Phonological Variation and Its Interfaces
University of Barcelona
2018/11/23

Backcopying driven by Surface Correspondence

Yifan Yang
University of Southern California
yangyifa@usc.edu

1. Introduction

(1) *Backcopying*: the case of reduplication where BASE conforms to REDUPLICANT rather than the reverse. Backcopying is a special prediction in classic Base-Reduplicant Correspondence Theory (henceforth ‘BRCT’, McCarthy and Prince 1995)

(2) A typical example of backcopying: Tagalog nasal substitution (McCarthy and Prince 1999:239)

Underlying	*Surface*
/paŋ-RED-pu:tul/	[pa-**m**u-**m**u:tul]

(3) The crucial mechanisms in BRCT that lead to backcopying: ***BR-Correspondence*** and ***parallelism***

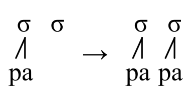
/paŋ-RED-pu:tul/	PHONO-CONSTRAINT	BR-IDENT	IO-FAITH
a. pam-pu-pu:tul	*! (mp)		
b. ☞ pa-**m**u-**m**u:tul			*
c. pa-**m**u-**p**u:tul		*!	

(McCarthy and Prince 1999:252)

(4) However, many subsequent theories of reduplication view backcopying as peculiar, and therefore, these theories do not equip with the mechanism that can predict backcopying. One such theory is Generalized Non-linear Affixation (GNA) (Bermúdez-Otero 2012).

(5) Generalized Non-linear Affixation (GNA) is a general morphological approach which proposes that all kinds of morphological processes can be viewed as the concatenation of nonlinear phonological representations. (Bermúdez-Otero 2012; see also Davis and Ueda 2002; Saba Kirchner 2010, 2013; Bye and Svenonius 2012; Trommer and Zimmermann 2014; Zimmermann 2015)

(6) In GNA, when a stem is affixed with an empty prosodic template, one possible outcome is reduplication:



- No RED morpheme; no BR-Correspondence assumed
- Not able to predict backcopying

(7) Two issues:
(a) Are there any robust cases of backcopying?
(b) Should a theory of reduplication predict backcopying and how?

(8) This work presents the data of Rapa Nui, an Austronesian language spoken on Easter Island (Chile), to shed light on the above issues.

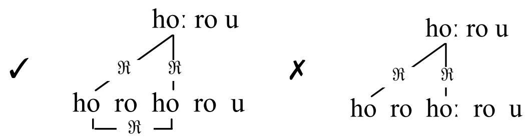
(9) A sketch of Rapa Nui reduplication (Kieviet 2017) (parentheses indicate foot boundary; moraic trochee)

	Base	***Reduplicated***		***Gloss of base***
		i. Left-edge copy	***ii. Right-edge copy***	
a.	ma.na.ʔu L L L	ma.(ˌna.ma).(ˈna.ʔu) (ˌma.na).ma.(ˈna.ʔu)	(ˌma:).(ˌna.ʔu).(ˈna.ʔu)	‘to think’
b.	ho:.ro.u H L L	**ho**.(ˌro.**ho**).(ˈro.u) (ˌ**ho**.ro).**ho**.(ˈro.u)	(ˌho:).(ˌro.u).(ˈro.u)	‘quick’
		ho ~ ho; syllable weight matches (*suspected backcopying*)		

(10)Claims and goals:

(a) The 'suspected backcopying' is not epiphenomenal (not due to metrical optimization)

(b) There is correspondence relation between [ho]~[ho] in the surface, and they are required to be identical:



(c) The patterns shown in Rapa Nui favors Generalized Non-linear Affixation (GNA) over BRCT.

(d) Though there is no BR-Correspondence in GNA, the general Surface Correspondence (Hansson 2001; Rose and Walker 2004; Bennett 2013, 2015) exists in reduplication, which can explain the patterns observed in Rapa Nui.

Goals

- present new interpretation of recently-published reduplication data;
- use MaxEnt HG to model the variations and therefore show the advantage of Surface Correspondence relationship in the analysis of reduplication

(11)Data Sources:

(a) *A Grammar of Rapa Nui* (Kieviet, 2017)

(b) *Programa Lengua Rapa Nui* Text Corpus (PLRN Corpus) (around 538,000 words)
- A portion of PLRN Corpus around 240,000 words[1]
- 12,450 reduplication tokens
- 746 unique reduplicated forms

(c) An online Corpus '*Rapa Nui Corpus*' [http://rapanui.polycorpora.org/]

(d) A Rapa Nui-French Dictionary: *Puka àkaero rapa, Lexique rapa-français avec glossaire français-rapa* (Tomite Reo Rapa, P. Kieviet and A. Kieviet 2006)

(e) An online English-Rapa Nui dictionary [http://kohaumotu.org/Rongorongo/Dictionary/index.html]

(12)Roadmap
- Section 2. Metrical structure and reduplication patterns in Rapa Nui
- Section 3. Backcopying driven by Surface Correspondence
- Section 4. Closing remarks
- [Appendix: Discussion of the alternative approaches]

2. Rapa Nui: metrical structure and reduplication

(13)Basics of Rapa Nui phonology

(a) Phonemic inventory

consonants	/p, t, k, ʔ, m, n, ŋ, h, v, ɾ, (f), (s)/
vowels	/a, i, u, e, o, aː, iː, uː, eː, oː/

(segments in parentheses only appear in loanwords)

(b) Syllable structure
- (C)V in general: onset is optional; coda is not allowed except in some loanwords
- No diphthong: e.g. *hoa* is syllabified as [ho.a]

2.1. Metrical structure

(14)The foot of Rapa Nui is moraic trochee.

