The emergent typology of reduplication: Universals and variations in learning biases.

Yang Wang (yangx.wang@utah.edu)

Department of Linguistics, University of Utah

1. Summary

Big picture questions

- What are the **analytic biases** (e.g., Wilson, 2006; Moreton and Pater, 2012) that guide (morpho)phonological learning?
- How do these biases reflect natural language **typology** (e.g. Moreton, 2008)?

Artificial grammar learning experiments with reduplication

When reduplicative patterns are only observed in relatively short forms
e.g., monosyllabic copying ['pif] ∼ ['pifpif] or ['pif] ∼ ['pipif]

- On what levels of **phonological abstraction** (e.g., syllables, feet, or prosodic words) are human learners biased to form reduplicative generalizations?
- What does the **hypothesis space** for a human learner look like given the input?

Findings and takeaway

- Human learners generalize in a manner that is sensitive to phonological abstractions characterizable **by the vocabulary of prosody**
  ↪ support for the Prosodic Morphology (McCarthy and Prince, 1986)

  Participants were guided by other theoretically grounded principles (see paper)
- **Spontaneous responses appeared to reflect the reduplicative typology.**
  - the most frequent forms were consistent with the typological trends
  - the variations in individually biased grammars reflected the attested variations.

2. The reduplicative typology

Cross-linguistic variation along many crucial dimensions (e.g., Inkelas and Downing, 2015)

- Dimension I: the phonological shape (i.e., how much to copy)
  1. Total (Indonesian; Austronesian; McCarthy and Cohn, 1998)

     bu.ku ∼ bu.ku-bu.ku *'book'* *'book-PL'*

     ma.ša.ra.kat ∼ ma.ša.ra.kat-ma.ša.ra.kat *'society'* *'society-PL'*
  2. Partial
     1. A bisyllabic foot (Diyari; Pama-Nyungan; Austin, 1981)

        pir.ta ∼ pir.ta-pir.ta *'tree'* *'DIM- tree'*

        wil.ha.pi.na wil.ha-wil.ha.pi.na *'old woman'* *'DIM-old woman'*
     2. A heavy syllable (Ilokano; Austronesian; Hayes and Abad, 1989)

        kut.tóŋ ∼ naka-kut-kut.tóŋ *'thin'* *'ADJ-INTENS-thin'*

        bu.téŋ ∼ naka-but-bu.téŋ *'afraid'* *'ADJ-INTENS-afraid'*
     3. A light syllable (Tonkawa; Coahuiltecan; Gouskova, 2007)

        to.po s ∼ to-to.po s *'I cut it'* *'REP-I cut it'*

        xej.tso s ∼ xe-xej.tso s *'I rub him'* *'REP-I rub him'*
  3. Infixation (Samoan; Austronesian; Broselow and McCarthy, 1983)

     alófa ∼ a-lo-lófa *'love'* *'love-PL'*

     saváli ∼ sa-va-váli *'walk'* *'walk-PL'*
- Dimension II: which part of the stem is copied if partially reduplicated
  1. Left-edge oriented (see above)
  2. Right-edge oriented (Manam; Austronesian; Lichtenberk, 1983)

     salaga ∼ salaga-laga *'be long'* *'long-SG'*

     sapara ∼ sapara-para *'branch'* *'having branches'*
  3. Infixation (Samoan; Austronesian; Broselow and McCarthy, 1983)

     alófa ∼ a-lo-lófa *'love'* *'love-PL'*

     saváli ∼ sa-va-váli *'walk'* *'walk-PL'*
- Other possible variations
  1. Vowel reduction (Palauan; Austronesian; Zuraw, 2002)

     tóɾð ∼ bəkə-təɾ-tóɾð *''frustration''* *'easily frustrated'*
  2. Templatic back-copying (?) (Guarijio; Uto-Aztecan; Austronesian; Caballero, 2006)

     toní ∼ to-tó *'to boil'* *'to start boiling'*

     muhíba ∼ mu-mú *'to throw'* *'to start throwing'*

*See the draft of the paper for more details!

3. Experiment

Poverty of the stimulus paradigm (Wilson, 2006 *et seq.*)

Participants: Engish speakers recruited from Prolific ($N_1 = 144$; $N_2 = 105$)

Familiarization: auditory input, no orthographic help

Listen to the singular.
Expt. 1 + 2: ['pif]

Listen to the plural.
Expt. 1:['pifpif] **Expt. 2:**['pipif]

- Four pairs of base and reduplicated forms
- Bases are all monosyllabic $C_1V_2C_3$ forms (e.g.,['pif]) with variegated segment choices
  - Expt. 1: reduplicant is $C_1V_2C_3$ (e.g.,['pifpif]) Expt. 2: reduplicant is C_1V_2 (e.g.,['pipif])

