Phonology (2025), **Vol. 42**:e7 1–30
doi:10.1017/S0952675725000041

PHONOLOGY

ARTICLE

Evidence for stress in Filipino text-setting

Kie Zuraw[1] and Paolo Roca[2]

[1]Department of Linguistics, University of California, Los Angeles, Los Angeles, CA, USA.
[2]School of Nursing, University of California, Los Angeles, Los Angeles, CA, USA.

Corresponding author: Kie Zuraw; Email: p.roca@ucla.edu

Received: 19 June 2023; **Revised:** 24 January 2024; **Accepted:** 4 March 2024

Keywords: Filipino; Tagalog; music; text-setting; stress

Abstract

Words in Tagalog/Filipino can be either penult-prominent or ultima-prominent. Scholars have been divided on whether the language has stress, or only phonemic vowel length in penults and default phrase-final prominence. Using a corpus of Original Pilipino Music, we find that both prominent penults and prominent ultimas are set to longer notes and stronger beats, even in phrase-medial position. We further find that among pre-tonic syllables, those that would plausibly attract secondary stress are mostly set to longer notes and stronger beats. Text-setting does not faithfully reflect differences in phonetic cues between the two types of prominence, nor is it sensitive to presumed phonetic differences between high and low vowels. We conclude that songwriters' text-setting decisions reflect phonological stress in Filipino, and that both penult-prominent and ultima-prominent words bear stress.

Contents

1. Introduction

We use data from Filipino pop music to address a long-standing debate about stress in Tagalog/Filipino, with three main findings.

First, we find that phonologically prominent syllables are set to longer and stronger musical notes. This is not surprising; many previous works on other languages have found that songwriters choose to begin stressed syllables on strong beats, and/or to give them long durations (to mention a few: Dell & Halle 2009; Hayes 2009; Proto 2013; Proto & Dell 2013; Temperley & Temperley 2013; Bellik 2019). However, to the best of our knowledge, this is the first study of phonological aspects of text-setting in a Philippine language, and possibly in any language of the Austronesian language family.[1]

Our second finding concerns the contrast between penultimate and final prominence. Nearly, all words in Tagalog/Filipino fall into one of two types: penult-prominent (prominence on the second-to-last syllable of the word), as in [ʔábot] 'power, capacity' and ultima-prominent (prominence on the last syllable of the word), as in [ʔabót] 'arrival'. As discussed below, some scholars treat the two types as qualitatively different, with penult-prominent words like [ʔábot] having underlying length (/ʔa:bot/) and ultima-prominent words like [ʔabót] having only default phrase-final prominence (Constantino 1965; Schachter & Otanes 1972; Soberano 1980; Himmelmann & Kaufman 2020). We find evidence instead for a stress analysis: penultimate and final prominent syllables are treated similarly in the music corpus we examine, even when phrase-medial, both being set to stronger beats and longer notes. As further support for the stress analysis, we find that text-setting prominence does not shift onto phrase-final enclitics; that syllable shape matters for text-setting only where it has been claimed to matter for stress; and that syllables plausibly predicted to bear secondary stress are also set to longer notes and stronger beats. We conclude that Tagalog/Filipino has stress in all words, not just the penult-prominent ones.

Our third finding is that text-setting appears to be sensitive to prominence at an abstract, phonological level, in that it does not mirror the phonetic cues of word prominence, nor is it sensitive to vowel height, even though in speech low vowels should have longer duration and greater loudness.

In the next section, we lay out the two types of word prominence and their phonetic cues, and give background information on the language and music. §3 describes our methods; §4 and §5 give results for duration and beat strength. §6 provides Monte Carlo tests of statistical significance. §7 and §8 give results for phrase-final enclitics and pre-tonic syllables. §9 argues that the text-setting in this corpus does not track phonetics. An html file showing the annotated R code that generated our figures and results is provided as Supplementary Material.

2. Background

2.1. Tagalog and Filipino

The terms 'Tagalog language' and 'Filipino language' are often used interchangeably. When a distinction is intended, 'Tagalog' generally refers to the language of the Tagalog ethnic group in the northern

[1] We have found some (ethno)musicological studies of Austronesian song genres, such as Goldsworthy (1995, Fijian), Moyle (2007, Takuu) and Yampolsky (2022, Southern Tetun), that include basic text-setting information, such as number of syllables per line and whether vowel elisions and vowel length changes are used to make lyrics fit the meter. Aloufi (2021) sketches a preliminary analysis of line length and rhythm in Pohnpeian songs using Optimality Theory (Prince & Smolensky 1993, 2004).

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

Table 1. *Two types of word in Filipino.*

	Penult-prominent	Ultima-prominent
Examples	[ʔábot] 'power, capacity' [bajáran] 'to be paid for'	[ʔabót] 'arrival' [bajarán] 'payment due'
Penult shape	Penult must be open (some loan exceptions)	Penult can be open or closed
Vowel duration	Penult long, ultima long	Penult short, ultima long
Intensity	Penult louder than ultima	Penult and ultima similar
Pitch accent in citation form	High pitch ends right after penult vowel	High pitch ends late in ultima vowel
Pitch accent in longer phrases	Accent may remain as in citation form	Accent may move onto phrase-final word or clitic

Philippines, whereas 'Filipino' is the national language of the Philippines, based on Tagalog and enriched with vocabulary from other Philippine languages, English, Spanish and elsewhere. 'Filipino' tends to refer to the language as spoken outside the Tagalog region, or in Philippine cities where Tagalog and non-Tagalog people interact; it can also refer to a prestige variety of the language used in national media (Nolasco 2007). Filipino, especially as used outside of the Tagalog region, has grammatical differences from Tagalog (Rubrico 2012; Demeterio & Dreisbach 2017).

We use the term 'Filipino' in this article, because the music we are analysing forms part of the Philippine national mass media; the linguistic sources we cite use the term 'Tagalog'.

2.2. Word prosody in Filipino

2.2.1. Two types of word. The great majority of words in Filipino fall into two classes: penult-prominent and ultima-prominent. Table 1 lists the phonetic properties that have been observed in Gonzalez (1970), Anderson (2006) and Klimenko *et al.* (2010);[2] examples are in (1). In the song corpus, there are about the same number of words of each kind (by type: 269 penult-prominent, 266 ultima-prominent; by token: 498 penult-prominent, 475 ultima-prominent). To avoid committing prematurely to a phonological analysis, we avoid the IPA stress notation (International Phonetic Association 1999) and instead place an acute accent over the prominent vowel.

(1) *More examples of each word type, sampled from the song corpus*

a. *Penult-prominent*

i.	ʔaʔawátin	'will be restrained'	vii.	magʔaʔálaj	'is offering, dedicating'
ii.	hinihíliŋ	'is being requested'	viii.	nagdúlot	'to cause, offer'
iii.	ʔibígin	'to be loved'	ix.	nakalípas	'past'
iv.	kahápon	'yesterday'	x.	nasánaj	'got used to'
v.	kanlúran	'west'	xi.	paglípas	'passing'
vi.	láŋit	'sky'	xii.	tiwálaʔ	'trust'

[2]Hwang *et al.* (2019) conducted a perception experiment of the minimal pair *bábad* 'to soak' and *babád* 'immersed', manipulating pitch, duration and intensity. If the penult was long, participants perceived *bábad*, regardless of ultima duration. If the penult was short, participants perceived *babád*, except that low pitch on the ultima caused slightly fewer *babád* responses. They conclude that listeners use primarily duration and secondarily pitch.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

b. *Ultima-prominent*

i.	ʔakó	‘I’	vii.	maŋgá	‘mango’
ii.	dalawá	‘two’	viii.	maŋá	(plural marker)
iii.	doʔón	‘there’	ix.	nilá	‘them’
iv.	hagkán	‘to be kissed’	x.	pintíg	‘beat’
v.	ʔinjó	‘your’	xi.	sadjáʔ	‘purpose’
vi.	mahalagá	‘valuable’	xii.	ʔuntíʔ	‘smallness’

While lexical words are generally at least disyllabic, there are several monosyllabic function words, for example, [ʔat] ‘and’, [ba] ‘(question particle)’, [din] ‘also’. There are loanwords with antepenultimate prominence, such as [ʔágila] ‘eagle’, from Spanish *águila*, which was the only such word in the song corpus we used. And there are loans with prominence on a closed penult, such as [bɾiljánte] ‘diamond’, from Spanish *brillante*.

The spectrograms in Figure 1, made with Praat (Boersma & Weenink 2017) from recordings in the online dictionary of Tagalog.com, illustrate a minimal pair in citation form.[3] The top example is [ʔábot] ‘power, capacity’, with penult and ultima vowels similarly long, penult louder than ultima (as can be seen by comparing the heights of the waveforms for [a] and [o]) and high pitch on the penult followed by a pitch fall between the two vowels, as shown by the pitch track (thick black trace overlaid on the spectrogram). The bottom example is [ʔabót] ‘arrival’, with a short penult vowel and long ultima vowel, penult and ultima similarly loud and a pitch fall late in the ultima.

Further evidence that final prominence is not associated with longer duration (prominent ultimas are not longer than non-prominent ultimas) comes from Reed’s (2022) acoustic study of reduplication.

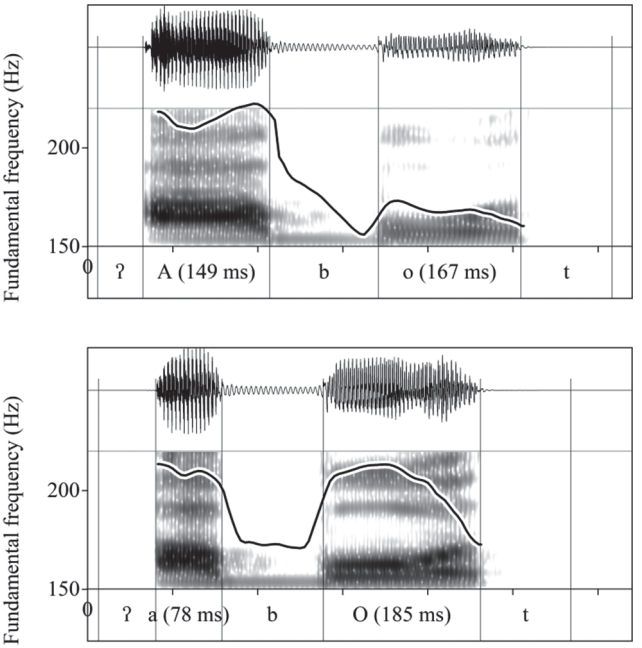


Figure 1. *Citation-form disyllables: penult-prominent [ʔábot] ‘power, capacity’ (top) and ultima-prominent [ʔabót] ‘arrival’ (bottom).*

[3]The two words are etymologically related.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

Table 2. *Approaches to word prominence in Filipino.*

	Penult-prominent [ʔábot]	Ultima-prominent [ʔabót]
Phonemic length analysis	/ʔa:bot/ → [ʔa:.bot] with intonational prominence on long vowel	/ʔabot/ → [ʔa.bot] with default phrasal prominence on final syllable
Phonemic stress analysis	/ˈʔabot/ → [ˈʔa:.bot]	/ʔaˈbot/ → [ʔa.ˈbot]

Tagalog has semi-productive copying of a stem's first two syllables, as in *ka-sáma* 'companion' *vs.* *ka-sáma-sáma* 'constant companion'. Impressionistically, both copies often have word-level prominence. Reed finds that when penult-prominent roots undergo copying of their first two syllables, the copied penult is also long, as in *d*[a:]*mi-dámi* 'quite a lot', from *dámi* 'quantity'. But when ultima-prominent roots are copied, the copy of the ultima is not long: *an*[o]-*anó* 'what-PL', from *anó* 'what'. Reed takes this difference as evidence for the underlying-length analysis, but it is also consistent with final stress not causing additional lengthening beyond word-final lengthening.

The facts for multi-word utterances are less clear. Schachter & Otanes (1972) describe most Tagalog intonation patterns as including a phrase-final pitch-accent; for some patterns, this accent's location depends on vowel length. The behaviour of these 'fixed-P2 patterns' (Schachter & Otanes 1972: 32) is illustrated in (2): for penult-prominent words, the pitch accent can optionally move to the phrase-final syllable when an enclitic is added; for ultima-prominent words, the pitch accent *must* move to the phrase-final syllable.[4]

(2) *Location of phrase-final pitch accent (bold)*

a.	sa**pá**tos	'shoe(s)'	sa**pá**to(s) ko *or* sapátos **ko**	'my shoes'
b.	da**mít**	'clothing'	damít **ko**	'my clothing'

The optionality for penult-prominent words, as well as the variety of intonation types and their overlapping semantics or usage, makes it difficult to say whether any given utterance falsifies these claims.

2.2.2. Previous analyses. Previous authors have fallen into two broad camps, summarised in Table 2.[5] For some (Constantino 1965; Schachter & Otanes 1972; Soberano 1980; Himmelmann & Kaufman 2020), there is a phonemic length contrast in open penults. This explains the phonetic duration facts: long vowels are pronounced with greater duration, as are all final syllables (leaving the [ʔa] of [ʔabót] short). This analysis must stipulate that long vowels attract pitch accent away from its default phrase-final location, and that a syllable following the pitch accent has lower amplitude. For these authors, stress, if it exists at all in the language, is the surface result of prominence due to vowel length or intonational pitch-accent. (Constantino 1965 uses underlying stress to encode exceptions such as /brilˈjante/ → [bril.ján.te] 'diamond', which has prominence on a closed penult.)