---

[1] Thanks are due to Dr. Palaus Kieviet for making this corpus available.

(15) Stress assignment

(a) When the last syllable of a word is bimoraic (containing a long vowel), the ultimate syllable forms a foot and receives primary stress; e.g. [maˈŋoː] ‘shark’

(b) When the last syllable is light, the penultimate syllable receives primary stress; e.g. [ˈnoho] ‘to sit’

(c) In both cases, the strong mora of the other feet receives secondary stress, e.g. [ˌhaŋuˈpotu] ‘younger child’, [ˌkereˈtuː] ‘pumice’.

(d) Variable positions of secondary stress in **five-mora** words (Kieviet 2017:45)

mora count	*orthography*	*transcription*	*gloss*
5	*vanavanaŋa*	va.ˌna.va.ˈna.ŋa ˌva.na.va.ˈna.ŋa	‘to chat’

(16) Foot construction and stress assignment in Rapa Nui can be attributed to a set of metrical constraints

PARSE	All moras must be parsed by feet. (cf. Kager 1999:153)
FTBIN	Feet are composed of two moras. (Kenstowicz 2007; cf. Kager 1999:161)
ALL-FT-L	Every foot stands at the left edge of the PrWd. (*violations counted in syllables*)
ALL-FT-R	Every foot stands at the right edge of the PrWd. (*violations counted in syllables*)
ALIGN-HEAD-R	The right edge of the head foot must be aligned with the right edge of the PrWd. (*violations counted in syllables*)

(17) The analysis is couched in MaxEnt Harmonic Grammar (Goldwater and Johnson 2003; Hayes and Wilson 2008, etc., see also Coetzee and Pater 2011) in order to capture the variation. The weights were obtained in MaxEnt Grammar Tool (Hayes et al. 2009). (Note: the frequency assigned in the following tableaux is hypothetical.)

a)

	ma.u.ku	*freq.*	FTBIN	HEAD-R	PARSE	ALL-FT-L	ALL-FT-R	
	weights		*19.97*	*13.96*	*4.89*	*1.48*	*1.48*	*H*
☞ a.	ma.(ˈu.ku)	*1*			-1	-1		*-6.37*
b.	(ˈma.u).ku	*0*		-1	-1		-1	*-20.33*
c.	(ma).(ˈu.ku)	*0*	-1			-1	-2	*-24.41*
d.	(ˌma.u).(ˈku)	*0*	-1			-2	-1	*-24.41*
e.	(ˌma).(ˈu.ku)	*0*	-1			-1	-2	*-24.41*

b)

	va.na.va.na.ŋa	*freq.*	FTBIN	HEAD-R	PARSE	ALL-FT-L	ALL-FT-R	
	weights		*19.97*	*13.96*	*4.89*	*1.48*	*1.48*	*H*
a.	(ˌva.na).va.(ˈna.ŋa)	*0.5*			-1	-3	-3	*-13.77*
b.	va.(ˌna.va).(ˈna.ŋa)	*0.5*			-1	-4	-2	*-13.77*
c.	(ˌva.na).(ˈva.na).ŋa	*0*		-1	-1	-2	-4	*-27.73*
d.	(ˌva).(na.va).(ˈna.ŋa)	*0*	-1			-4	-6	*-34.77*
e.	(ˌva.na).(va).(ˈna.ŋa)	*0*	-1			-5	-5	*-34.77*
f.	(va).(ˌna.va).(ˈna.ŋa)	*0*	-1			-4	-6	*-34.77*
g.	(ˌva.na).(ˈva.na).(ŋa)	*0*	-1	-1		-6	-4	*-48.73*

2.2. Reduplication

(18) There are two types of reduplication with different shapes and functions.

