Choosing just the right amount of over-application in Texistepec Popoluca

Ehren M. Reilly[1]

Johns Hopkins University

The process of reduplication in Texistepec Popoluca (Zoquean) is over-applied to morpho-phonemic alternations that result from inflectional prefixation. The felicity of over-application varies depending on both the prefix and the phonology of the reduplicated stem. Furthermore, a survey of 15 speakers revealed that speakers vary greatly in terms of which phonological types of stems are subject to over-application of reduplication. Although there is variation between speakers, each individual in the survey consistently employed a single strict grammar. Interestingly, the set of attested grammars reveals a universal ranking among the Base-Reduplicant identity constraints that are responsible for over-application fall. Across all attested grammars, they adhere to a stringency hierarchy. The acquisition of this pattern of grammars, and the role of the stringency hierarchy are investigated using computational modeling.

1. Introduction

Texistepec Popoluca is a Zoquean language spoken in Texistepec, Veracruz, Mexico. As in several other Zoquean languages, certain agreement prefixes in this language fuse or coalesce with the stems to which they are attached, producing a wide range of regular morpho-phonemic alternations. These agreement prefixes are often used on reduplicated verb stems, and the morpho-phonemic alternations interact with the reduplication process in interesting ways. The morpho-phonemic processes that result from prefixation always apply to the first copy of the reduplicated verb stem, which is linearly adjacent to the prefix. Under certain conditions these process may over-apply to the second copy.

Primarily, this paper serves to present the results of a fairly thorough investigation of the specific conditions under which these morpho-phonemic alternations do or do not

[1] This work has benefited greatly from the input and tutelage of Adam Albright, Sara Finley, Paul Smolensky, Adam Wayment, and of course the awesome corps of HUMDRUM and HOWL participants. I am most deeply indebted to my collaborator, survey confederate, and friend Carmen Román Telesforo.

over-apply in reduplication. These results comprise a set of data which is quite complicated, at least on the surface. The felicity of over-application varies depending on both the prefix and the phonology of the stem to which it attaches. What is more, Texistepec Popoluca speakers differ in terms of the conditions under which they over-apply reduplication. While these three parameters allow for many potential generalizations about when to over-apply, given the morpho-phonemics of this language, the actual set of generalizations attested in my survey was small and systematically restricted. I argue that the set of grammars of over-application attested in my survey is governed by hierarchy of increasingly specific constraints, along two different dimensions. I adopt a conventional OT analysis of over-application, and then characterize the set of possible grammars in terms of a *specific* » *general* stringency hierarchy of Base-Reduplicant identity constraints.

Within the group of speakers surveyed, there was extensive variation between speakers, but, surprisingly, almost no variation in the performance of each individual speaker. In addition to describing the grammars known by the Texistepec Popoluca speakers I surveyed, I will discuss a learning problem that I believe these data present. I ask the question of how an otherwise homogeneous small community of speakers could have acquired so varied a set of individually rigid grammars. The conditions under which the current adult speakers learned their language were sub-optimal, due to declining language use and bilingualism with Spanish. I hypothesize that a community of speakers presented with a more sparse set of learning data would be more likely to learn this unusual distribution of grammars. In order to test this hypothesis, I conducted a number of learning simulations using the Gradual Learning Algorithm (Boersma and Hayes, 2001). The results of these simulations suggest that a community of learners presented with sparse data is more likely to learn a varied set of strict grammars, just like the community of Texistepec Popoluca described above. Given a sufficiently large and robust set of examples to learn from, speakers would probably have learned to match the overall statistical pattern of variation in their environment, but under the impoverished learning conditions each speaker adopted a single, invariant grammar, often different from their peers' grammars.

2. Texistepec Popoluca morpho-phonemics

Person agreement in Texistepec Popoluca is expressed by an ergative and an absolutive series of agreement markers.[2] The table in (1) presents these agreement series as two lists of affixes and clitics.

[2] In addition to transitive subjects, the "ergative" series marks possession on nouns, and agrees with the subjects of intransitive verbs only in the imperfective aspect. The "absolutive" series is used to agree with subjects of intransitive clauses (in the non-imperfective aspects), direct objects, and the subjects of copular constructions. See Reilly (submitted) for an alternative to the labels "ergative" and "absolutive".

(1) Texistepec Popoluca agreement series (Reilly, 2002, 2004)[3]

	Ergative	Absolutive
1st Exclusive	/n-/	/k+/
1st Plural Inclusive	/ta+n-/	/tɛ+/
2nd	/j-n-/	/k+j-/
3rd	/j-/	Ø

While the segments /k/, /j/ and /n/ each occur in several cells, assigning a unique set of features to each segment is not a simple matter. A more succinct but fairly un-intuitive way to think about the inflections in (1) is as in (2).

(2) Texistepec Popoluca agreement series, simplified

/n-/	= {1st,2nd} Ergative	/ta+ /	= Plural Inclusive (Ergative)
/k+/	= {1st,2nd} Absolutive	/tɛ+/	= Plural Inclusive (Absolutive)
/j-/	= 2nd	/j-/	= 3rd Ergative [2]

Because the absolutive /k+/ clitic and plural inclusive clitics are not involved in any interesting morpho-phonemic alternations, they will not be discussed here. So, the morpho-phonological segments of interest for the remainder of this paper will be the prefixes /n-/ and /j-/. The data in (3) reflect two types of alternations; those conditioned on the presence or lack of /j-/, and those conditioned on the presence or lack of /n-/.