For other authors (Bloomfield 1917; Blake 1925; Ramos 1981; French 1988, 1991; Avery & Lamontagne 1995; Sabbagh 2014; Richards 2017), there is a phonemic stress contrast. This analysis must stipulate that both stressed syllables and word-final syllables are lengthened, but not additively, so that stressed, word-final syllables are not extra-long. Like the vowel-length analysis, the stress analysis stipulates that a syllable following the pitch accent has lower amplitude.

[4] A special case: when an ultima-prominent word ending in /ʔ/ is followed by a consonant-initial enclitic, the /ʔ/ deletes and the preceding vowel lengthens, according to Schachter & Otanes: /malíʔ na/ → [mali: na] 'wrong now'. In these cases, the pitch accent can either stay on [li:] or move to [na].

[5] Many of these authors treat word-initial [ʔ] as inserted, and therefore would treat the words in Table 2 as underlyingly /a/-initial.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

We now review three phonological phenomena relevant to length or stress, concluding that they can be analysed under either approach.[6] Thus, new data are needed in order to distinguish between the two approaches.

The first phenomenon is that closed penults cannot bear prominence (with exceptions, mainly in loans): **hágkan* is not allowed. Under the length analysis, this is straightforward: as in many languages, a syllable can have either a long vowel or a final consonant, but not both, ruling out *[haːg.kan]. A stress analysis can stipulate that stress must fall on one of the last two moras of the word (with final consonants extrametrical – that is, not bearing a mora). As shown in (3), stress (the bottom-most × mark on the grids) cannot fall on a consonant mora (ruling out (3d) *[haǵ.kan]), nor can it fall on a mora that is not one of the last two (ruling out (3e) *[hág.kan]).[7] This is reminiscent of many trochaic languages' ban on words that end with a heavy penult and a light ultima (see Hayes 1995 on trochaic shortening; Zuraw 2018).

(3) *Underlying-stress analysis: stress must fall within last two moras*

The second phenomenon is prominence shift in verbs. As shown in (4), prominence shifts one syllable to the right when a suffix is added. Exceptional loan verbs with prominence on a closed penult retain that prominence and gain another on the final syllable.

(4) *Prominence shifts in verbs*

a. *Penult-prominent*
Unsuffixed: ʔi-súlat 'to be written to'
Suffixed: sulát-in 'to be written'
i. Length analysis: /suːlat+in/ → [sulaːtin] (length moves)
ii. Stress analysis: /ˈsulat+in/ → [suˈlatin] (stress moves)

b. *Ultima-prominent*
Unsuffixed: d⟨in⟩asál 'was prayed about'
Suffixed: dasal-ín 'to be prayed about'
i. Length analysis: /dasal+in/ → [dasalin] (no change)
ii. Stress analysis: /daˈsal+in/ → [dasaˈlin] (stress moves)

c. *Exceptional closed prominent penult*
Unsuffixed: ʔi-kwénto 'to be narrated'
Suffixed: kwéntu-hán 'to be told a story to'
i. Length analysis: /kweːnto+han/ → [kweːntuhan] (no change)
ii. Stress analysis: /ˈkwento+han/ → [ˌkwentuˈhan] (stress added to ultima)

Under a length analysis, ultima-prominent words require no explanation; they have no long vowel underlyingly, and continue to have no long vowel after suffixation. For penult-prominent words, some

[6]There is a fourth phenomenon that will not be discussed in detail: Avery & Lamontagne (1995) appeal to stress, specifically avoidance of stress clash and stress lapse, in their analysis of infix placement in loans that begin with two consonants (on which see also Zuraw 2007). It is difficult to imagine a length-based analysis of their data.

[7]This version of the stress analysis assigns only one mora to any vowel, even those that are phonetically lengthened, such as the /a/ in the first syllable of (3b).

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

form of prosodic faithfulness can be invoked, whereby length moves in order to remain penultimate; see Shryock (1993) and Crosswhite (1998) for analyses of similar phenomena in other languages. Prosodic faithfulness has to be overridden by some mechanism that anchors exceptional vowel length in a closed syllable, as in [kweːntuhan]; the final syllable is far enough from the long vowel to get phrase-final prominence.

Under a stress analysis, a similar prosodic faithfulness mechanism is needed to keep stress penultimate when a penult-prominent root is suffixed, and the same mechanism applies to ultima-prominent roots, keeping stress final (see French 1988; Sabbagh 2004). For the exceptional words, again some mechanism keeps the exceptional stress in place, and an additional stress is added to the final syllable to avoid stress lapse.

The third phenomenon is prominence shift in some suffixed nouns. Depending on the morphology, several patterns are possible (see Schachter & Otanes 1972: 98–102). Both penult-prominent and ultima-prominent words can, when suffixed, become either penult-prominent or ultima-prominent, with or without lengthening of the root-initial vowel, and possibly other vowels. A sampling is given in (5). Under any account, there must be morpheme-specific prosodic requirements and optionality. A length account requires length to be deleted, moved, or added; a stress account requires stress to be moved or not moved and pre-tonic length to sometimes be added (see Sabbagh 2004 and Hagberg 2006 for stress-based accounts of part of the pattern).

(5) *Prominence shifts in nouns*

a. *Penult-prominent becomes penult-prominent*
Unsuffixed: páŋit ‘ugly’
Suffixed: ka-paŋít-an ‘ugliness’
i. Length analysis: /ka+paːŋit+an/ → [kapaŋiːtan] (length moves)
ii. Stress analysis: /ka+ˈpaŋit+an/ → [kapaˈŋitan] (stress moves)

b. *Penult-prominent becomes ultima-prominent*
Unsuffixed: gúlaj ‘vegetables’
Suffixed: gulaj-án ‘vegetable garden’
i. Length analysis: /guːlaj+an/ → [gulajan] (length deletes)
ii. Stress analysis: /ˈgulaj+an/ → [gulaˈjan] (stress moves)

c. *Ultima-prominent becomes penult-prominent*
Unsuffixed: hiɾám ‘borrow’
Suffixed: hiːɾám-an ‘place for borrowing’
i. Length analysis: /hiɾam+an/ → [hiːɾaːman] (penult and initial lengthen)
ii. Stress analysis: /hiˈɾam+an/ → [hiːˈɾaman] (initial lengthens)

d. *Ultima-prominent becomes ultima-prominent*
Unsuffixed: puló? ‘island’
Suffixed: ka-pulu?-án ‘archipelago’
i. Length analysis: /ka+pulo?+an/ → [kapulu?an] (no change)
ii. Stress analysis: /ka+puˈlo?+an/ → [kapuluˈ?an] (stress moves)

While these length-based and stressed-based accounts have their strengths and weaknesses, both are workable, and these primary phonological data are thus not decisive.

2.3. Original Pilipino Music (OPM)

OPM stands for Original Pilipino Music. The term can refer to any Philippine pop music, but there is a stylistic core of music that is most likely to be labelled OPM. Arceo-Dumlao (2017) is a rich collection of interviews with OPM songwriters, focusing on topics such as how an individual got into the music

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

business, the inspiration for a song, or how a singer and songwriter met. There is little discussion of songwriting mechanics such as word choice, but the interviews do provide insight into the songwriting process. Some famous songs were written, music and lyrics, by one person in one day. In other cases, lyrics were written for a melody that had already been composed by someone else. And in yet other cases, the lyrics were written first and then another songwriter composed the melody.

Other literature on OPM includes engineering projects to classify songs into sub-genres, identify mood, distinguish OPM from K-pop, recommend songs automatically or predict hits (Deja *et al.* 2016; Mital *et al.* 2019; Abisado *et al.* 2021; Sulit 2022; Monterola *et al.* 2009); humanistic studies of the history, culture and politics of OPM (Lockard 1996; Maceda 2007; Gabrillo 2018; Domingo 2021; Peña 2021; Cayabyab 2021; Prudente 2021; Shunwei & Jia 2022; Nagai 2022; Gaillard 2022); a study of cover performances on social media (Anacin *et al.* 2021); and social-science studies of music preference (Boer *et al.* 2013; Limjuco *et al.* 2014) and of attitudes towards code-switching in OPM (Bareng 2019). We have identified two linguistic studies of OPM. Alegado *et al.* (2021) analyses instances of English in OPM according to their length and syntax (the corpus we examine here did not happen to contain any code-switches into English). Sumalinog *et al.* (2021) discuss OPM lyrics' use of Swardspeak, 'the vernacular language or code used by Filipino gay men in the Philippines and in the diaspora' (Manalansan 2003: 46). Strikingly, more than half of the literature we identified was published since 2020.

We failed to find linguistic analyses of text-setting – the relationship between text and notes – in music of any Philippine language. The closest publication we could find was Anderson (2015), a guide to performing Tagalog Kundiman songs, which notes several instances where it is musically effective for duration, beat strength and/or pitch to correlate with stress.

3. Methods and predictions

3.1. Song corpus

We found sheet music online for 19 usable songs. Nine songs were purchased from the composer Aldy Santos's Web site, Aldy Sheet Music (aldysheetmusic.com), and ten songs from MuseScore (musescore.com), a site that allows users to upload their own transcriptions. (We found but excluded another seven songs: one was in 3/4 time; one was translated from Visayan, and song translators are working under different constraints; and five were each in a markedly different style from the core 19 songs.) The Appendix Table A1 lists the songs.

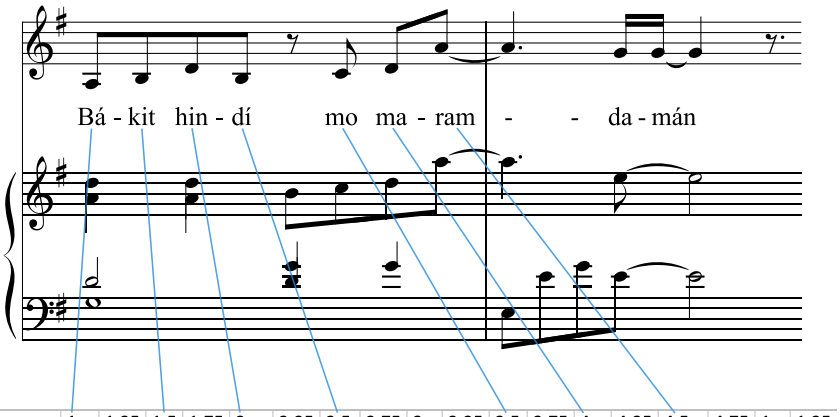
All songs were in the 4/4 time signature, which means that each measure has four beats. In this time signature, each beat is a quarter note (or crotchet). For those unfamiliar with musical notation, what is important about this time signature is that one can count along to the music in a repeating pattern of *1–2–3–4, 1–2–3–4, 1–2–3–4, ….*

3.2. Coding

After listening to recordings and correcting the sheet music where necessary (this was rare), we hand-converted each piece of sheet music into a spreadsheet. In Figure 2, we show a fragment from 'Akin ka na lang' by Francis Salazar, as performed by Morisette (with accent marks added). The opening words are *Bákit hindí? mo maramdamán…* ('Why don't you feel…'). Each column in the spreadsheet stands for one sixteenth-note of duration, with sixteen columns per 'measure' of music. The rows include a repeating metrical grid (the rows with × marks; Liberman 1975; Lerdahl & Jackendoff 1981, [1983] 1996), to guide us to the correct cell for data entry, and rows to enter information about each syllable.

As mentioned above, each measure is counted as *1–2–3–4*; these numbers appear on the 'beat' row of the spreadsheet. We assume, following the usual convention for music in the 4/4 time signature, that the strongest position in the measure is the 1, or downbeat; we show this by giving the downbeats the tallest columns in the metrical grid, with five × marks. The second-strongest position is the 3, which

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press



beat	1	1.25	1.5	1.75	2	2.25	2.5	2.75	3	3.25	3.5	3.75	4	4.25	4.5	4.75	1	1.25
whole	x																x	
half	x								x								x	
quarter	x				x				x				x				x	
eighth	x		x		x		x		x		x		x		x		x	
sixteenth	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
text	Ba-	_	kit	_	hin-	_	diQ	_	0	0	mo	_	ma	_	ram-	_	_	_
stress	1		0		0		1				0		0		0			
length	L		C		C		C				S		S		C			
syll_position	2		1		2		1				1		4		3			
line_number	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
repeat																		
notes																		

Figure 2. *Transcription fragment, from 'Akin ka na lang'.*

we give four × marks. The next-strongest positions are the 2 and the 4, with three × marks each. If a musician wants to count to the music more finely, dividing the measure into eight equal parts, they can count *1–and–2–and–3–and–4–and,* with an *and* falling in the middle of each beat. These *and*s are numbered 1.5, 2.5, 3.5 and 4.5 in the 'beat' row, and are the next-strongest positions, with two × marks each. A musician can count even more finely, dividing the measure into 16 equal parts, often as *1–ee–and–a–2–ee–and–a–3–ee–and–a–4–ee–and–a*. These *ee*s and *a*s, numbered 1.25, 1.75, 2.25, 2.75, etc., are the weakest positions, with one × mark each.