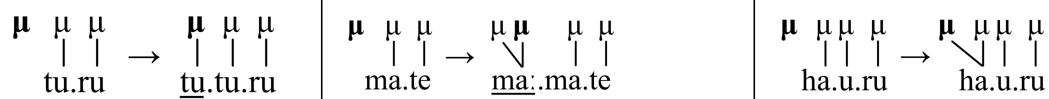
- *Plural reduplication*: reduplication is used to indicate plurality of verbs.
- *Intensifying reduplication*: intensification, repetition, emphasis, or conversion (noun→verb)

Plural reduplication

(19) Mora affixation results in either reduplication or lengthening

	Base	***Reduplicated***	***Gloss***	***Note***
a.	tu.ru	<u>tu</u>.tu.ru	'to go down (pl.)'	LL → LLL (copy a syllable)
b.	ma.te	<u>ma:</u>.ma.te	'to die (pl.)'	LL → HLL (copy a syllable and lengthening)
c.	ha.u.ru	ha:.u.ru	'to sleep (pl.)'	LLL → HLL (lengthening)

(19a) (19b)* (19c)



Reduplication only | *Reduplication and lengthening* | *lengthening*

[***Note**: (19a) and (19b) can be attributed to the competition between **metrical wellformedness** and **DEP-*μ***. Based on the corpus data, however, the variation between (19a) and (19b) seems to be lexical. The pattern shown in (19b) will be set aside in the following analysis.]

(20) The variation shown in (19) favors **Generalized Non-linear Affixation (GNA)** over Base-Reduplicant Correspondence Theory (BRCT):

(a) By adopting BRCT, two underlying forms are assumed: a RED morpheme (for 19a, 19b) and a mora (for 19c).
(b) By adopting GNA, the underlying representation for the plural marker is reduced to a single mora, and both surface realizations can be derived.
(c) The choice between reduplication (as in 19a) and lengthening (as in 19c) is related to metrical wellformedness.

(21) INTEGRITY: No elements of S_1 has multiple correspondents in S_2. (McCarthy and Prince 1995:124)

(22) *turu*

	μ + turu	*freq.*	FTBIN	HEAD-R	PARSE	INTEG	ALL-FT-L	ALL-FT-R	
	weights		*19.97*	*13.96*	*4.89*	*2.89*	*1.48*	*1.48*	*H*
☞	a. tu.(ˈtu.ru)	*1*			-1	-2	-1		*-12.15*
	b. (ˈtu:.ru)	*0*	-1						*-19.97*
	c. (ˈtu:).ru	*0*		-1	-1			-1	*-20.33*
	d. (tu).(ˈtu.ru)	*0*	-1			-2	-1	-2	*-30.19*

(23) *hauru*

	μ + hauru	*freq.*	FTBIN	HEAD-R	PARSE	INTEG	ALL-FT-L	ALL-FT-R	
	weights		*19.97*	*13.96*	*4.89*	*2.89*	*1.48*	*1.48*	*H*
☞	a. (ˌha:).(ˈu.ru)	*1*					-1	-2	*-4.44*
	b. (ˌha.ha).(ˈu.ru)	*0*				-2	-2	-2	*-11.70*

(24) A note on the shape of reduplicant in GNA:

- The current analysis does not consider the candidate that does not copy the onset, such as [<u>u</u>.tu.ru].
- The copy of onset can be ensured by high-ranked {RIGHT, LEFT}-ANCHOR(S_1, S_2); see Bye and Svenonius (2012: 456-458, 463-466) for discussion.

Intensifying reduplication:

(25) Functions and form:

- Intensification, repetition, or conversion (derivational)
- It can be viewed as the affixation of two moras: /μμ/

(26) Examples: *Monosyllabic or disyllabic base (bimoraic)*

	Base	***Reduplicated***	***gloss***
a.	pa: /μμ/ H	pa:.pa: H H	'to fold repeatedly'
b.	ho.a /μμ/ L L	ho.a.ho.a L L L L	'to throw various things'

(27) Examples: *Trisyllabic base*

a) Some LLL and HLL bases can undergo either left-edge copy or right-edge copy, as in (27a) and (27d);
b) Other words like (27b, c, e, f, g) only have one reduplication pattern (lexical variation).
c) Left-edge copy: the output shape is **L L L L L(H)**
d) Right-edge copy: the output shape is **H L L L L**

Base			***Reduplicated***		***Gloss of the base***
			left-edge copy	***right-edge copy***	
L L L base	a.	ma.na.ʔu +/μμ/ L L L	ma.na.ma.na.ʔu L L L L L	ma:.na.ʔu.na.ʔu H L L L L	'to think'
	b.	pu.hi.a +/μμ/ L L L	pu.hi.pu.hi.a L L L L L		'to fly'
	c.	ka.pu.a +/μμ/ L L L		ka:.pu.a.pu.a H L L L L	'mist'
H L L base	d.	ho:.ro.u +/μμ/ H L L	ho.ro.ho.ro.u L L L L L	ho:.ro.u.ro.u H L L L L	'quick'
	e.	ma:.ro.a +/μμ/ H L L	ma.ro.ma.ro.a L L L L L		'to stand'
	f.	ha:.ʔe.re +/μμ/ H L L		ha:.ʔe.re.ʔe.re H L L L L	'to stroll'
L L H base	g.	ʔa.ʔu.e: +/μμ/ L L H	ʔa.ʔu.ʔa.ʔu.e: L L L L H		'to cry'