(3) Allomorphs under four possible combinations of /n-/ and /j-/

Prefix	'to come'	'chicken'	'to fire'	'to pass'	'to say'	'to rise'
/Ø-/	mɛɲ	pɛjoʔ	tuh	hak	nɨm	kɛʔm
/n-/	mɛɲ	m͡bɛjoʔ	n͡duh	hãk	nɨm	ŋ͡gɛʔm
/j-/	biɲ	pijoʔ	t͡ʃuh	hjak	dʲim	kiʔm
/n-j-/	miɲ	m͡bijoʔ	n͡dʒuh	hj̃ãk	ɲim	ŋ͡giʔm
(UR)	/miɲ/	/pijoʔ/	/tuh/	/hak/	/nɨm/	/kiʔm/

I will refer to the alternations between forms in /n-/ and forms in Ø- as the "nasalization" alternations. I will refer to the alternations between forms in /j-/ and forms in Ø- as the "palatalization" alternations. The following two sections describe how the palatalization and nasalization alternations apply to various types of stems.

2.1. Nasalization and De-Nasalization

Under nasal prefixation, voiceless stops /p/, /t/, and /k/ become prenasalized stops /m͡b/, /n͡d/ and /ŋ͡g/, respectively (4a-c). Nasalization spreads across glides /w/ and /j/ to

[3] The symbol "-" indicates an affix-stem boundary, and "+" a boundary between clitic and host.

nasalize both the consonants and the following vowel (4d,e). Similar nasal spreading to the vowel occurs when the stem's initial consonant is a glottal /h/ or /ʔ/, though in this case the consonants do not become nasal (4f,g).[4] The nasal prefix is expressed on sibilants /s/ and /ʃ/ by voicing these segments rather than actually nasalizing them (4h,i). For reasons that will become apparent later on, this set of processes in (4) will be referred to as nasalization Type Y.

(4) Type Y Nasalizations

	Process		UR	→ Output	Gloss
a.	/n-**p**/	→ [**m͡b**]	/n-**p**ɛt/	→ [**m͡b**ɛt]	'I'm sweeping it'
			/nj-piŋ/	→ [**m͡b**iŋ]	'You're picking it up'
b.	/n-**t**/	→ [**n͡d**]	/n-**t**uh/	→ [**n͡d**uh]	'I'm firing it'
			/n**j**-**t**umeːɲ/	→ [**n͡dʒ**umeːɲ]	'Your money'
	/n-**t͡s**/	→ [**n͡dz**]	/n-**t͡s**eʔ/	→ [**n͡dz**ɛʔ]	'I'm washing it'
	/n-**t͡ʃ**/	→ [**n͡dʒ**]	/n-**t͡ʃ**uːʃ/	→ [**n͡dʒ**uːʃ]	'My jay (bird)'
c.	/n-**k**/	→ [**ŋ͡g**]	/n-**k**uʔt/	→ [**ŋ͡g**uʔt]	'I'm sucking it'
			/nj-**k**ɨʔ/	→ [**ŋ͡g**iʔ]	'Your hand'
d.	/n-**wV**/	→ [**w̃Ṽ**]	/nj-**wi**ʔk/	→ [**w̃ĩ**ʔk]	'You're eating'
e.	/n-**jV**/	→ [**j̃Ṽ**]	/n-**jo**s/	→ [**j̃õ**s]	'My work'
f.	/n-**hV**/	→ [h**Ṽ**]	/n**j**-h**a**k/	→ [h**j̃ã**k]	'You're passing it'
g.	/n-**ʔV**/	→ [ʔ**Ṽ**]	/n-ʔ**a**ʔm/	→ [ʔ**ã**ʔm]	'I'm watching it'
h.	/n-**s**/	→ [**z**]	/n-**s**ɨk/	→ [**z**ɨk]	'My bean'
			/nj-**s**os/	→ [**ʒ**os]	'You're cooking it'
i.	/n-**ʃ**/	→ [**ʒ**]	/n-**ʃ**iʃ/	→ [**ʒ**ɛʃ]	'My meat'
			/nj-**ʃ**iʃ/	→ [**ʒ**iʃ]	'Your meat'

Nasalized vowels and glides, prenasalized stops, and voiced sibilants occur only as a result of /n-/ prefixation.

---

[4] It is impossible to detect acoustically whether the velum is open, but there is some evidence that, at an abstract phonological level, these potentially nasalized glottals do not have the [+VOICE] feature that is implicit in nasality. The clitic /k+/ becomes [g] when it follows a vowel and precedes a voiced consonant (5a,b). This /k/-voicing appears before other morphophonemically nasalized segments but not /h/ (5c).

(5)	a.	/ʔu+k+ᴺ-sat͡s/	[ʔu**g**zat͡s]	'I'm scrubbing you.'
	b.	/ʔu+k+ᴺ-kap/	[ʔu**g**ŋ͡gap]	'I'm carrying you (on my shoulder).'
	c.	/ʔu+ k+ᴺ-hak/	[ʔu**k**hãk]	'I'm cutting you.'

Until a recent sound change, which was unique to Texistepec Popoluca, the pairs [b]~[m] and [d]~[n] were in systematic complementary distribution, as they still are in many related languages (Wichmaan, 1995). The standard Zoqueanist analysis posits a single phoneme /m/ and /n/ for each pair. Texistepec has a marginal oral vs. nasal contrast for these segments, but this contrast is neutralized for almost all members of the paradigm. Unsurprisingly, /b/, /d/ and /dʲ/ become nasal stops under /n-/ prefixation.

