Emphatic vowel length in Cebuano

Kevin Samejon
Notre Dame of Marbel University, Koronodal City
samekevin.ling@gmail.com

Abstract

Phonetic length has been generally observed to indicate emphasis also as in the Cebuano utterance, *Lamiiiiii kaayo ang sud-an*, trans. 'The dish is veeeeeery delicious', a pragmatic use of phonetic length. Following the observation of an earlier experiment in English, the present study investigated this phonetic behavior on Cebuano vowels. An experimental production task was administered. Responses were recorded, annotated, described, and acoustically analyzed. Results revealed positive correlation between emphasis levels and phonetic duration but no other significant comparisons across levels of emphasis were found. These suggest that, based on the articulatory facility of speakers, they can produce and seem to prefer binary phonetic length distinction, and find it inconsequential to maintain clear durational distinction on other levels of emphasis.

Keywords: *vowel length, phonetic duration, emphatic speech, Cebuano*

1. Introduction

In many languages of the world, phonetic length indicates lexical contrasts, i.e., it signals different meanings between two similarly spelled words.[*] This phonetic lengthening is possible for both consonants and vowels. Whether it is due to articulatory or phonological reason, a binary length distinction seems to be preferred. Languages such as Malayalam (Local & Simpson, 1999), Toba Batak (Cohn, Ham, & Podesva, 1999), Pattani Malay (Abramson, 1987), and Itawis in the Philippines (Bollas, 2013) exhibit binary length contrasts. Mixe (Hoogshagen, 1959), a language of Southern Mexico, is even observed to have clear ternary contrast in its vowels. However, many linguists are still skeptical of other languages that insist three-way phonetic length distinctions (e.g., Estonian; Prince, 1980). Literature has shown that this type of phonetic behavior can be explained through phonological and morphological processes (Lehiste, 1970; Chen, 1970; Shryock, 1993; Ladefoged & Maddieson, 1996; Riduoane, 2010). At present, a growing number of literature are describing how phonetic length is also discriminated in non-length contrastive languages (Braver, Dresher, & Kawahara, 2014).

[*] I am grateful to the reviewers of the 19th International Congress of Phonetic Sciences, the audience at the Linguistics Society of the Philippines International Conference 2019, Dr. David Weenink for assisting me with the completion of the Praat script needed for this project, and the journal article reviewers of the earlier versions of the manuscript. Moreover, thanks to the volunteer research assistants of NDMU-SHS for completing the first passes of data annotation. The author declares that there is no financial or personal interest, or belief that could affect the objectivity of the results of this project. Specifically, funding is secured, and independently and personally raised by the author from his salary.

For Japanese, lexical contrast via phonetic length is possible on both consonants and vowels (Tsujimura, 2007). With regard to the function of vowel length, a recent study by Kawahara and Braver (2013) provides clues on emphatic lengthening patterns in speech that go beyond the common binary distinction. In producing emphasis, Japanese speakers seem to "potentially make six-way durational [vowel] distinctions" (Kawahara & Braver, 2013; p. 146). The ability to produce these patterns is not only limited to languages with binary phonetic length distinction (Braver et al., 2014). For example, in a study in English, a language that does not have contrastive vowel length, speakers were able to similarly produce the fine-grained durational contrast as the Japanese speakers. In both studies of Kawahara and Braver (2013) and Braver et al. (2014), lengthened phonetic production has been a feature of emphasis. Given the novelty of this approach in this context, the present study endeavors to contribute to and expand this inquiry by investigating Cebuano, an Austronesian language from the Philippines.

1.1 Cebuano vowel length

Cebuano, or commonly "Bisaya", is one of the most widely spoken languages of the Philippines with 15 million speakers (Ethnologue, n.d.). Wolff (1972) recorded three vowels and 18 consonants in the Cebuano language.

Table 1

Phonetic inventory of the Cebuano language (Wolff, 1972).

Consonants					
	Bilabial	*Apico-alveolar*	*Alveolar*	*Dorsal*	*Glottal*
VOICELESS STOPS	p	t	c	k	ʔ
VOICED STOPS	b	d	j	g	
NASALS	m	n		ŋ	
SPIRANTS		s			h
LIQUIDS	w	l, r	y		

Vowels		
High- or Mid-front	*Low Central*	*High- or Mid-back*
i	a	u

Vowel length in Cebuano is readily observable. Bunye and Yap (1971), and later Shryock (1993), indicated that Cebuano possesses vowel length distinction affected by morphological processes, word stress, and context, among others, as shown below.