The other rows encode linguistic information. *Ba*, for example, is entered in the 'text' row, in the column where it begins (the downbeat, identified as 1 in the 'beat' row). Because *Ba* is set to an eighth note, it extends over two columns; the underscore in the next cell indicates continuation. The 1 in the 'stress' row indicates that *Ba* is prominent.[8] The L in the 'length' row indicates that this syllable has a long vowel (in this case, predictable from being a stressed open penult); closed syllables are coded as C, and open syllables with short vowels as S. The 'syll_position' row shows 2, indicating that *Ba* is a penult (second from the end of the word). The 'line_number' row shows that all the syllables depicted here are in the first line of the song. The line is not a repeat, and we made no special notes (so the 'repeat' and 'notes' rows are blank). Filipino spelling does not indicate prominence, pre-tonic length, or word-final glottal stops; we relied on a combination of dictionary entries and one author's native-speaker knowledge for the 'stress' and 'length' rows, and for the Qs indicating word-final glottal stops in the 'text' row.

An R script (R Core Team 2021) reads and processes these spreadsheets.

[8] We mark only a word's main prominent syllable with 1. §7 tests the idea of predictable secondary stress earlier in the word.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

3.3. Exclusions

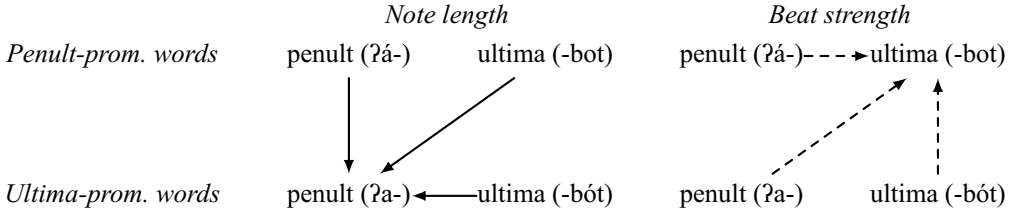
We excluded 223 syllables because they belonged to a word that fell entirely or in a part on a triplet (a division of a note into three equal parts instead of two or four), because we had no principled way to classify the prominence of the second and third sub-beats of a triplet. All the songs we coded had four beats per measure, but some had short passages with a different number of beats per measure, and we excluded 11 syllables because they belonged to words that fell partly or wholly within such passages. We excluded repeated lines, so that our data set would not appear (in plots and statistical analyses) to be bigger and more consistent than it really is. Finally, we excluded the one word in the corpus with antepenultimate stress, *ágila* 'eagle', a Spanish loan.

3.4. Predictions

We are comparing four syllable types: prominent and non-prominent penults and ultimas. Our absolute null hypothesis is that all four types are assigned to notes of similar length, and beats of similar strength.

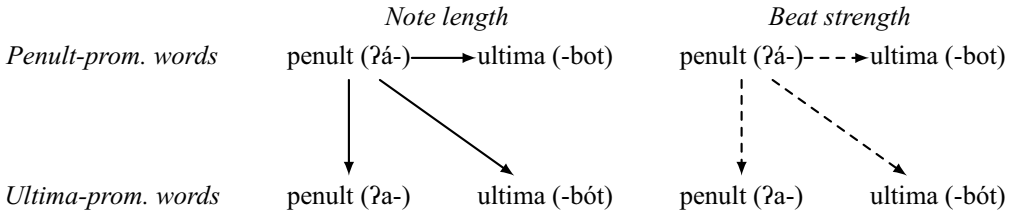
The next closest to a null hypothesis is that OPM text-setting purely reflects the phonetics of word prosody, and tells us nothing about its phonology. In that case, non-prominent penults should be assigned to shorter notes than all the rest, as illustrated by the arrows on the left in (6), which start from the syllable type predicted to be set to longer notes, and point to the syllable type predicted to be set to shorter notes. The musical equivalent of loudness is less direct, but there is a tendency for loudness to signal beat strength (Lerdahl & Jackendoff [1983] 1996: 17–18, 78–79). Non-prominent ultimas should then be assigned to weaker beats than all the rest, as illustrated on the right in (6), but with dashed arrows to show that the predictions are less direct, starting from the syllable type predicted to be set to stronger beats, and pointing to the syllable type predicted to be set to weaker beats.

(6) *Predictions if text-setting reflects phonetics of speech only*



There are two non-null hypotheses: that generated by the vowel-length theory, and that generated by the stress theory. As illustrated in (7), the vowel-length theory predicts that prominent penults, which contain a phonemically long vowel, should be set to longer notes than all other syllable types. Versions of the vowel-length theory that assign predictable stress to those phonemically long vowels also predict that prominent penults should be assigned to stronger beats than the rest (since cross-linguistically, stressed syllables tend to fall on strong beats, as stated in §1). Dashed lines are again used for the beat-strength predictions, to reflect that they are made by only one version of the vowel-length theory.

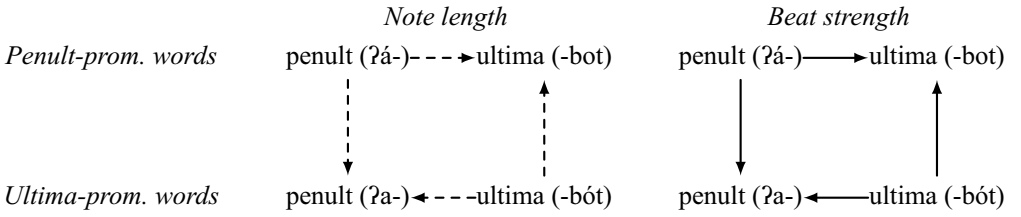
(7) *Predictions of vowel-length analysis*



https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

The hypothesis generated by the stress theory is that prominent syllables, because they are stressed, should be assigned to stronger beats than non-prominent syllables are, as shown in (8). There is less research on note length in uncontroversial stress languages, with some evidence that German stressed syllables are set to longer notes (Girardi & Plag 2022). The arrows for note length are dashed to show that this prediction is less clear.

(8) *Predictions of stress analysis*



4. Duration

4.1. Quantifying duration

Duration was quantified in quarter-note beats: a whole note (or semibreve, a note that lasts one measure) has a duration of 4 beats, a half note (minim) has a duration of 2, a quarter note (crotchet) 1, an eighth note (quaver) 0.5 and a sixteenth note (semiquaver) 0.25.

4.2. Duration results

There were 498 penult-prominent words and 475 ultima-prominent words analysed.

4.2.1. All final two syllables. The bean plots in Figure 3, made in R using the `beanplot` package (Kampstra 2008), show results for the final two syllables of all usable words. On the left are the penult-prominent words, like *ʔábot*, and on the right are the ultima-prominent words, like *ʔabót*. The left side of each pair, coloured orange, represents the smoothed distribution of duration for the penult in each type of word, and the right side, coloured sky blue, represents the distribution of duration for the ultima. The four horizontal line segments show the mean of each distribution.[9]

In the penult-prominent words, on the left, the penult tends to have a shorter duration (0.7 beats on average) than the ultima (1.0 beats). This might seem unexpected, but recall that in speech, final syllables tend to be long regardless of stress. In ultima-prominent words, the gap is bigger: penults have an average duration of 0.4 beats and ultimas 1.5. Overall then, ultimas are long, but less so in penult-prominent words.

Another way of looking at this plot is to compare the two orange distributions to each other: penults are slightly longer when prominent (0.7 on the left > 0.4 on the right). And comparing the two sky-blue distributions to each other, ultimas are longer when prominent (1.5 on the right > 1.0 on the left).

[9]Because some distributions here are skewed, we considered placing these lines at the medians. But because our duration and beat strength are coarse-grained, the medians ended up being misleading in two ways. First, distributions that are rather different sometimes had the same median. For example, the plot below shows that the mean durations for penults and ultimas of penult-prominent words are 0.7 beats and 1.0 beats. Their medians, however, are the same (0.5 beats); the longer tail for the ultimas is not enough to pull the median up to the next possible value, 0.75 beats. Second, distributions that are similar sometimes had rather different medians. The subtle difference between a distribution where almost half of the tokens are downbeats, and one where just over half of the tokens are downbeats, produces a correspondingly subtle difference in means, but a large difference in medians.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

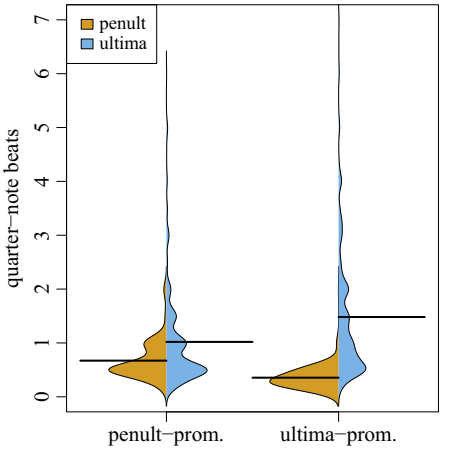


Figure 3. *Duration of final two syllables of all words.*

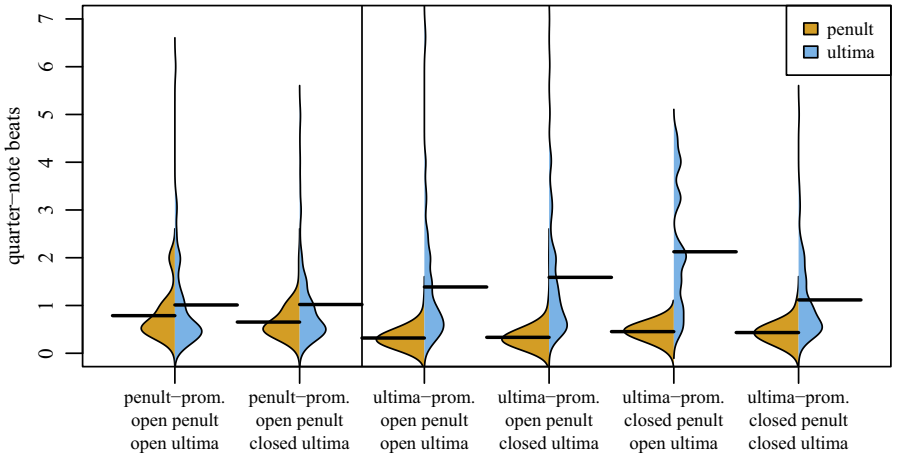


Figure 4. *Duration, broken down by syllable shape.*

The plot in Figure 4 shows that syllable shape has little consistent effect. Whether a penult-prominent word has a closed or an open penult, ultimas tend to be slightly longer than penults. And regardless of the syllable shapes in an ultima-prominent word, the ultima is substantially longer than the penult.

4.2.2. Separating out phrase- and line-final words. The very long durations seen for some ultimas are mostly line-final syllables, reflecting the musical tendency to place a long note at the end of a line. Splitting up the results into line-final *versus* line-medial words will allow us to see whether prominent ultimas are still long even when not line-final.

Furthermore, given various authors' observations about how intonational prominence can track the end of the phrase, rather than the end of the word, we should further split line-medial words into phrase-medial and phrase-final, to see whether prominent ultimas are still long even when phrase-medial. (We assume that line-final words are always phrase-final.)

To determine phrase boundaries, we follow Schachter & Otanes's (1972: 36) description of where optional pauses may occur; we take these optional pause locations to represent phrase boundaries. Schachter & Otanes state that phrase boundaries never occur after a proclitic or before an enclitic, and

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

are more likely to occur at large syntactic breaks than at small ones, such as between a modifier and the word it modifies. They give the following examples:

(9) *Example sentences (normal spelling), with slashes showing phrase boundaries* (Schachter & Otanes 1972: 36)

a. **Nakita** mo ba / si **Maria** / **kahapon**
see you Q / CASE Maria / yesterday
'Did you see Maria yesterday?'

b. **Bibigyan** din **sila** / ng **pagkain** / sa bagong **paaralan**
will.be.given also they / CASE food / CASE new school
'they will be given some food at the new school too'

c. **Magaganda** raw / ang **mga bulaklak** / kung **tag-ulan**
beautiful reportedly / CASE PL flower / when rainy.season
'They say that the flowers are beautiful in the rainy season'

d. Kay **Pedro** / ang **aklat** / na nasa **ibabaw** / ng **mesa**
CASE Pedro / CASE book / LINKER in underneath / CASE table
'The book on top of the table is Pedro's'

We operationalised phrasehood by having a look-up list of enclitics, based mainly on Kaufman (2010).[10] If a word was followed by one of these enclitics, it was considered phrase-medial.[11] If not, then it was considered phrase-final. In the examples above, our procedure correctly codes all the boldface words, but leaves *bagong* miscoded as phrase-final. (Monosyllabic words are ignored here because they are excluded from our results.) Following this example and Schachter & Otanes's description, we decided that two-word modifier–modified pairs connected by the linker suffix *-ng,* like *bagong paaralan,* should be hand-coded as belonging to the same phrase; we identified 19 such sequences in the song corpus.