(28) Closer examination of **(27a)** and **(27d)**

		left-edge copy	*right-edge copy*
(27a)	ma.na.ʔu +/μμ/ L L L	ma.na.ma.na.ʔu L L L L L	ma:.na.ʔu.na.ʔu H L L L L
		expected	***base lengthened***
(27d)	ho:.ro.u +/μμ/ H L L	ho.ro.ho.ro.u L L L L L	ho:.ro.u.ro.u H L L L L
		base shortened (backcopying?)	***expected***

(29) Summary for intensifying reduplication:

a) The position of the copied string is variable (either at left edge or right edge)
   - The variable positions can result from a set of constraints: O-CONTIGUITY, ALIGN(STEM, R, PRWD, R), and ALIGN(STEM, L, PRWD, L), which will be discussed later.

b) The variation is accompanied by the modification of the base.

2.3. Interim summary and hypotheses

(30) For plural reduplication, the variable patterns favor Generalized Non-Linear Affixation over classic Base-Reduplicant Correspondence Theory.

(31) For intensifying reduplication, the variable patterns are complicated

- Five-mora words: variable secondary stress positions
- Trisyllabic words: variable reduplicated forms

(32) Example: **LLL** word /manaʔu, *μμ*/, 'to think'

variable reduplicant		
left-edge copy	*right-edge copy*	
ma.(ˌna.ma).(ˈna.ʔu)	(ˌma:).(ˌna.ʔu).(ˈna.ʔu)	*variable stress pattern*
(ˌma.na).ma.(ˈna.ʔu)		

(33) Example: **HLL** word /va:naŋa, *μμ*/, 'to talk'

variable reduplicant		
left-edge copy	*right-edge copy*	
va.(ˌna.va).(ˈna.ŋa)	(ˌva:).(ˌna.ŋa).(ˈna.ŋa)	*variable stress pattern*
(ˌva.na).va.(ˈna.ŋa)		

(34) The problems in the examples above:

(a) What motivates mora epenthesis for LLL words in right-edge copy?
e.g. /ma.na.ʔu/, *μμ* → [**ma**:.na.ʔu.na.ʔu]

(b) Why there is no mora epenthesis in left-edge copy?
e.g. /ma.na.ʔu/, *μμ* → [**ma**.na.ma.na.ʔu], instead of *[**ma**:.na.ma.na.ʔu]

(c) What motivates mora deletion for HLL words in left-edge copy? ('suspected backcopying')
e.g. /va:.na.ŋa/, *μμ* → [va.na.**va**.na.ŋa]

(35) Two hypotheses:

Hypothesis-1: Base lengthening/shortening in (34a) and (34c) are epiphenomenal, possibly due to metrical optimization.

Hypothesis-2: Base lengthening in (34a) is due to metrical optimization but base shortening in (34c) is not epiphenomenal; extra mechanism is responsible for the patterns.

3. Backcopying driven by Surface Correspondence

3.1. Proposal: Surface Correspondence in reduplication

(36) Given the patterns shown in (33) and (34), it is possible that there exists a correspondence relation in the surface:

e.g. /va:naŋa/ left-edge copy

✓ (a) **va**.na.**va**.na.ŋa — vowel length (syllable weight) matches
✗ (b) **va**.na.**va**:.na.ŋa — vowel length does not match
✗ (c) **va**:.na.**va**.na.ŋa — vowel length does not match
✗ (d) **va**:.na.**va**:.na.ŋa — vowel length matches, but non-optimal metrical structure

(37) However, reduplication in GNA does not have the notions 'BASE' and 'RED', and therefore no BR-Correspondence between strings.

(38) This work borrows the notion 'Surface Correspondence':

- Originally proposed for consonant harmony in Walker (2000a, 2000b, 2001) as 'consonantal correspondence'
- Further developed in Hansson (2001), Rose and Walker (2004, Agreement by Correspondence; 'ABC') and Bennett (2013, 2015).

(39) Two types of constraints

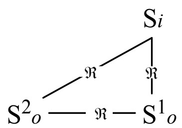
(a) Correspondence-requiring constraints (SCORR constraints), e.g. CORR-C↔C constraint in Rose and

Walker (2004), later reformulated as CORR[αF] in Bennett (2013, 2015).

(b) IDENT constraints that require identity between the segments in correspondence, e.g. IDENT-CC[F], IDENT-VV, etc; named as Limiter Constraints in Bennett (2013, 2015).

(40) In the current proposal, no correspondence-requiring constraints such as CORR or REDUP (Zuraw 2002) are needed for reduplication.

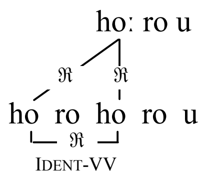
(41) Correspondence can be established naturally in reduplication, even without BASE and RED:



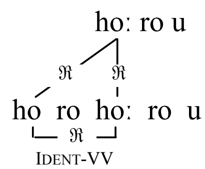
- For a string S_i in the input, if there are multiple strings S_o^1 and S_o^2 in the output such that $S_o^1\ \mathfrak{R}\ S_i$ and $S_o^2\ \mathfrak{R}\ S_i$, there must be $S_o^1\ \mathfrak{R}\ S_o^2$.
- IDENT constraints will require the similarity between elements (S_o^1 and S_o^2) in correspondence.

