$$\begin{array}{ccccccccccc} |\alpha| & \rightarrow & v & \rightarrow & \varepsilon & \rightarrow & l & \rightarrow & a & \rightarrow & |\omega| \\ & & & & \downharpoonleft\upharpoonright & & & & \downharpoonleft\upharpoonright & & \\ & & & & p_1 & & & & p_2 & & \end{array}$$

(60) *Linearized structure* *(level-2)*

$|\alpha| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l \rightarrow a_1 \rightarrow p_2 \rightarrow a_2 \rightarrow |\omega|$

In (59), there is a single instance of /ɛ/, which is a token built out of the unique type /ɛ/ from (56). In (60), however, there are two instances of /ɛ/ built out of the same /ɛ/ in (59). Both are representations of the token /ɛ/ at the next level. This multiplication of /ɛ/ happens because the only way for the system to map both immediate relations in which /ɛ/ participates is to actually create two distinct instantiations of /ɛ/ out of the same single token of /ɛ/ present in the pre-linearized structure, such that each instantiation of that token will represent a distinct relational property of it, given that those properties couldn't otherwise be represented at the surface phonetic level, by virtue of a *bare output condition* of the A-P interface, which makes it physiologically impossible to pronounce both [p] and [l] simultaneously immediately after pronouncing [ɛ]. We call each distinct representation of a token an *occurrence* of that token. Therefore, the two distinct instances of /ɛ/ shown in (60) are *occurrences* of the single /ɛ/ shown in (59), which, in turn, is a *token* of the *type* /ɛ/ from the inventory in (56).

In sum, at, the level of deep morpho-phonological structure the relevant formatives participating in precedence relations are tokens (of segment types), whereas, at the level of surface morpho-phonological structure, the relevant formatives participating in precedence relations are occurences (of tokens). Moreover, the linear order requirement holds only at the surface morpho-phonological structure. In the next section, we demonstrate how Multiprecedence representations are built, focusing on the process of introduction of new immediate precedence relations that constitute LdP.

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.

(61) 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

In other words: $\forall\ \Sigma, \exists p \mid P(\text{last}(\Sigma), p)\ \&\ P(p,(\text{nucleus}(\Sigma))$, 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 segment, there must be a token of the segment/p/ that is "available" for the application of rule (61). Moreover, we assume that there it is a new token of /p/ for each syllable in the word. In *Appendix I*, we demonstrate that a morphological structure for LdP with a single /p/ attached to every syllable will not be linearizable.

Application of (61) to an input like (62) is shown in (63):

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

(63)
$$\begin{array}{ccccccccccc} |\alpha| & \rightarrow & v & \rightarrow & \varepsilon & \rightarrow & l & \rightarrow & a & \rightarrow & |\omega| \\ &&&& \downarrow\uparrow &&&& \downarrow\uparrow && \\ &&&& p_1 &&&& p_2 && \end{array}$$

Application of (61) to an input like (64) is shown in (65):

(64) $|\alpha| \rightarrow a \rightarrow r \rightarrow t \rightarrow e \rightarrow |\omega|$

(65)
$$\begin{array}{ccccccccccc} |\alpha| & \rightarrow & a & \rightarrow & r & \rightarrow & t & \rightarrow & e & \rightarrow & |\omega| \\ &&& \nwarrow\swarrow &&&&& \downarrow\uparrow && \\ &&& p_1 &&&&& p_2 && \end{array}$$

3.4.3 Linearization of Multiprecedence Structures

To understand the linearization algorithm, we must introduce the following notation:

(66) 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

$|\alpha|$ = The start symbol.

$|\omega|$ = The terminal symbol.

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.

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

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

c: *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 $[P(x,y) \in I]$.

d: *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 (67a-c) should be fairly clear, brief discussion is necessary on the properties of Axiom (67d). 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 (67d) 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 (67d) 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*). That is, segments may be represented more than once (yielding multiple occurrences out of the same token) only if necessary. Without this condition, there would be no way to prevent the system from taking 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) be also represented in the output.

