Morphology

Trade-off between morphological marking and usage of tones

Jerry Chih-Chun Chang

Department of Computational Linguistics
University of Tübingen

Typology

Content

Research Question

Morphology

Is there a trade-off between morphological marking and usage of tones?

I.e., languages that extensively use tonal marking tend to have less morphological markers.

Tone languages

Tone (Crystal, 2011)

the distinctive pitch level of a syllable

- level tone: a relatively flat pitch at a particular level
- contour tone: a pitch rising or falling over the duration of the syllable

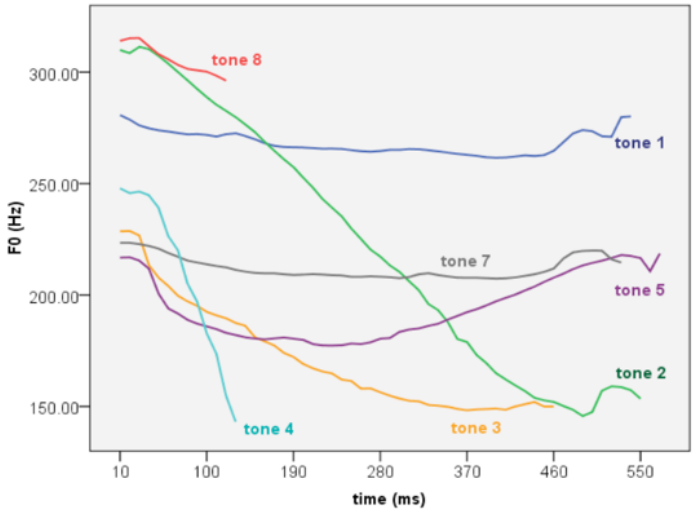
Tone languages

where a tone carried by a syllable/word is an essential feature of the meaning of that word

- grammatical, e.g. Mashi (Atlantic-Congo) (Kutsch Lojenga, 1993) nagánja'I counted' [nágánja] 'I will count'
- lexical, e.g. Taiwanese Southern-Min (Sino-Tibetan)

Tone languages

Taiwanese lexical tones[1][2] (Pan, 2008)



High level	55 (1)	/kun/	軍	‘army’
High falling	51 (2)	/kun/	滾	‘boiling’
Low falling	21 (3)	/kun/	棍	‘baton’
Low checked	21 (4)	/kut/	掘	‘plow’
Low rising	24 (5)	/kun/	裙	‘skirt’
Mid level	33 (7)	/kun/	郡	‘prefecture’
High checked	51 (8)	/kut/	滑	‘slippery’

[1] Checked tones have a voiceless stop coda and via tone sandhi surface as different tones from their unchecked counterparts.

[2] Tone 6 is merged with tone 2

Exponence

On the level of formatives

Exponence

the number of categories grouped in one formative, e.g. number, case, person, or tense (Bickel and Nichols, 2007)

- monoexponential (separative): encodes one category, e.g. English plural -s
- polyexponential (cumulative): encodes more than one category, e.g. Russian genitive (with number marked)

Synthesis

On the level of words

Synthesis

the number of formatives and lexemes in one word (Bickel and Nichols, 2007)

- analytic: allows little affixation, e.g. English future auxiliary 'will'; *more beautiful* (periphrastic)
- synthetic: allows affixation of one lexeme and a handful of formatives , e.g. English past tense -ed; *happier* (synthetic)
- polysynthetic: allows affixation of more than one lexemes and usually multiple formatives

Turkish (Bickel and Nichols, 2007)
tan-ɨş-ɨl-a-ma-dɨk-lar-ɨn-dan-dɨr
know-RECIP-CAUS-PASS-POT-NEG-NZR-PL-3POSS-ABL-3COP
'It is because they cannot be introduced to each other.'

Nominal Categories: Case

Case

"system of marking dependent nouns for the type of relationship they bear to their heads" (Blake, 2001)

Number of Cases (Case Inventory)

languages differ in terms of the kinds and numbers of cases they use (Velupillai, 2012)

- small inventory, e.g. English, 2
- medium inventory, e.g. German, 4

Types of case (Velupillai, 2011)

- Syntactic roles:nominative, accusative, dative, genetive, ...
- Semantic roles: ergative, absolutive, benefactive, comitative, ...

Database

WALS: The World Atlas of Language Structures
http://wals.info/

Levels of Tones

Maddieson (2013)
Non-tonal

No tones (307), e.g. German

without tones

Tonal

Simple tone system (132), e.g. Norwegian

with only a two-way basic contrast, usually between high and low levels

Complex tone system (88), e.g. Mandarin

with a more complex set of contrasts

Levels of Case Exponence

Bickel and Nichols (2013a)

Monoexponential case (71), e.g. Spanish

Polyexponential case (16), e.g. Russian (genitive + number)

- Case + number (8)
- Case + referentiality (specifying host NP as topics or its definiteness of reference) (6)
- Case + TAM (tense-aspect-mood) (2)

No case (75), e.g. Thai

Levels of Case Inventory

Iggesen (2013)

Number of Cases

- No morphological case-marking (100), e.g. Dutch
- 2 case categories (23), e.g. Irish
- 3 case categories (9), e.g. Modern Greek
- 4 case categories (9), e.g. Icelandic
- 5 case categories (12), e.g Latvian
- 6-7 case categories (37), e.g Polish
- 8-9 case categories (23), e.g. Burmese
- 10 or more case categories (24), e.g. Estonian

