

Prosodic conditioning: An instrumental production study of Tagalog **o/u** variation

Darlene Intlekofer
Linguistics Program, The CUNY Graduate Center
Committee: Jason Bishop (Chair), Dianne Bradley, Robert Vago
Second Qualifying Paper
May 1, 2015

1

Overview

- ❖ **o** and **u** alternate in the native Tagalog lexicon
- ❖ However, there is free variation in reduplicants*, e.g., /puno-puno/ 'overflowing'
[puno-puno] ~ [punu-puno]
- ❖ This study instrumentally investigates this free variation and probes for an effect of prosody.

2

2

Outline

1. Introduction

- ❖ The data: Tagalog back vowels
- ❖ Zuraw's (2009) study
- ❖ Present study: Questions and predictions

2. The Experiment

- ❖ Methods
- ❖ Analysis 1 and results 1
- ❖ Analysis 2 and results 2

3. General Discussion

- ❖ Research questions and predictions
- ❖ Limitations and future research

3

3

Tagalog Vowel Inventory

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

Minimal pairs in loans:

<i>lu</i>		<i>lo</i>	
u <i>so</i> 'fad'		o <i>so</i> Sp. 'bear'	
u <i>tas</i> 'hole'		o <i>tas</i> Sp. 'boots'	
u <i>ro</i> 'think'		o <i>ro</i> Sp. 'choir'	
u <i>ka</i> Sp. 'fountain'		o <i>ka</i> Sp. 'vowel'	

4

4

Tagalog Back Vowels

- ❖ In the native lexicon, however, **o** and **u** are in complementary distribution:
 - [o] surfaces in the final syllable; raising is blocked
 - /halo/ 'mix'
 - [halo] cf. *[halu]
 - [u] surfaces in suffixed words; raising is obligatory*
 - /halo+in/ 'to mix together'
 - [haluin] cf. *[haloin]

5

5

Tagalog Optionality

- ❖ The alternation is not straightforward in reduplicated words.
 - /halo-halo/ 'ice dessert'
 - [halu-halo] cf. * [halu-halu]
 - [halo-halo] cf. * [halo-halu]

6

6

Zuraw's (2009) Study

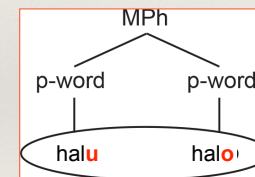
- ❖ Case study: **o/u** alternation
 - Web corpus (Zuraw, 2006) of ~20 million
 - Investigated the effects of frequency
 - Looked at the rate of **u** spellings
- ❖ Findings
 - Lexical frequency effects in unsuffixed reduplicants: more "u" spellings in higher frequency words.
 - Reduplicative identity effect:
 - /halo-halo+an/ "halu-haluan" cf. *"halo-haluan"
 - 'very well mixed'

7

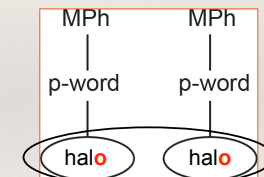
7

Zuraw's (2009) Study, cont.

- ❖ Prosody-based explanation
 - Lexical frequency influences which prosodic structure is accessed, e.g., /halo-halo/ 'ice dessert'



- Single prosodic unit
- Higher frequency item
- "u" surfaces



- Separate prosodic units
- Lower frequency items
- "o" surfaces

8

8

Summary of o/u

❖ Straightforward alternation

- [o] surfaces when it's final in the word
- [u] surfaces in suffixed forms

❖ Optionality (Zuraw, 2009)

- Compound-reduplicated words
- Found that it correlates with frequency, but not when the reduplicants are suffixed
- Conditioned by a prosodic structure assignment that is sensitive to lexical frequency

9

9

The Current Study

❖ No existing data that:

- Describe the phonetic details of the Tagalog o/u optionality
- Support Zuraw's (2009) findings
- Provide support for the presence of a prosodic boundary

❖ Research questions and predictions

- Q1: Is there variation? Does [o] or [u] surface?
- Q2: Is there gradience?
- Q3: Are there frequency influences?
- Q4: Is there evidence for a relation with prosody?