An example derivation of the Língua do Pê for the input in (68) is provided in (69a-i).[10]

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

(69) a: *step 1*

By Axiom (67a), the relation $P(|\alpha|,v)$ is identified as the starting point of the mapping procedure, and, once $P(|\alpha|,v)$ is read off the input structure, it is represented in the output structure by the concatenation of an occurrence of /v/ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v$

b: *step 2*

The relation $P(v,\varepsilon)$ is read off the input structure and it is represented in the output structure by the concatenation of an occurrence of /ε/ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon$

c: *step 3*

The relations $P(\varepsilon,l)$ and $P(\varepsilon,p_1)$ 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. The *Axiom of Choice of Path* (67d) demands that the next segment be the candidate be /p_1/, because /p_1/ transitively precedes /l/ (cf. the sub-path /$p_1 \rightarrow \varepsilon \rightarrow l$/) and not vice-versa.[11]

$|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow p_1$

[10] Note that 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*.

[11] Notice that neither / $\varepsilon \rightarrow l$/ nor / $\varepsilon \rightarrow p_1$/ are present in the output yet, so all other requirements of Axiom (67d) are trivially satisfied.

d: *step 4*

The relation $P(p_1,\varepsilon)$ is read off the input structure and it is represented in the output structure by the concatenation of another occurrence of /ε/ (created out of the same token) at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2$

e: *step 5*

The system then comes back to the same bifurcation point of step 3. Once again it has to decide whether /l/ or /p_1/ will immediately follow /ε/ in the output. Axiom (67d) demands that the next segment be the candidate be /l/, because, at this point, concatenating another occurrence of /p_1/ at the end of the string would violate both conditions (i) and (ii) of Axiom (67d), since $P(\varepsilon,p_1)$ has already been mapped onto the output (contrary to what the first conjunct of condition (i) demands), and $P(\varepsilon,l)$ has not yet been mapped onto the output (contrary to what condition (ii) demands). Therefore, the relation $P(\varepsilon,l)$ is read off the input structure and it is represented in the output structure by the concatenation of an occurrence of /l/ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l$

f: *step 6*

The relation $P(l,a)$ is read off the input structure and it is represented in the output structure by the concatenation of an occurrence of /a/ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l \rightarrow a_1$

g: *step 7*

The system again faces a situation where it has to choose between two alternative paths. The relations $P(a,|\omega|)$ and $P(a,p_2)$ are both read off the input structure. At this point the system has to decide whether $|\omega|$ or /p_2/ will immediately follow /a/ in the output. Axiom (67d) demands that the next segment be the candidate be /p_2/, because /p_2/ transitively precedes $|\omega|$ (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: *step 8*

The relation $P(p_2,a)$ is read off the input structure and it is represented in the output structure by the concatenation of another occurrence of /a/ (built out of the same token) at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l \rightarrow a_1 \rightarrow p_2 \rightarrow a_2$

i: *step 9*

The system then comes back to the same bifurcation point of step 3. Once again it has to decide whether $|\omega|$ or $/p_2/$ will immediately follow /a/ in the output. The *Axiom of Choice of Path* (67d) demands that the next segment be the candidate be $|\alpha|$, because, at this point, concatenating another occurrence of $/p_2/$ at the end of the string would violate both conditions (i) and (ii) of Axiom (67d), since $P(a,p_2)$ has already been mapped onto the output (contrary to what the first conjunct of condition (i) demands), and $P(a,|\omega|)$ has not yet been mapped onto the output (contrary to what condition (ii) demands). Therefore, the relation $P(a,|\omega|)$ is read off the input structure and it is represented in the output structure by the concatenation of the terminal symbol $|\omega|$ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon_1 \rightarrow p_1 \rightarrow \varepsilon_2 \rightarrow l \rightarrow a_1 \rightarrow p_2 \rightarrow a_2 \rightarrow |\omega|$

By Axiom (67b), the linearization procedure terminates at this point, and the output representation is ready to be delivered to the interpretive component.

The reader can verify that this Linearization algorithm will succesfully yield the output structures of any Língua do Pê input created by the rule in (61).

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:

(70) Salvador BP Nasalization, *general definition*

Structural Description: $\forall x, \forall y \mid [P(x,y) \in I\,] \;\&\; V(x) \;\&\; C_{[+nasal]}(y)$,

Structural Change: x becomes [+nasal]

When applied to (71a), this rule yields (71b).