(1) *Verbal suffixations*
- a. [ˈlaːlis] 'argue' -> [la.ˈliː.s-an] 'argue-local.passive.punctual.future'
- b. [ˈhaːtag] 'give' -> [ha.ˈtaː.g-an] 'give-local.passive.punctual.future
- c. [ˈpajpaj] 'fan' -> [paj ˈpaː.j-i] 'fan-local.passive.punctual.subjunctive'

(2) *Stress*
- a. [ˈdaː.pit] 'place (n.)' -> [da.ˈpit] 'invite'
- b. [ˈbuː.hiʔ] 'let go' -> [bu.ˈhiʔ] 'alive'
- c. [ˈiː.num] 'drink heavily' -> [i.ˈnum] 'drink (v.)'

(3) *Context*
a. [baːj] ‘house’; [baːj] vocative for a male friend
b. [saː] ‘living room’; [saː] PREP
c. [uː] ‘head’; [uː] ‘yeah’

Data set (1; Shryock, 1993) shows how stress affects vowel length in Cebuano, but with an additional condition, that is, when the word undergoes verbal suffixation. Data set (2; Defenders of Indigenous Languages of the Archipelago, n.d.) demonstrates how stress assignment can modify vowel length in a syllable. In general, stressed primary syllables undergo lengthening, while placing stress on the second syllable removes vowel length altogether in order to accomodate intended meaning, further showing a contrastive feature. Data set (3; Bollas, 2013) includes classic examples already-lengthened vowels and how meaning changes based on context. For the latter data set, any Cebuano speaker will soon recognize that vowel lengthening in this case involves yet another process. I now briefly discuss this process below.

The omission of /l/ is common in spoken Cebuano, and different scholars use different terms to refer to this process. The outputs are lengthend vowels specifically on the penultimate open syllable’s vowel, e.g., [ˈŋa.lan] ‘to name’ to [ŋaːn], [tu.ˈlu] ‘three’ to [tuː], [ˈdi.liʔ] ‘no’ to [diːʔ]. In an extensive list of Cebuano phonetic rules (Defenders of Indigenous Languages of the Archipelago, n.d.) this process is considered both syncope and apocope. However, in Newton’s (1991) generative approach to Cebuano phonology, he views syncope in Cebuano differently, that is, devoid of the *l*-omission. The term “*l*-dropping”, preferred in this present discussion, is first observed and used by Wolff (1972). This intervocalic *l*-dropping, i.e., V_1lV_1, mainly occurs when V_1 is either /a/ or /u/. Thus, when /l/ is dropped V_1 seems to be lengthened as it tries to compensate for the syllabic space occupied by /l/. This can be theoretically explained using the metrical system of Cebuano (Newton, 1991), but a deeper acoustic examination of this phenomenon may establish this. Furthermore, it is possible to drop /l/ between different vowels, /a/ and /u/, and /u/ and /a/, where the the phoneme /l/ changes to [w], e.g., [la.ˈlum] ‘deep’ to [la.ˈwum] or in rapid speech just [lawm̩], or [ˈdu.laʔ] ‘play’ to [ˈdu.waʔ]. Again these all happen in spoken Cebuano. This innovation from /l/ to [w] cannot occur, however, for /i/ pairs, e.g., the /i/-/a/ pair [ˈdi.laʔ] ‘tongue’ to *[di.waʔ], or the /u/-/i/ pair [ˈtu.liʔ] ‘to circumcise’ to *[ˈtu.wiʔ].

Evidently, vowel lengthening is not foreign to the Cebuano language. Phonetic length is mostly associated with syllable properties, metrical stress, and the morphology of the language (Bunye & Yap, 1971; Zorc, 1993). Still, similar to Japanese, English, and other languages in the world, Cebuano is observed to express emphatic utterances via lengthening.

1.2 Emphatic utterance

In Cebuano, the sentence “*I am very hungry*” can be expressed emphatically in two grammatical ways:

(4)	*Gutom*	*kaayo*	*ko.*
	gu.ˈtu:m	ka.ˈʔa.ju	ku
	hungry	INT	1SG.ERG
	I am very hungry.		

(5)	*Hastang*	*gutom*	*-a*	*na*	*nako.*
	has.ˈta:ŋ	gu.ˈtu.ma		na	ˈna.kuʔ
	INT	hungry	-DPS	already	1SG.ERG
	I am very hungry (already).				

To highlight, the difference between these examples is the assignment of lengthened vowels and syllables, i.e., /V:/. Cebuano may lengthen the vowel in the final stressed syllable of an adjective in (4), in this case [gu.ˈtu:m], but not the intensifier as in (6) below, as with English, or the word *gutom* as it is syntactically ordered in (7; length indicator colons in the transcriptions visualize this). In the syntax and utterance of (5), it is always the vowel in the final syllable of an intensifier that is lengthened, i.e., [has.ˈta:ŋ] vis-à-vis with (7). Still, there are variants for the intensifier word in (5) including *haskang, askang, astang*, and *arang*. The non-/ŋ/ ending variants *asta* and *aska* might also be possible utterances but not in the case in example (5).