(cf. Struijke 2001; Stanton and Zukoff 2016, to appear).

(42) For the case of Rapa Nui:



violate MAX-μ
satisfies IDENT-VV-LENGTH



satisfies MAX-μ
violates IDENT-VV-LENGTH

- **IDENT-VV- LENGTH**: Assign a violation if the vowels that stand in correspondence are associated with a different number of moras.

(43) The effect of IDENT-VV-LENGTH (weights are assigned manually)*

	/ho:.ro.u/, $\mu\mu$	ID-VV- LENGTH	DEP-SEG	INTEGRITY	MAX-μ	DEP-μ	
	weights	*2*	*2*	*1*	*1*	*1*	*H*
☞ a.	ho$_x$ ro$_y$ ho$_x$ ro$_y$ u			-4	-1		*-5*
b.	ho$_x$ ro$_y$ ho:$_x$ ro$_y$ u	-1		-4			*-6*
c.	ho:$_x$ ro$_y$ ho$_x$ ro$_y$ u	-1		-4	-1	-1	*-8*
d.	ho ro ho: ro u		-4				*-8*

[***Note:** One candidate, [ho:.ho:.ro.u], is not considered here. This candidate has perfect foot structure and does not violate any faithfulness constraints except for INTEGRITY. One explanation could be that this language does not prefer adjacent heavy syllables. According to Kieviet (2017:39-40), only 16 out of 4644 words are HHLL, and the other words that contain adjacent heavy syllables, e.g. LHH, LLHH, and HHH, have very low type frequency.]

3.2. Modelling variations in Rapa Nui

(44) In order to test whether IDENT-VV has advantage or not, two MaxEnt models are designed:

- *Model-1* (No IDENT-VV): only a set of constraints for metrical structure and reduplication; IDENT-VV is not included. (*Hypothesis-1*)
- *Model-2* (With IDENT-VV): a set of constraints for metrical structure and reduplication plus the proposed Surface Correspondence relationship. (*Hypothesis-2*)

(45) ***Constraints*** included in the models

Metrical structure	*Reduplication*	*IO-Faith*	*Limiter Constraint*
PARSE FTBIN ALL-FT-L ALL-FT-R ALIGN-HEAD-R	INTEGRITY O-CONTIGUITY ALIGN(STEM, R, PRWD, R) ALIGN(STEM, L, PRWD, L)	DEP-μ MAX-μ	IDENT-VV-LENGTH

- Reduplication constraints

O-CONTIGUITY	The portion of output string standing in correspondence with the input string form a contiguous string. ("No intrusion", e.g. xz→x**y**z) (McCarthy and Prince 1995:123)
ALIGN(STEM, R, PRWD, R)	Align the right boundary of the stem with the right boundary of the prosodic word. (*violations counted in segments*)
ALIGN(STEM, L, PRWD, L)	Align the left boundary of the stem with the left boundary of the prosodic word. (*violations counted in segments*)

Violation profile:

ma.na.ʔu, *μμ*	O-CONTIG	ALIGN-R	ALIGN-L
a. ma.na.ma.na.ʔu		****	
b. ma.na.ʔu.na.ʔu			****
c. ma.na.ʔu.na.ʔu	*		
d. ma.na.ma.na.ʔu	*		

(46) ***Inputs:*** for each model, 8 inputs with various candidates are designed; for each possible winning candidates, a hypothetical observed probability is assigned. For example:

Input /manaʔu, *μμ*/, 'to think'

possible winner	*reduplication*	*stress position*	*probability*
(ˌmaː).(ˌna.ʔu).(ˈna.ʔu)	right-edge copy 0.5	1	0.5
ma.(ˌna.ma).(ˈna.ʔu)	left-edge copy 0.5	0.5	0.25
(ˌma.na).ma.(ˈna.ʔu)		0.5	0.25

(47) There are two reasons why the probability is hypothetical in the current work:

(a) The variable stress positions cannot be directly observed in orthography;

(b) No matter the output is in the shape of LLLLL or HLLLL, the base can be either LLL or HLL. Given the accessible materials so far, not all the bases can be identified.

