

# Prosodic indicators of phrase structure in Tagalog transitive sentences

Henrison Hsieh, McGill University  
henrison.hsieh@mail.mcgill.ca

AFLA 23, Tokyo University of Foreign Studies  
11 June 2016

## 1 Introduction

- Tagalog Phrase Structure
  - No consensus in analysis despite the amount of work devoted to it
  - Wide range of different approaches and viewpoints
    - \* e.g., Aldridge 2004, Guilfoyle et al. 1992, Kaufman 2009, Kroeger 1993, Rackowski and Richards 2005, etc.
  - Interrelated phenomena: voice, case marking, extraction restrictions
  - Conflicting evidence: constituency tests, binding, definiteness/specificity
- Main point of interest: Constituency
  - What parts of a Tagalog sentence (if any) form tighter units?
  - Does this correlate with any property of the syntactic objects involved?
    - \* Linear position, verb form, thematic role, case marking, etc.
- This study
  - Experimental study looking at prosodic information
  - Main finding: Verbs are durationally shorter in certain configurations
  - Largest determining factor: case marking on the following argument
  - Other findings: Experimental confirmation of word order preferences, duration facts for the first argument

## Roadmap

- §2. Background
- §3. Experimental methods
- §4. Results and discussion
- §5. Summary and conclusion

## 2 Background

### 2.1 Tagalog basics

- Verb-initial language that exhibits so-called Philippine-type alignment
  - Verbal morphology correlates with the thematic role of the syntactically prominent argument in the clause
- **Case Markers:**
  - Two markers that appear on core arguments: *ang*, *nang*<sup>1</sup>
  - *Ang* marks the syntactically prominent argument in a clause; called variously: subject, topic, focus, trigger, pivot, etc.
  - *Nang* marks other core arguments ( $\approx$  elsewhere case)
- **Voice Morphology:**<sup>2</sup>
  - Two voice forms considered: **Agent Voice (AV)**, **Patient Voice (PV)**
  - In Agent Voice: The agent is syntactically prominent (in 1a)
  - In Patient Voice: The patient/theme is syntactically prominent (in 1b)

<sup>1</sup>This marker is standardly spelled *ng* in the orthography. However, to avoid potential confusion for this talk, I use a spelling that better reflects its phonological form: *nang*.

<sup>2</sup>Various labels are used in the literature to refer to this system. Common alternatives are “topic” and “focus”. For this talk, I use the term “voice” pretheoretically.

- (1) a. K<um>ain **ang bata** nang isda.  
 <AV>ate    ANG child NANG fish  
 'The child ate fish.'                      *Agent Voice; ang-marked Agent*
- b. K<in>ain nang bata **ang isda**.  
 <PV>ate    NANG child ANG fish  
 'The child ate the fish.'                      *Patient Voice; ang-marked Patient*

- Some degree of **word-order variability**; Compare (1) to (2)

- (2) *Verb-Object-Agent orders*
- a. K<um>ain nang isda **ang bata**.  
 <AV>ate    NANG fish ANG child  
 'The child ate fish.'
- b. K<in>ain **ang isda** nang bata.  
 <PV>ate    ANG fish NANG child  
 'The child ate the fish.'

*cf. (1a)*

*cf. (1b)*

## 2.2 Constituency?

- Do the different word orders reflect different structures?
- Does the voice system figure in to this? How?
- In the literature:
  - “Traditional” VP consisting of verb and thematic patient (partially Rackowski and Richards 2005 and Aldridge 2004)
  - Verb and the *nang*-marked phrase(s) form a constituent (notably Kroeger 1993 and Kaufman 2009)
  - The two DPs form a constituent excluding the verb (partially Aldridge 2004)
- **Main Idea:** Investigate the prosodic properties of “transitive” (two-argument) sentences in Tagalog in various configurations to see if any systematic differences can be found.