Levels of Synthesis

Bickel and Nichols (2013b)

Inflectional Synthesis of the Verb

- 0-1 category per word (5), e.g. Vietnamese
- 2-3 categories per word (24), e.g. Burmese
- 4-5 categories per word (52), e.g. Indonesian
- 6-7 categories per word (31), e.g. Arabic (Egyptian)
- 8-9 categories per word (24), e.g. Parahã
- 10-11 categories per word (7)
- 12-13 categories per word (2)

Averaged categories per word

0.5, 2.5, ..., 12.5 as degrees of synthesis

Affixation

Dryer (2013)

Prefixing vs. Suffixing in Inflectional Morphology

- Little affixation (2 or less affixing index) (141), e.g. Tagalog
- Strongly suffixing (more than 80% suffixing) (406), e.g. Korean
- Weakly suffixing (60% - 80% suffixing) (123), e.g.Latvian
- Equal prefixing and suffixing (147), Arabic (Iraqi)
- Weakly prefixing (60% - 80% prefixing) (94), e.g. Swahili
- Strongly prefixing (more than 80% prefixing) (58), e.g. Xhosa

Partitioned into

- little affixation
- moderate affixation

Results

Method

Using **Linear Regression** to model the relationships between **predictors** and **dependent variables**

Predictor: Tonality

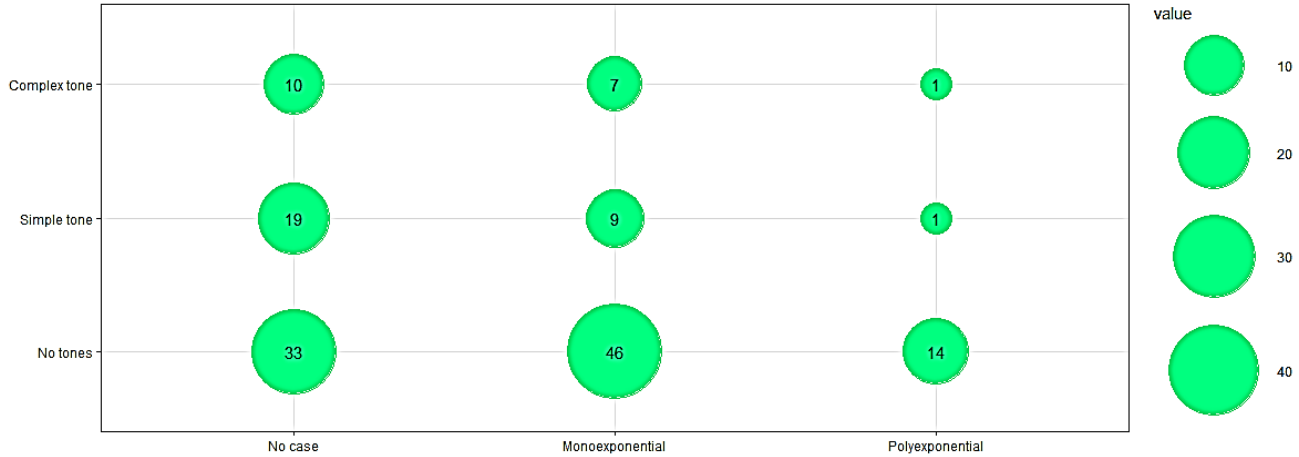
- binary: non-tonal vs. tonal
- three-way contrast: none, simple, complex

Dependent variables: Morphological marking

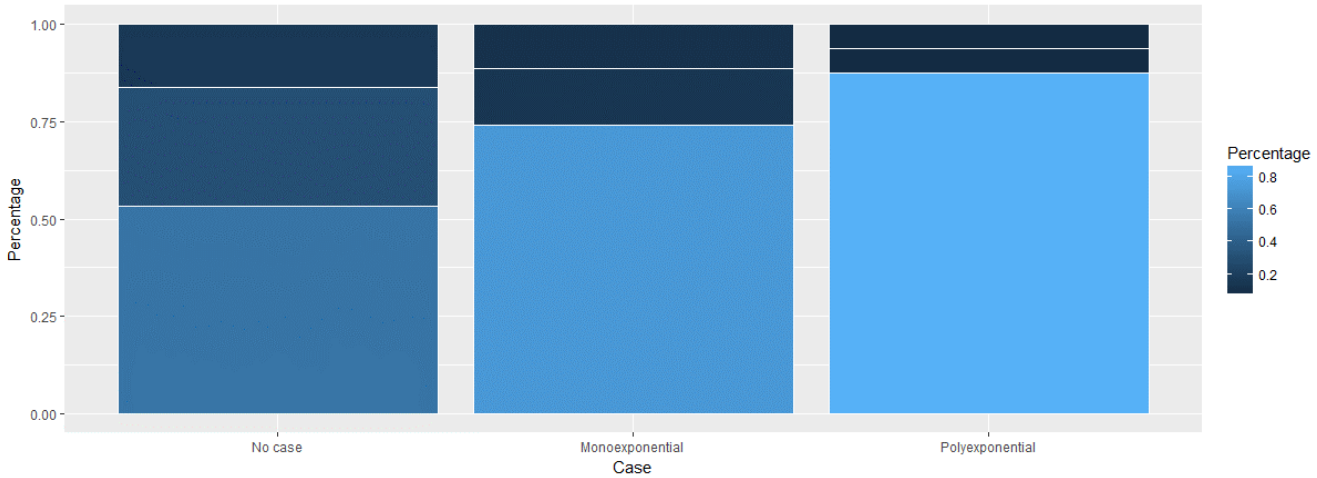
- case exponence (none, mono-, poly-)
- case numbers (none, 2, 3, ..., 10 and more)
- synthesis (0.5, 2.5, ..., 12.5)
- affixation (little, moderate)