(71) a: $|\alpha| \rightarrow k \rightarrow a_1 \rightarrow m \rightarrow i \rightarrow z \rightarrow a_2 \rightarrow |\omega|$

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

Notice that those simple cases are compatible with another slightly different formulation of the rule, as shown in (72):

(72) 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 basic difference between (70) and (72) 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, (70) states that "**AN** immediate preceder of the nasal consonant becomes nasalized" whereas (72) 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 [kẽ.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 obviously the one with in (72) with iota, as it applies in a subset of the environments in with (70) applies.

Although forms like [ˈkẽ.mɐ] could be due to either (70) or (72), when morphologically complex structures involving multiprecedence arise, (70) and (72) 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 (72) over (70).

Adopting (72), the rule in (61) 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.

First, the LdP rule in (61) converts /kamiza/ (73a) into (73b):

(73) a: $|\alpha| \rightarrow k \rightarrow a_1 \rightarrow m \rightarrow i \rightarrow z \rightarrow a_2 \rightarrow |\omega|$

b:
$$\begin{array}{ccccccccccccccc} |\alpha| & \rightarrow & k & \rightarrow & a_1 & \rightarrow & m & \rightarrow & i & \rightarrow & z & \rightarrow & a_2 & \rightarrow & |\omega| \\ & & & & \downarrow\uparrow & & & & \downarrow\uparrow & & & & \downarrow\uparrow & & \\ & & & & p_1 & & & & p_2 & & & & p_3 & & \end{array}$$

Next, the nasalization rule in (72) applies. Notice that there is only one preceder of /m/, so the iota condition is met.

(74) $$\begin{array}{cccccccccccccccc} |\alpha| & \to & k & \to & \tilde{a}_1 & \to & \mathbf{m} & \to & i & \to & z & \to & a_2 & \to & |\omega| \\ & & & & \downarrow\uparrow & & & & \downarrow\uparrow & & & & \downarrow\uparrow & & \\ & & & & p_1 & & & & p_2 & & & & p_3 & & \end{array}$$

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. Finally, the Linearization procedure yields (75):

(75) $$|\alpha| \to k \to \tilde{a}_1 \to p_1 \to \tilde{a}_2 \to m \to i_1 \to p_2 \to i_2 \to z \to a_4 \to p_3 \to a_4 \to |\omega|$$

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 [kɐ̃.pɐ̃.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 of which is 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.

(76)	a:	kɐ̃.ˈmi.zɐ	kɐ.ki.mi.ki.za.ki	'shirt'
	b:	kɐ̃.ni.ˈbaw	kɐ.ki.ni.ki.baw.ki	'cannibal'
	c:	fɐ̃.ˈmi.ʎɐ	fɐ.ki.mi.ki.ʎa.ki	'family'
	d:	ʒɐ̃.ˈnɛ.lɐ	ʒɐ.ki.nɛ.ki.la.ki	'window'
	e:	ka.fũ.ˈnɛ	ka.ki.fu.ki.nɛ.ki	'caress'
	f:	kɐ̃.ˈme.lʊ	kɐ.ki.me.ki.lo.ki	'camel'
	g:	pɐ̃.nɛ.ˈtõ.nʲɪ	pɐ.ki.nɛ.ki.to.ki.ne.ki	'pannetone bread'
	h:	ta.ˈksĩ.mɛ.tɾʊ	ta.ki.ˈksi.ki.mɛ.ki.tɾo.ki	'taxi-meter'
	i:	sis.ˈtẽ.mɐ	sis.pis.te.pe.ma.pa	'system'

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 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.

4.2 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 Correspondence model of LdP involved the following output-oriented constraints:

(77) 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*.

(78)

/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 Lingua do Ki under the ranking shown in (78). The real problem is not with Base-Reduplicant Correspondence, which is trivially satisfied and hence irrelevant in the case of infixation, but with opacity of underapplication of nasalization. We have abstracted away from the constraints that yield the actual iterative infixing reduplication pattern that constitutes the LdP game itself here.

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 (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 (80).