(6)	*#Gutom*	*kaayo*	*ko.*
	gu.ˈtum	ka.ˈʔa:.ju	ku
	hungry INT		1SG.ERG
	I am veeery hungry.		

(7)	*#Hastang*	*gutom*	*-a*	*na*	*nako.*
	has.ˈta:ŋ	gu.ˈtu:.ma		na	ˈna.kuʔ
	INT	hungry	-DPS	already	1SG.ERG
	I am very huuungry (already).				

The morphophonemic and prosodic merits of the above constructions of (6) and (7) deserve a separate paper, and will not be covered here, e.g., why lengthening the indicated syllables seems unnatural in these utterances (see Kawahara, 2007)? On another note, it can be argued that one word utterances can also indicate emphasis, e.g., [u:] 'yeah' for [ˈu. ʔu] 'yes'. The present study's concern, however, is vowel lengthening variations on adjectives as shown in (4).

1.3 Present study

There are observed differences in levels of emphatic length production between length contrastive and non-length contrastive languages, and speakers appear to be equally sensitive to fine-grained emphatic length distinctions regardless of the fact (Braver et al., 2014). To date, there is no known study that documented emphatic lengthening in any of the Philippine languages apart from generalized observation. This study positions itself as the first exploratory investigation of the phonetic length distinction of emphasis among Philippine language speakers, specifically among Cebuano speakers.

2 Methods

This study explored Cebuano speakers' ability to produce fine-grained durational distinctions. The experiment will have supported or rejected the notion of cross-linguistic length sensitivity in emphatic utterances.

This section details the methodological component of the study.

2.1 Project approval

This independent experimental study was first deliberated with the Notre Dame of Marbel University's Faculty Research Council members. Having successfully addressed ethical, procedural, and technical concerns of the council, I obtained approval for the immediate implementation of this project.

2.2 Participants

Seven female volunteers, average age of 19, who are native speakers of Cebuano, speakers since age 2, participated in the study. A letter was sent to the school and an approval was secured. Participants were senior high school students of a local high school in General Santos City. Their time was compensated with Php 50.00 and a meal.

2.3 Stimuli

Six adjectival target words were chosen for this study, all of which were familiar to the participants. The stimuli were grouped according to the final vowels they are orthographically written, i.e., *i*, *a*, and *o*. All target words, however, phonetically end with a glottal stop [ʔ]. Chosen words were disyllabic bearing a CV.CVC structure, with stress on the second syllable.

Table 2

List of target words.

[i]	***[a]***	***[u]***
[la.ˈmiʔ] 'delicious'	[la.ˈtaʔ] 'rotten'	[da.ˈkuʔ] 'big'
[pu.ˈtiʔ] 'white'	[ba.ˈsaʔ] 'wet'	[ba.ˈhuʔ] 'smelly'

Every *i-, a-*, and *o*-final words were framed first in a non-emphatic sentence, e.g., *Lami kaayo ang sud-an*. To explore emphatic length production distinctions, each target word was placed in a frame sentence while varying their orthographic length ranging from level 1 through 5, i.e., *lamiii, lamiiii, lamiiiii, lamiiiiii,* and *lamiiiiiii*. Noticeably, level 1 starts with three orthographic *i*s. This precaution has been undertaken because, for example, the initial word *lami*, in *Lami kaayo ang sud-an* may be interpreted as "to prepare something very well/deliciously" when written as *Lamii kaayo ang sud-an*. Deciding whether to produce the emphatic [la.ˈmiːʔ] or [la.ˈmiʔi] for the stimuli *Lamii kaayo ang sud-an* poses a challenge. Hence, level 1 for all adjective stimuli begins with three orthographic iterations of the final vowel to help participants process the word as adjectives and not verbs, consequently disambiguating between emphasis and commands; the former being the focus of the present study. Words that end with the vowel *-a* and *-o* were also modified accordingly as seen in Table 3.

There was a total of 36 stimuli (3 vowels * 2 adjectives * 6 emphasis levels). Each was assigned a number for transcription and analysis purposes.

2.4 Procedures

Each session comprised a simple sociolinguistic background questionnaire, and a production task. Cebuano was used to interact with the participant throughout the stages of the experiment. Recordings were done in the quietest place on campus. Participants were informed about the

purpose of the experiment, which is documenting various levels of emphasis in Cebuano. After filling out consent forms and other IRB-compliant documentary requirements, participants individually took the experiment. Each was instructed to read every sentence as naturally as possible.