We assumed that line breaks occurred where displayed on the lyrics website Musixmatch (https://www.musixmatch.com/). Some of the songs rhyme, providing further support for the line-break locations, but we did not use rhyming or other criteria to make changes from Musixmatch's line breaks.

Figure 5 illustrates this three-way breakdown. As we move from phrase-medial to phrase-final to line-final, durations get longer, especially for ultimas. But the key result still holds in all three environments: prominent penults are longer than non-prominent penults; and prominent ultimas are longer than non-prominent ultimas.

4.3. Duration summary

We have seen that, as predicted by both the vowel-length analysis and the stress analysis, prominent penults are set to longer notes than non-prominent penults. As predicted by only the stress analysis, prominent ultimas are also set to longer notes than non-prominent ultimas. This is true even for phrase-medial position, where, in speech, ultima-prominent words may lose their intonational prominence.

We defer discussion of statistical significance to §6. A regression model, included in the Supplementary Material, finds that ultimas are longer, prominent syllables are longer and prominent ultimas

[10] Our list of enclitics comprises *ako, ka, siya, kami, kata, kita, tayo, kayo, sila, ko, mo, niya, ta, ninyo, nila, kita, na, pa, nga, talaga, po, ho, pala, yata, sana, nawa, ba, baga, namin, natin,din, rin, man, naman, lang, lamang, daw* and *raw,* plus the combination of each of these with *-ng, 't* and *'y.* Because *na* has non-enclitic homophones, we hand-coded all 52 instances of *na* in the song corpus, marking 16 of them as non-enclitic.

[11] Excluding function words that should not be phrase-final, namely the plural marker *mga,* the preposition-like *nasa* and linker-suffix-bearing numerals, demonstrative pronouns and personal pronouns, which in our corpus were *aking, akong, ating, inyong, isang, itong, iyong, kayong, kitang, nating, nilang, nitong, siyang* and *tayong*. These words were always coded as phrase-medial, except for one token of *kayong* 'you' that was line-final.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

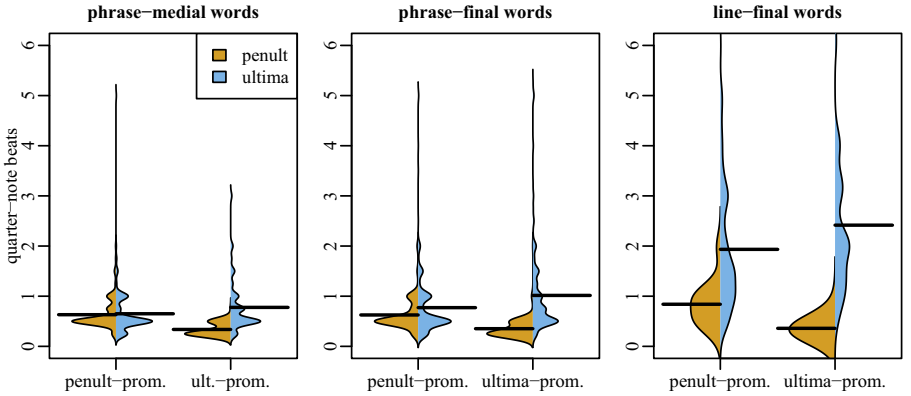


Figure 5. *Note duration broken down by position.*

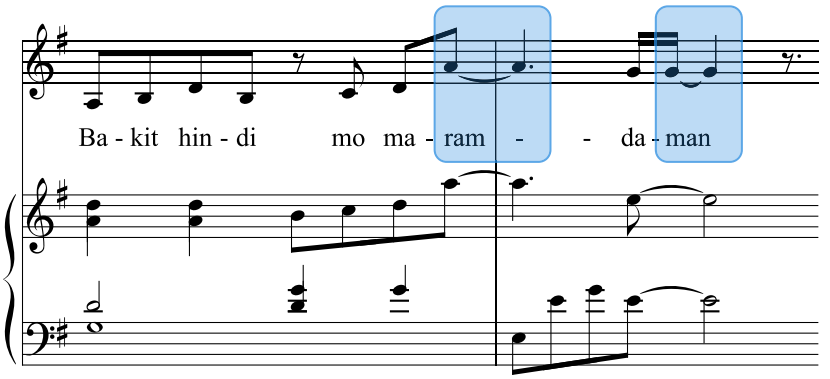


Figure 6. *Examples of syncopation in first line of 'Akin ka na lang'.*

are the longest of all. In §6, we argue that regression may not be conservative enough, and provide an alternative method for assessing significance.

5. Beat strength

5.1. Measuring beat strength: syncopation

OPM has extensive anticipatory syncopation (Temperley 1999; Tan *et al.* 2019), meaning that beats that, according to various expectations, should count as strong, begin slightly early. Two examples are shown in Figure 6: the syllable *ram* begins on the 'and' (second half) of a beat, the second-weakest position, but musically it behaves as though it began on the following measure's first beat, which is the strongest position; the syllable *man* begins on the last sixteenth note of a beat, the weakest position, but musically behaves as though it began on the third beat of the measure, the second-strongest position. Temperley's (1999) solution is to move each syncopated note forward, so that it counts as beginning later than it really does.

To operationalise this, we identified, for each syllable, the strongest beat that it *contains*: for *ram*, beat 1 of a measure, and for *man* beat 3 of a measure. One danger is that if a note is very long, it will always end up counting as strong, because it eventually goes on long enough to include a strong position. Therefore, we only looked for the strongest beat contained in the first 1.25 beats of the note. This is enough for a note that begins on the last beat of the measure (or even one sixteenth-note earlier) to count

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

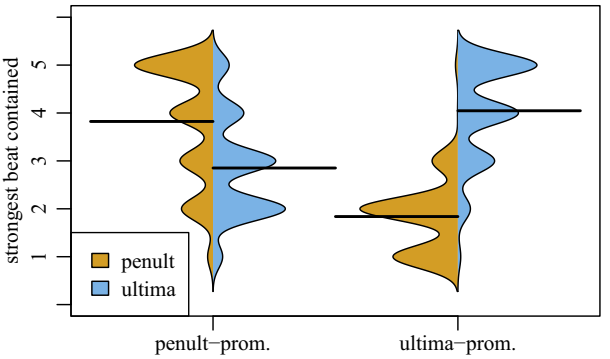


Figure 7. *Beat strength (correcting for syncopation) of final two syllables of all words.*

as beginning on a downbeat, if it lasts long enough.[12] Our procedure 'corrected' 49% of sixteenth notes, 27% of eighth notes and 11% of quarter notes to a stronger position.

5.2. Beat strength results

As with duration, there were 498 penult-prominent words and 475 ultima-prominent words analysed.

5.2.1. All final two syllables. Results are plotted in Figure 7. The plot is the same as in Figure 3, except that the vertical axis measures beat strength. The strongest value, 5, is for a note that starts on or contains a downbeat (beat 1 of the measure); the next-strongest, 4, is for a note that starts on or contains beat 3 of a measure; 3 is for a note that starts on or contains beat 2 or 4; 2 is for a note that starts on or contains at most the second half of a beat, and 1 is for a note that starts on and contains at most the second or fourth quarter of a beat.

On the left, we see that in penult-prominent words, average strength is somewhat higher in penults than in ultimas; the modal penult of such words is set to a downbeat, while their modal ultima is set to the second half of a beat (the peak of the distribution is at 2). On the right, for ultima-prominent words, ultimas are on average much stronger (modally downbeats) than penults (modally second half of a beat). Thus, beat strength matches prominence, especially for ultima-prominent words.

Just as we did for duration, we can also divide the results by syllable shape, as shown in Figure 8, with little effect: regardless of syllable shape, penult-prominent words have a stronger penult, and ultima-prominent words have a stronger ultima.

5.2.2. Separating out phrase- and line-final words. Just as with duration, we separate out phrase- and line-final words. As shown in Figure 9, for phrase-medial words, penult-prominent and ultima-prominent words look symmetrical, but moving to phrase-final and then line-final, penult-prominent words show less and less of a strength difference between penults and ultimas. Nevertheless, in all three environments, prominent syllables are stronger than non-prominent syllables within the same word

[12]Tan *et al.* (2019) implement a different procedure to correct for anticipatory syncopation: a syllable on a weak *n*th note is considered as though it actually began on the next *n*th note, as long as no other syllables begin on or before that next *n*th note. (Working with English-language music, Tan *et al.* apply this correction only to stressed syllables, but we ignored stress so as not to bias the results to make stressed syllables appear stronger.) For example, the *ram* of *maramdaman* begins on the weak eighth note of a beat (i.e., the second of the two eighth notes in that beat), so it would be considered as if it began on the next eighth note, which is the downbeat of the next measure. As with our strongest-beat-contained procedure, *ram* is thus treated as though it really began on a downbeat. Unlike in our procedure, the *di* of *hindi* would also be promoted, from the second half of a beat to the beginning of the measure's third beat. This is because our procedure requires the note to continue into a position to be promoted to it, and Tan *et al.*'s does not. Only 1.3% of syllables had different corrected strengths under the two procedures, so we did not pursue a reanalysis of the data under Tan *et al.*'s procedure.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

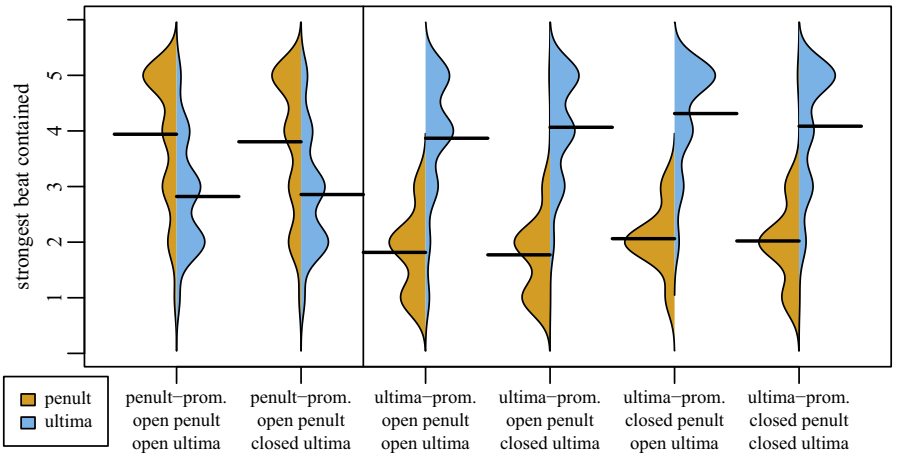


Figure 8. *Beat strength broken down by syllable shape.*

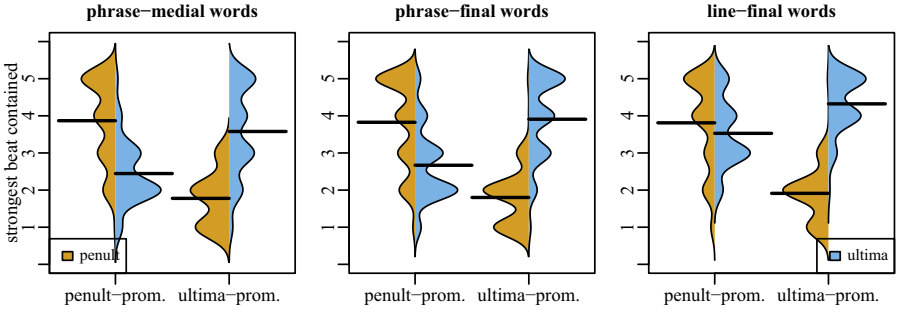


Figure 9. *Beat strength broken down by position.*

type; prominent penults are stronger than non-prominent penults; and prominent ultimas are stronger than non-prominent ultimas.

5.3. Beat strength summary

The results again support the stress analysis. As both the vowel-length analysis and the stress analysis predict, prominent penults are set to stronger beats than both non-prominent penults and non-prominent ultimas. But as predicted only by the stress analysis, prominent ultimas are also set to stronger beats than both non-prominent ultimas and non-prominent penults. This holds even phrase-medially, where, in speech, ultima-prominent words may lose their intonational prominence.

A regression model, included in the Supplementary Material, finds that prominent ultimas and penults are set to the strongest beats, and non-prominent penults to the weakest. A more conservative method of assessing statistical significance is given now.

6. Monte Carlo tests for significance

Imagine a language where most words have final stress, and a musical style where most lines of music end in a long note. Even if songwriters make no effort to align stress with long notes, stressed syllables will still tend to be placed on long notes, because of line-final syllables. In other words, the simple null hypothesis that the regression models mentioned above use, ‘stressed and unstressed syllables are

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

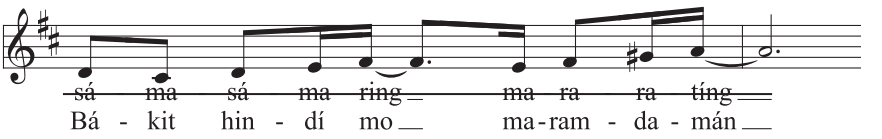


Figure 10. *Lyrics randomly assigned to a new line of music.*

set to notes of equal duration', won't do, because the data are inherently biased to falsify that null hypothesis. We want to construct a null hypothesis that already includes such biases. One way to do this is to randomly re-combine text and music. In some studies, this scrambled, null-hypothesis corpus is constructed by drawing text from prose (the 'Russian method' for poetry; Ryan 2011; see references in Hayes 2013), but it is not clear what would be suitable prose to use in the case of OPM. Instead, we follow Gunkel & Ryan (2011) in scrambling the lines of lyrics *within* our corpus of songs.