## 3 Experimental methods

### 3.1 Stimuli

- Consisted entirely of verb-initial sentences with two arguments
- 2 × 2 × 2 design (8 conditions), crossing...
  - **Verb Form:** Agent Voice vs Patient Voice
  - **Order of arguments:** *ang*-first or *nang*-first
  - Presence or absence of **adjectives** on both arguments
- Chose verbs with interchangeable arguments to keep nouns in-place

Table 1: Sample Experimental Item

Verb	Det	Adjective	Noun	Det	Adjective	Noun
'killed'		'brave'	'whale'		'ferocious'	'shark'
P<um>atay	<i>ang</i>		balyena	<i>nang</i>		pating
P<um>atay	<i>nang</i>		balyena	<i>ang</i>		pating
P<in>atay	<i>ang</i>		balyena	<i>nang</i>		pating
P<in>atay	<i>nang</i>		balyena	<i>ang</i>		pating
P<um>atay	<i>ang</i>	matapang na	balyena	<i>nang</i>	mabangis na	pating
P<um>atay	<i>nang</i>	matapang na	balyena	<i>ang</i>	mabangis na	pating
P<in>atay	<i>ang</i>	matapang na	balyena	<i>nang</i>	mabangis na	pating
P<in>atay	<i>nang</i>	matapang na	balyena	<i>ang</i>	mabangis na	pating

- 16 experimental items (sets of 8 tokens varied as described above)<sup>3</sup>
- Total 128 sentences, no fillers used

### 3.2 Procedure

- 16 native speakers of Tagalog
  - 18 – 45 y.o.
  - From Manila and the surrounding provinces, living in Manila

<sup>3</sup>See the appendix for a summary of the experimental items

- Self-paced production task carried out via Psychtoolbox (Matlab package)
- For each stimulus sentence, participants were instructed to:
  - First read a sentence silently to familiarize
  - Initiate recording by pressing a key, then read the sentence aloud
  - Terminate recording by pressing a key again
  - Rate the naturalness of the sentence on a 1 (worst) – 7 (best) scale
- Whole stimulus set presented to each participant in a pseudorandom order (no consecutive sentences from the same item or condition)
- Each sentence was presented with one of four predetermined frames to anticipate late starts and early stops of the recording:

(3) Alam mo? Pinatay nang balyena ang pating. Yun ang kwento sa akin.  
 ‘Did you know? The whale killed the shark. That’s the story I was told.’

### 3.3 Data processing

- **Truncation:**
  - Leading and trailing silence manually removed from each sound file with the assistance of a Praat (Boersma and Weenink 2013) script
  - Bad sound files (disfluencies, stuttering, etc) were excluded
- **Annotation:** Automatic annotation of word and phone boundaries using the Prosodylab Forced Aligner (Gorman et al. 2011)
- **Measurement:** Acoustic measures extracted for each of seven words of interest (verb, *ang*, *nang*, both adjectives, both nouns) via Praat script
- **Analysis:**
  - Mixed effects linear regression models (R lmerTest package)
  - Help filter out some of the by-item and by-participant variability

## 4 Results and discussion

### 4.1 Naturalness Rating

- Best: PV Verb followed by *nang*-marked agent
- Worst: PV Verb followed by an *ang*-marked patient
- AV constructions have freer word order, in a sense