(79) *Salvador BP nasalization*:

$V \rightarrow [+nasal] \; / \; \_\_ \; N$

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

a:	/ka.mi.za/	(Underlying Representation)
b:	kẽ.mi.za	(application of (79))
c:	kẽ.pa.mi.pi.za.pa	(application of LdP formation)
d:	kẽ.pẽ.mi.pi.za.pa	(application of (79) 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:

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

a:	/ka.mi.za/	(Underlying Representation)
b:	kẽ.mi.za	(application of (79))
c:	kẽ.ki.mi.ki.za.ki	(application of LdK formation)
d:	* kẽ.kĩ.mi.ki.za.ki	(application of (79) again)

Recall that speakers who fall into the Opaque group actually 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. We begin by reviewing the morphological operations of iterative infixation that yield LdK in Section 4.4.1. In Section 4.4.2 we show how the linearization axioms developed in 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 in Javanese low-vowel rounding.

4.4.1 How to Build a Lingua do Ki Structures

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

(82) For every syllable Σ:

a: Add a new immediate precedence relation between the x, the last segment of Σ and the segment /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 (83a), the rule (82) yields the structure in (83b).

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

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

This is one of the cases where it becomes clear why the edge symbols ($|\alpha|$ and $|\omega|$) are crucial. At first blush, it seems that word boundaries can be defined without such primitives. Indeed, this is possible. For example, 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 in relational terms, 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 (82) above cannot be defined in a trivial fashion, as the infixation of the last instance of /ki/ in a structure like (84) would require a different (and most likely, disjunctive) rule than the one above in (82), given that there is no "segment that /a/ immediately precedes" in the structure.

(84) a: $v \to \varepsilon \to l \to a$

b:
$$\begin{array}{cccc} v \to & \varepsilon \to & l \to & a \\ \downarrow & \uparrow & \downarrow & \uparrow \\ k_1 \to & i_1 & k_2 \to & i_2 \end{array}$$

The same reasoning applies to the starting symbol $|\alpha|$. For instance, consider another dialect of LdP (mentioned in Section 2.1, though not one we studied) that involves the insertion of the fixed syllable /pe/ right before each syllable, so that [ˈvɛ.la] becomes [pe.vɛ.pe.la].

In a system with boundary symbols, the rule can be trivially formalized as follows:

(85) For every syllable Σ:

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

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

(86) a: $|\alpha| \to v \to \varepsilon \to l \to a \to |\omega|$

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

Without boundary symbols, however, the infixation of the first instance of /pe/would require a different rule, given that there is no "segment that immediately precedes /v/" in the input structure.

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

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

We thus adopt both the formalization of LdK in (K) and a representation that explicitly includes the begin and end symbols as participants in immediate-precedence relations.

4.4.2 The Linearization of Infixes

A sample derivation of the *Língua do Ki* output in (83b) from the input in (83a) – repeated below in (88) –, is provided in (89).[12]

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

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

The linearization procedure, as developed in Section 3.4.3, will apply to (88b) in a step-by-step derivation as follows:

(89) a: *step 1*

By Axiom (67a), the relation P(|α|,v) is identified as the starting point of the mapping procedure, and once P(|α|,v) is read off the input structure, it is represented in the output structure by the concatenation of an occurrence of /v/ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v$

b: *step 2*

The relation P(v,ε) is read off the input structure and it is represented in the output structure by the concatenation of an occurrence of /ε/ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon$

[12] Note that 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.*

c: *step 3*

The relations $P(\varepsilon,l)$ and $P(\varepsilon,k_1)$ 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. Axiom (67d) (*Choice of Path*) demands that the next segment be the candidate be /p_1/, because /k_1/ transitively precedes /l/ (cf. the sub-path /$k_1 \rightarrow i_1 \rightarrow l$/) and not vice-versa.[13]

$|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow k_1$

d: *step 4*

The relation $P(k_1,i_1)$ is read off the input structure and it is represented in the output structure by the concatenation of an occurrence of /i_1/ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow k_1 \rightarrow i_1$

e: *step 5*

The relation $P(i_1,l)$ is read off the input structure and it is represented in the output structure by the concatenation of /l/ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow k_1 \rightarrow i_1 \rightarrow l$

f: *step 6*

The relation $P(l,a)$ is read off the input structure and it is represented in the output structure by the concatenation of /a/ at the end of the string being incrementally built, as dictated by Axiom (67c).