We first compute several measures for the real data, such as the mean duration of prominent penults' notes minus the mean duration of non-prominent penults' notes (all the measures are listed in (10)). Then, we convert each musical line into a pattern representing the number of syllables, and the locations where a disyllabic or longer word ends. For example, *Bákit hindí? mo maramdamán* 'why can't you feel it' has the structure σ σ | σ σ | σ σ σ σ σ |, with | representing the ends of words, not including monosyllabic *mo* 'you'. These word boundaries are important because, as we saw in §4.2, word-final syllables tend to be given a long duration regardless of whether the word is penult- or ultima-prominent. Since we are interested not in that effect, but rather in the difference between penult- and ultima-prominent words, we want our null hypothesis to include word-final lengthening.

Our script randomly selects a line from the corpus that has the same structure, such as *Sáma-sáma ŋa mararatíŋ* 'will also arrive together'. (We generally coded words with two-syllable reduplication, like *sáma-sáma* 'together', as two separate words.) In this example, the selected line is from a different song ('Tagumpay nating lahat', written by Gary Granada and performed by Lea Salonga). The script combines the old line's lyrics with the new line's musical notes, as shown in Figure 10.

We recompute the measures of interest on the new, scrambled corpus, and repeat the scrambling procedure 1,000 times, to obtain the distribution of values that we would expect to see under the null hypothesis, following Kessler (2001); Martin (2007, 2011); Hayes *et al.* (2009). Each plot in Figure 11 is for one of the measures in (10). The grey bars are a histogram of the measure's values in the shuffled corpora, and the solid blue line is the value in the actual corpus. The estimated *p* value for each measure is the proportion of shuffled corpora that lie to the right of the solid blue line – how often we'd expect to see such an extreme result by chance. For many measures, the histogram does not overlap with the solid blue line at all, meaning that the estimated *p* value is less than 0.001. The measures and *p* values are given in (10), and depicted graphically in (11).[13]

(10)
a. **mean log duration of prominent penults – mean log duration of non-prominent ultimas,** $p < 0.001$
b. **mean log duration of prominent penults – mean log duration of non-prominent penults,** $p < 0.001$
c. **mean log duration of prominent ultimas – mean log duration of non-prominent penults,** $p < 0.001$
d. **mean log duration of prominent ultimas – mean log duration of non-prominent ultimas,** $p = 0.038$
e. mean log duration of prominent penults – mean log duration of prominent ultimas, $p = 0.006$
f. mean log duration of non-prominent ultimas – mean log duration of non-prominent penults, $p < 0.001$
g. **mean beat strength of prom. penults – mean str. of non-prom. ultimas,** $p < 0.001$
h. **mean beat strength of prom. penults – mean str. of non-prom. penults,** $p < 0.001$

[13] Measures (10a)–(10d) are restricted to non-line-final words.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

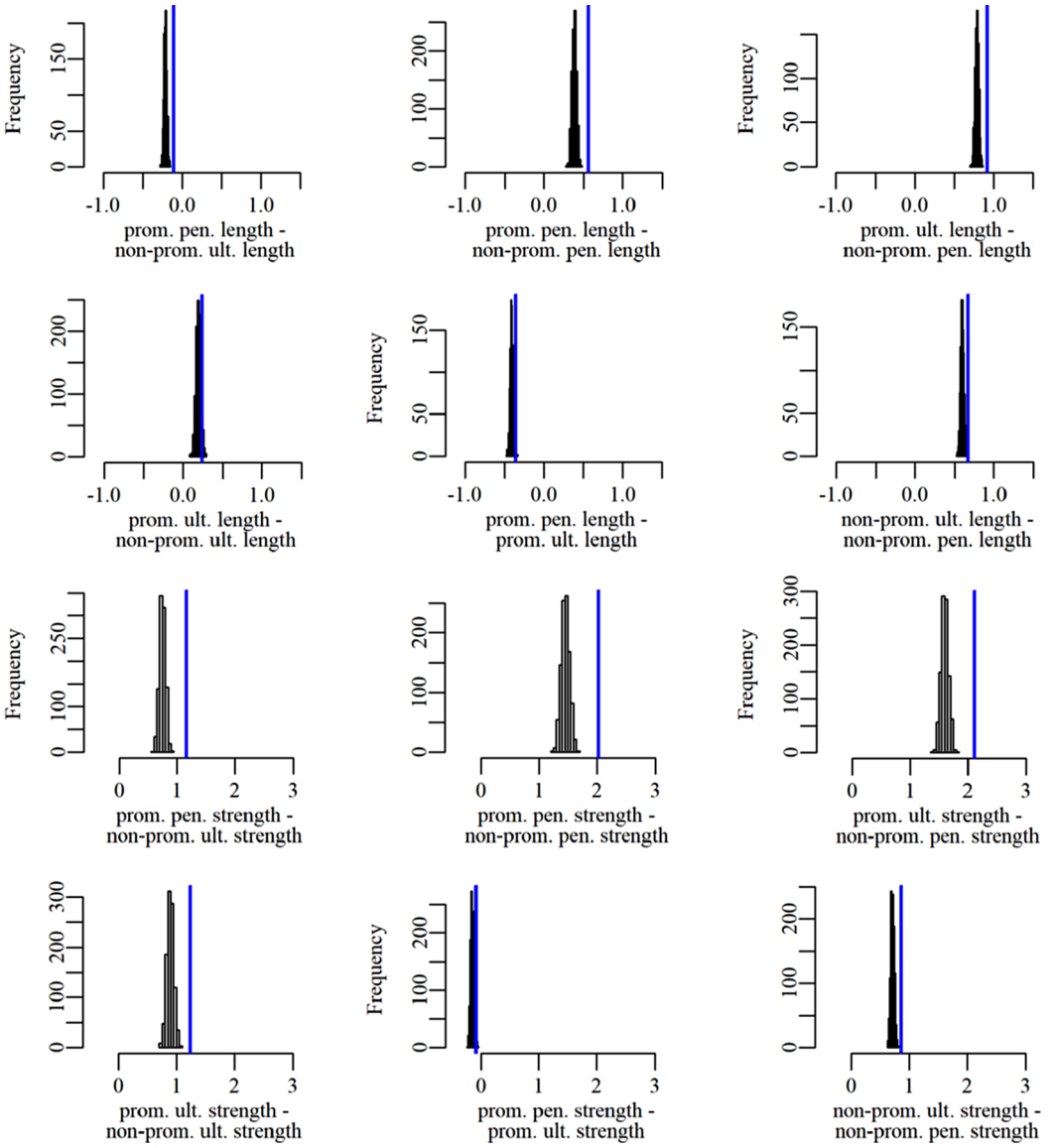


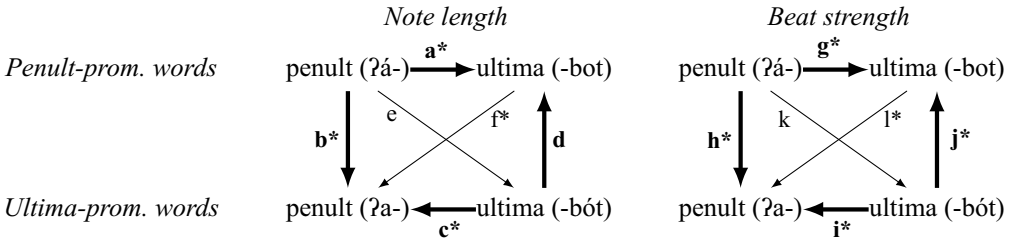
Figure 11. *Monte Carlo results.*

i. **mean beat strength of prom. ultimas – mean str. of non-prom. penults,** $p < 0.001$
j. **mean beat strength of prom. ultimas – mean str. of non-prom. ultimas,** $p < 0.001$
k. mean beat strength of prom. penults – mean beat str. of prom. ultimas, $p = 0.015$
l. mean beat strength of unprom. ultimas – mean bt str. of unprom. penults, $p < 0.001$

Measures (10a)–(10d) and (10g)–(10j), shown in bold, are those predicted by the stress analysis to be positive, or in any case greater than expected by chance; the other measures are included for completeness, even if no analysis predicts them to be non-zero. Since there are 12 measures being taken, if we require $p < 0.05$ to reject the null hypothesis, a Bonferroni correction (Dunn 1961) adjusts that threshold to $p < 0.004$. The measures meeting that significance criterion are marked with * in (11).

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

(11) *Graphical depiction of measures in (10)*



Except for (10d), then, the predictions of the stress analysis given in §3.4 are supported. The vowel-length analysis and the stress analysis overlap in correctly predicting (10a), (10b), (10g) and (10h) to be positive. The vowel-length analysis alone predicts (10e) and (10k) to be positive, which was not strongly supported. We conclude that our results support the stress analysis over the vowel-length analysis.

For those who prefer a regression analysis, despite its drawbacks, the html file in the Supplementary Material includes regression models that reach similar conclusions.

7. Phrase-final enclitics

Recall from §2.2.2 that some scholars hold that apparent word-final prominence is really phrase-final prominence, because when an ultima-prominent word is put into a phrase, its prominence shifts. For example, ultima-prominent *damít* 'clothes' may appear to have final stress in citation form, but when it is combined with the enclitic *ko* 'my' to form *damit ko* 'my clothes', this 'stress' now falls on *ko*. By contrast, the *pa* of penult-prominent *sapátos* 'clothes' remains (or at least can remain) 'stressed' in *sapatos ko*.

To test this idea, we examine enclitics like *ko*. If apparent word-final prominence is really phrase-final prominence, *ko* should be set to longer and stronger notes when it follows ultima-prominent words (because prominence should shift onto it) than when it follows penult-prominent words.

We extracted all monosyllabic enclitics from the corpus, and counted them as phrase-final if they were followed by a content word. This yielded 143 total tokens of phrase-final monosyllabic enclitics, of which 101 were after penult-prominent words and 42 after ultima-prominent words.

Looking first at note duration, shown in Figure 12, we see that monosyllabic enclitics, whether line-medial or line-final, are not set to longer notes when they follow an ultima-prominent word (yellow) than when they follow a penult-prominent word (green).

As for beat strength, shown in Figure 13, monosyllabic enclitics are set to a *weaker* beat when they follow an ultima-prominent word than when they follow a penult-prominent word, the opposite of what

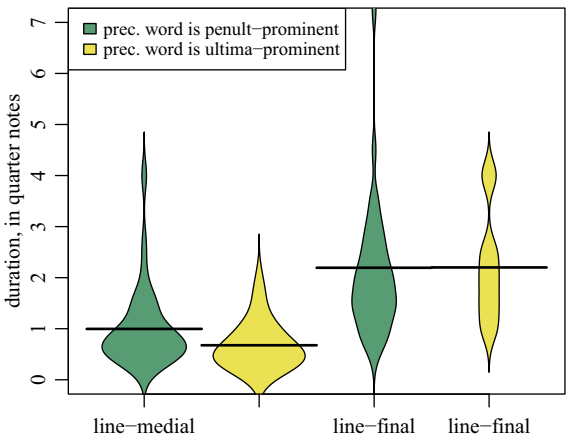


Figure 12. *Phrase-final clitic duration.*

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

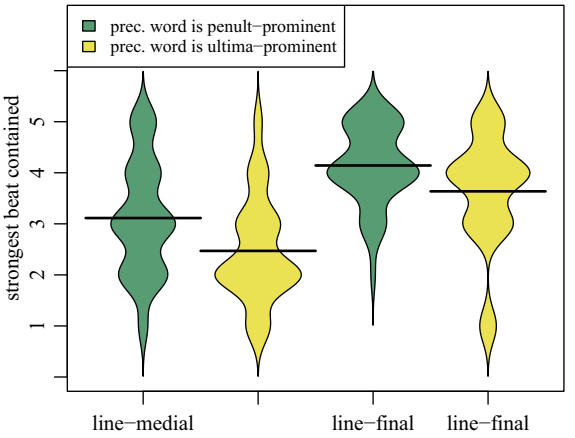


Figure 13. *Phrase-final clitic beat strength.*

phrasal prominence-shifting predicts. This is probably explained by the fact that weak and strong beats tend to alternate in music. When a prominent ultima, like the *mit* in *damít*, gets set to a strong beat, a following syllable will tend to be set to a weak beat; whereas when a prominent penult, like the *pa* in *sapátos*, is set to a strong beat, an enclitic two syllables later can also be set to a strong beat.

Using the same Monte Carlo method as in §6 (plots in Supplementary Material), we found that enclitics were not set to significantly longer notes after ultima-prominent words than after penult-prominent ($p = 0.55$), nor to significantly stronger beats ($p = 0.82$).

The text-setting data thus contradict the idea that ultima prominence is illusory. Even if intonational pitch accent often moves onto an enclitic in speech, songwriters keep prominence on the word ultima.