(6) Type X Nasalizations

	Process	UR	→ Output	Gloss
a.	/n-**b**/ → [**m**]	/n-**b**ɛt͡s/	→ [**m**ɛt͡s]	‘I am dancing’
		/nj-**b**awɛɲ/	→ [**m**jawɛɲ]	‘You are dreaming’
b.	/n-**d**/ → [**n**]	/n-**d**ɨk/	→ [**n**ɨk]	‘I am going’
		/nj-**d**aj/	→ [**ɲ**aj]	‘Your light’
c.	/n-**dʲ**/ → [**ɲ**]	/n-**dʲ**iːw /	→ [**ɲ**ɛːw]	‘My chile’
		/nj-**dʲ**oːn /	→ [**ɲ**oːn]	‘Your mushroom’

What is interesting is what happens when contrastively nasal segments are inflected with non-nasal prefixes or clitics: they “de-nasalize”, neutralizing their contrast with oral stops. Only in the un-inflected, un-prefixed form does the contrast surface. In (7), these uniquely contrastive forms are in bold. Notice that these are the only rows in which one can distinguish between the initial consonants of contrastive pairs at the same place of articulation. Forms in which the contrast in neutralized are indicated by a dashed line between the cells.

(7) Nasal/Oral contrast preserved only in un-prefixed forms

Prefix	a. ‘to moo’ /muh/	‘to wet’ /buh/	b. ‘to sprout’ /naj/	‘light’ /daj/
/Ø-/	**muh**	**buh**	**naj**	**daj**
/n-/	muh	muh	naj	naj
/n-j-/	mjuh	mjuh	ɲaj	ɲaj
/j-/	bjuh	bjuh	dʲaj	dʲaj
/k+/	kbuh	kbuh	kdaj	kdaj
IMPER **/-ɨʔ/**	**muːhɨʔ**	**buːhɨʔ**	**naːjɨʔ**	

What we see in Table 4 is that in the forms with only an imperative suffix and/or a null prefix, a contrast emerges between oral and nasal stops, one which does not occur elsewhere in the paradigm. Interestingly, /j-/ and even proclitics like /k+/ can effect de-nasalization on nasal-initial stems. One innovative member of the paradigm, the Ø-form, reflects the new contrast, but the other paradigm members have not changed.

2.2. Palatalization and De-Palatalization

The prefix /j-/ either effects a mutation on the stem or metathesizes into the stem. All stems in Texistepec Popoluca are CV(C).[5] There are no complex onsets, except in overtly inflected forms. On a stem whose initial consonant is not coronal, /j-/ metathesizes into the stem, as in (8). I will call this metathesis process Type D Palatalization.

(8) Type D Palatalization: /j-[–COR]/→ [Cj]

a.	/j-kuʔt/	→ [kjuʔt]	'He sucks it'
b.	/k+j-boŋ/	→ [kbjoŋ]	'You slept'

On stems with initial coronal consonants, initial stops and affricates become (or remain) palato-alveolar affricates. Initial fricatives become palato-alveolar. These will be referred to as Type C palatalizations in (9).

(9) Type C Palatalizations: /j-[+COR]/→ [Č]

a.	/j-sa͡ts/	→ [ʃa͡ts]	'She scrubs it'
b.	/k+j-doʔ/	→ [kdʲoʔ]	'It burned you'
c.	/n-j-naj/	→ [ɲaj]	'You're sprouting'
d.	/j-t͡ʃuːʃ/	→ [t͡ʃuːʃ]	'His jay'
e.	/j-t͡seʔ/	→ [t͡ʃeʔ]	'She washes it'

There is also a vowel fronting process affecting /ɨ/ under /j-/ prefixation. When the vowel /ɨ/ appears after a coronal consonant, this process occurs in combination with coronal place assimilation (Type C). This vowel process will be called Type B.

(10) Type B Palatalization: /j-Cɨ /→ [Ci]

a.	/j-tɨp/	→ [tʃip]	'He spear-fishes'
b.	/k+j-wɨʔt/	→ [kwiʔt]	'She strapped you down'
c.	/n-j-nɨm/	→ [ɲim]	'You speak'

An unusual alternation, which I will call Type A, affects the vowel /i/. This process changes underlying /i/ to [ɛ], as in (11a-d). This "de-palatalization" occurs in all contexts *except* when the prefix /j-/ is present (11e-g). This prefix has the effect of preserving a contrast between underlying /i/ and /ɛ/ that is otherwise neutralized.

(11) Type A Palatalization: /i/→[e] except with /j-/ prefix

a.	/kij/	→ [kɛj]	'He signaled'	(/Ø-/ prefix only)
b.	/k-piŋ/	→ [kpɛŋ]	'It picked me up'	(/k+/ clitic only)
c.	/n-wiʔk/	→ [w̃ɛ̃ʔk]	'I am eating'	(/n-/ prefix only)
d.	/j-ʔaʔm-kiʔm/	→ [ʔjaʔm-kɛʔm]	'It looks upwards'	(/i/ in postpound)

[5] Also, a nuclear glottal stop may follow the vowel, and certain codas permit a final [s] or [ʃ].

e. /**j**-kij/ → [k**i**j] 'He signals'
f. /k-**j**-piŋ/ → [kp**i**ŋ] 'It picked you up'
g. /n-**j**-wiʔk/ → [w̃**ĩ**ʔk] 'He is eating'

Wichmann (1994) and Reilly (2002) analyze the [i]~[ɛ] alternation in terms of underspecification. Several colleagues have suggested that this alternation might also be explained in terms of morphologically controlled constraint ranking, à la Anttila (2002), but a problem for such an account is the fact that underlying /i/ surfaces as [ɛ] in the second syllable of words prefixed with /j-/, such as (11d). No effort will be made in this paper to develop a proper analysis of any of the morpho-phonemic alternations. Of interest for the remainder of the paper will be the ways in which the mutations resulting from these processes are over-applied in reduplication. The six types of morpho-phonological processes described in this section are summarized in (12).