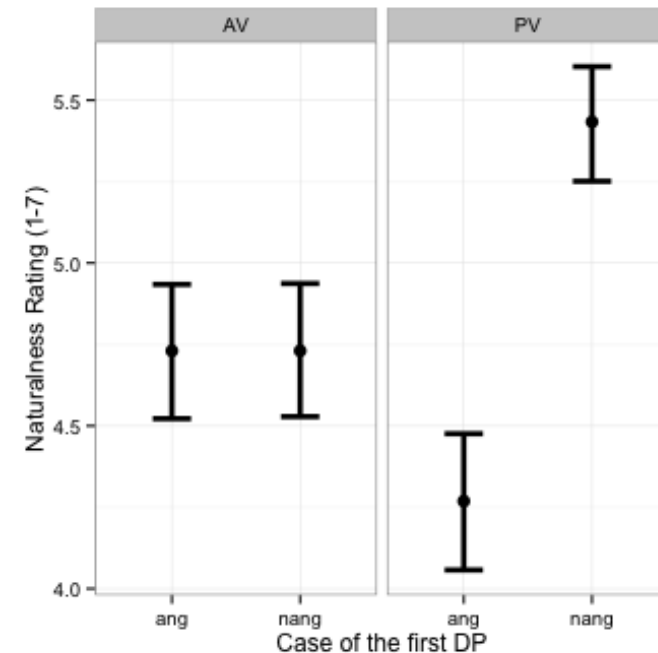


Figure 1: Naturalness Ratings by Voice and Argument Order

## 4.2 Verb duration

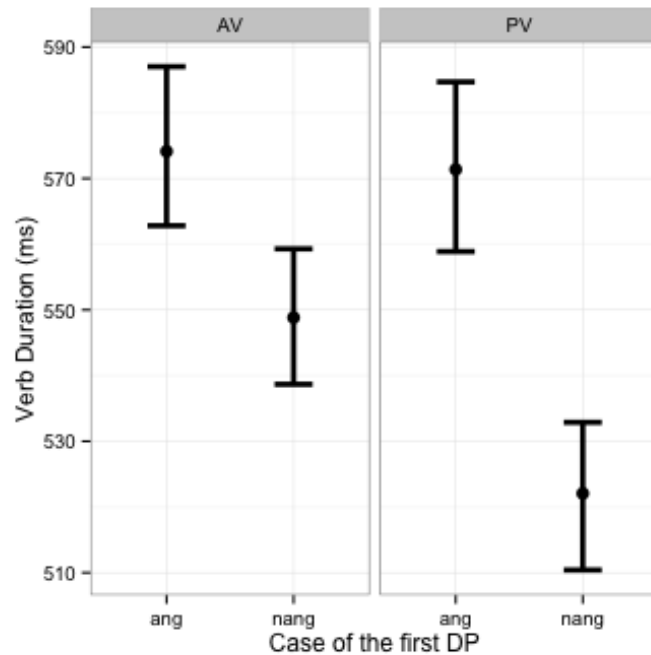


Figure 2: Absolute verb duration by voice and argument order

### • Main Findings

- Largest effect: Shorter before *nang*-marked arguments ( $p < 0.001$ )
- Next largest: PV verbs ( $p < 0.05$ ) and verbs in sentences without adjectives ( $p < 0.001$ ) are shorter
- Significant interaction of voice and argument order: Difference between argument orders is larger in PV than in AV ( $p < 0.05$ )
- Effect of naturalness rating was not significant ( $p = 0.121$ )

### • Other Details

- Effect of voice became non-significant when verb duration was normalized by phone length ( $p = 0.825$ )

- Variable effects on verb duration based on experimental item
  - \* Naturalness Rating: Likely due to the varying pragmatic naturalness of the different items
  - \* Voice Form: Possibly due to the specific forms chosen