8. Pre-tonic syllables

8.1. Introduction

Filipino morphology can create a long vowel earlier in the word – before the penult, that is – though this is often variable or optional. The most common morphological source of long vowels is verb aspect reduplication. As illustrated in (12), aspect reduplication produces a copied, prefixed syllable that typically contains a long vowel. (The infixes *-um-* and *-in-* represent voice and aspect.) All but ten of the pre-tonic long vowels in the song corpus come from aspect reduplication.

(12) *Long vowels in aspect reduplication*

a.	t⟨um⟩akbó	‘ran’	ta:-takbo	‘will run’
b.	t⟨in⟩akbó	‘was run to’	ta:-takbu-hín	‘will be run to’
c.	s⟨um⟩úlat	‘wrote’	su:-súlat	‘will write’
d.	s⟨in⟩ulát-an	‘was written on’	su:-sulát-an	‘will be written on’

French (1988) and French (1991) give partly conflicting descriptions of secondary stress, and French (1991) calls for acoustic analysis of secondary stress – which as far as we know has still not been carried out – to clarify the picture. We will focus here on French’s claims about the types of words that are well attested in the song corpus. French’s two accounts agree that aspect reduplicants, like those shown in (12), receive secondary stress; for example, French would transcribe ‘will write’ as [ˌ**su**ˈsulat]. French (1988) claims that closed syllables in prefixes generally attract secondary stress, as in [**màg**-pa-ka-ʔáral] ‘study intensely’ (and does not address closed, pre-tonic root syllables, as in the penult of [ta:-**tak**bó]). French (1988) further claims that a closed prefix syllable will not receive secondary stress if a following

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

Table 3. *Number of open-syllable observations for analysing pre-tonic length.*

	Syllable				
	7th-last	6th-last	5th-last	4th-last	3rd-last
Short vowel, open syllable	0	3	32	106	263
Short vowel, closed syllable	0	0	5	32	67
Long vowel, open syllable	1	1	1	29	78

prefix syllable itself has secondary stress (the context found in aspect reduplication), as in [**mag-ˌpa**-pa-ka-ˈʔaral] 'will study intensely', where the prefix /pa/ has undergone aspect reduplication. The two works make conflicting claims about default locations of secondary stress when prefixes are all open syllables with no aspect reduplication.

While acknowledging that much remains to be determined about Filipino secondary stress, we extract two hypotheses from these descriptions. First, pre-tonic syllables that are closed or have long vowels, as in words like [**sa:**-sabí-hin] 'will be said' and [**nag**-simulá?] 'began', should tend to be treated as having secondary stress, and thus be set to longer notes and stronger beats than pre-tonic syllables that are open and have a long vowel, as in [**ka**-ʔibíg-an] 'friend'. Second, looking just at antepenults, stress clash avoidance should weaken or eliminate this effect when the next syllable is a prominent penult, so that the antepenult in a word like [**pag**-ʔíbig] 'love' or [ma-**pa**-páwiʔ] 'will come to an end' would not be set to particularly long notes or strong beats, despite being closed or having a long vowel, because the following syllable is prominent.

For this part of the analysis, we used words of three or more syllables. Table 3 shows how many tokens were found in each position. Because there were so few observations of fifth-, sixth- and seventh-to-last syllables, they are not included in the analysis.

8.2. Pre-tonic duration

The bean plots in Figures 14 and 15 show the distribution of note duration for pre-tonic syllables that are short-vowelled and open, short-vowelled and closed, or long-vowelled and open. (There were no long closed tokens.) In fourth-to-last syllables (Figure 14), we see that songwriters have assigned the short open syllables to the shortest note durations. In contrast to the duration data for penults and ultimas above in Figure 4, syllable shape does matter here: the closed syllables pattern with the long-vowel syllables – although, not surprisingly given the small amount of data, the differences are not significant.[14] This is consistent with French's (1991) contention that both closed syllables and long-vowelled syllables (aspect reduplicants) attract secondary stress, and goes against the otherwise appealing notion that the reason a closed penult cannot be prominent is that it can't have a long vowel: even though these closed pre-tonic syllables have short vowels, they appear to be receiving prominence.[15]

For third-to-last syllables, in Figure 15, the data are further divided according to whether the following syllable is prominent (penult-prominent word, as in *li:lípas* 'will elapse') or not (ultima-

[14]See Supplementary Material for Monte Carlo plots. For fourth-to-last syllables, we compared closed to short-open and long to short-open, in both note length and beat strength. Long *vs.* short-open note duration was the most promising, with $p = 0.031$, but this does not survive any correction for multiple comparisons. The Supplementary Material also contain regression models of pre-tonic note length and beat strength, which largely support the trends seen in the plots: in fourth-to-last syllables, long-vowel syllables are set to longer notes and stronger beats, closed syllables are set to longer notes (but not stronger beats); in third-to-last syllables, closed syllables are set to longer notes, especially if the following syllable is not stressed, and long vowels and lack of stress clash both predict stronger beats.

[15]The reason closed penults don't attract secondary stress in ultima-prominent words is, under French's analysis, that Tagalog avoids stressing two syllables in a row (stress clash) where possible.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

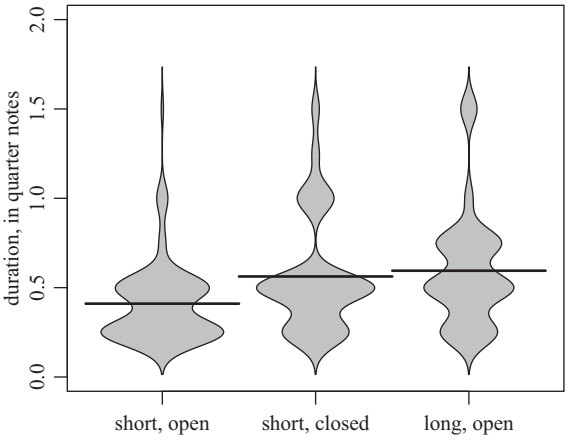


Figure 14. *Note duration in fourth-to-last syllables.*

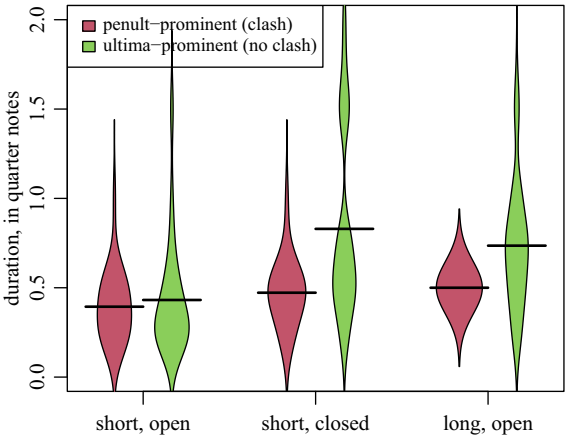


Figure 15. *Note duration in third-to-last syllables.*

prominent, as in *pu:putí?* 'will turn white'). The only syllables to receive longer note duration are closed and long-vowelled syllables that are not followed by a prominent penult, and thus not subject to stress clash. The difference between on the one hand long-vowel syllables not subject to stress clash and on the other hand short-vowel syllables was significant ($p < 0.001$).

8.3. Pre-tonic strength

The results for beat strength are less perfectly in line with our secondary-stress predictions, but still broadly support them. In fourth-to-last position, long-vowelled syllables – but not closed syllables – trend towards being set to stronger beats, as shown in Figure 16.

In third-to-last syllables (Figure 17), the syllables set to the strongest beats are those that we predict to have secondary stress: closed and long-vowelled syllables in ultima-prominent words (no stress clash). There also appears to be a difference within the short open syllables between those that are followed by a prominent syllable and those that are not. It could be that short open syllables prefer to bear secondary stress if followed by an unstressed syllable (Blake 1925; Avery & Lamontagne 1995). There is also a plausible musical explanation for this: unlike note length, beat strength alternates in the underlying musical structure. Because prominent penults and ultimas tend to be assigned to strong beats, there will

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

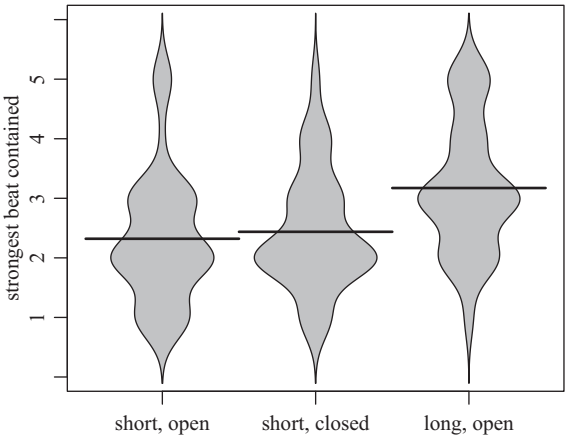


Figure 16. *Beat strength in fourth-to-last syllables.*

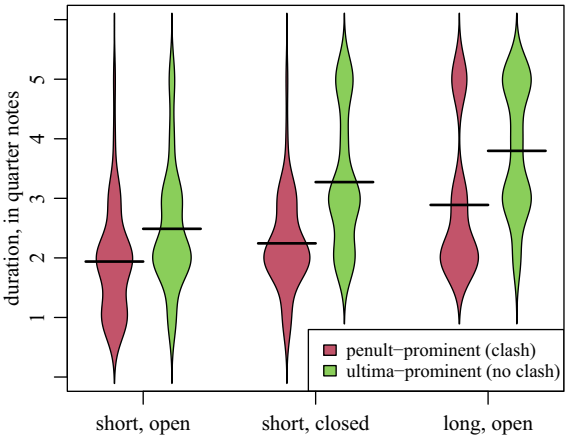


Figure 17. *Beat strength in third-to-last syllables.*

thus be a tendency for an antepenult preceding a prominent penult to be weak, and for an antepenult preceding a prominent ultima to be strong. In the case of short-vowelled open syllables, this musical tendency creates a small difference; in the closed and long-vowelled syllables the musical tendency combines with stress clash avoidance to create a bigger difference. The difference between, on the one hand, long-vowelled syllables not subject to stress clash and, on the other, short-vowelled syllables was significant ($p < 0.001$).

8.4. Pre-tonic syllables summary

We have seen that closed or long-vowelled pre-tonic syllables are set to longer notes and stronger beats, as long as the following syllable is not the tonic (as in a word like *li:lípas* 'will elapse (time)'). This supports French's (1991) contention that pre-tonic closed syllables and long vowels attract secondary stress, subject to some stress clash avoidance. We saw earlier that open *versus* closed syllable shape in penults and ultimas, which does not affect stress (except that closed penults may not bear stress), was not important for note length and beat strength. Thus, syllable shape seems to matter for text-setting only where it has been claimed to matter for stress.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

Table 4. *Summary of phonetic and musical properties of last two syllables of words.*

	Duration in speech	Duration in song corpus	Loudness in speech	Beat strength in song corpus
Penult-prominent (ʔábot)	Long–long	Medium–medium	Loud–quiet	Strong–weak
Ultima-prominent (ʔabót)	Short–long	Short–long	Loud–loud	Weak–strong

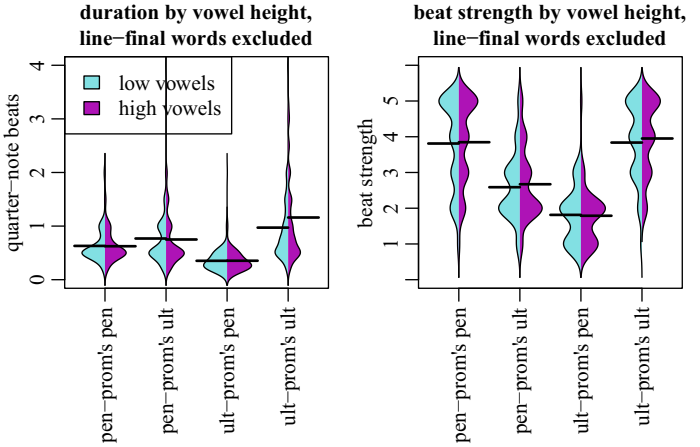


Figure 18. *Note duration and beat strength by vowel height.*

9. OPM text-setting does not track phonetics

While text-setting partly tracks the phonetics of duration and loudness (see §2.2.1), there are some mismatches. In speech, the last two syllables of penult-prominent words have two long vowels, and those of ultima-prominent words have a short and then a long vowel; music was a rough match to this (excluding line-final words), as summarised in Table 4, except that prominent ultimas were set to longer notes than either non-prominent ultimas or prominent penults, as predicted by the stress analysis. In speech, penult-prominent words have a loud penult and quiet ultima, which was reflected in beat strength, but prominent ultimas, which in speech have similar loudness to their non-prominent penults, were set to stronger beats, again as predicted by the stress analysis.

We also looked at vowel height, on the assumption that the low vowel /a/ should be longer and louder in speech than the high vowels /i, u/ (though the acoustic results on this in Gonzalez 1970 are not straightforward). The plots in Figure 18 show the note durations (left plot) and beat strengths (right plot) of the last two syllables of penult-prominent and ultima-prominent words. Rather than pairing the distributions of penult and ultima for each word type, the four syllables are all separated out, and each pair of distributions is for a low vowel (left) and a high vowel (right). Within each pair, the left and right distributions are almost identical, with no indication that songwriters assign low vowels to longer notes or stronger beats.