Table 3

List of stimuli.

No emphasis	Level 1	Level 2	Level 3	Level 4	Level 5
Lami *kaayo ang sud-an.*	Lamiii…	Lamiiii…	Lamiiiii…	Lamiiiiii…	Lamiiiiiii…
Puti *kaayo ang iro.*	Putiii…	Putiiii…	Putiiiii…	Putiiiiii…	Putiiiiiii…
Lata *kaayo ang kan-on.*	Lataaa…	Lataaaa…	Lataaaaa…	Lataaaaaa…	Lataaaaaaa…
Basa *kaayo imong likod.*	Basaaa…	Basaaaa…	Basaaaaa…	Basaaaaaa…	Basaaaaaaa…
Dako *kaayo iyang mata.*	Dakooo…	Dakoooo…	Dakooooo…	Dakoooooo…	Dakooooooo…
Baho *kaayo ang kanal.*	Bahooo…	Bahoooo…	Bahooooo…	Bahoooooo…	Bahooooooo…

The production task was composed of two parts: practice and experimental session. In the practice session, each participant was given a sentence on every adjective from different emphasis levels, which makes a total of six practice sentences. This was done to confirm that the participant recognizes all the words and understand the task. They were also given time to ask questions before the experimental session.

The experimental session consisted of seven blocks. In each block, respondents repeat all 36 randomized sentences displayed on PowerPoint on black text and white background. Randomization is generated through Research Randomizer (Urbaniak & Plous, 2018). This resulted to a total of 252 tokens for each round of experiment. Participants were given a one-minute break after each block, which they may decide to take or not. Audio recordings were captured using Jabra UC Voice 550 MS Duo Lync optimized corded headset, and a computer running Audacity at a 16-bit resolution with sampling rate of 22,050 Hz. Each participant completed the session in less than 45 minutes. Few tokens were skipped and were still mispronounced even with the considerations made to prevent this as explained in the previous section.

2.5 Acoustic analysis

All target words end with a glottal stop. This provided a clearer measurement of the end of emphasized vowels.

Durational boundaries were marked and extracted using Praat (Boersma & Weenink, 2018) using waveform and spectrogram as guides. Three unique passes were administered to confirm that acoustical boundaries were accurate. A sample of the segmentation procedure is shown in Figure 1.

2.6 Statistics

Similar to the work of Kawahara and Braver (2013) and Braver et al. (2014), a Pearson correlation (*r*) was used to measure correlation between the five emphasis levels and duration, i.e., "no

emphasis" stimuli is excluded. A linear regression analysis on the emphatic conditions followed to determine the increase in vowel duration for each level.

Given the multiplicity of comparisons of every condition in this experiment (5 emphasis levels * 3 vowels * 7 participants), a pair-wise comparison was not pursued to avoid Type I error. However, as learned from Braver et al. (2014), an independent *t*-test for each speaker via Bonferroni adjustment of significance level to $\alpha = .01$ in each successive pair of comparison (5 levels for each speaker; 0.5/5) was used. In addition, error bars were employed to aid visualization of this variance. All statistical treatment and figures were processed using SPSS.

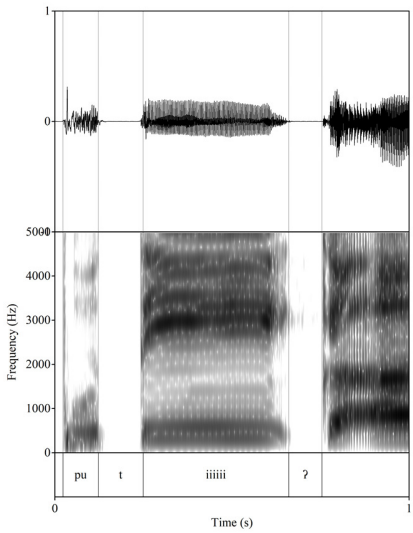

Figure 1. Segmentation procedure for annotating and extracting vowel duration.

3 Results

Each speaker has varying distinctions on emphatic vowel length. To highlight these differences, individual speaker results are discussed. The speaker that demonstrated clear graduated levels of emphasis is presented first, based on the *r*-values. Also, a summary of results and comparison among speakers are provided.