## 4.3 First noun duration

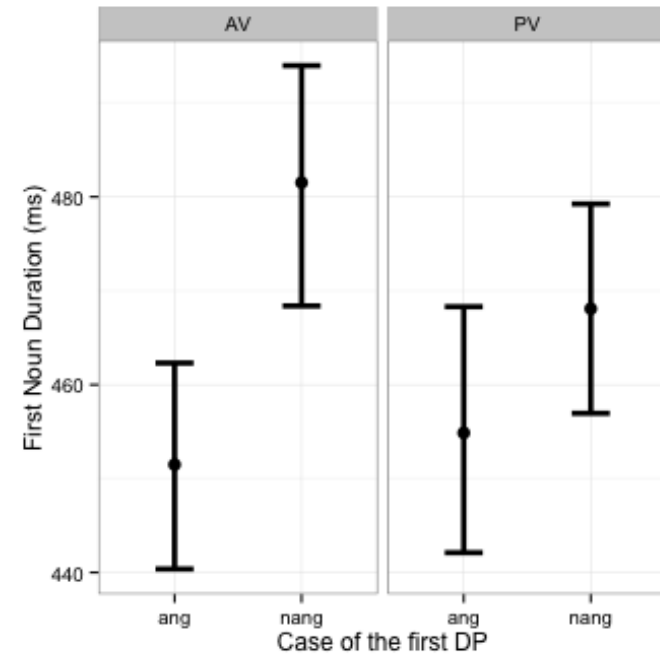


Figure 3: Absolute duration of the first noun by Voice and Argument Order

### • Main Findings

- Largest Effect: Shorter when no adjectives were present ( $p < 0.001$ )
- Smaller Effect: Longer when *nang*-marked ( $p < 0.01$ )
- Smallest Effect: Shorter with a higher rating ( $p < 0.05$ )
- Interaction between voice and arg. order not significant ( $p = 0.384$ )

## 4.4 Discussion

- Implications for constituency
  - Findings from verb duration indicate a stronger, more uniform phrasal boundary before an *ang*-marked argument
  - Findings from first noun duration suggest a stronger boundary after *nang*-marked DP
    - \* Indicative of the verb and the *nang*-marked argument forming a constituent when adjacent to each other
    - \* Unsure how to interpret the shorter duration for the *ang*-marked arguments in this position
  - Possible tighter constituency between verb and *nang*-marked agent than with a *nang*-marked patient?
- Case marking vs Thematic role
  - Ran models using argument order as determined by thematic role instead of case marking
  - Verbs were longer when followed by the theme, and first nouns were longer when they were a theme.
  - Effect was smaller (with verb duration) or non-significant (with first noun duration) in comparison to using case marking
  - Models using thematic role had a larger effect of the interaction between voice and argument order
  - Will have to perform a more thorough model comparison

## 5 Conclusion

- Results corroborate one competing claim in the literature that verbs in Tagalog form tighter constituents with the *nang*-marked arguments
- Prosodic evidence for a thematic-role-based constituency is more tenuous
- Future work: Pitch, ruling out possible phonetic explanations of the data

## 6 Acknowledgements

I would like to thank Lisa Travis, Michael Wagner, and Morgan Sonderegger for their advice and support at various stages of this project, as well as the audience at ETI<sub>3</sub> at McGill, various members of the McGill Linguistics department, and reviewers for helpful comments on this work. I also thank Lisa for funding the trip to Manila on which this data was gathered.

## References

- Aldridge, Edith. 2004. Antipassive and specificity in Tagalog. In *Proceedings of AFLA 11*, volume 11, 1–14.
- Boersma, Paul, and David Weenink. 2013. Praat: doing phonetics by computer. Computer program.
- Gorman, Kyle, Jonathan Howell, and Michael Wagner. 2011. Prosodylab-Aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics* 39:192–193.
- Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis. 1992. Spec of IP and spec of VP: Two subjects in Austronesian languages. *Natural Language & Linguistic Theory* 10:375–414.
- Kaufman, Daniel. 2009. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35:1–49.
- Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog*. Dissertations in Linguistics. Center for the Study of Language and Information.
- Rackowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36:565–599.