Results

Case exponence



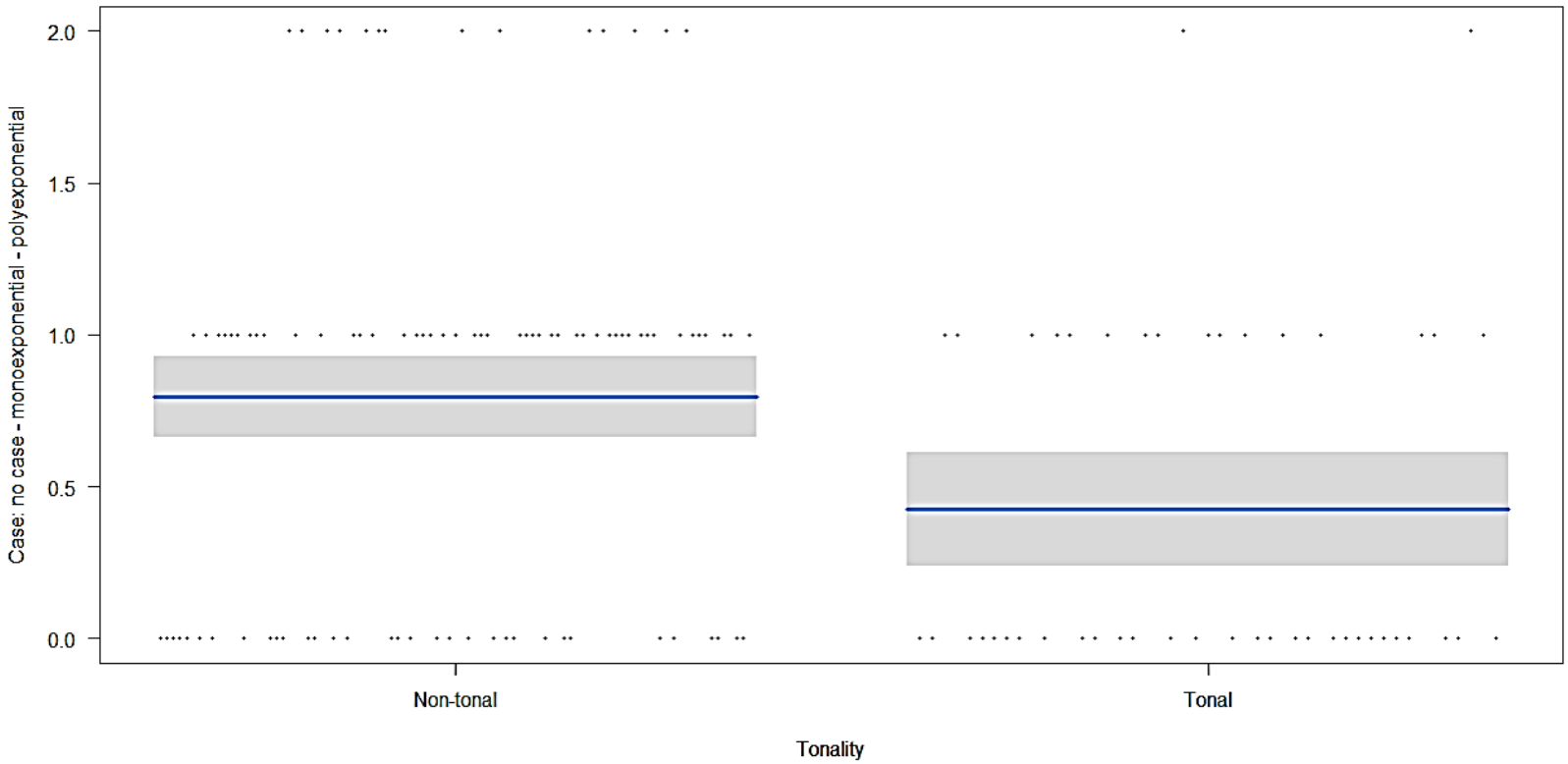
Make-up: no tone - simple tone system - complex tone system