Although we found that OPM text-setting does not track phonetic detail, there is one area we found where it does track surface rather than underlying phonology. When a [ʔ]-final word is phrase-medial, the [ʔ] usually deletes, and the preceding vowel lengthens in compensation, as in, using Schachter & Otanes's (1972: 16) length-based notation, [lu:toʔ] 'cooked' *vs.* [lu:tu: ba] 'cooked?' and [hindiʔ] 'no' *vs.* [hindi: ba] 'no?'. Within prominent ultimas followed by a consonant-initial enclitic, we found that underlyingly glottal-final ultimas, like the /diʔ/ in 'no', are set to somewhat longer notes and stronger beats than other prominent ultimas are, presumably reflecting their surface lengthening. (Plots are provided in the Supplementary Material.)

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

10. Conclusion

This study has found that prominent penults and prominent ultimas are both set, in a corpus of OPM songs, to longer notes and stronger beats, both phrase-medially and phrase-finally – and that these text-setting tendencies are not simple reflections of duration and loudness in speech. Text-setting seems to reflect stress at the word level, and not merely phrasal prominence: when an ultima-prominent word is followed by a phrase-final enclitic (e.g., *damít ko*), many authors have observed that intonational prominence tends to shift onto the enclitic, but as we saw in §6, it is the content word's ultima (*mít*) that is musically prominent, not the enclitic (*ko*). Furthermore, while syllable shape (open *vs.* closed) did not affect text-setting of penults and ultimas, it did affect text-setting of pre-tonic syllables, which is where French (1991) has claimed that syllable shape affects stress. All this is evidence in favour of analysing Filipino as having stress, even though the stress is realised differently in different positions in speech, with stressed penults having greater duration than unstressed, and stressed ultimas having greater loudness than unstressed (in addition to possible intonational differences). As we discussed in §2.2.2, standard phonological data were insufficient to decide between the length-driven and the stress-driven analyses. We believe that the musical data here provide the first straightforward evidence in favour of one analysis, the stress-driven one.

If the basic phonological data are not decisive for phonologists, how is it that songwriters have converged on treating Filipino as having stress? It is possible that there's something in the basic data that no phonologists have noticed, but which is decisive for children learning the language. Or cases like Filipino could be telling us that, faced with ambiguous data of the Filipino type, learners are biased to acquire a lexicon and grammar with stress.

Our findings echo those of Domene Moreno & Kabak (2022) for Turkish songs. In Turkish, as in Filipino, it has been proposed that words with non-final prominence bear true stress, while words with final prominence bear only phrase-final accent. Domene Moreno & Kabak measure beat strength and melodic peakhood in a song corpus. They find that, in Western European-style children's songs, linguistically prominent syllables receive more of both types of musical prominence, with no difference between penultimate prominence and final prominence. Like us, they take this as evidence for word-level stress in Turkish.

Domene Moreno & Kabak found that songs they analysed in the Makam style did not give musical prominence to either type of Turkish prominent syllable. This raises the question of whether Western European-style Turkish children's songs and OPM songs are both showing influence from English-language pop music's tendency to align musical prominence and stress. This is possible, but does not explain away their or our results, because songwriters influenced by English songs would still have to decide what counts, in their language, as the equivalent to English's stress. And in these Turkish and Filipino corpora, the songwriters have decided to treat both final and non-final prominent syllables as needing to be musically prominent.

The one interpretation of our data that could be consistent with an underlying-length analysis is effectively an empty one, where, before any phonology applies, an underlying length contrast gets converted into surface stress for all content words, both in words that have an underlying long vowel (/ʔa:bot/ → [ˈʔa:.bot]) and in words that do not, which receive final stress (/ʔabot/ → [ʔa.ˈbot]). Without direct access to speakers' underlying representations, the availability of a deeper level of analysis with length only, cannot be refuted by any data. More broadly, data alone cannot rule out an analysis of any phenomenon where a feature that the phonology appears to be sensitive to is actually the (un-neutralised) reflex of a different underlying feature, though there could be cross-linguistic or theoretical justifications for such an analysis. We do not, however, find any support for underlying length in the text-setting data, which appears to be sensitive only to stress.

We end with a methodological note on the usefulness of musical data for low- and medium-resource languages. Filipino could be considered a medium-resource language. Unlike for most of the world's languages, there are corpora and engineering tools, either available or in development: see Jakubíček *et al.* (2013), Go & Nocon (2017), Go *et al.* (2017), Lazaro *et al.* (2009), Ang *et al.* (2014), and many

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

others. But the extent of these resources is small compared to what exists for English, Korean, French and other languages with well-funded public and private research infrastructure.[16] Our song corpus consists of 1,662 words in total. A spoken corpus of that size would be too small for studying stress correlates, with too many sources of noise (speech rate, inherent duration and loudness of vowels, etc.). But in songs, we have access to songwriters' categorical decisions about duration and strength, which makes the data clean enough for clear patterns to emerge. We originally coded and analysed just nine songs, and the main patterns were already there; adding the remaining ten songs made us more confident in the results, but didn't change them. A small corpus of songs, even a number as small as what a research team could transcribe themselves from listening to recordings, can thus be useful for gaining insight into the phonology of a lower-resource language, as long as the object of study occurs with sufficient density in songs. In our case, most of the syllables in a song provided relevant data, so the density of observations per song was high.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0952675725000041.

Data availability statement. The annotated R code that generated our figures and results, and some additional statistical analysis is available in the Supplementary Material.

Acknowledgements. For their feedback on various parts of this project, we thank participants in the UCLA phonology seminar, the UC Berkeley Phorum, the University of Ottawa linguistics colloquium, the Linguistic Society of the Philippines RIPPLE series (Researches, Insights, and Perspectives on the Philippine Linguistics Enterprise), the Manchester Phonology Meeting and the Linguistic Society of America annual meeting. We especially thank Peter Avery, François Dell, Bruce Hayes, Larry Hyman, Dan Kaufman, Stephanie Reed, Hannah Sande and the editors and reviewers of *Phonology* for their comments.

Funding statement. This research was supported by a grant from the UCLA Academic Senate's Committee on Research.

Competing interests. The authors declare no competing interests.

A. Songs

Table A1 lists the songs used in the corpus. Because the sheet music we used is made by listeners, where possible we list the performer whose performance the sheet music is based on, as well as the composer, when known.

Table A1. *Songs in the corpus.*

Song title	Performer (composer)	Approx. year of composition	Number of usable penult–ultima pairs
Akin ka na lang	Morisette (Francis Salazar)	2014	28
Bakit ngayon ka lang	Freestyle (Aaron Paul & Ogie Alcasid)	1993	32
Diamante	Morisette (Jungee Marcelo)	2016	50
Handog	Sarah Geronimo (Florante)	1978	63
Iduyan mo	Basil Valdez (Ryan Cayabcab)	1980	22
Iisa pa lamang	Joey Albert	1990	40
Ikaw	Sharon Cuneta (Louie Ocampo & George Canseco)	1993	47

[16]The multi-language DALI corpus of pop song recordings (Meseguer-Brocal *et al.* 2018) has three songs coded as Filipino, but on inspection two of those were Japanese songs that had been mislabelled, so there was only one Filipino song in the corpus, not enough to analyse. The multi-language Vocadito corpus (Bittner *et al.* 2021) has about 300 syllables of folk-song and nursery-rhyme Filipino songs, but provides only duration and requires considerable hand-correction if duration at the syllable level is desired. The Smule data sets of amateur karaoke performances (Smule, Inc. 2018a, b) no doubt include valuable data, but the raw data would require considerable processing to be usable for our purposes.

Table A1. *(Continued).*

Song title	Performer (composer)	Approx. year of composition	Number of usable penult–ultima pairs
Ikaw	Yeng Constantino	2015	66
Ikaw at ako	Moira dela Torre & Jason Marvin Hernandez	2019	66
Kahit maputi na ang buhok ko	Sharon Cuneta (Rey Valera)	1977	42
Kanlungan	Noel Cabangon (Rene Bonconan)	2009	79
Kapag ako ay nagmahal	Jolina Magdangal (Larry Hermoso)	1999	46
Kung di rin lang ikaw	December Avenue feat. Moira dela Torre	2018	55
Maging sino ka man	Erik Santos (Rey Valera)	1979	49
Ngayon at kailanman	Ariel Rivera (George Canseco)	1977	70
Pagkat ang Diyos ay pag-ibig	Erik Santos	2011	47
Pangarap ko ang ibigin ka	Regine Velasquez	2006	55
Tagumpay nating lahat	Lea Salonga (Gary Granada)	2010	69
Tuloy pa rin	Neocolours	1990	49
Total			**975**

References

Abisado, Mideth B., Mardyon B. Yongson & Ma Ian P. Los Trinos (2021). Towards the development of music mood classification of original Pilipino music (OPM) songs based on audio and lyrics keyword. In *ICSET 2021: 5th International Conference on E-Society, E-Education and E-Technology*. New York: Association for Computing Machinery, 87–90.

Alegado, Joni Rose, Abigail M. Labaya, Pia S. Lirio & Ruthier S. Rivera (2021). *A linguistic analysis of Tagalog-English code switching in OPM love songs*. Ms., Polytechnic University of the Philippines.

Aloufi, Aliaa (2021). Ponapean song meter in Optimality Theory. *Advances in Language and Literary Studies* **12**, 46–50.

Anacin, Carljohnson, David Baker & Andy Bennett (2021). Mimicking the mimics: problematizing cover performance of Filipino local music on social media. *Media, Culture & Society* **43**, 1414–1430.

Anderson, Quiliano Niñeza (2015). *Kundiman love songs from the Philippines: their development from folksong to art song and an examination of representative repertoire*. PhD dissertation, University of Iowa.

Anderson, Victoria B. (2006). Lexical stress without postlexical head-marking: evidence from Tagalog. *JASA* **120**, 3092–3092.

Ang, Federico, Rowena Cristina Guevara, Yoshikazu Miyanaga, Rhandley Cajote, Joel Ilao, Michael Gringo Angelo Bayona & Ann Franchesca Laguna (2014). Open domain continuous Filipino speech recognition: challenges and baseline experiments. *IEICE Transactions on Information and Systems* **97**, 2443–2452.

Arceo-Dumlao, Tina (2017). *Himig at titik: a tribute to OPM songwriters*. Makati City: Inquirer Books.

Avery, Peter & Greg Lamontagne (1995). Infixation <and metathesis> in Tagalog. Paper presented at the Annual Meeting of the Canadian Linguistic Association.

Bareng, Jeanmarc M. (2019). Code switching in original Pilipino music (OPM). *Ascendens Asia Journal of Multidisciplinary Research Abstracts* **3**, 685.

Bellik, Jennifer Ann (2019). *Vowel intrusion in Turkish onset clusters*. PhD dissertation, University of California, Santa Cruz.

Bittner, Rachel M., Katherine Pasalo, Juan José Bosch, Gabriel Meseguer-Brocal & David Rubinstein (2021). Vocadito: a dataset of solo vocals with f_0, note, and lyric annotations. In *Proceedings of the 22nd International Society for Music Information Retrieval Conference*, 3 pp.

Blake, Frank Ringgold (1925). *A grammar of the Tagalog language*. New Haven, CT: American Oriental Society.

Bloomfield, Leonard (1917). *Tagalog texts with grammatical analysis*. Urbana, IL: University of Illinois.