3.1 Individual speaker

As shown in Figure 2, Cebuano Speaker (CS) 1 has the highest Pearson *r*-value of correlation between emphasis and duration at $r = .47$ ($p < .001$), a statistically significant correlation. CS1 shows only a clear distinction between level 0 (no emphasis; $M = 142.47$, $SD = 52.69$) vs. 1 ($M = 296.82$, $SD = 55.37$), $t(80.00) = 12.93$, $p < .001$, and level 4 ($M = 360.21$, $SD = 96.085$) vs. 5 ($M = 430.21$, $SD = 108.181$), $t(78.00) = 3.06$, $p < .01$, but there is no significant level of comparison with other levels when analyzed via non-paired *t*-test (see Appendix C). Using linear regression analysis, each level of emphasis for CS1 can be distinguished for each 30 ms increase.

While CS2 (Fig. 3) has Pearson *r*-value of $r = .29$ ($p < .001$), larger than the last five Cebuano speakers, it demonstrates similar non-distinction of emphasis level as the rest. Non-paired *t*-test further reveals that there are no significant comparisons across levels except for level 0 ($M = 120.50$, $SD = 24.91$) vs. 1 ($M = 182.19$, $SD = 36.264$), $t(70.70) = 9.01$, $p < .001$. A great deal of fluctuation is observed from levels 1 through 5, where 5 ($M = 212.76$, $SD = 48.48$) is even recorded

having a lower duration than level 4 (M = 360.21, SD = 96.09), t(80.00) = 0.57, p > 0.01. This is subsequently revealed by a very small regression coefficient of 8 ms.

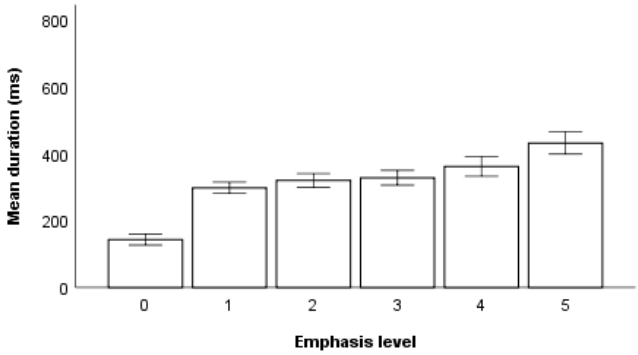

Figure 2. Results for CS1. r = .47, regression coefficient = 30 ms.
(Error bars all throughout represent 95% CI.)

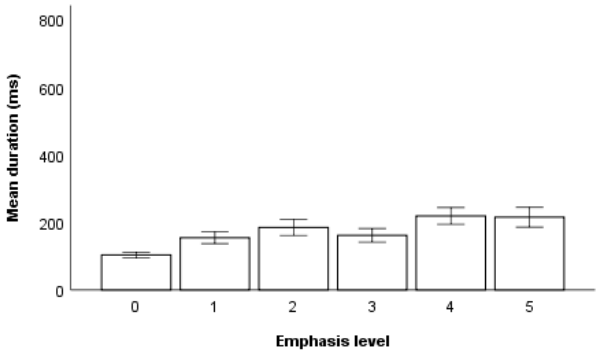

Figure 3. Results for CS2. r = .29, regression coefficient 8 ms.

CS3 and CS4 have very close r-values result at r = 0.27 and 0.26, a significant correlation at p < .001. Using non-paired t-test, CS3 level 0 (M = 103.45, SD = 27.90) and level 1, (M = 154.45, SD = 58.60), t(80.00) = 7.22, p < .001, and CS4 level 0 (M = 131.30, SD = 55.63) and level 1 (M = 239.22, SD = 78.40), t(51.77) = 5.01, p < .001, clearly distinguish emphasis level 0 and 1 yet did not show a statistically significant distinction with other paired levels. Overlaps in error bars across levels 1 through 5 show little changes in duration (see Fig. 4 and Fig. 5). However, it can be noticed that CS4 has a fairly stable durational increase for each level shown by error bars in Fig. 5, but not enough to be statistically significant based on the conducted t-test. Regression coefficient values, for which speakers determine increase in emphasis level, are at a very small measure of 15 ms for CS3, and 19 ms for CS4.

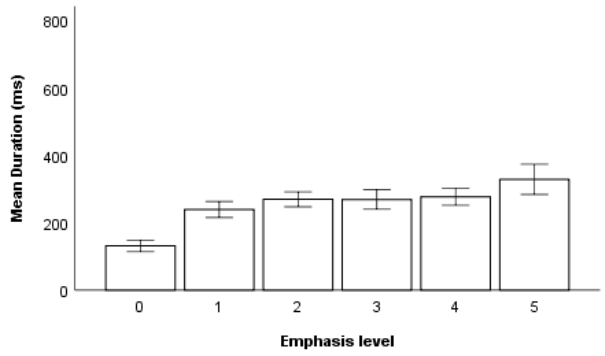

Figure 4. Results for CS3. $r = .27$, regression coefficient = 15 ms.