Results

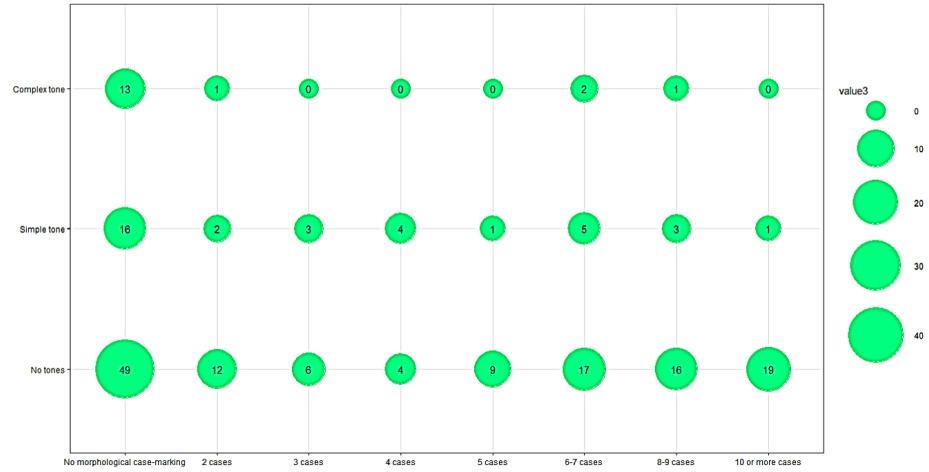
Case exponence: significant

Coefficients: Estimate Std. Error t value Pr($>|t|$)
(Intercept) 0.79570 0.06758 11.775 <2e-16 ***
tonalityTonal -0.37017 0.11663 -3.174 0.00186 **
Residual standard error: 0.6517 on 138 degrees of freedom

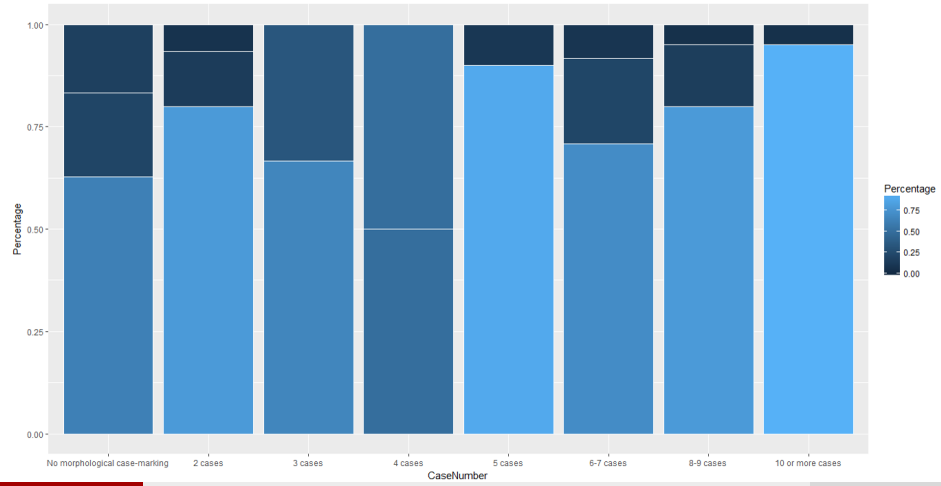


Results

Case numbers



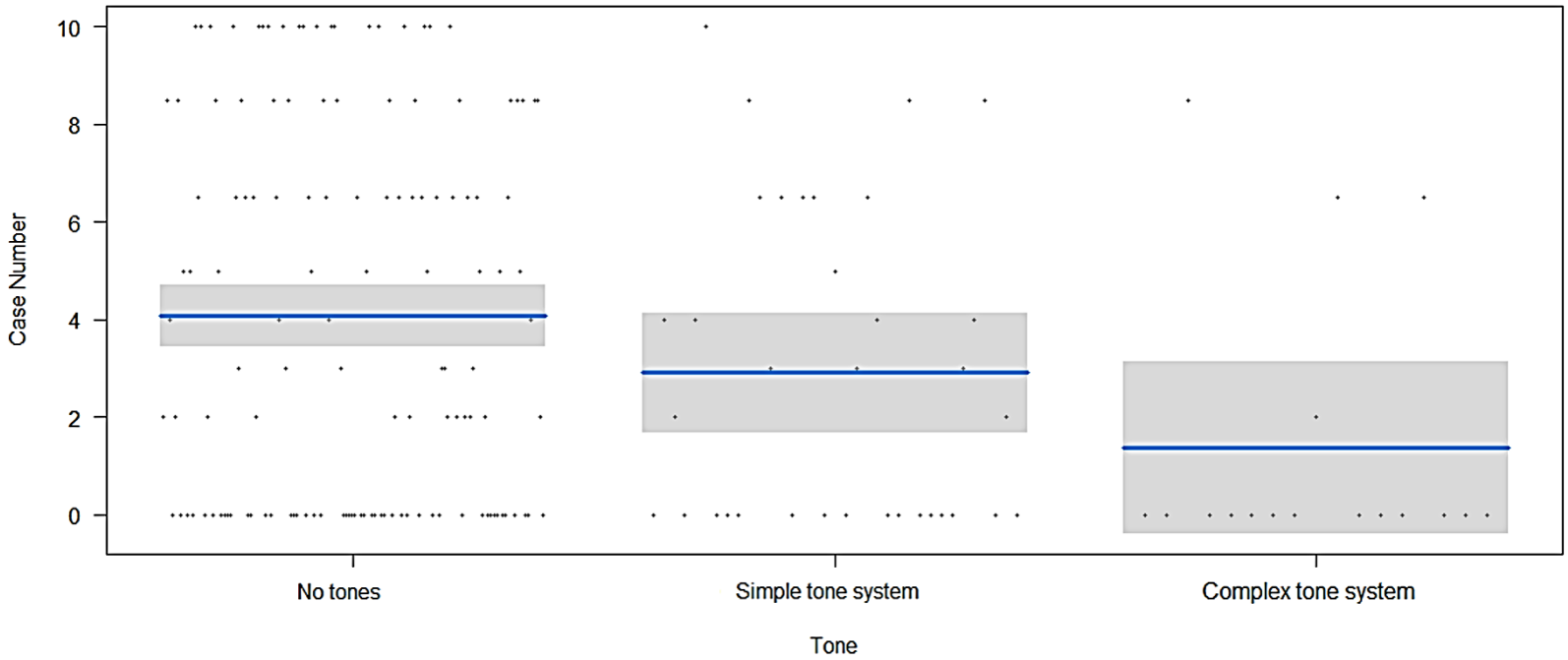
Make-up: no tone - simple tone system - complex tone system