Boer, Diana, Ronald Fischer, Ma Luisa Gonzalez Atilano, Jimena Garay Hernández, Luz Irene Moreno Garcia, Socorro Mendoza, Valdiney V. Gouveia, Jason Lam & Eva Lo (2013). Music, identity, and musical ethnocentrism of young people in six Asian, Latin American, and Western cultures. *Journal of Applied Social Psychology* **43**, 2360–2376.

https://doi.org/10.1017/S0952675725000041 Published online by Cambridge University Press

Boersma, Paul & David Weenink (2017). Praat: doing phonetics by computer. Version 6.0.29. Available at http://www.praat.org/.
Cayabyab, Krina (2021). The (de-) and (re-) mythification of OPM: decentring a popular music sign. In Adil Johan & Mayco A. Santaella (eds.) *Made in Nusantara: studies in popular music*. New York: Routledge, 45–54.
Constantino, Ernesto (1965). The sentence patterns of twenty-six Philippine languages. *Lingua* **15**, 71–124.
Crosswhite, Katherine (1998). Segmental vs. prosodic correspondence in Chamorro. *Phonology* **15**, 281–316.
Deja, Jordan Aiko, Kim Blanquera, Carlo Eliczar Carabeo & Jo Rupert Copiaco (2016). Genre classification of OPM songs through the use of musical features. In Shin-ya Nishizaki, Masayuki Numao, Jaime D.L. Caro & Merlin Teodosia C. Suarez (eds.) *Theory and practice of computation: proceedings of Workshop on Computation: Theory and Practice WCTP2014*. Singapore: World Scientific, 77–88.
Dell, François & John Halle (2009). Comparing musical textsetting in French and in English songs. In Andy Arleo & Jean-Louis Aroui (eds.) *Towards a typology of poetic forms*. Amsterdam: Benjamins, 63–78.
Demeterio, Feorillo A. , III & Jeconiah Louis Dreisbach (2017). Disentangling the Rubrico and Dolalas hypotheses on the Davao Filipino language. *Recoletos Multidisciplinary Research Journal* **5**, 1–15.
Domene Moreno, Christina & Barış Kabak (2022). Prominence alignment in English and Turkish songs: implications for word prosodic typology. In Scharinger & Wiese (2022), 223–258.
Domingo, Luis Zuriel P. (2021). Korean pop music a threat to contemporary Filipino identity? Globalization, nation, and interrogation in Philippine culture and identity. *Asia Review* **11**, 247–265.
Dunn, Olive Jean (1961). Multiple comparisons among means. *Journal of the American Statistical Association* **56**, 52–64.
French, Koleen Matsuda (1988). *Insights into Tagalog reduplication, infixation and stress from nonlinear phonology*. Dallas, TX: Summer Institute of Liguistics and University of Texas at Arlington.
French, Koleen Matsuda (1991). Secondary stress in Tagalog. *Oceanic Linguistics* **30**, 157–178.
Gabrillo, James (2018). Rak en rol: the influence of psychedelic culture in Philippine music. *Rock Music Studies* **5**, 257–274.
Gaillard, J.C. (2022). On the production of hybrid urban space(s) in *post*-colonial cities: Manila in the music of Eraserheads. *Geografiska Annaler: Series B, Human Geography* **105**, 305–320.
Girardi, Elena & Ingo Plag (2022). Metrical mapping in text-setting: empirical analysis and grammatical implementation. In Scharinger & Wiese (2022), 191–221.
Go, Matthew Phillip & Nicco Nocon (2017). Using Stanford part-of-speech tagger for the morphologically-rich Filipino language. In Rachel Edita Roxas (ed.) *Proceedings of the 31st Pacific Asia Conference on Language, Information and Computation (PACLIC) 31*. Manila: National University, 81–88.
Go, Matthew Phillip, Nicco Nocon & Allan Borra (2017). Gramatika: a grammar checker for the low-resourced Filipino language. In *TENCON 2017 – 2017 IEEE Region 10 Conference*. 471–475.
Goldsworthy, David (1995). Continuities in Fijian music: meke and same. *Yearbook for Traditional Music* **27**, 23–33.
Gonzalez, A. (1970). Acoustic correlates of accent, rhythm, and intonation in Tagalog. *Phonetica* **22**, 11–44.
Gunkel, Dieter & Kevin M. Ryan (2011). Hiatus avoidance and metrification in the Rigveda. In Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.) *Proceedings of the 22nd annual UCLA Indo-European Conference*. Bremen: Hempen, 53–68.
Hagberg, Lawrence Raymond (2006). *An autosegmental theory of stress*. Dallas, TX: SIL International.
Hayes, Bruce (1995). *Metrical stress theory: principles and case studies*. Chicago, IL: University of Chicago Press.
Hayes, Bruce (2009). Textsetting as constraint conflict. In Andy Arleo & Jean-Louis Aroui (eds.) *Towards a typology of poetic forms: from language to metrics and beyond*. Amsterdam: Benjamins, 43–61.
Hayes, Bruce (2013). Milton, Maxent, and the Russian method. Paper presented at the M@90 Workshop on Metrical Structure, Massachusetts Institute of Technology, September 2013. Available at https://youtu.be/IRRyxW6xPeg.
Hayes, Bruce, Kie Zuraw, Zsuzsa Cziráky Londe & Peter Siptár (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Lg* **85**, 822–863.
Himmelmann, Nikolaus P. & Daniel Kaufman (2020). Austronesia. In Carlos Gussenhoven & Aoju Chen (eds.) *The Oxford handbook of prosody*. Oxford: Oxford University Press, 370–383.
Hwang, Hyun Kyung, Naonori Nagaya & Julián Villegas (2019). Cue weighting in the perception of Tagalog stress. *JASA* **146**, 3052.
International Phonetic Association (1999). *Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý & Vít Suchomel (2013). The TenTen corpus family. In Andrew Hardie & Robbie Love (eds.) *Corpus linguistics 2013 abstract book*. Lancaster: UCREL, 125–127.
Johan, Adil & Mayco A. Santaella (eds.) (2021). *Made in Nusantara: studies in popular music*. New York: Routledge.
Kampstra, Peter (2008). Beanplot: a boxplot alternative for visual comparison of distributions. *Journal of Statistical Software* **28**, 1–9.
Kaufman, Daniel (2010). *The morphosyntax of Tagalog clitics: a typologically driven approach*. PhD dissertation, Cornell University.
Kessler, Brett (2001). *The significance of word lists*. Stanford, CA: Center for the Study of Language and Information.
Klimenko, Sergey B., Maria Paz C. San Juan & Jem R. Javier (2010). Stressed out with stress: perceptual recognition of acoustic correlates of stress in Tagalog. *Paper presented at the 1st Philippine Conference-Workshop on Mother Tongue Based Multilingual Education*, Capitol University, Cagayan de Oro City.

Lazaro, Lito Rodel S., Leslie L. Policarpio & Rowena Cristina L. Guevara (2009). Incorporating duration and intonation models in Filipino speech synthesis. In *Proceedings: APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 annual summit and conference*, 45–49.
Lerdahl, Fred & Ray S. Jackendoff (1981). On the theory of grouping and meter. *The Musical Quarterly* **67**, 479–506.
Lerdahl, Fred & Ray S. Jackendoff ([1983] 1996). *A generative theory of tonal music*. Cambridge, MA: MIT Press. Reissue.
Liberman, Mark Yoffe (1975). *The intonational system of English*. Master's thesis, Massachusetts Institute of Technology.
Limjuco, Renan P., Ronald Jay S. Ticudo & Helenne U. Pregua (2014). Determinants of music type preference of university students in Davao City. *UIC Research Journal* **20**, 177–189.
Lockard, Craig A. (1996). Popular musics and politics in modern Southeast Asia: a comparative analysis. *Asian Music* **27**, 149–199.
Maceda, Teresita Gimenez (2007). Problematizing the popular: the dynamics of Pinoy pop (ular) music and popular protest music. *Inter-Asia Cultural Studies* **8**, 390–413.
Manalansan, Martin F. (2003). *Global divas: Filipino gay men in the diaspora*. Durham, NC: Duke University Press.
Martin, Andrew (2011). Grammars leak: modeling how phonotactic generalizations interact within the grammar. *Lg* **87**, 751–770.
Martin, Andrew Thomas (2007). *The evolving lexicon*. PhD dissertation, University of California, Los Angeles.
Meseguer-Brocal, Gabriel, Alice Cohen-Hadria & Geoffroy Peeters (2018). DALI: a large dataset of synchronized audio, lyrics and notes, automatically created using teacher-student machine learning paradigm. In *Proceedings of the 19th International Society for Music Information Retrieval (ISMIR) conference*, 431–437.
Mital, Matt Ervin G., Rogelio Ruzcko N.M.I. Tobias, Argel A. Bandala, Robert Kerwin Billones & Elmer P. Dadios (2019). Utilization of genetic algorithm in classifying Filipino and Korean music through distinct windowing and perceptual features. In S.K. Niranjan, Ajay Rana & Himdweep Khurana (eds.) *Proceedings of the Fourth International Conference on Contemporary Computing and Informatics (iCI 2019)*. Piscataway, NJ: IEEE, 121–126.
Monterola, Christopher, Cheryl Abundo, Jeric Tugaff & Lorcel Ericka Venturina (2009). Prediction of potential hit song and musical genre using artificial neural networks. *International Journal of Modern Physics C* **20**, 1697–1718.
Moyle, Richard (2007). Taking five – quintuple metre in Takū *tuki* songs. In Richard Moyle (ed.) *Oceanic music encounters: the print resource and the human resource: essays in honour of Mervyn McLean*. Auckland: Department of Anthropology, University of Auckland, 123–132.
Nagai, Hiroko (2022). Embrace our color: nationalism in Philippine popular music. In Hafzan Zannie Hamza, Clare Suet Ching Chan & Lena Farida Hussain Chin (eds.) *Proceedings of the 4th International Music & Performing Arts Conference: trending digital virtual and capital*. Tanjong Malim: Sultan Idris Education University, 32–37.
Nolasco, Ricardo Ma. (2007). Filipino and Tagalog, not so simple / how to value our languages. Blog post. Archived at https://web.archive.org/web/20170730230733/http://svillafania.philippinepen.ph/2007/08/articles-filipino-and-tagalog-not-so.html.
Peña, Verne de la (2021). Songs for and of the youth: mapping trends in Philippine popular music, 1900–2000. In Johan & Santaella (2021), 92–100.
Prince, Alan & Paul Smolensky (1993). Optimality Theory: constraint interaction in generative grammar. Technical Report 2, Rutgers Center for Cognitive Science. Subsequently published as Prince & Smolensky (2004).
Prince, Alan & Paul Smolensky (2004). *Optimality Theory: constraint interaction in generative grammar*. Oxford: Blackwell.
Proto, Teresa (2013). Singing in German: text-setting rules and language rhythm. In Galatà Vincenzo (ed.) *Multimodalità e multilingualità: la sfida più avanzata della comunicazione orale*. Rome: Bulzoni, 9–10.
Proto, Teresa & François Dell (2013). The structure of metrical patterns in tunes and in literary verse. Evidence from discrepancies between musical and linguistic rhythm in Italian songs. *Probus* **25**, 105–138.
Prudente, Felicidad A. (2021). Colonialism and identity: a short history of popular music in the Philippines. In Johan & Santaella (2021), 35–44.
R Core Team (2021). R: a language and environment for statistical computing. Available at https://www.R-project.org.
Ramos, Teresita V. (1981). *Tagalog structures*. Honolulu, HI: University of Hawaii Press.
Reed, Stephanie (2022). Duration in Tagalog reduplication as evidence for phonemic vowel length. Paper presented at LabPhon 18, online, June 2022.
Richards, Norvin (2017). Some notes on Tagalog prosody and scrambling. *Glossa* **2**, article no. 21.
Rubrico, Jessie Grace U. (2012). *Indigenization of Filipino: the case of the Davao City variety*. Ms., University of Malaya. Available at https://www.languagelinks.org/onlinepapers/Indigenization-of-Filipino.pdf.
Ryan, Kevin M. (2011). *Gradient weight in phonology*. PhD dissertation, University of California, Los Angeles.
Sabbagh, Joseph (2004). *Stress shift and prosodic correspondence in Tagalog. Presented at the Western Conference on Linguistics (WECOL)*, University of Southern California.
Sabbagh, Joseph (2014). Word order and prosodic-structure constraints in Tagalog. *Syntax* **17**, 40–89.
Schachter, Paul & Fe T. Otanes (1972). *Tagalog reference grammar*. Berkeley, CA: University of California Press.
Scharinger, Mathias & Richard Wiese (eds.) (2022). *How language speaks to music: prosody from a cross-domain perspective*. Berlin: De Gruyter.
Shryock, Aaron (1993). A metrical analysis of stress in Cebuano. *Lingua* **91**, 103–148.
Shunwei, Liu & Li Jia (2022). Establishment of Philippine popular music industry. *Multicultural Education* **8**, 60–82.

Smule, Inc. (2018a). DAMP-MVP: Digital archive of mobile performances – Smule multilingual vocal performance 300x30x2. Data set. Available at https://doi.org/10.5281/zenodo.2747436.

Smule, Inc. (2018b). DAMP-VSEP: Digital archive of mobile performances – vocal separation. Data set. Available at https://doi.org/10.5281/zenodo.3553059.

Soberano, Rosa (1980). *The dialects of Marinduque Tagalog*. Number 69 in Pacific Linguistics Series B. Canberra: Department of Linguistics, School of Pacific Studies, Australian National University.

Sulit, Jimson Montilla (2022). *Content-based original Pilipino music (OPM) recommender system centered on mood*. MICS capstone project, University of the Philippines.

Sumalinog, Deo Jamael M., Ronald S. Salid, Exel Rose D. Sarino & F. Amante (2021). Exploring gay lingo in some selected OPM songs. *Asian Journal of Advanced Multidisciplinary Researches* **1**, 1–4.

Tan, Ivan, Ethan Lustig & David Temperley (2019). Anticipatory syncopation in rock: a corpus study. *Music Perception* **36**, 353–370.

Temperley, David (1999). Syncopation in rock: a perceptual perspective. *Popular Music* **18**, 19–40.

Temperley, Nicholas & David Temperley (2013). Stress-meter alignment in French vocal music. *JASA* **134**, 520–527.

Yampolsky, Philip (2022). Poetic text and melodic text: text-setting in two song traditions of Timor. *Asian Music* **53**, 80–126.

Zuraw, Kie (2007). The role of phonetic knowledge in phonological patterning: corpus and survey evidence from Tagalog infixation. *Lg* **83**, 277–316.

Zuraw, Kie (2018). Beyond trochaic shortening: a survey of Central Pacific languages. *Lg* **94**, e1–e42.