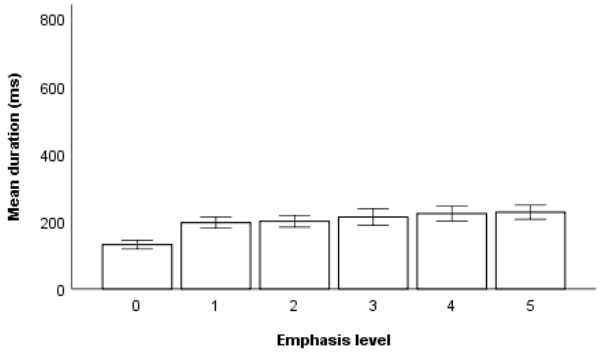

Figure 5. Results for CS4. $r = .26$, regression coefficient = 19 ms.

CS5 (Fig. 6) and CS6 (Fig. 7) have noticeable fluctuations from level 1 through 4 similar to CS2. Pearson *r*-values show a statistically significant correlation of $r = .22$ for CS5 and $r = .25$ for CS6 at $p < .001$. These two speakers, CS5 level 0 ($M = 159.03$, $SD = 54.18$) and level 1 ($M = 259.51$, $SD = 73.43$), $t(80.00) = 7.08$, $p < .001$, and CS6 level 0 ($M = 218.13$, $SD = 71.56$) and level 1 ($M = 316.03$, $SD = 83.28$), $t(80.00) = 5.72$, $p < .001$, seem to launch from a clear level 0 vs. 1 ($p < .001$) distinction and followed by indistinguishable levels 1, 2, 3, 4. At level 5 the speakers reach their max durations but to no avail of establishing clear distinctions even at level 4 vs. 5, i.e., CS5 level 4 ($M = 286.13$, $SD = 90.29$) and level 5 ($M = 316.30$, $SD = 90.75$), $t(80.00) = 1.51$, $p > .01$, and CS6 level 4 ($M = 356.19$, $SD = 95.51$) and level 5 ($M = 366.23$, $SD = 114.87$), $t(78.00) = 0.42$, $p > .01$. Appendix C records that when analyzed via non-paired *t*-test, there is no statistically significant difference across these levels. Regression coefficients indicate a 14 ms durational increase for CS5 and 15 ms for CS6 for each step of emphasis level taking into account durational variability.

Lastly, CS7 (Fig. 8) has the lowest Pearson *r*-value at $r = .18$ ($p < .001$). A non-paired *t*-test shows that, similar to previous speakers except for CS1, CS7 only and clearly distinguishes level 0 ($M = 131.12$, $SD = 42.02$) vs. 1 ($M = 196.14$, $SD = 53.30$), $t(80.00) = 6.15$, $p < .001$, and no other comparison shows significance (see Appendix C). Linear regression analysis also reveals a very minute level of increase for each emphasis level at 9 ms.

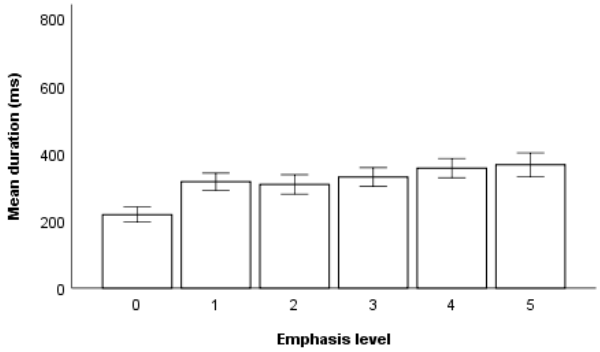

Figure 6. Results for CS5. *r* = .25, regression coefficient = 14 ms.

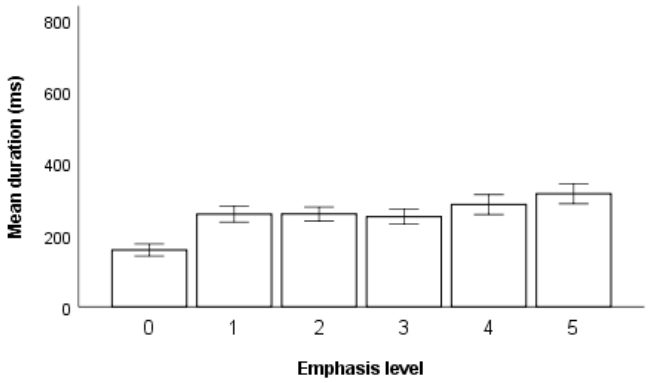

Figure 7. Results for CS6. *r* = .22, regression coefficient = 15 ms.