(12) Morpho-phonological processes X-Y and A-D

NASALIZATION		PALATALIZATION	
Process	Description	Process	Description
X	/n-b/ → [m]	**A**	/j-Ci/ → [Ci], /Ci/ → [Cɛ]
	/n-d/ → [n]		
	/(⌐Ø,⌐n)-m/ → [b]	**B**	/j-Cɨ/ → [Ci]
	/(⌐Ø,⌐n)-n/ → [d]		
Y	/^{N}-p/ → [m͡b]	**C**	/j-s/→ [ʃ]
	/^{N}-t/ → [n͡d]		/j-t/→ [t͡ʃ]
	/^{N}-k/ → [n͡g]		/j-d/→ [d^{j}]
	/^{N}-{ʔ,h}V/ → [{ʔ,h}Ṽ]		/j-j/ → [d^{j}]
	/^{N}-{w,j}V/ → [{w̃,j̃}Ṽ]		/j-n/→ [ɲ]
	/^{N}-{s,ʃ}/ → [{z, ʒ}]		/j-t͡s/→ [t͡ʃ]
		D	/j-[–COR]/→ [Cj]

3. Over-application of reduplication

Texistepec Popoluca has "full" reduplication, meaning that the reduplicant reflects the entire base, even when the base is a non-canonical or multi-syllabic stem. This may seem a simple enough formula, except that there are two competing notions of the *base* for reduplication. Is it the underlying form of the stem that is reduplicated or the often morpho-phonemically altered surface form? If the reduplicant is faithful to the base's underlying stem, it may be quite distinct from the surface form of the base, making it difficult to identify the form as reduplicated (13a). However, if the form is faithful to the (potentially mutated) surface form of the base, then it risks copying segments or mutations that belong to a separate morpheme from the one being reduplicated (13b). In the latter situation, the reduplication process is said to "over-apply" to this additional material. Texistepec Popoluca reduplication often over-applies to the agreement prefixes.

In the standard OT analysis of reduplication, a family of correspondence constraints known as BASE-REDUPLICANT IDENTITY (BR-ID) require the copy to be identical to the surface form of the base, while Input-Output Faithfulness constraints (IO-F) prohibit the reduplicant from deviating from the underlying form of the reduplicated morpheme (McCarthy and Prince, 1995). Tableau (13) illustrates the violation profile of two possible outputs in Texistepec Popoluca. If BR-ID dominates, the reduplicant will match the surface form of the base. If IO-F dominates, the reduplicant will match the input form of the base.

(13) BASE-REDUPLICANT IDENTITY and INPUT-OUTPUT FAITHFULNESS

'You go pecking all around' /n-j-<u>tɨʔks</u>-RED-hoʔj/	BASE-RED IDENT	I-O FAITH
a. n͡dʒiʔks-tɨʔks-hoʔj	*	
b. n͡dʒiʔks-n͡dʒiʔks-hoʔj		*

3.1. Variation in overapplication: Who over-applies and when?

The only two prefixes that are subject to over-application are /j-/ and /n-/, and the felicity of overapplication varies depending on the prefix and the phonological context. Certain morpho-phonemic mutations are preserved in the reduplicant, and others are not. Of great interest is the fact that the set of mutations that are over-applied varies greatly between speakers. This section will describe the variation, and present the results of a survey of Texistepec Popoluca speakers' pronunciations of a set of reduplicated and inflected words.

In §2, I divided the possible morpho-phonemic alternations into six categories: nasalizations X and Y, and palatalizations A, B, C and D. These distinctions reflect the variation among Texistepec Popoluca speakers in the over-application of reduplication. X is distinguished from Y, because many speakers overapply the processes in X, but not those in Y. The constraint ranking for such a speaker is illustrated in (14) and (15).[6]

(14) Constraint ranking for a speaker who overapplies X but not Y

'I go pecking all around' /n-<u>tɨʔks</u>-RED-hoʔj/	BASE-RED IDENT (X)	I-O FAITH	BASE-RED IDENT (Y)
☞ a. n͡dɨʔks-tɨʔks-hoʔj		✓	*
b. n͡dɨʔks-n͡dɨʔks-hoʔj		*!	✓

[6] I use the symbol "✓" to indicate the active—rather than vacuous—satisfaction of a constraint.

(15) Constraint ranking for a speaker who overapplies X but not Y

‘I go bounding all around’ /n-bɨʔm-RED-hoʔj/	BASE-RED IDENT (X)	I-O FAITH	BASE-RED IDENT (Y)
a. mɨʔm-bɨʔm-hoʔj	*!	✓	
☞ b. mɨʔm-mɨʔm-hoʔj	✓	*	

Similar distinctions motivate the divisions between A, B, C and D. The over-application of palatalization is essentially independent from the over-application of nasalization, and there may be differences between speakers on neither, one, or both of these parameters.[7] The ranking in (16) and (17) enforces the over-application of C, but not D, as well as both X and Y.

(16) Ranking for a speaker who overapplies C, X, and Y, but not D

‘You’re beating it over and over’ /n-j-<u>daks</u>-RED-dɛʔ/	B-R ID (X)	B-R ID (C)	B-R ID (Y)	I-O FAITH	B-R ID (D)
a. ɲaks-daks-dɛʔ	*!	*!		✓	
b. ɲaks-dʲaks-dɛʔ	*!	✓		*	
c. ɲaks-naks-dɛʔ	✓	*!		*	
☞ d. ɲaks-ɲaks-dɛʔ	✓	✓		**	

Since this stem is a candidate for both X and C, and the BR-ID constraints favoring X and C are ranked above IO-F, the winning candidate is the one that overapplies both the X and C mutations. In (17) the stem is a candidate for Y and D.