10

10

Q1: Variation?

❖ Unsuffix reduplicants

- High frequency items: more [u] productions
- Mid frequency items: variable
- Low frequency items: more [o] productions

❖ Suffixed reduplicants

- All [u] productions (recall the reduplicative identity effect)

11

Q2: Gradience?

- ❖ A large number of words have o spellings most of the time, the 2nd largest group have u most of the time, but there's also a lot of within-item variation.

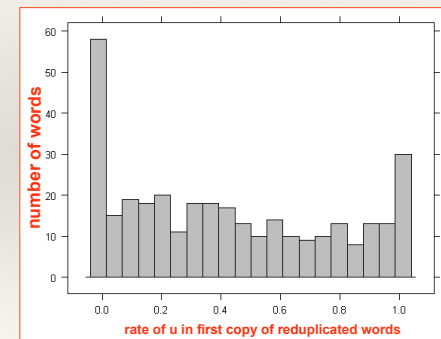


Fig. 8 (Zuraw, 2009) reproduced

12

Q2: Gradience?, cont.

- ❖ A comparison of the rate of **u**-use in first copy vowels (L) versus the rate of tapping (R)
- ❖ Fig 1: Binary choice between the tap or the stop; very little within-item variation

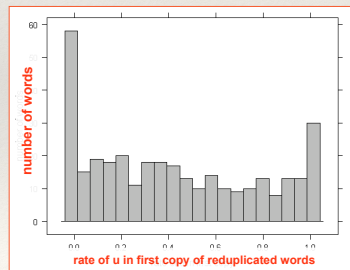


Fig. 8 (Zuraw, 2009) reproduced

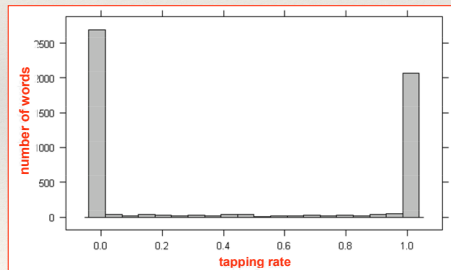


Fig. 1 (Zuraw, 2009) reproduced

13

13

Q3: Frequency Effects?

- ❖ As frequency gets lower, the preference for **o** increases. In the higher frequency ranges, the rate of **u**-use is more evenly distributed.

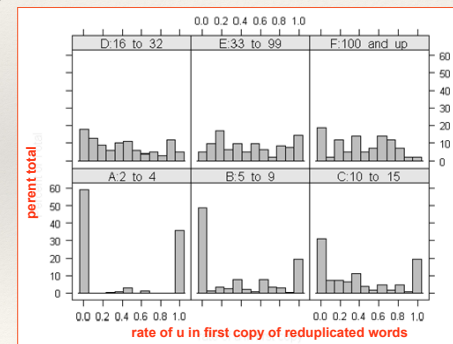
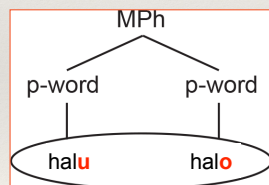


Fig. 9 (Zuraw, 2009) reproduced

14

Q4: Relation with Prosody?*

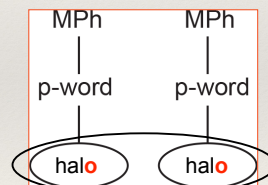
- ❖ The prosodic structure affects the vowel variant.
- ❖ The prosodic boundary in the item with first-copy **[o]** should be bigger than the boundary for first-copy **[u]**.
- ❖ Segmental durations: final lengthening and initial strengthening



- Single prosodic unit
- Higher frequency item
- "u" surfaces

15

15



- Separate prosodic units
- Lower frequency items
- "o" surfaces

Outline

1. Introduction

- ❖ The data: Tagalog back vowels
- ❖ Zuraw's (2009) study
- ❖ Present study: Questions and predictions