Results

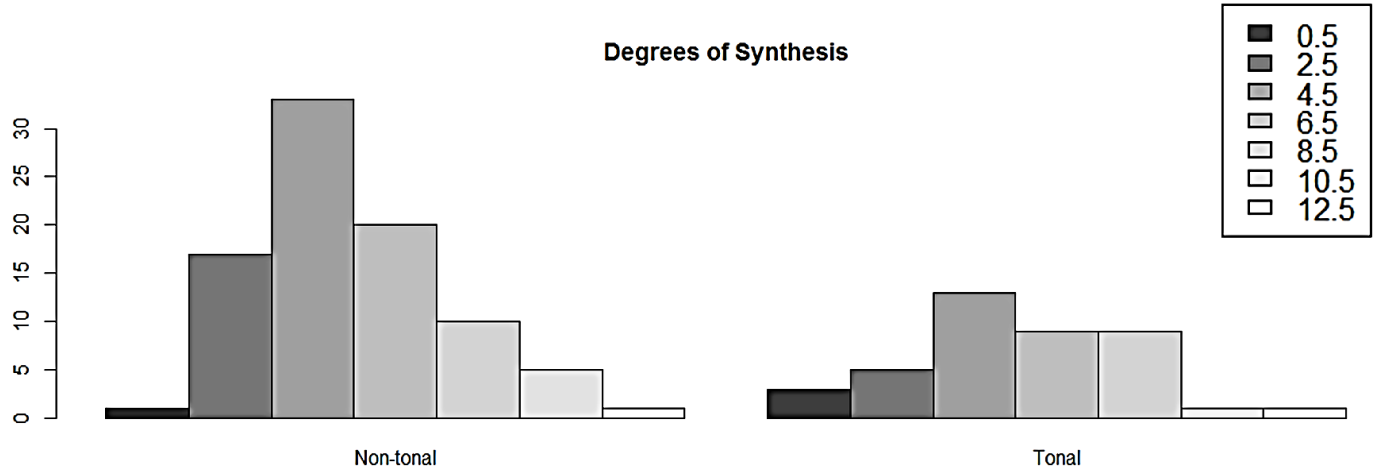
Case numbers: significant

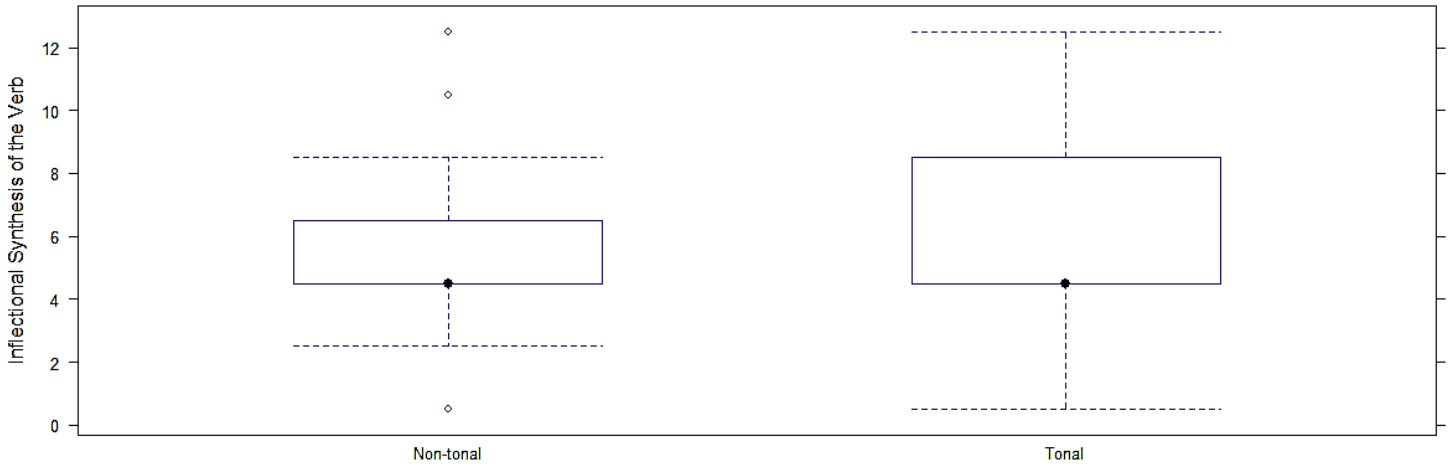
Coefficients: Estimate Std. Error t value Pr($>|t|$)
(Intercept) 4.0871 0.3198 12.781 $<$ 2e-16 ***
ToneComplex tone system -2.7048 0.9467 -2.857 0.00478 **
ToneSimple tone system -1.1728 0.6985 -1.679 0.09487
Residual standard error: 3.674 on 181 degrees of freedom
p-value: 0.008748