Testing: auditory input, free spontaneous production responses

Testing types	*Shapes*	*Examples*	*# Seg.*	*# σ*
FAMILIAR	$_{\sigma}C_1V_2C_3$	['noʊg]	3	1
DISYLLABIC CV	$_{\sigma}C_1V_2.C_3V_4C_5$	['ti.kɐp]	5	2
DISYLLABIC CVC	$_{\sigma}C_1V_2C_3.C_4V_5C_6$	['dɛb.gɪv]	6	2
TRISYLLABIC	$_{\sigma}C_1V_2.C_3V_4.C_5V_6C_7$	['ti.fæ.pəs]	7	3
PENTASYLLABIC	$_{\sigma}C_1V_2.C_3V_4.C_5V_6.C_7V_8.C_9V_{10}C_{11}$	[ˌpi.sæ.'goʊ.bɛ.kʊt]	11	5

- Four trials for each of the five testing types (20 trials in total); tested together, order randomized

Possible hypotheses: Expt. 1

Hypothesis	DISYLLABIC CV ['ti.kɐp]	DISYLLABIC CVC ['dɛb.gɪv]	TRISYLLABIC ['ti.fæ.pəs]	PENTASYLLABIC [ˌpi.sæ.'goʊ.bɛ.kʊt]
total	**'ti.kɐp**-'ti.kɐp	**'dɛb.gɪv**-'dɛb.gɪv	**'ti.fæ.pəs**-'ti.fæ.pəs	**ˌpi.sæ.'goʊ.bɛ.kʊt**-ˌpi.sæ.'goʊ.bɛ.kʊt
$_{wd}$[FT	**'tik**/**'tikɛ**-'ti.kɐp	**'dɛb**/**'dɛb.gɪ**-'dɛb.gɪv	**'tif**/**'tifæ**-'ti.fæ.pəs	**ˌpis**/**ˌpisæ**-ˌpi.sæ.'goʊ.bɛ.kʊt
$_{wd}[\sigma_{\mu\mu}$	**tik**-'ti.kɐp	**dɛb**-'dɛb.gɪv	**tif**-'ti.fæ.pəs	**pis**-ˌpi.sæ.'goʊ.bɛ.kʊt
$\sigma_{\mu\mu}]_{wd}$	'ti.kɐp-**kɐp**	'dɛb.gɪv-**gɪv**	'ti.fæ.pəs-**pəs**	ˌpi.sæ.'goʊ.bɛ.kʊt-**kʊt**
$'\sigma_{\mu\mu}$	**tik**-'ti.kɐp	**dɛb**-'dɛb.gɪv	**tif**-'ti.fæ.pəs	ˌpi.sæ.-**goʊb**-'goʊ.bɛ.kʊt

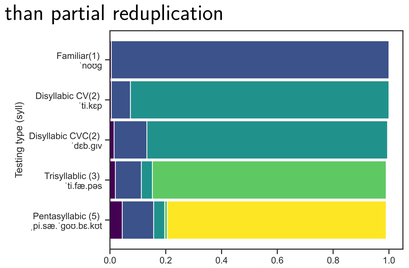
Possible hypotheses: Expt. 2

Hypothesis	DISYLLABIC CV ['ti.kɐp]	DISYLLABIC CVC ['dɛb.gɪv]	TRISYLLABIC ['ti.fæ.pəs]	PENTASYLLABIC [ˌpi.sæ.'goʊ.bɛ.kʊt]
WD + NOFINALCODA	**'ti.kɐ**-'ti.kɐp	**'dɛb.gɪ**-'dɛb.gɪv	**'ti.fæ.pə**-'ti.fæ.pəs	**ˌpi.sæ.'goʊ.bɛ.kʊ**-ˌpi.sæ.'goʊ.bɛ.kʊt
$_{wd}$[FT+NOCODA	**'ti.kɐ**-'ti.kɐp	**'dɛb.gɪ**-'dɛb.gɪv	**'ti.fæ**-'ti.fæ.pəs	**ˌpi.sæ**-ˌpi.sæ.'goʊ.bɛ.kʊt
$_{wd}[\sigma_{\mu}$	**ti**-'ti.kɐp	**dɛ**-'dɛb.gɪv	**ti**-'ti.fæ.pəs	**pi**-ˌpi.sæ.'goʊ.bɛ.kʊt
$\sigma_{\mu}]_{wd}$	'ti-**kɐ**-kɐp	'dɛb-**gɪ**-gɪv	'ti.fæ.-**pə**-pəs	ˌpi.sæ.'goʊ.bɛ-**kʊ**-kʊt
$'\sigma_{\mu}$	**ti**-'ti.kɐp	**dɛ**-'dɛb.gɪv	**ti**-'ti.fæ.pəs	ˌpi.sæ-**goʊ**-'goʊ.bɛkʊt

4. Results I: the universals

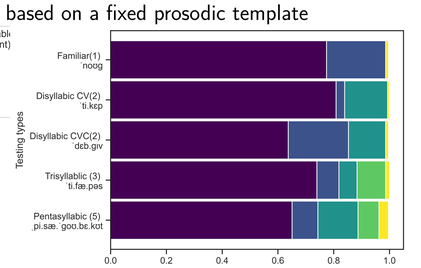
Expt 1: Total reduplication

Typology: total reduplication is more frequent than partial reduplication



Expt 2: fixed light syllable copying with more variations

Typology: partial reduplication is characterized based on a fixed prosodic template



*Figures: the averaged proportion of affix shape conditioned on testing type

*The most frequent shapes were verified by Bayesian mixed multinomial logistic regression with by-subject random effects.