2. The Experiment

- ❖ Methods
- ❖ Analysis 1 and results 1
- ❖ Analysis 2 and results 2

3. General Discussion

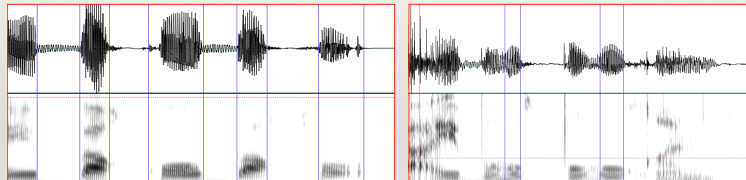
- ❖ Research questions and predictions
- ❖ Limitations and future research

16

16

Methods: Stimuli, cont.

- ❖ A production task: reduplicated words and their corresponding suffixed forms.



/bako-bako/
'rough'

Sentence B
1st repetition

/buhok-buhok+an/
'a lot of hair'

Sentence A
1st repetition

17

17

Methods: Stimuli, cont.

- ❖ Selected from Zuraw's (2006) web corpus
 - Subset tagged for frequency and rate of u-use
 - Provided to J. Bishop by K. Zuraw and used for this study.
- ❖ Selection based on the frequencies ranges:
 - Low (2-9); Mid (10-18); High (19-32)
- ❖ Additional criteria:
 - Tagalog roots selected only in attested forms, vetted by 2 non-participants*
 - Fairly easy to segment, i.e., vowels flanked by obstruents and nasals

18

18

Methods: Stimuli, cont.

- ❖ Example stimuli from each frequency range

Frequency	Baseline condition	Test condition
Low	buhuk-buhukan 'play or fake hair'	buhuk-buhok ~ buhok-buhok 'lots of hair'
	lutu-lutuin 'to overcook'	lutu-luto ~ luto-luto 'cooked'
Mid	bungkus-bungkusin 'to make into a wad'	bungkus-bungkus ~ bungkos-bungkos 'bunch/wad'
	yapus-yapusin 'to hug tightly'	yapus-yapus ~ yapos-yapos 'act of hugging tightly'
High	dugu-duguin 'to make bloody'	dugu-dugo ~ dugo-dugo 'very bloody'
	butu-butuhan 'a lot of bones'	butu-buto ~ buto-buto 'bones'

19

19

Methods: Stimuli, cont.

- ❖ 60 test items
 - 10 compound reduplicated words and their suffixed forms for each of the 3 frequency ranges
- ❖ 48 filler items
 - 24 compound reduplicated words and their suffixed forms for each of the 3 frequency ranges

20

20

Methods: Carrier Sentences*

❖ Design of production task

- Reduce effects of orthography
- Encourage naturalistic speech production
- Increase token count

Sentence A:

Ang unang salita ay [tatlo], at ang pangalawang salita ay [____].

(The first word is [three], and the second word is [____].)

Sentence B:

Ang paborito kong salita ay [____].

(My favorite word is [____].)

21

21

Methods: Participants

- ❖ 13 female native Tagalog speakers
- ❖ Mean age: 36 years
- ❖ LOR: 0 to 25 years
- ❖ AoA: 14-22 years
- ❖ Different hometowns (8 represented)
- ❖ 11 use a 2nd home language; one has 2 additional home languages; 2 reported English as their home language
- ❖ All received monetary compensation

22

22

Methods: Test Phase

- ❖ 240 items per speaker: two repetitions of the 60 (test and control) items x2/slide
- ❖ Stimuli were pseudo-randomized
- ❖ Break period in between trial repetitions
- ❖ Self-paced

23

23

Methods: Analysis 1

- ❖ Goal: Provide data for Q1, Q2, Q3
- ❖ Phonetic transcriptions were used as the measure of analysis
- ❖ Vowel labeling in reduplicants
 - 4 categories — [u], [o], “?”, “other” — for all first-copy vowels in both unsuffixed and suffixed compounds
 - Productions from sentences A and B were combined and analyzed for this study

24

24

Methods: Analysis 1, cont.