## Appendix: Summary of stimuli

	AV form	PV form	English	Adjective 1	Noun 1	Adjective 2	Noun 2	Frame
1.	<i>pumatay</i>	<i>pinatay</i>	'killed'	<i>matapang na</i> 'brave'	<i>balyena</i> 'whale'	<i>mabangis na</i> 'fearsome'	<i>pating</i> 'shark'	(A)
2.	<i>pagdala</i>	<i>dinala</i>	'brought'	<i>itim na</i> 'black'	<i>pusa</i> 'cat'	<i>puting</i> 'white'	<i>daga</i> 'rat'	(B)
3.	<i>kumain</i>	<i>kinain</i>	'ate'	<i>matandang</i> 'old'	<i>lalaki</i> 'man'	<i>malaking</i> 'big'	<i>buwaya</i> 'crocodile'	(C)
4.	<i>humipo</i>	<i>hinipo</i>	'touched'	<i>makulit na</i> 'persistent'	<i>sanggol</i> 'baby'	<i>maamong</i> 'tame'	<i>aso</i> 'dog'	(D)
5.	<i>nangiliti</i>	<i>kiniliti</i>	'tickled'	<i>mabait na</i> 'kind'	<i>doktor</i> 'doctor'	<i>masayang</i> 'happy'	<i>bata</i> 'child'	(A)
6.	<i>bumibili</i>	<i>binibili</i>	'buying'	<i>matabang</i> 'fat'	<i>lapu-lapu</i> '(fish species)'	<i>masiglang</i> 'lively'	<i>talaba</i> 'oyster'	(B)
7.	<i>nangurot</i>	<i>kinurot</i>	'pinched'	<i>pikuning</i> 'upsettable'	<i>nars</i> 'nurse'	<i>maasamang</i> 'wicked'	<i>pasyente</i> 'patient'	(C)
8.	<i>humuli</i>	<i>hinuli</i>	'caught'	<i>matalinong</i> 'smart'	<i>lobo</i> 'wolf'	<i>maliit na</i> 'small'	<i>tigre</i> 'tiger'	(D)
9.	<i>kumagat</i>	<i>kinagat</i>	'bit'	<i>pulang</i> 'red'	<i>ahas</i> 'snake'	<i>mabagal na</i> 'slow'	<i>pagong</i> 'turtle'	(A)
10.	<i>bumati</i>	<i>binati</i>	'greeted'	<i>matangkad na</i> 'tall'	<i>guro</i> 'teacher'	<i>masipag na</i> 'hardworking'	<i>estudyante</i> 'student'	(B)
11.	<i>nanggulat</i>	<i>ginulat</i>	'surprised'	<i>galit na</i> 'angry'	<i>unggoy</i> 'monkey'	<i>malungkot na</i> 'sad'	<i>ibon</i> 'bird'	(C)
12.	<i>nagluto</i>	<i>linuto</i>	'cooked'	<i>mabahong</i> 'smelly'	<i>manok</i> 'chicken'	<i>dilaw na</i> 'yellow'	<i>baboy</i> 'pig'	(D)
13.	<i>nanuntok</i>	<i>sinuntok</i>	'punched'	<i>maruming</i> 'dirty'	<i>ipis</i> 'cockroach'	<i>malinis na</i> 'clean'	<i>langgam</i> 'ant'	(A)
14.	<i>bumangga</i>	<i>binangga</i>	'crashed into'	<i>bagong</i> 'new'	<i>sasakyan</i> 'car'	<i>magarang</i> 'extravagant'	<i>dyip</i> 'jeepney'	(B)
15.	<i>nagbebenta</i>	<i>binebenta</i>	'selling'	<i>malakas na</i> 'strong'	<i>pabo</i> 'turkey'	<i>malinis na</i> 'clean'	<i>maya</i> 'sparrow'	(C)
16.	<i>nananakot</i>	<i>tinatakot</i>	'scares'	<i>masungit na</i> 'grumpy'	<i>bayawak</i> 'monitor lizard'	<i>magandang</i> 'beautiful'	<i>paniki</i> 'bat'	(D)

### Sentence frames

A. Alam mo? ... Yun ang kwento sa akin.

"Did you know? ... That's the story I was told."

B. May nalaman ako. ... Ang galing!

"I found something out. ... Cool!"

C. Sabihin ko daw sa iyo. ... OK?

"I was told to tell you. ... OK?"

D. Sabi daw: ... Totoo kaya?

"They say: ... I wonder if it's true."