$|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow k_1 \rightarrow i_1 \rightarrow l \rightarrow a$

[13] Notice that neither /$\varepsilon \rightarrow l$/ nor /$\varepsilon \rightarrow p_1$/ are present in the output yet, so all other requirements of Axiom (49d) are trivially satisfied.

g: *step 7*

The system again faces a situation where it has to choose between two alternative paths. The relations $P(a,|\omega|)$ and $P(a,k_2)$ are both read off the input structure. At this point the system has to decide whether $|\omega|$ or $/k_2/$ will immediately follow /a/ in the output. The *Axiom of Choice of Path* (67d) demands that the next segment be the candidate be $/k_2/$, because $/p_2/$ transitively precedes $|\omega|$ (cf. the sub-path $/k_2 \rightarrow i_2 \rightarrow |\omega|/$) and not vice-versa.

$$|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow k_1 \rightarrow i_1 \rightarrow l \rightarrow a \rightarrow k_2$$

h: *step 8*

The relation $P(k_2,i_2)$ is read off the input structure and it is represented in the output structure by the concatenation of $/i_2/$ at the end of the string being incrementally built, as dictated by Axiom (67c).

$$|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow k_1 \rightarrow i_1 \rightarrow l \rightarrow a \rightarrow k_2 \rightarrow i_2$$

i: *step 9*

The relation $P(i_2,|\omega|)$ is read off the input structure and it is represented in the output structure by the concatenation of the terminal symbol $|\omega|$ at the end of the string being incrementally built, as dictated by Axiom (67c).

$$|\alpha| \rightarrow v \rightarrow \varepsilon \rightarrow k_1 \rightarrow i_1 \rightarrow l \rightarrow a \rightarrow k_2 \rightarrow i_2 \rightarrow |\omega|$$

By Axiom (67b), the linearization procedure terminates at this point, and the output representation is ready to be delivered to the interpretive component.

Notice that when the mapping is complete, there have been two immediate precedence relations left unmapped, namely $P(\varepsilon,l)$ and $P(a,|\omega|)$. This is not "harmful" at all since (i) all segments have been mapped, and (ii) the correct predictions are being made for all cases, 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 given multiprecedence representations.

The reader can verify that this Linearization algorithm will succesfully yield the output structures of any LdK input created by the rule in (82).

4.4.3 The Failure of Nasalization

Given the nasalization rule in (79) above – repeated below as (90), defined in terms of a *unique preceder*, it now becomes clear why underapplication occurs in LdK structures, as shown below.

(90) *formal definition*

Structural Description: $\forall y, \iota x \mid [P(x,y) \in I\,] \;\&\; V(x) \;\&\; C_{[+nasal]}(y),$

Structural Change: x becomes [+nasal]

(91) a: *input*

$$|\alpha| \rightarrow k \rightarrow a_1 \rightarrow n \rightarrow e \rightarrow t \rightarrow a_2 \rightarrow |\omega|$$

b: *infixation*

$$\begin{array}{ccccccccccccccc} |\alpha| & \rightarrow & k & \rightarrow & a_1 & \rightarrow & n & \rightarrow & e & \rightarrow & t & \rightarrow & a_2 & \rightarrow & |\omega| \\ &&&& \downarrow && \uparrow && \downarrow && \uparrow && \downarrow && \uparrow \\ &&&& k_1 & \rightarrow & i_1 && k_2 & \rightarrow & i_2 && k_3 & \rightarrow & i_3 \end{array}$$

c: *nasalization fails to apply, since both* /a_1/ *and* /i_1/ *precede* /n/

$$\begin{array}{ccccccccccccccc} |\alpha| & \rightarrow & k & \rightarrow & a_1 & \rightarrow & n & \rightarrow & e & \rightarrow & t & \rightarrow & a_2 & \rightarrow & |\omega| \\ &&&& \downarrow && \uparrow && \downarrow && \uparrow && \downarrow && \uparrow \\ &&&& k_1 & \rightarrow & i_1 && k_2 & \rightarrow & i_2 && k_3 & \rightarrow & i_3 \end{array}$$

d: *linearization*

$$|\alpha| \rightarrow k \rightarrow a_1 \rightarrow k_1 \rightarrow i_1 \rightarrow n \rightarrow e_1 \rightarrow k_2 \rightarrow i_2 \rightarrow t \rightarrow a_2 \rightarrow k_3 \rightarrow i_3 \rightarrow |\omega|$$