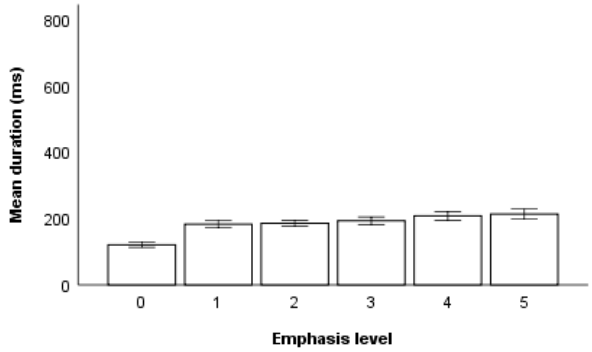

Figure 8. Results for CS7. *r* = .18, regression coefficient 9 ms.

3.2 Summary

Table 4 provides a summary of each Cebuano speaker (CS) data. Pearson *r*-value, regression coefficient to indicate durational increase for each step of emphasis, and the recorded maximum range of lengthened vowels are reported.

Table 4.

Pearson r-value, regression coefficients, and maximum vowel/rhyme duration.

Cebuano Speaker	*r*	Coefficient (ms)	Max duration (ms)
1	.47	30	643
2	.29	8	344
3	.27	15	487
4	.26	19	752
5	.25	14	575
6	.22	15	634
7	.18	9	530

All speakers showed a positive correlation between emphasis and length duration ($p < .001$). Speakers' *r*-values have a range $r = .47$ from CS 1 and $r = .18$ from CS 7. All produced a clear distinction, a jump, between level 0 (no emphasis) and 1 with surprisingly small durational gaps on other levels.

Only CS1 (Fig. 2), CS4 (Fig. 5), and CS7 (Fig. 8) revealed a consistent progression and an increase from level 0 through 5. However, these speakers still failed to produce a clear length distinction as revealed in the results of non-paired *t*-tests, and the overlap of error bars on levels of emphasis for each speaker in a way visualize this. Essentially, all speakers did not demonstrate a clear distinction. Most of the speakers even have noticeable reversals, specifically level 2 for CS5 (Fig. 6), level 3 for CS3 (Fig. 4) and CS6 (Fig. 7), and level 3 and 5 for CS2 (Fig. 3).

In general, a correlation between *r*-values and maximum duration is weak. CS1 has the highest *r*-value but not the highest maximum duration while CS2 has the second highest *r*-value but it has the lowest maximum duration. The highest maximum duration is CS4 but its *r*-value is not the highest overall.

4 Discussion

4.1 Summary

This study documents the first record of empathic lengthening behavior on Cebuano vowels. All the speakers hardly distinguished levels of emphasis and seemed to subscribe to the usual two-way durational distinction of short/long or no emphasis/emphasis, except for CS1 that identifies a three-way durational distinction. The implication of this study is that Cebuano speakers, unlike Japanese (Kawahara & Braver, 2013) and English (Braver et al., 2014), do not produce fine-grained durational distinctions for emphasis nor they possess such an articulatory control to do so. Also, this limitation is seemingly significant because it now starts to point Cebuano emotive utterance research to considerations other than length.

4.2 Preference to two-way length distinction

As shown in the results, Cebuano speakers are well aware of null emphasis and emphasized sentences, yet such native distinctions start to break apart from the second until the fifth level of emphasis. This is not the case of other non-vowel length contrastive language such as English. For English, speakers appear to clearly produce and distinguish up to six levels of phonetic length (Braver et al., 2014). Furthermore, such distinctions have also been observed by Kawahara and Braver (2013; 2014) on Japanese. The apparent sensitivity of Cebuano speakers on cues other than length may have limited the speaker's ability to successfully distinguish length as shown in this

study (see Zorc, 1993). While adding more participants and drawing data from conversational and uncontrolled utterances are desirable, for now, results suggest that there seem to be a preference to a two-way vowel length contrast for emphatic utterances, i.e., emphasis and no-emphasis.

4.3 Orthographic counting

It is also worth noting that using orthographic replication as indicator of emphasis in the stimuli does not interfere with the aims of the study. Simply, if participants count orthographic iterations of vowels, we may then expect linear and exact correlations between levels of emphasis and duration. Moreover, if counting has been employed by the speakers, we cannot expect significant jump from level 0 and level 1, or even incongruous variation on other emphasis levels, given the number of the added orthographic vowels for each level (see Braver et al., 2014).