(48) Results

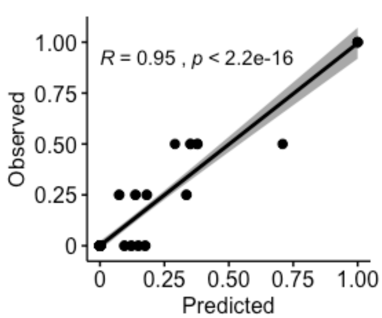
Constraint weights

Without IDENT-VV		**With IDENT-VV**	
PARSE	4.90	PARSE	4.60
FTBIN	13.18	FTBIN	14.23
ALL-FT-L	1.94	ALL-FT-L	0.30
ALL-FT-R	1.05	ALL-FT-R	0.26
ALIGN-HEAD-R	10.02	ALIGN-HEAD-R	9.08
INTEGRITY	0	INTEGRITY	0
O-CONTIG-STEM	9.89	O-CONTIG-STEM	11.92
ALIGN(STEM, R, PRWD, R)	0	ALIGN(STEM, R, PRWD, R)	0
ALIGN(STEM, L, PRWD, L)	0.06	ALIGN(STEM, L, PRWD, L)	0.803
DEP-*μ*	0	DEP-*μ*	0
MAX-*μ*	0.69	MAX-*μ*	0
		IDENT-VV[long]	11.84

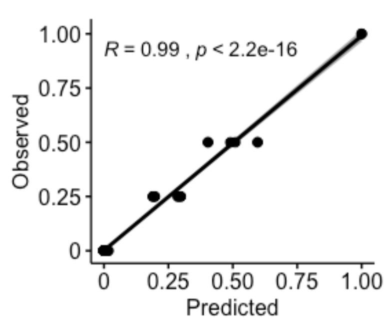
Predictions

	Input	*Candidates*	*Observed*	*Predicted* Without IDENT-VV	*Predicted* With IDENT-VV
a)	ma.ŋo:	ma.(ˈŋo:)	1	1.000	1.000
		(ˈma).ŋo:	0	0	0
		... 3 more are omitted	…	…	…
b)	ma.u.ku	ma.(ˈu.ku)	1	1.000	1.000
		(ˈma. u).ku	0	0	0
		... 3 more are omitted	…	…	…
c)	ke.re.tu:	(ˌke.re).(ˈtu:)	1	1.000	1.000
		ke.re.(ˈtu:)	0	0	0
		... 2 more are omitted	…	…	…
d)	va.na.va.na.ŋa	va.(ˌna.va)(ˈna.ŋa)	0.5	0.291	0.490
		(ˌva.na).va.(ˈna.ŋa)	0.5	0.709	0.510
		(ˌva.na).(ˈva.na).ŋa	0	0	0
		... 4 more are omitted	…	…	…
e)	ma.na.ʔu, *μμ*	(ˌma:).(ˌna.ʔu).(ˈna.ʔu)	0.5	0.351	0.404
		ma.(na ˌ.ma).(ˈna.ʔu)	0.25	0.138	0.287
		(ˌma.na).ma.(ˈna.ʔu)	0.25	0.336	0.298
		ma.(ˌna.ʔu).(ˈna.ʔu)	0	0.176	0.012
		... 6 more are omitted	…	…	…
f)	ma:.ta.ki, *μμ*	(ˌma:).(ˌta.ki).(ˈta.ki)	0.5	0.378	0.404
		ma.(ˌta.ma).(ˈta.ki)	0.25	0.075	0.287
		(ˌma.ta).ma.(ˈta.ki)	0.25	0.182	0.298
		(ˌma).(ˌta.ma).(ˈta.ki)	0	0.148	0
		ma.(ˌta.ki).(ˈta.ki)	0	0.095	0.012
		... 8 more are omitted	…	…	…
g)	ma.na.u, *μμ*	(ˌma:).(ˌna.u).(ˈna.u)	0.5	0.351	0.596
		ma.(ˌna.ma).(ˈna.u)	0.25	0.138	0.190
		(ˌma.na).ma.(ˈna.u)	0.25	0.336	0.197
		ma.(ˌna.u).(ˈna.u)	0	0.176	0.017
		... 6 more are omitted	…	…	…
h)	ho:.ro.u, *μμ*	(ˌho:).(ˌro.u).(ˈro.u)	0.5	0.378	0.596
		ho.(ˌro.ho).(ˈro.u)	0.25	0.075	0.190
		(ˌho.ro).ho.(ˈro.u)	0.25	0.182	0.197
		(ˌho:).(ˌro.ho).(ˈro.u)	0	0.149	0
		ho.(ˌro.u).(ˈro.u)	0	0.095	0.017
		... 8 more are omitted	…	…	…

(49) The results improved when IDENT-VV-LENGTH is included in the grammar.



a. Without IDENT-VV; r^2 = *0.895*



b. With IDENT-VV; r^2 = *0.986*

Summary

(50) The variations shown in intensifying reduplication result in base lengthening or base shortening, which helps demonstrate the advantage of Surface Correspondence (IDENT-VV-LENGTH) in reduplication.

(51) Recall the questions in (34) and the hypotheses

(a) What motivates mora epenthesis for LLL words in right-edge copy?
e.g. /ma.na.ʔu/, *μμ* → [**ma**ː.na.ʔu.na.ʔu]

(b) Why there is no mora epenthesis in left-edge copy?
e.g. /ma.na.ʔu/, *μμ* → [**ma**.na.ma.na.ʔu], instead of *[**ma**ː.na.ma.na.ʔu]

(c) What motivates mora deletion for HLL words in left-edge copy? ('suspected backcopying')
e.g. /vaː.na.ŋa/, *μμ* → [va.na.**va**.na.ŋa]

Two hypotheses:

✗ *Hypothesis-1:* Base lengthening/shortening in (34a/51a) and (34c/51c) are epiphenomenal, possibly due to metrical optimization.