The crucial step is step (91c): the multiprecedence representation of infixes is such that *both* the lexical segment /a/ (which undergoes ordinarily nasalization in [kãmiza] 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 which do not involve infixes at all.

4.4.1 Underapplication of Low Vowel Rounding in Javanese

A concrete case can be found in Javanese[14], which exhibits a rule of low-vowel rounding (Dudas 1976: 206).

(92)	a. mejɔ 'table'	meja-ku 'my table'	meja-ne	'his table'
	b. jiwɔ 'soul'	jiwa-ku 'my soul'	jiwa-ne	'his soul'
	c. dɔngɔ 'color'	dɔnga-ku 'my color'	dɔnga-ne	'his color'

Dudas (1976) provides the rule in (93), which we formulate in precedence-based terms in (94)

(93) $a \rightarrow ɔ \;/\; \_\_|\omega|$

(94) Structural Description: y such that $|\alpha|$ (y), ιx | [P(x,y) $\in$ I] & low(x)
Structural Change: x becomes [+round]

In reduplication, the rule in (94) overapplies, as shown in (95). This is a consequence of the graph in (96)

(95)	a:	mejɔ	'table'
	b:	mejɔ-mejɔ	'tables'
	c.	dɔngo	'prayer'
	d.	dɔngo-dɔngo	'prayers'

(96)
$$\begin{tikzcd} {|\alpha|} \arrow[r] & m \arrow[r] & e \arrow[r] & ʤ \arrow[r] & a \arrow[r] \arrow[lll, bend left] & {|\omega|} \end{tikzcd}$$

Since /a/ remains the unique preceder of $|\omega|$ in (96), it is expected that the low-vowel rounding rule of (94) should apply. It is interesting to consider the interaction of reduplication with suffixation, however. The rule of low-vowel rounding *underapplies* in this case, as can be seen in (97):

(97)	a:	meja-meja-ne	'his tables'.
	b:	dɔnga-dɔnga-ne	'his prayers'

When one examines the multiprecedence representation of (97), which contains both reduplication and suffixation, the underapplication of low-vowel rounding is in fact entirely to be expected, given the rule in (94)

[14] We heartily thank Eric Raimy and Bill Idsardi for bringing this case to our attention and acknowledge them in proposing its analysis.

(98) |α| → m → e→ ʤ → a → |ω|
↓ ↑
n → e

The rule in (94) cannot apply to (98), 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 results here too, from the interaction of the iota operator with a multiprecedence representation.

Before concluding, we will note that the cases of total reduplication such as (95b,d) provide additional evidence for taking word boundaries to be defined in terms of abstract symbols, specified as 'first' and 'last', that are in bona-fide immediate-precedence relations. Consider again the structure in (98) above, but without the boundary symbols.

(99) a: *input to total reduplication*
m → e → ʤ → a

b: *output from total reduplication*
m → e → ʤ → 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 (99a), /m/ and /a/ can be identified as the first and the last segments, respectively. However, the same doesn't hold for (99b). 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. First of all, there is no way in which an automaton or proof-system can get started, since no 'first symbol' can be detected.

Second of all, even if the first segment could be identified, another problem arises with respect to the loop. After the linearization procedure reaches the state /a/, Axiom (67c) 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 introduced in the output, the automaton wouldn't know when 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. Notice, however, that this is not enough to derive the facts. If that were true, then we would expect the linearized from of (99b) – repeated below as (100) – to be (101b) rather than (101a), since (101b) realizes all the segments and relations from the input.