Results

Synthesis

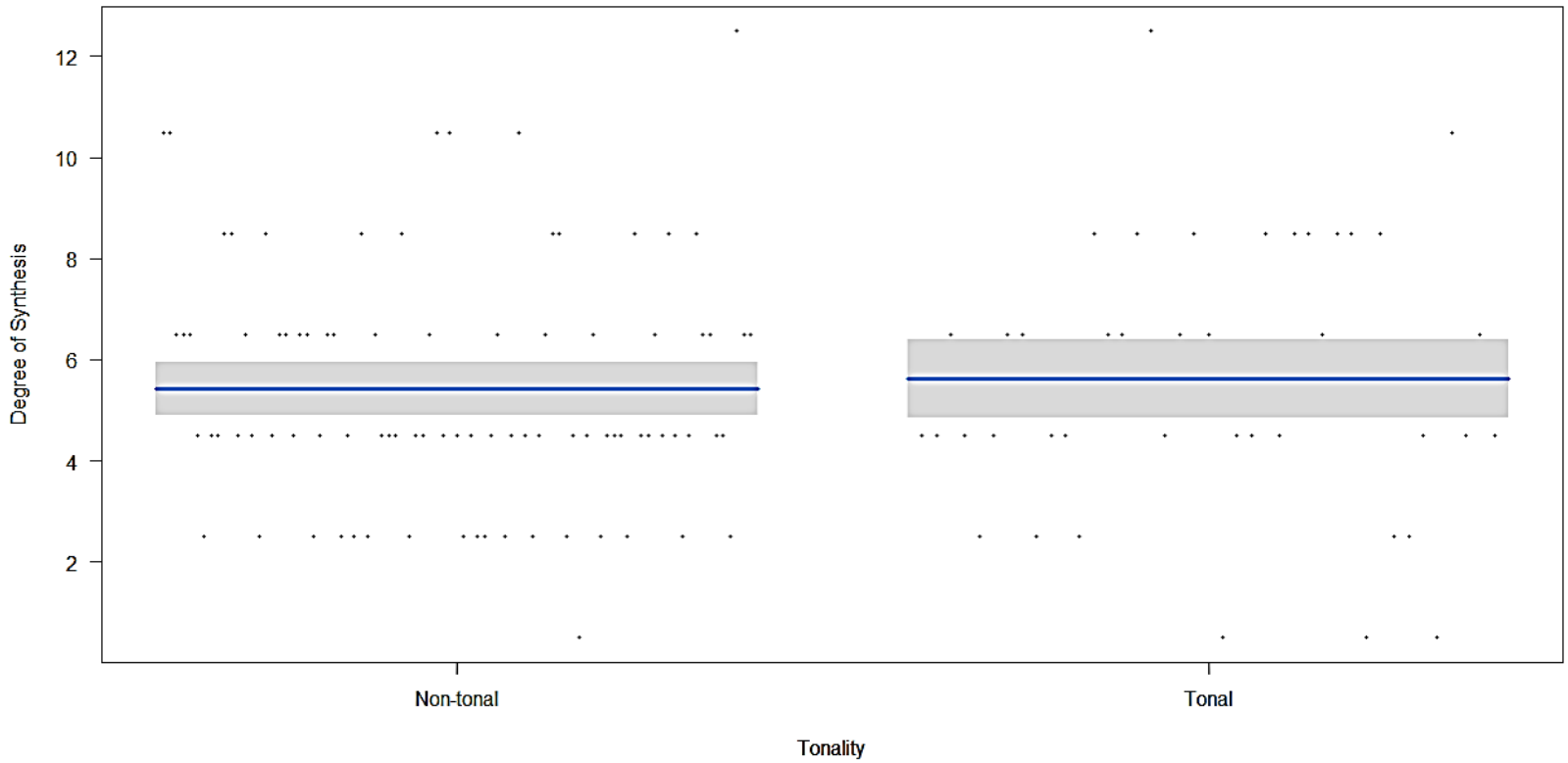




Results

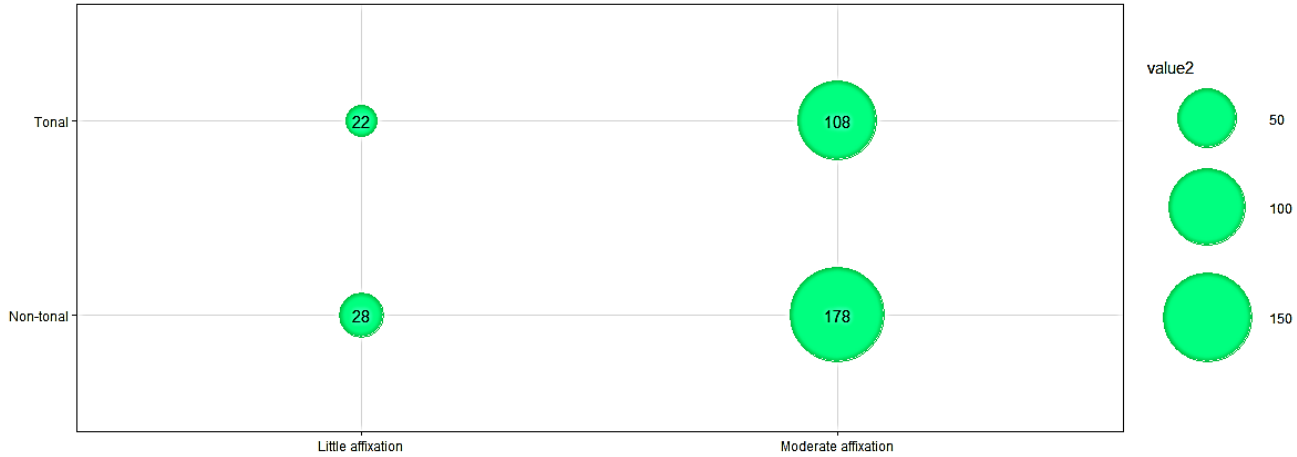
Synthesis:not significant

Coefficients: Estimate Std. Error t value Pr($>|t|$)
(Intercept) 5.4195 0.2686 20.177 <2e-16 ***