✓ *Hypothesis-2:* Base lengthening in (34a/51a) is due to metrical optimization but base shortening in (34c/51c) is not epiphenomenal; extra mechanism is responsible for the patterns.

3.3. Further evidence on Surface Correspondence

(52) Recall monosyllabic and disyllabic bases in intensifying reduplication:

	Base	*Reduplicated*	*gloss*
a.	paː	paː.paː	'to fold repeatedly'
b.	ho.a	ho.a.ho.a	'to throw various things'

(53) Since the template for intensifying reduplication is bimoraic, some candidates are like [ho.a.aː]

	/ho.a/, *μμ*	ID-VV-LENGTH	INTEGRITY	
	weights	*3*	*1*	*H*
☞	a. $ho_x.a_y.\underline{ho_x.a_y}$		-3	*-3*
	b. $ho.a_x.\underline{a_x}ː$	-1	-1	*-4*

4. Closing remarks

(54) *Summary*

- **Empirically**, this work presents new interpretation of recent-published data and argues that it is a case of backcopying. **Theory-wise**, the advantage of Surface Correspondence is shown by MaxEnt models.
- In BRCT, the crucial mechanism that lead to backcopying include **BR-Correspondence** and **parallelism**. Some theories such as GNA that does not predict backcoping due to the lack of BR-Correspondence.
- However, BR-CORR can be viewed as a special case of SCORR. The establishment of Surface Correspondence is natural in reduplication, and it is not morphologically required (i.e. not required by BASE and RED).
- The 'GNA + SCORR' retains the superiority of GNA when analyzing the reduplicative allomorphy in an economical way (plural reduplication), but also offers account for the backcopying in Rapa Nui.

(55) *The role of variation*

(a) The variation shown in plural reduplication (mora affixation results in either reduplication or lengthening) shows the advantage of GNA over BRCT.

(b) The variation of intensifying reduplication is accompanied by base modification (shortening and lengthening), which shows a case of backcopying and the advantage of correspondence relation in reduplication.

(56) *Questions remain open …*

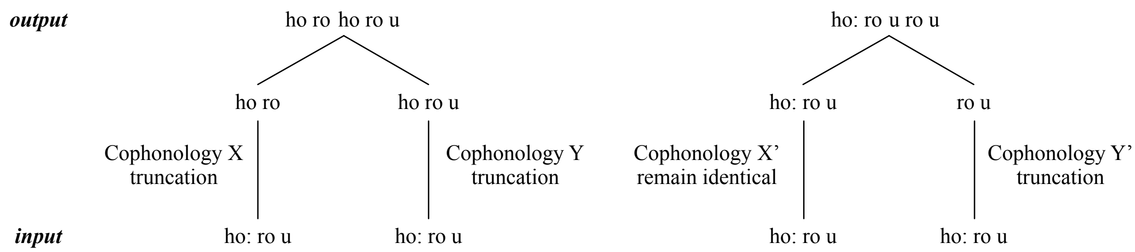
- Does the correspondence relation in reduplication come for free (supplied by GEN) or at the expense of certain constraints (e.g. CORR)?
- If backcopying exists (the base conforms to the copied string), then why it is typologically rare? (the influence between surface strings is asymmetrical)

[Appendix. Alternative approaches]

(1) Approaches that assume RED morpheme, including **classic BRCT** (McCarthy and Prince 1995) and **Generalized Template Theory** (**GTT**, Urbanczyk 1996, 2006; McCarthy and Prince 1999), are not able to capture both reduplication or lengthening.

(2) **Serial Templatic Satisfaction** (**STS**, McCarthy, Kimper and Mullin, 2012) is in Harmonic Serialism, which lacks BR-Correspondence and parallelism, and therefore cannot predict backcopying.

(3) **Morphological Doubling Theory** (**MDT**, Inkelas and Zoll 2005) claims that morphological reduplication is the consequence of morpheme doubling, and different cophonologies can target 'BASE' and 'REDUPLICANT' respectively. The variable patterns in Rapa Nui could be analyzed in the following ways:



A problem in the MDT analysis is what motivates these cophonologies. For example, why Cophonology X specifically deletes /u/ and shorten /ho:/ remain unclear.

Acknowledgement

Thanks to Rachel Walker, Stephanie Shih, and all the participants of PhonLunch at USC, for their valuable comments. Special thanks are due to Dr. Kieviet for allowing me to access the corpus of Rapa Nui. All errors remain my own.

References

Bennett, William G. 2013. *Dissimilation, consonant harmony, and surface correspondence*. PhD dissertation. Rutgers The State University of New Jersey-New Brunswick.