(100) *hypothetical input to the rule of low-vowel rounding*

m → e → ʤ → a

(101) *potential outputs from the rule of low-vowel rounding, given input (100)*

a: m → e → ʤ → a → m → e → ʤ → a (actually attested form)

b: m → e → ʤ → a → m (wrongly predicted form)

In summary, the total reduplication pattern of Javanese provide another formal argument for the necessity of abstract start |α| and end |ω| symbols as crucial to a coherent linearization algorithm.

4.5 Summary of Approaches to Underapplication in LdK

In summarizing this section, it is worth considering again the essential fact of underapplication in LdK, as repeated in (102):

(102) kẽ.ˈmi.za ka.ki.mi.ki.za.ki

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 by appeal 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, who proposes that infixation is the result of metathesis, and Yu 2003, who 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 [kakimikazi], 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 [kakimikazi] 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 (103):

(103) Unattested Pattern across the Two Games
  a. Overapplication of LdK output: kẽ. kĩ.mi.ki.za.ki
  b. Underapplication of LdP output: ka.pa.mi.pi.za.pa

The pattern in (103) is correctly ruled out by our representations. (103a) 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 (103a)), 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 (103b). This pattern is thus impossible under the 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.

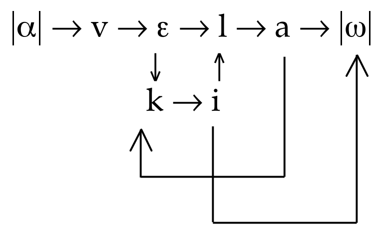
Finally, we hope to have shown that pretonic nasalization in Salvador BP is, at least for some speakers, the consequence of a rule-governed process, and not simply the result of memorizing static patterns in the lexicon. The fact that two invented language games reveal very different results for whether 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.na] happens to have a pretonic nasal vowel in its un-infixed and un-reduplicated form, it ain't necessarily so.

Appendix: Why not a Single /p/ or a Single /ki/

In this appendix we address the question of why the rules of LdP and LdK must involve a distinct infix for each syllable, by showing that an alternative, in which a single "p" or "ki" is simultaneously affixed to all syllables, yields an unlinearizable result[15].

(0) *Notation*
a: P(x,y) = x immediately precedes y
b: P'(x,y) = x transitively precedes y

(1) *Axiom of Choice of Path*
∀x,y,z, if the input structure contains P(x,y) & P(x,z), the system chooses to represent P(x,y) if and only if P'(y,z) & ¬ P'(z,y).

(2)


(3) a: *step 1*
The relation P(|α|,v) is read off the input structure and it is represented in the output structure being incrementally built.

$|\alpha| \rightarrow v$

b: *step 2*
The relation P(v,ε) is read off the input structure and it is represented in the output structure being incrementally built.

$|\alpha| \rightarrow v \rightarrow \varepsilon$

c: *step 3*
The relations P(ε,l) and P(ε,k) are both read off the input structure. At this point the system has to decide whether /l/ or /k/ will immediately follow /ε/ in the output. The *Axiom of Choice of Path* demands that the next segment be the candidate that transitively precedes the other. Notice, however, that both /k/ transitively precedes /l/ (cf. the sub-path /k→i→l/) and /l/ transitively precedes /k/ (cf. the sub-path /l→a→k/). Since the axiom presupposes that the transitive precedence be also asymmetric (i.e.

[15] We thank Klaus Abels for fruitful discussion of this possibility.

it does not allow /k/ and /l/ transitively preceding each other), the linearization algorithm gets stuck at /ε/ and the mapping procedure is aborted.

(4) $|\alpha| \rightarrow \mathrm{v} \rightarrow \varepsilon \rightarrow \mathrm{l} \rightarrow \mathrm{a} \rightarrow |\omega|$

p

(5) a: *step 1*

The relation P(|α|,v) is read off the input structure and it is represented in the output structure being incrementally built.

$|\alpha| \rightarrow \mathrm{v}$

b: *step 2*

The relation P(v,ε) is read off the input structure and it is represented in the output structure being incrementally built.