(17) Ranking for a speaker who overapplies C, X, and Y, but not D

‘You’re gulping over and over’ /n-j-<u>kuʔt</u>-RED-dɛʔ/	B-R ID (X)	B-R ID (C)	B-R ID (Y)	I-O FAITH	B-R ID (D)
a. n͡gjuʔt-kuʔt-dɛʔ			*!	✓	*
b. n͡gjuʔt-kjuʔt-dɛʔ			*!	*	✓
☞ c. n͡gjuʔt-n͡guʔt-dɛʔ			✓	*	*
d. n͡gjuʔt-n͡gjuʔt-dɛʔ			✓	**!	✓

If the BR-ID constraint favoring the over-application of a particular mutation is ranked above IO-F, then that mutation is preserved in the reduplicant. Otherwise it is not, and the reduplicant is faithful to the underlying form of the stem. Given this simple generalization, and given that there are six BR-ID constraints, which may each be ranked

[7] Statistically speaking, these scales are not “independent”. However, the interaction is a small, non-absolute, numerical effect—not fully implicational or categorical. (See fn.8).

either above or below IO-F, we predict a typology of $2^6 = 64$ possible rankings. Having identified the existence of considerable inter-speaker variation in the over-application pattern, we can ask how widespread the variation is, and which of these 64 possible rankings are actually possible or attested.

To answer this question, I conducted a survey of 15 Texistepec Popoluca speakers. With the help of a Popoluca-speaking confederate, I had each participant in the survey pronounce 29 different words, which were designed to reveal which processes were being over-applied. The items were designed such that the confederate could ask a question using a non-reduplicated, non-overtly-inflected form of the stem, and through circumlocution encourage the participant to produce the reduplicated form. An average of only 3.4 items per speaker out of 29 (11.7%) did not yield a codable response.

The results of this survey were quite surprising. Each speaker's responses were almost always consistent with a single constraint ranking—there was vanishingly little free variation within each speaker. However, among the 15 speakers surveyed, nine different invariant grammars were attested. These grammars fell into a highly constrained distribution, illustrated in (18).

(18) Nine strict rankings attested in survey

Processes over-applied	Ranking that yields that pattern of outputs
Ø$^{(1\text{ speaker})}$	**IO-F** » {BR-ID$_X$, BR-ID$_Y$, BR-ID$_A$, BR-ID$_B$, BR-ID$_C$, BR-ID$_D$}
X$^{(1)}$	BR-ID$_X$ » **IO-F** » {BR-ID$_Y$, BR-ID$_A$, BR-ID$_B$, BR-ID$_C$, BR-ID$_D$}
A$^{(1)}$	BR-ID$_A$ » **IO-F** » {BR-ID$_X$, BR-ID$_Y$, BR-ID$_B$, BR-ID$_C$, BR-ID$_D$}
XA$^{(2)}$	{BR-ID$_X$, BR-ID$_A$} » **IO-F** » {BR-ID$_Y$, BR-ID$_B$, BR-ID$_C$, BR-ID$_D$}
XYA$^{(2)}$	{BR-ID$_X$, BR-ID$_Y$, BR-ID$_A$} » **IO-F** » {BR-ID$_B$, BR-ID$_C$, BR-ID$_D$}
XAB$^{(3)}$	{BR-ID$_X$, BR-ID$_A$, BR-ID$_B$} » **IO-F** » {BR-ID$_Y$, BR-ID$_C$, BR-ID$_D$}
XYAB$^{(2)}$	{BR-ID$_X$, BR-ID$_Y$, BR-ID$_A$, BR-ID$_B$} » **IO-F** » {BR-ID$_C$, BR-ID$_D$}
XYABC$^{(2)}$	{BR-ID$_X$, BR-ID$_Y$, BR-ID$_A$, BR-ID$_B$, BR-ID$_C$} » **IO-F** » BR-ID$_D$
XYABCD$^{(1)}$	{BR-ID$_X$, BR-ID$_Y$, BR-ID$_A$, BR-ID$_B$, BR-ID$_C$, BR-ID$_D$} » **IO-F**

3.3. Base-Reduplicant Identity constraints along a stringency hierarchy

Is there a reason why this survey found these nine grammars and not others? If we were to treat each of the palatalization constraints as being independent from the other palatalization constraints, and similarly for the nasalization constraints, this pattern of results would be a remarkable coincidence. Every speaker who over-applies Y also over-applies X. A similar scale of implication holds across D→C→B→A. This is to say that while the ranking of IO-F ranges across the entire spectrum, there is a universal ranking among the palatalization constraints and among the nasalization constraints. If we take the expression $\alpha \underline{\gg} \beta$ to mean "a constraint α either dominates or shares a stratum with β", then all of grammars attested in the survey adhere to the generalizations in (19).[8]

(19) a. BR-ID$_X$ $\underline{\gg}$ BR-ID$_Y$

b. BR-ID$_A$ $\underline{\gg}$ BR-ID$_B$ $\underline{\gg}$ BR-ID$_C$ $\underline{\gg}$ BR-ID$_D$

Is there any reason why a set of constraints ought to be subject to such restrictions as (19)? I argue that the best explanation for this pattern is that the BR-ID constraints are ranked along a stringency hierarchy similar to the sort proposed by de Lacy (2002). More specific constraints always dominate or share a stratum with more general ones. This will require us to assert that X is more general than Y, and that A is more general than B, and so on. To see that this assertion is not as stipulative as it seems will require us to further formalize our definitions of the BR-ID constraints, which until now have carried the arbitrary labels X-Y and A-D. A less arbitrary treatment of these constraints is given in (20), where the BR-ID constraint for nasalization and for palatalization each lie on a specificity continuum.

(20)

	Nasalization Processes		Palatalization Processes	
X.	BR-ID$_{[b,d,m,n]}$	A.	BR-ID$_{[i]}$	↑ More specific
Y.	BR-ID	B.	BR-ID$_{[-CONSONANTAL]}$	
		C.	BR-ID$_{[+CORONAL, \pm CONSONANTAL]}$	
		D.	BR-ID$_{[\pm CORONAL, \pm CONSONANTAL]}$	↓ More general

Supposing that the set of possible rankings of the constraints governing each of these processes is restricted by specificity along the scale in (20), the set of grammars found in my survey are probably a fairly representative sample. It is possible that some other speakers make different distinctions along one of these scales. Type Y, for example, would permit a variety of finer-grained distinctions. However, I predict that even if a speaker were to permit over-application of only a specific subset of Y, they would necessarily over-apply all processes in X.