❖ Acoustic analysis

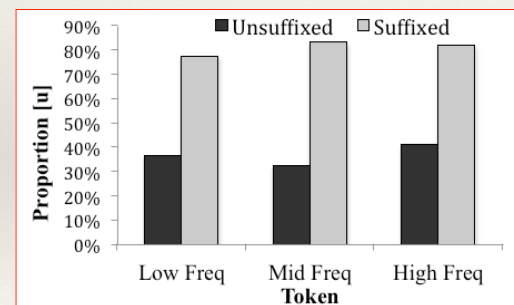
- Properties that correlate with their assignment to [u], [o], and “?” — mainly F1; also duration
- The intervals were defined by a strong F2, marking the vowel's onset and offset; this interval also defined the vowel's duration.
- F1 values were extracted from a stable region from the interval's midpoint using a Praat script

25

25

Results: Analysis 1

❖ Assignment of vowel variants: Overall pattern



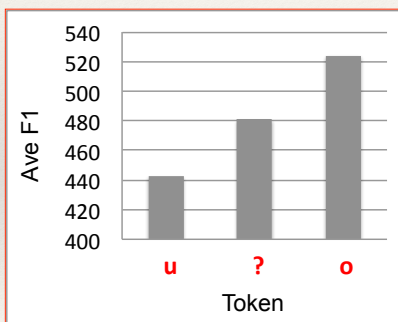
- Some variation in this vowel for both unsuffixed and suffixed forms
- [u] is more common in the suffixed versus the unsuffixed forms ($p < .001$).
- There was no main effect of frequency on vowel raising ($p > .10$)
- Frequency was not a reliable predictor in unsuffixed reduplicants: HIGH had more [u] tokens (relative to the suffixed items), but this was significant only when compared to MID and not to LOW.

26

26

Results: Analysis 1, cont.

❖ Acoustic properties of ambiguous vowels: F1



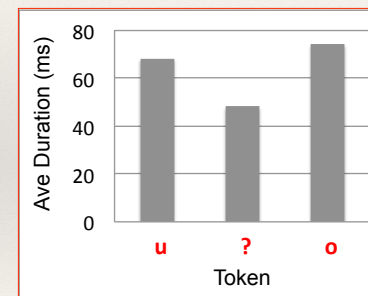
- Vowels assigned ? had significantly larger F1 values than [u], and significantly smaller F1 values than [o].

27

27

Results: Analysis 1, cont.

❖ Acoustic properties of ambiguous vowels: Duration



- Vowels assigned ? had a significantly shorter duration compared to both [u] and [o].

28

28

Methods: Analysis 2

- ❖ **Goal: Provide data on Q4**
- ❖ **Measure of analysis: Segmental durations**
 - Final lengthening and initial strengthening
 - Subset of unsuffixed reduplicants; Cs were obstruents
 - Tokens produced with a large pause between the two copies were excluded

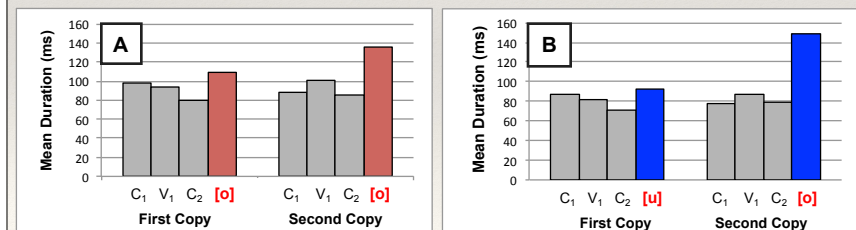
Reduplicant	Gloss	Reduplicant	Gloss
/bago-bago/	'new; more recent'	/buto-buto/	'bones'
/bako-bako/	'rough'	/dugo-dugo/	'bloody'
/buko-buko/	'node'	/puno-puno/	'overflowing'

29

29

Results: Analysis 2

- ❖ **Final lengthening**
 - First-copy [o]s (A) are longer than first-copy [u]s (B), based on transcriptions in Analysis 1
 - Relative effect: first-copy [u]s are significantly shorter compared to their second-copy vowel counterparts than [o] is compared to *its* counterpart.