$|\alpha| \rightarrow \mathrm{v} \rightarrow \varepsilon$

c: *step 3*

The relations P(ε,p) and P(ε,l) are both read off the input structure. At this point the system has to decide whether /l/ or /p/ will immediately follow /ε/ in the output. The *Axiom of Choice of Path* demands that the next segment be the candidate that transitively precedes the other. Notice, however, that both /p/ trasitively precedes /l/ (cf. the sub-path /p→ε→l/) and /l/ trasitively precedes /p/ (cf. the sub-path /l→a→p/). Since the axiom presupposes that the transitive precedence be also asymmetric (i.e. it does not allow /k/ and /l/ transitively preceding each other), the linearization algorithm gets stuck at /ε/ and the mapping procedure is aborted.

References

Arregi, Karlos and Isabel Oltra-Massuet. 2005. Stress-by-Structure in Spanish. *Linguistic Inquiry* 36.1: 43-84.

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

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

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

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

Cedergren, Henrietta & David Sankoff. 1974. Variable rules: Performance as a Statistical Reflection of Competence. *Language* 50.2: 333-355.

Chomsky, Noam. 1975 [1955]. *The Logical Structure of Linguistic Theory*. Chicago: University of Chicago Press.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

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

Fitzpatrick, J. & A. Nevins 2002, Phonological Occurences: relations and copying. Cambridge: MIT, ms. Available at http://ling.auf.net/lingBuzz/000005

Fitzpatrick, J. & A. Nevins 2002, 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.

Guy, Gregory. 1981. Linguistic Variation in Brazilian Portuguese: Aspects of the Phonology, Syntax and Language History. Doctoral dissertation. University of Pennsylvania.

Halle, Morris & Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge, MA: MIT Press.

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

Harrison, David and Abigail Kaun. 2000. Pattern-Responsive Lexicon Optimization. In NELS 30. GLSA, University of Massachusetts, Amherst.

Iba, Aaron & Andrew Nevins. 2004. A Selectionist Learner for Parametrically Ambiguous Reduplicants. Paper presented at the LSA Meeting, Boston.

Idsardi, William & Eric Raimy. 2006. Remarks on Language Play. Ms., University of Maryland and Swarthmore College.

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

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

Manzini, Rita & Kenneth Wexler. 1987. Parameters, Binding Theory, and Learnability. *Linguistic Inquiry* 18.3:413-444.

Mattoso Câmara Jr., Joaquim. 1970. *Estrutura da Língua Portuguesa*. Petrópolis: Vozes.

McCarthy, John & Alan Prince. 1995. Faithfulness and Reduplicative Identity. In *University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory,* 249–384. GLSA, University of Massachusetts, Amherst.

Mester, R. Armin. 1986. *Studies in Tier Structure*. Doctoral dissertation,University of Massachusetts,Amherst.

Myers, Scott. 1991. Persistent Rules. *Linguistic Inquiry* 22: 315-344.

Nevins, Andrew. 2005. Overwriting does not Optimize in Nonconcatenative Word Formation. *Linguistic Inquiry* 36.2

Nevins, Andrew. 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.

Nobiling, Oskar. 1903. Die Nasalvokale in Portugiesischen. *Die Neuren Sprachen* 11: 129-153.

Ohala, Manjari. 1983. *Aspects of Hindi Phonology*. Delhi: Motilal Banarsidass.

Ohala, Manjari & John Ohala. 1991. Nasal epenthesis in Hindi. *Phonetica* 1991: 207-220.

Partee, Barbara H., Alice ter Meulen & Robert E. Wall. 1993. *Mathematical Methods in Linguistics*. Dordrecht: Kluwer Academic Publishers.

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

Raimy, Eric. 2000. Remarks on Backcopying. *Linguistic Inquiry* 31.3:541-552.

Raimy, Eric. 2000. *The Phonology and Morphology of Reduplication*. Berlin: Mouton de Gruyter.

Rubach, Jerzy. 1977. Nasalization in Polish. *Journal of Phonetics*, 5: 17-25.

Russell, Bertrand. 1905. On Denoting. *Mind* 14:479-493.

Russell, Bertrand. 1919. *Introduction to Mathematical Philosophy*. New York: The Macmillan Co.

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

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

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

Yu, Alan. 2003. The phonology and morphology of infixation. Doctoral Dissertation, University of California, Berkeley.

Authors' Contact information:
nevins –at- fas.harvard.edu
maxguimaraes –at ufpr.br