4. Learning a grammar of reduplication

It is surprising to find speakers who vary so much from one another, but who were quite consistent about their individual generalizations. After the initial survey, I attempted to collect speakers' judgments about the felicity of alternative pronunciations. However, only one Texistepec Popoluca speaker whom I worked with, my consultant and confederate in the survey, was ever able to be made aware of his own pronunciation of the reduplicant. For example, my confederate and I spent a good deal of time trying to

[8] These generalizations also predict the possibility of grammars that over-apply XABCD, XABC, ABCD, ABC, AB, and XY respectively. What is interesting about these is that they are the least "balanced" between palatalization and nasalization, being lenient with one and restrictive with the other. Due to this strong interaction, the scales are not statistically "independent".

help a husband and wife notice that while he was very stingy about overapplication (X only), she was much more lenient (XYAB). A similar situation was found between two cohabitating widower brothers who over-applied XA and XYABCD, respectively.

It is easy to describe the systematicity of the inter-speaker variation in terms of a stringency hierarchy, but the issue of why the modern community of speakers learned such a diverse set of grammars is a somewhat different matter. If, as I assume, two brothers close in age or two neighbor children are exposed to roughly the same set of examples from which to learn their grammars, the current situation in Texistepec reflects the acquisition of different grammars by different speakers, given essentially the same input. How did this come to pass?

4.1. The learning problem

I believe a number of factors may have contributed to this phenomenon. Internal to the language, it should be noted that the reduplicants in question never bear stress, which is always born by the first copy in reduplicated structures. (In the rare case that the base for reduplication is disyllabic, the first syllable of the reduplicant bears secondary stress). Also, though reduplication is a fairly frequent construction, the vast majority of reduplication in texts applies to members of a distinct lexical category of ideophonic roots, which are typically used to describe sounds, movements, and perceptual effects of animals, body parts and inanimate objects. As a result, reduplicated stems almost always have 3rd person agreement morphology, limiting the set of forms available as input to predominantly Ø- and /j-/ prefixed forms.

Over-application of reduplication is, in a certain sense, a morphologically gratuitous process, since it does not convey additional information. The benefit of overapplication lies in the way it facilitates the identification of a word as reduplicated, as opposed to a compound. If the base is very different from the reduplicant on the surface, it may not be obvious that they are two copies of the same stem. In Texistepec Popoluca, the outputs of the processes in X, A, B, and marginally C would themselves be well-formed stems, while the outputs of Y and D would not be. The viability of both the underlying and surface forms of the base as stems presents a particularly difficult challenge to the hearer, who must consider equally plausible analyses in which the reduplicant is a copy of the base or, alternatively, a postpound. Consider the form in (21), which has two possible morphologically analyses, given the ranking IO-F » BR-ID_{B}. In analysis (a) *pɨŋ* is a postpound on *piŋ*, while in analysis (b) *pɨŋ* is a reduplicant of the stem, faithful to the stem's UR.

(21) A single word-form rendered ambiguous by optional over-application

a.	piŋ-<u>pɨŋ</u>-hoʔj /j-piŋ-<u>pɨŋ</u>-hoʔj/ 3ERG-*gather*-*<u>fall</u>*-AMBLTV	b.	piŋ-<u>pɨŋ</u>-hoʔj /j-pɨŋ-<u>RED</u>-hoʔj/ 3ERG-*gather*-<u>RED</u>-AMBLTV

For sentences like (21b), it may have been functionally beneficial to the hearer to have heard two identical copies of *piŋ*. However, in most cases, failure to over-apply does not introduce any ambiguity. The choice between *n͡gjuʔt-**kuʔt**-dɛʔ* and *n͡gjuʔt-**n͡gjuʔt**-dɛʔ* has very little functional significance, and the alternation between these is probably less salient for Texistepec Popoluca learners than alternations that uniquely carry crucial information. These factors alone do not explain anything, but they likely enhance the effects of the extra-linguistic obstacles to the acquisition of reduplication.

A number of extra-linguistic factors have clearly had a hand in creating the unusual distribution of grammars in Texistepec. This language is moribund, spoken by only about 150 surviving speakers in a town of several thousand citizens. The remaining speakers are all in their late fifties or older. The dominant language in the town of Texistepec is Spanish, and all but the oldest surviving speakers learned Popoluca concurrently with their acquisition of Spanish. It is therefore likely that speakers' being entrenched in their different rigid patterns of reduplication is a result of the slightly impoverished learning environment, and the fact that the frequency with which they use Popoluca has decreased gradually over the course of their lifetimes.

Under normal circumstances, we would expect that speakers in a fairly homogeneous learning environment would learn whatever patterns are present in that environment. If their input is consistent with a single, rigid grammar, they learn that grammar. If there is variation, they are likely to adopt such variation. In the case of Texistepec Popoluca, this environment would contain considerable variation in the extent of over-application. I believe that the failure of Texistepec Popoluca speakers to match the statistical patterns of variation in the reduplication data that they have heard over the course of their lives can be explained in terms of a combination of the factors discussed above. I hypothesize that a community of speakers presented with fewer, less salient data, spread more thinly over the time course of acquisition, will be more likely to learn a variety of distinct, rigid grammars. In contrast, a community with a richer learning environment will probably converge on similar grammars, each matching the overall variation that is present in their communal linguistic experience. Such speakers would be less likely to each over-generalize to different invariant grammars.