tonalityTonal 0.2024 0.4746 0.427 0.67
Residual standard error: 2.505 on 126 degrees of freedom

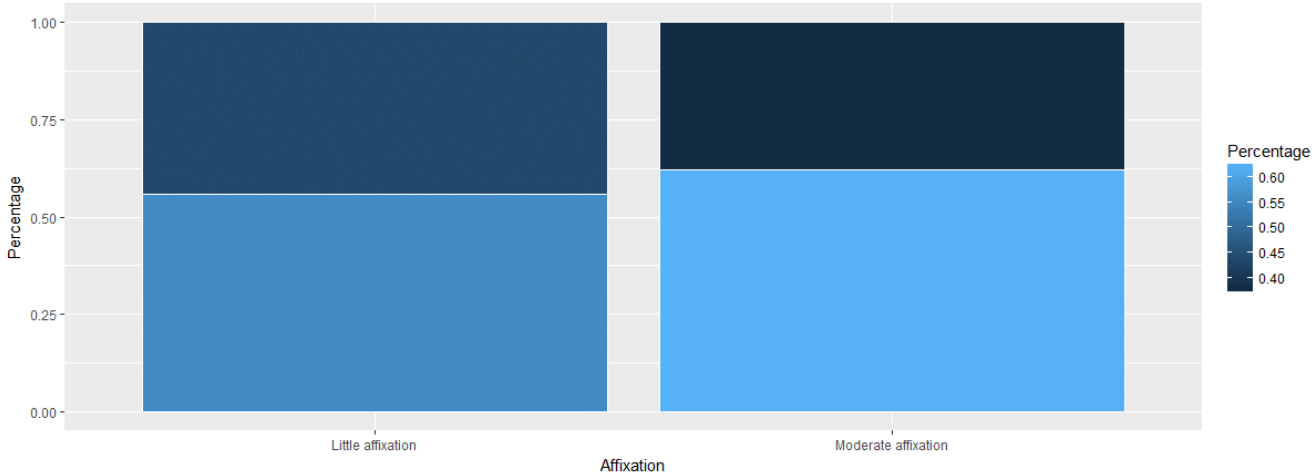


Results

Affixation



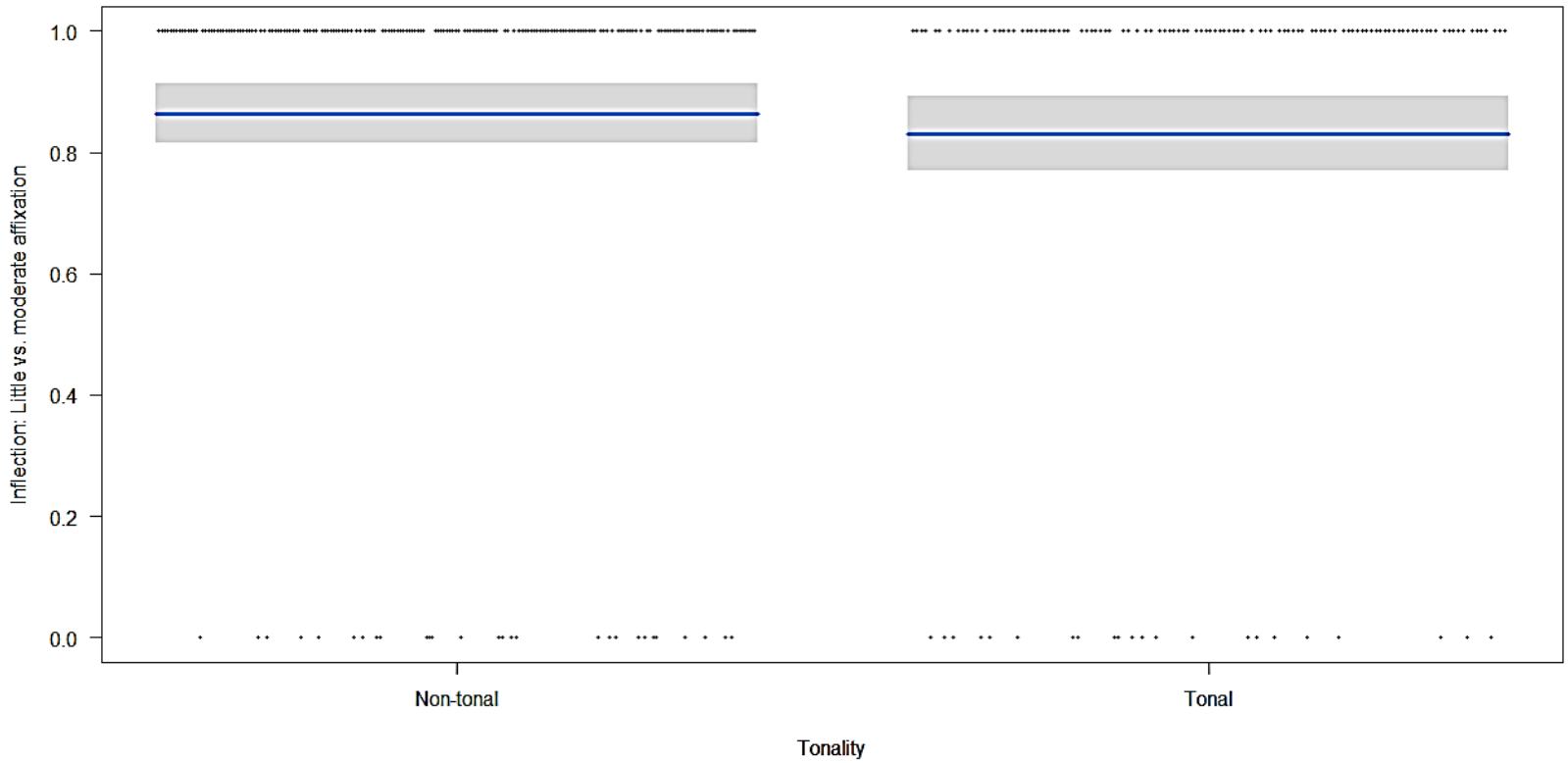
Make-up: no tone - simple tone system - complex tone system