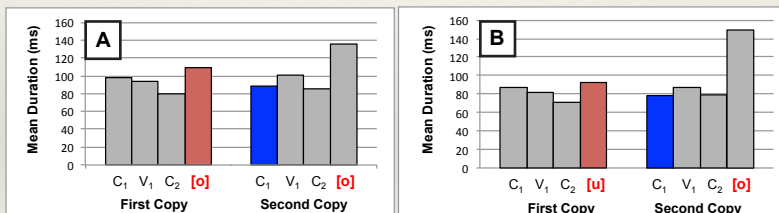


30

30

Results: Analysis 2, cont.

- ❖ **Initial strengthening**
 - Consonants following [o] productions were slightly longer; however, this was not found to be significant ($p > .10$).



- ❖ **Discarded tokens**
 - 26 tokens discarded for first-copy [o]s and only 2 for first-copy [u]s.
 - Marginal significance for pauses following [u] versus [o] ($p = .07$).

31

31

Outline

1. Introduction
 - ❖ The data: Tagalog back vowels
 - ❖ Zuraw's (2009) study
 - ❖ Present study: Questions and predictions
2. The Experiment
 - ❖ Methods
 - ❖ Analysis 1 and results 1
 - ❖ Analysis 2 and results 2
3. General Discussion
 - ❖ Research questions and predictions
 - ❖ Limitations and future research

32

32

Q1: Variation

❖ Is there evidence for variation? Does /u/ or /o/ surface? **YES.**

- Results from Analysis 1 are in line with Zuraw's (2009) findings: unsuffixed reduplicants like /halo-halo/ showed far fewer [u] tokens than for suffixed forms like /halo-halo+an/.
- Though variation was found in suffixed forms, it was trending in the expected direction.

33

33

Q2: Gradience

❖ Is there gradience? **YES**

- Analysis 1 showed that this perceptually in-between vowel was also acoustically in between [u] and [o].

34

34

Q3: Frequency Effects

❖ Are there frequency influences on this gradience? **NOT REALLY.**

- No strong relationship between variation and frequency for both suffixed forms (in line with Zuraw) and unsuffixed forms (not in line with Zuraw).
- Suggestive trends, however: there were more [u] tokens in the high frequency range compared to the mid range, but not compared to the low range.

35

35

Q4: Relation with Prosody

❖ Is there evidence for a relation with prosody? **MOSTLY YES**

- ❖ Final lengthening of first-copy [o] vowels compared to [u] vowels suggests that [o]s tend to occur before a larger boundary than [u] productions.
- ❖ A significant relative effect suggested a difference in the overall prosodic structure of the compound.

36

36

Q4: Relation with Prosody, cont.

❖ Is there evidence for a relation with prosody?

MOSTLY YES

- Greater pauses following [o] productions in the dropped tokens, suggesting a larger prosodic boundary.
- Initial strengthening. **NO**
 - No statistically significant evidence for any initial strengthening effects that correlated with vowel production in the compounds.

37

37

Discussion, cont.

❖ Conclusion

- Attested o/u optionality in native Tagalog reduplicants.
- The current paper investigated Zuraw's proposal of lexically-sensitive prosodic structures using data from a web corpus
- Findings provided partial support
- If prosody story is right, then frequency doesn't fit in so neatly.

38

38

Discussion, cont.

❖ Contributions:

- Variation is described in the literature but not as much for compound reduplicants.
- New instrumental production data
- Influence of prosody on segmental alternations

39

39

Discussion, cont.

❖ Limitations and future research:

1. Intended to be a first look. More that can be done with the data that was collected.
2. No existing model of Tagalog sentence prosody to mark boundary tones
3. Using a second and third transcriber
4. Using a different pool of participants to rule out dialectal differences and language attrition

40

40

Thank you!