4.2. Modeling variation in the acquired grammars

This hypothesis is very much in agreement with results of evolutionary dynamical system modeling work of language change. Niyogi and Berwick (1997) define a language community as a triple ($\mathcal{G}$, $\mathcal{P}$, $\mathcal{A}$), where $\mathcal{G}$ is a grammar formalism (i.e., a space of possible grammars), $\mathcal{A}$ is a non-deterministic learning algorithm for grammars in $\mathcal{G}$, and $\mathcal{P}$ is a set of primary linguistic data. Assuming the primary linguistic data come from a "parent" generation g_n, the grammars learned by the subsequent generation g_{n+1} are the results of a function $\mathcal{A}(\mathcal{P}_n)$. They show that if g_n is perfectly homogeneous and $\mathcal{P}_n$ is infinitely large, all speakers in g_{n+1} will converge on the same target grammar, identical to g_n. However, as the number of relevant data in $\mathcal{P}_n$ decreases, the number of speakers in g_{n+1} who misconverge increases. That is, as the same algorithm is applied to an increasingly small set of data, it becomes more likely that the algorithm will yield a heterogeneous set of results. Noise is similarly a possible source of misconvergence and heterogeneity in a linguistic dynamical system. These effects are, in principle, true for any ($\mathcal{G}$, $\mathcal{P}$, $\mathcal{A}$), but see Niyogi and Berwick (1997) for additional discussion.

Another possibility, not considered by Niyogi and Berwick (1997) is that if $\mathcal{G}$ allows for individual speakers' grammars to be probabilistic and have within-speaker variation, then speakers exposed to a sparse or noisy set of data $\mathcal{P}$ might each learn the same probabilistic grammar, one which matches the statistical distribution of tokens in $\mathcal{P}$. It is striking the extent to which this did not happen among the Texistepec Popoluca speakers in my survey. Can the small set of data explain for speakers' acquisition of different invariant grammars as opposed to the same variant grammar, as I hypothesized above?

This is any empirical question, which I tested using a modeling approach similar to but simpler than that of Niyogi and Berwick (1997). To test my hypothesis, I conducted a simulation using the Gradual Learning Algorithm (GLA) for OT grammars (Boersma and Hayes, 2001). This is a computational learning algorithm designed to simulate the acquisition of subtle statistical patterns of free variation. My goal was to simulate in the GLA the special conditions that affected the acquisition of Texistepec Popoluca by the speakers in my survey, and see if different iterations of the GLA could be made to learn different strict grammars given the same varied input. Essentially, I conducted a test to see if a hypothetical new generation of speakers exposed to the disparate data produced by the current one would, as a community, become fixed on a variety of invariant grammars (a pattern like that of the current speakers), or if they would all learn roughly the same grammar with free variation.

A representative corpus of underlying words was constructed, each prefixed with /n-j-/, along with each of the four possible surface forms of that word, corresponding to each of /n-/ and /j-/ being optionally over-applied. In order to determine the relative frequencies for each of these surface forms, I sampled the forms that would be used by each of the 15 speakers in my survey. For example, since only one of the 15 speakers was found to over-apply XYABCD, the frequency for a form with over-application of D and Y was 6.67%. The competing variant with D over-applied, but not Y, would have a 0% frequency, since no speakers overapply only XABCD. Two sample data given to the GLA are shown in (22), formatted according to Hayes et al. (2003).

(22) Sample data given to GLA: ('You are passing back and forth', 'You are mewing')

Input	RED Output	% freq.	IO-F	X	Y	A	B	C	D
nj-hak-<u>RED</u>-dɛʔ	hak	53.33%			*				*
	hãk	40%	*						*
	hjak	0%	*		*				
	hj̃ãk	6.67%	**						

nj-ɲiːw-RED-dɛʔ	dʲɛːw	6.67%		*	*	*	*	*	*
	ɲɛːw	6.67%	*			*	*	*	*
	dʲiːw	6.67%	*	*	*				
	ɲiːw	80%	**						

Not surprisingly, the GLA run with a large number of epochs always learns a grammar with variation, given a corpus of input like this. The grammar that is learned tends to match the statistical patterns of the overall corpus of data very well. The resulting ranking values tend to end up along a gradual continuum such as (23).

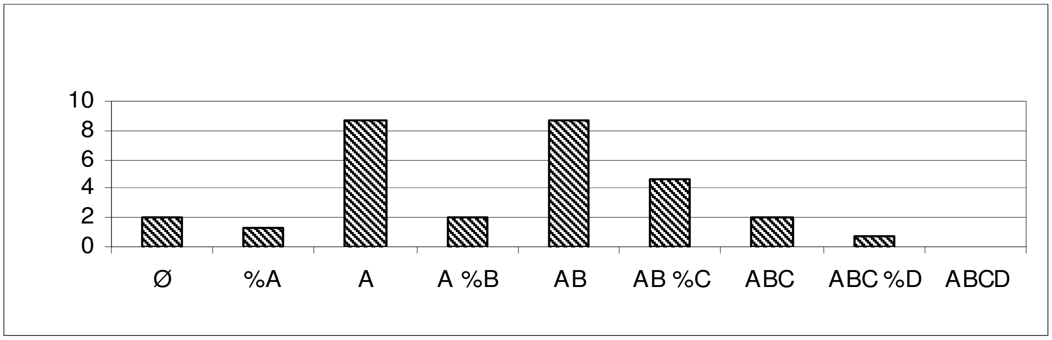
(23) Constraint: BR-ID$_X$ » BR-ID$_Y$ » BR-ID$_A$ » IO-F » BR-ID$_B$ » BR-ID$_C$ » BR-ID$_D$
Ranking Value: 90.368 89.823 89.219 88.402 88.380 84.218 72.802

Due to the proximity of these constraints, their rankings varied with each invocation of the grammar, and so the forms produced by this grammar contain free variation. This is not the sort of grammar that any current Texistepec Popoluca speaker speaks, individually. The grammar in (23) reflects a sort of "average" grammar over the community of speakers. This is presumably what a Popoluca-learning child would learn under normal conditions, since her experience would be an amalgam of different speakers' speech. In order to model the conditions under which the current generation learned, the GLA's parameters must be modified.