5. Results II: the individual variations

Expt. 1: example individual grammars

- **word-final heavy syllable copying** (to a varying degree, N = 15)

  'zi.vɪb 'zi.vɪb-vɪb
  'tɛf.kʊp 'tɛf.kʊp-kʊp
  'gɑ.və.dus 'gɑ.və.dus-dus
  pi.sæ.'goʊ.bæ.kʊt pi.sæ.'goʊ.bæ.kʊt-kʊt
- **Templatic Backcopying** (to a varying degree, N = 3)

  'gɑ.və.dus 'gɑv-gɑv
  pi.sæ.'goʊ.bæ.kʊt 'pis-pis (3/4 for one participant)

Expt. 2: example individual grammars

- **Bisyllabic trochaic foot copying**: (N = 10)

  ˌpi.sæ.'goʊ.bæ.kʊt ˌpi.sæ-ˌpi.sæ.'goʊ.bæ.kʊt
- **Total + No word final coda** (N = 2, all categorical)

  'zi.vɪb 'zi.vɪ-'zi.vɪb
  'tɛf.kʊp 'tɛf.kʊ-'tɛf.kʊp
  'gɑ.və.dus 'gɑ.və.du-'gɑ.və.dus
  ˌpi.sæ.'goʊ.bæ.kʊt ˌpi.sæ.'goʊ.bæ.kʊ-ˌpi.sæ.'goʊ.bæ.kʊt
- **Infixation** (to a varying degree, N = 14)

  ˌpi.sæ.'goʊ.bæ.kʊt ˌpi.sæ.-ˌgou.bæ-'goʊ.bæ.kʊt
- **Vowel reduction** (to a varying degree, N = 20)

  'zi.vɪb zə-'zi.vɪb
  'tɛf.kʊp tə-'tɛf.kʊp
  'gɑ.və.dus gə-'gɑ.və.dus
  ˌpi.sæ.'goʊ.bæ.kʊt pə-ˌpi.sæ.'goʊ.bæ.kʊt

*For more detailed investigations on the individual variations, see the draft of the paper.

6. Discussions

- Human learners generalize reduplicative patterns in a manner that is sensitive to phonological abstractions characterizable by the vocabulary of the **prosody**.
- The **systematicity** in participants' responses → the possibility of the *poverty of stimulus* design as a **sampler** of the learner's hypothesis space conditioned on the input, at least for reduplication learning.
- The **diversity** of possible linguistic structures may also have its root in learning, reflected by the great variety of possible analyses ***individual*** human learners are biased towards.

7. Future directions

- A **large-scale corpus** of AGL experiments as a benchmark for **quantitative** predictions
- **Computational modeling** to capture the general trends and variations (current work in progress)

Acknowledgements

Many thanks to Bruce Hayes, Tim Hunter, Colin Wilson, Kie Zuraw, Claire Moore-Cantwell, Laurel Perkins, Megha Sundara, Iza Sola-Llonch, Canaan Breiss, Lily Xu, the audience of the UCLA Comp/Psycholing Seminar, UCLA Colloquium, Utah Comp Ling working group, and the anonymous reviewers for their comments, feedback and insights. Thanks to my RAs, Mariana Cui, Jacob Hanna, Jenessa Lathrop, Shawdi Sani, Alexandria Zarko and Boyi Zheng, for making this possible.

QR code to a dissertation chapter

QR code to this poster (with references).

Reference

Peter Austin. A grammar of Diyari, South Australia. Cambridge University Press Cambridge, 1981.

Ellen Broselow and John McCarthy. A theory of internal reduplication. 1983.

Gabriela Caballero. "Templatic backcopying" in Guarijio abbreviated reduplication. Morphology, 16:273–289, 2006.

Maria Gouskova. The reduplicative template in Tonkawa. Phonology, 24(3):367–396, 2007.

Bruce Hayes and May Abad. Reduplication and syllabification in ilokano. Lingua, 77(3-4): 331–374, 1989.

Sharon Inkelas and Laura J Downing. What is reduplication? Typology and analysis part 1/2: The typology of reduplication. Language and linguistics compass, 9(12):502–515, 2015.

Frantisek Lichtenberk. A grammar of Manam. Oceanic linguistics special publications, (18):i–647, 1983.

John J McCarthy and Abigail Cohn. Alignment and parallelism in Indonesian phonology. 1998.

John J. McCarthy and Alan S. Prince. Prosodic morphology. Ms., University of Massachusetts, Amherst, and Brandeis University, Waltham, Mass., 1986.

Elliott Moreton. Analytic bias and phonological typology. Phonology, 25(1):83–127, 2008.

Colin Wilson. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. Cognitive science, 30(5):945–982, 2006.

Kie Zuraw. Vowel reduction in Palauan reduplicants. In Proceedings of the 8th Annual Meeting of the Austronesian Formal Linguistics Association [AFLA 8], pages 385–398, 2002.