Results

Affixation :not significant

Coefficients: Estimate Std. Error t value Pr($>|t|$)
(Intercept) 0.86408 0.02485 34.779 <2e-16 ***
tonalityTonal -0.03331 0.03994 -0.834 0.405
Residual standard error: 0.3566 on 334 degrees of freedom



Discussions

Results

The trade-off between **morphological marking** and **usage of tones**

is supported by

- Case ($p < 0.05$)
- Number of cases ($p < 0.05$)
- Affixation (not significant)

is NOT supported by

- Synthesis (not significant): sampled non-tonal languages are more than twice as many as tonal-languages.

Discussions

Issues

1. Data sizes collapse after merged.

Tone (527) + Exponence of Case (162) $\rightarrow$ 140
Tone (527) + Number of Cases (237) $\rightarrow$ 184
Tone (527) + Synthesis (145) $\rightarrow$ 128
Tone (527) + Affixation (969) $\rightarrow$ 336

2. The representation of each category is sometimes skewed.

3. Categorization of tones and morphological cases should be improved.

Discussions

Fix

1. Collection with more databases
2. Finer categorization.

Tonal complexity: atonal - simple - moderately complex - complex
in terms of:

- number of tones
- level tone
- contour tone
- lexical (global) vs. grammatical (occasiaonl)
- tone sandhi (allomorphs of tones)

Affixation: little or no affixation - moderate affixation - rich affixation

Discussions

To do list

1. **Depth vs. Breadth**:
Refine the scope or perspective of morphological marking, e.g. **exponence of case** as how dense morphologcail marking is, and **number of cases** as how prevalent case marking is.

2. Use of other databases, e.g. PHOIBLE (Moran et al, 2014).

3. How about relationships of types of affixation (i.e. prefixing and suffixing) and functions of tones (i.e. lexical vs. grammatical), as a sub-hypothesis?

Summary

1. **Case** is a good indicator for tonality in terms of both depth (exponence) and breadth (inventory size).

2. **Synthesis** says little about tonality.

3. **Affixation**, with refined categorization, may offer significant support to our hypothesis.

Questions?

References I

Balthasar Bickel and Johanna Nichol. "Exponence of Selected Inflectional Formatives". In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013a. URL: http://wals.info/chapter/21.

Balthasar Bickel and Johanna Nichols. "Inflectional morphology". In: *Language typology and syntactic description* 3.2 (2007), pp. 169–240.

Balthasar Bickel and Johanna Nichols. "Inflectional Synthesis of the Verb". In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013b. URL: http://wals.info/chapter/22.

References II

B.J. Blake. *Case*. Cambridge Textbooks in Linguistics. Cambridge University Press, 2001. ISBN: 9780521014915. URL: https://books.google.de/books?id=FGAMJ7YAly0C.

D. Crystal. *Dictionary of Linguistics and Phonetics*. The Language Library. Wiley, 2011. ISBN: 9781444356755. URL: https://books.google.de/books?id=3ZPQVuSgDAkC.

Matthew S. Dryer. "Prefixing vs. Suffixing in Inflectional Morphology". In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. URL: http://wals.info/chapter/26.

References III


Oliver A. Iggesen. "Number of Cases". In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. URL: http://wals.info/chapter/49.


C. Kutsch Lojenga. "The writing and reading of tone in Bantu languages". In: *Notes on Literacy* (1993).


Ian Maddieson. "Tone". In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. URL: http://wals.info/chapter/13.


Steven Moran, Daniel McCloy, and Richard Wright, eds. *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2014. URL: http://phoible.org/.

References IV

Ho-Hsien Pan. "Focus and Taiwanese unchecked tones". In: *Topic and Focus*. Springer, 2008, pp. 195–213.

Viveka Velupillai. *An introduction to linguistic typology*. John Benjamins Publishing, 2012.