One simple modification to the GLA is to change the learning schedule and the rate at which plasticity decrements. I modified the learning schedule so that plasticity decayed exponentially with respect to the time course of the presentation of the leaning data. This had the effect of dramatically reducing the impact of each new block of data on the learner's grammar. Data presented earlier did much more to shape the grammar than data presented later. Since the GLA presents the data in a random order, this meant that different data occurred during the crucial early portion of the data set for different iterations of the algorithm. This modification was meant to simulate the sparseness of the relevant data, and the gradual decline of the linguistic community over the lifetimes of these speakers. The period of time during which plasticity declines and the grammar becomes fixed was accelerated, in order to model the fact that, during the childhoods of my speakers, the crucial data came too slowly to keep up with the natural process by which the grammar becomes less flexible.

This perturbation of the GLA did, in fact, achieve the hypothesized result. A GLA simulation that receives less data as its plasticity decays is more likely to get stuck in a local error minimum. However, since the order of presentation is random, the local error minimum where the GLA gets stuck can be different each time the algorithm is run. The chart in (25) depicts the results of running the GLA 40 times on palatalization data, with an initial plasticity of .5, decaying exponentially with each epoch, over 500 epochs. Each column lists the processes that overapply for that grammar, with processes that apply only sometimes indicated by a '%'. Unlike the grammar learned in (23), most of these grammar do not contain free variation (on the char, they do not bear a '%').

(24) Number of GLA Grammars out of 40 that over-apply each process



As this chart shows, the most prominent generalizations are not ones with variation, but A and AB, without variation. The input to the learner was consistently varied, but of this generation of 40 learners exposed to this input, 22 acquired a grammar with no variation. The impact of early experience strongly biased one ranking, and exposure to the full range of variation came too late to counteract this effect.

4.3. The role of the stringency relation in acquisition

The above simulations all employed an *a priori* ranking among the constraints, such that the stringency hierarchy was enforced. A pattern similar to that of the attested set of grammars emerged from these simulations. However, simulations conducted without imposing this *a priori* ranking invariably failed to yield these same results. In these simulations, the GLA learned grammars with internal variation, and these grammars often resulted in constraint rankings on some invocations that violated the stringency hierarchy.

A likely explanation for this result is that the stringency hierarchy is inherent in the content of the constraints, as per de Lacy (2002). Without this stringency hierarchy, it is not possible to learn the set of grammars that the surveyed speakers know. The set of grammars that are possible without the stringency ranking likely includes impossible grammars. An inaccurate characterization of the grammar space may have given the GLA too hard a task to accomplish. This result can be viewed as confirmation that the stringency hierarchy attested in the survey is not accidental. The acquisition of a set of grammars adhering to this hierarchy is actually quite unlikely unless the universal ranking is given.

5. Conclusion

This paper has served two purposes. First and foremost, I have sought to thoroughly describe the morpho-phonology of agreement prefixes and reduplication in Texistepec Popoluca. I presented the results of a 15-speaker survey of over-application tendencies. In addition, I have considered the results of this survey from the perspective of language change and morpho-phonological learning. The distribution of the grammars of

Texistepec Popoluca speakers poses some unique questions about learning under sub-optimal conditions. I argued that language death and the low salience of reduplicants in Texistepec Popoluca helped give rise to this unusual distribution of grammars. My simulation of this learning problem is fairly cursory, but the data presented here raise a number of important questions about phonological learning: Are some phonological processes less salient than others? What causes learners to sometimes learn the variation in their input, and at other times to regularize free alternations? Is it possible for different individuals to learn different grammars from the same input? The process of reduplication in Texistepec Popoluca raises these and a number of other questions, which may be addressed in future research.

References

Anttila, Arto. 2002. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* (20):1-42.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* (32):45-86.

de Lacy, Paul. 2002. The Formal Expression of Markedness. Doctoral dissertation. University of Massachusetts, Amherst.

Hayes, Bruce, Bruce Tesar, and Kie Zuraw. 2003. OTSoft 2.1, software package.

McCarthy, John J., and Prince, Alan. 1995. Faithfulness and Reduplicative Identity. In *UMOP 18*, eds. Jill Beckman, Suzanne Urbanczyk and Laura Walsh Dickey, 249-384. Graduate Linguistics Student Association: Amherst, Mass.

Niyogi, Partha and Robert Berwick. 1997. A dynamical systems model for language change. *Complex Systems* (11), pp. 161-204

Reilly, Ehren. 2002. A Survery of Texistepec Popoluca Verbal Morphology. Undergraduate thesis, Carleton College, Northfield, MN.

Reilly, Ehren. 2004. Ergativity and Agreement Splits at the Syntax/Phonology Interface. M.A. thesis. Johns Hopkins University.

Reilly, Ehren. (Submitted for publication). Morphological and phonological sources of split ergative agreement.

Reilly, Ehren, Terrence Kaufman, and Kathy Bereznak. (In preparation). Texistepec Popoluca Lexical Database. (6,500 items). Johns Hopkins University and the University of Pittsburgh.

Wichmaan, Søren. 1995. *The Mixe-Zoquean Languages of Mexico*. Salt Lake City: University of Utah Press.

Wichmann, Søren. 1994. Underspecification in Texistepec Popoluca phonology. *Acta Linguistica Hafniensia* (27):267-285.

Department of Cognitive Science
Johns Hopkins University
3400 N. Charles St.
Baltimore, MD 21218

reilly@cogsci.jhu.edu