41

References

- Anttila, A. (2012). Modeling phonological variation. In A.C. Cohn, C. Fougerson, and M. Huffman (Eds.), *The Oxford Handbook of Laboratory Phonology*, 76-91. Oxford, England: Oxford University Press.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). "Package lme4". Version 1.0-5 (10/25/2013). Available at <http://lme4.r-forge.r-project.org/>.
- Bybee, J. (2000). The phonology of the lexicon: evidence from lexical diffusion. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (65-85). Stanford, CA: CSLI Publications.
- Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer (Version 5.3.11) [Computer program]. Available from <http://www.praat.org>.
- Chan Yap, G. (1973). *Hokkien Chinese borrowings in Tagalog*. Manila: Ateneo de Manila University dissertation.
- Cho, T., Sun-Ah, J., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, 30(2), 193-228. Cho, T., Yoon, Y., & Kim, S. (2014). Effects of prosodic boundary and syllable structure on the temporal realization of CV gestures in Korean. *Journal of Phonetics*, 44, 96-109.
- Coetzee, A. 2009. Phonological variation and lexical frequency. In A.Schardl, M. Walkow, & M. Abdurrahman (Eds.), *NELS 38*(1), 189-202. Amherst, MA: GLSA.
- Coetzee, A. 2009. An integrated grammatical/non-grammatical model of phonological variation. In Y. Kang, et al., eds. *Current Issues in Linguistic Interfaces 2*, 267-294. Seoul: Hankookmunhwasa.

42

42

References, cont.

- English, L. 1986. *Tagalog-English Dictionary*. Manila: National Store.
- Fougerson, C. & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America*, 101:3728-3740
- French, K.M. (1988). *Insights into Tagalog reduplication, infixation and stress from nonlinear phonology*. V. Poulter & W. Merrifield (Eds.). Dallas: Summer Institute of Linguistics and University of Texas at Arlington.
- Jun, S. (1993). *The Phonetics and Phonology of Korean Prosody*. (Unpublished doctoral dissertation). The Ohio State University, Columbus, Ohio.
- Keating, P., Cho, T., Fougerson, C., Hsu, C. (2003, Fall). *Domain-initial articulatory strengthening in four languages*. Invited departmental colloquium presented at University of California, Santa Barbara, Santa Barbara, CA.
- Patrick, P.L. (1992). Creoles at the intersection of variable processes: -t, d deletion and past-marking in the Jamaican neselect. *Language Variation and Change*, 3:171-189.
- R Development Core Team (2014). R: a language and environment for statistical computing. Version 3.1.2. Vienna: R Foundation for Statistical Computing. Available at <http://www.r-project.org>.
- Ramos, T., & Cena, R.M. (1990). *Modern grammatical explanations and exercises for non-native speakers*. Honolulu, HI: University of Hawaii Press.

43

43

References, cont.

- Schachter, P., & Otones, F.E. (1972). *Tagalog Reference Grammar*. Berkeley, CA: University of California Press.
- Turk, A., Nakai, S., Sugahara, M. (2006). *Acoustic segment durations in prosodic research: a practical guide*. In Sudhoff, S., Lenertova, D., Meyer, R., Pappert, S., Augurzy, P., Meinek, I., Richter, N., & Schliesser, J. (Eds.), *Methods in empirical prosody research* (1-27). New York: Mouton de Gruyter.
- Wolff, J.U. (1989). Malay Borrowings in Tagalog. In: *Southeast Asian History and Historiography* (Festschrift for D.G.E. Hall), C.D. Cowan and O.W. Wolters (Eds.), 345-367. Ithaca and London: Cornell University Press.
- Zuraw, K. (2006). Using the web as a phonological corpus: a case study from Tagalog. EACL-2006: Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics/Proceedings of the 2nd International Workshop on Web As Corpus. Pp. 59-66.
- Zuraw, Kie (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog. *Language*, 83, 277-316.
- Zuraw, K. (2009). Frequency influences on rule application within and across words. *Proceedings of CLS (Chicago Linguistic Society)* 43.

44

44