Bennett, William G. 2015. Assimilation, dissimilation, and surface correspondence in Sundanese. *Natural Language & Linguistic Theory* 33(2). 371–415.

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Trommer (ed.). The Morphology and Phonology of Exponence. Oxford: Oxford University Press.

Bye, Patrik and Peter Svenonius. 2012. Non-concatenative morphology as epiphenomenon. In Trommer (ed.). The Morphology and Phonology of Exponence. Oxford: Oxford University Press.

Coetzee, Andries W. & Joe Pater. 2011. The Place of Variation in Phonological Theory. In John Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The Handbook of Phonological Theory*, 401–434. Oxford, UK: Wiley-Blackwell.

Davis, Stuart and Isao Ueda. 2002. Mora augmentation processes in Japanese. Journal of Japanese Linguistics 18, 1–23.

Goldwater, Sharon and Mark Johnson (2003). Learning OT constraint rankings using a Maximum Entropy model. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.) Proceedings of the Stockholm Workshop on Variation within Optimality Theory. Stockholm: Stockholm University. 111–120

Hansson, Gunnar Ólafur. 2001. Theoretical and typological issues in consonant harmony. PhD thesis, University of California, Berkeley.

Hayes, Bruce and Colin Wilson 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39. 379–440.

Hayes, Bruce, Colin Wilson and Ben George 2009. Maxent grammar tool. Software package. Available (November 2018) at http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/.

Inkelas, Sharon and Cheryl Zoll. 2005. Reduplication: Doubling in Morphology. Cambridge: Cambridge University Press.

Kager, Réne. 1999. Optimality Theory. Cambridge University Press

Kenstowicz, Michael. 2007. Salience and similarity in loanword adaptation: a case study from Fijian. *Language*

Sciences (Issues in English Phonology) 29(2). 316–340.
Kieviet, Paulus. 2017. *A Grammar of Rapa Nui*. Language Science Press. Berlin.
McCarthy, J.J., 2010. Agreement by correspondence without CORR constraints. *Ms. University of Massachusetts, Amherst*.
McCarthy, John J. and Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill N. Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds). *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. 249-384. Amherst, MA: GLSA.
McCarthy, John J. and Alan Prince. 1999. Faithfulness and Identity in ProsodicMorphology', in R. Kager, H. van der Hulst, and W. Zonneveld (eds.), The Prosody-Morphology Interface. Cambridge: Cambridge University Press, 218–309.
McCarthy, John J., Wendell Kimper & Kevin Mullin. 2012. Reduplication in Harmonic Serialism. *Morphology* 22(2). 173–232.
Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. Language 80: 475–531.
Saba Kirchner, Jesse. 2010. Minimal Reduplication. Ph.D. dissertation. University of California Santa Cruz.
Shih, Stephanie S, and Sharon Inkelas. 2014. A subsegmental correspondence approach to contour tone (dis)harmony patterns. In *Proceedings of the 2013 Annual Meeting on Phonology*, ed. by John Kingston, Claire Moore-Cantwell, Joe Pater, and Robert Staubs. Washington, D.C.: Linguistic Society of America.
Stanton, J. and Zukoff, S., 2016. Prosodic effects of segmental correspondence. *Proceedings of CLS*, *51*, pp.501-515.
Stanton, J. and Zukoff, S., *to appear*. Prosodic identity in copy epenthesis: Evidence for a correspondence-based approach. *Natural Language and Linguistic Theory*.
Struijke, Caro. 2014. *Existential Faithfullness: A Study of Reduplicative TETU, Feature Movement and Dissimulation*. Routledge.
Tomite Reo Rapa, Palaus. Kieviet and Antje Kieviet. 2006. *Puka àkaero rapa, Lexique rapa-français avec glossaire français-rapa.* SIL International.
Trommer, Jochen and Eva Zimmermann. 2014. Generalized mora affixation and quantity-manipulating morphology. *Phonology* 31(2014). 463-510.
Urbanczyk, Suzanne. 1996. Patterns of reduplication in Lushootseed. PhD dissertation. University of Massachusetts, Amherst.
Urbanczyk, Suzanne. 2006. Reduplicative form and the Root-Affix Asymmetry. *Natural Language & Linguistic Theory* 24(1). 179–240.
Walker, Rachel. 2000a. Long-distance consonantal identity effects. In West Coast Conference on Formal Linguistics (WCCFL). Vol. 19, 532–545.
Walker, Rachel. 2000b. Yaka nasal harmony: spreading or segmental correspondence? In Berkeley Linguistics Society (BLS). Vol. 26, 321–332.
Walker, Rachel. 2001. Consonantal correspondence. In Workshop on the lexicon in phonetics and phonology. Vol. 6 of Papers in experimental and theoretical linguistics, 73–84. Edmonton: University of Alberta.
Zimmermann, Eva. 2015. The power of a single representation: morphological tone and allomorphy. *Morphology* 26(3–4). 269–294.
Zuraw, Kie. 2002. Aggressive reduplication. *Phonology* 19:395-439