4.4 Dispersion Theory

The non-distinction of the five levels of emphasis and length contrast among Cebuano speakers may relate closely to what linguists call Dispersion Theory (Liljencrants & Lindblom, 1972). The theory includes a constraint-based approach in identifying and perceiving phonetic contrasts, which resembles principles from Optimality Theory. In relation to the present study, since there is not much movement in the vowel space for each lengthened syllable vowel, plus multiple lengthening is not naturally contrasted in the Cebuano language, non-distinction of graduated level of emphasis can be expected. While this may be readily observable for such languages, perception studies may actually be pursued, even for Cebuano and English.

Moreover, even with a very low regression coefficient of 30 ms in the present study, compared to other parallel studies, e.g., Kawahara & Braver (2013), using this value as basis for a perception task may be enough as human listeners are sensitive to durational differences of as short as 12.5 ms (Näätänen, Paavilainen, Alho, Reinikainen, & Sams, 1989). The question remains, however, whether this much nuance can be perceived as a reliable and most preferred cue for emphasis distinction or not.

5 Limitations

5.1 Syllable structure

Different phones interact distinctively with others. In the case of this study, it must be known that while the most careful attention was given in placing durational boundaries, the nasal sound /m/ and the fricative /s/ have been the challenging ones to pin down, for example in the stimuli *Basaaaa kaayo imong likod*. Here, it can be difficult to set a boundary for the start and end of /s/ and /a/ respectively. At the moment, this is addressed based on the intensity of striations of the particular segment's F1 and F2.

5.2 "Naturalness" of data

The experimental, laboratory nature of this study may cause applied linguists and other researchers in other fields to raise concerns about the naturalness of the recorded data. The key challenge for this project is isolating and targeting a particular phonetic behavior, but the current body of literature suggests that the methodological preference in the present study is reliable up to a point (Kawahara, 2007). As mentioned earlier, participants were asked to produce the stimuli as 'naturally' as possible. Nevertheless, future iteration or versions of this project may greatly benefit from the methodological strengths of other research fields including applied linguistics and language documentation, among others.

6 Conclusion

There are other linguistic cues for emphasis that may be employed by speakers for emphasis such as pitch and loudness. Length, on the other hand, does not seem to be as exploited in Cebuano as in other languages like Japanese and English. However, the diverging results of the present study from earlier findings can still neither establish nor nullify the crosslinguistic capacity of speakers to produce fine-grained durational distinction. There may still be more linguistic, speaker, and participant number factors that affect the results.

Nevertheless, for Cebuano, spoken and emphatic utterances like "*Putiiii kaayo ang iro*" and "*Putiiiii kaayo ang iro*", are treated without much discrimination, acoustically speaking. Considering other prosodic indices in future studies may provide a better view of the acoustic and articulatory configurations of emphatic speech in terms of the most preferred cues used by speakers in emphasizing words to indicate greatness in intensity, weight, shape, or size inter alia.

References

Abramson, A. S. (1987). Word-initial consonant length in Pattani Malay. *Haskins Laboratories Status Report on Speech Research*, *SR-91*, 143-147.

Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.40, retrieved 11 May 2018 from http://www.praat.org/

Bollas, A. (2013). *Comparative analysis on the phonology of Tagalog, Cebuano and Itawis*. Unpublished manuscript.

Braver, A., Dresher, N., & Kawahara, S. (2014). The phonetics of emphatic vowel lengthening in English. In A. Albright, & M. A. Fullwood (Eds.), *Proceedings of the 2014 Annual Meeting on Phonology* (Vol. 2). Cambridge, MA, USA: MIT.

Bunye, M., & Yap, E. (1971). *Cebuano grammar notes*. Honolulu, HI: University of Hawaii Press.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica, 22*, 129–159.

Cohn, A., Ham, W., & Podesva, R. (1999). The phonetic realization of singleton–geminate contrasts in three languages of Indonesia. In J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 587–590). San Francisco, CA, USA: University of California.

Defenders of Indigenous Languages of the Archipelago. (n.d.). *Cebuano Phonetics and Orthography*. Retrieved from https://dila.ph/cebuano%20phonetics%20and%20orthography.pdf

Ethnologue. (n.d.). Retrieved September 14, 2018, from https://www.ethnologue.com/language/ceb

Hoogshagen, S. (1959). Three contrastive vowel lengths in Mixe. *STUF-Language Typology and Universals*, *12*(1–4), 111–115.

Kawahara, S. (2007). *The emergence of phonetic naturalness* (Doctoral dissertation). Retrieved from Scholarwords at UMass Amherst. (Accession No. AAI3289261)

Kawahara, S., & Braver, A. (2013). The phonetics of multiple vowel lengthening in Japanese. *Open Journal of Modern Linguistics*, *3*(2), 141–148.

Kawahara, S., & Braver, A. (2014). Durational properties of emphatically lengthened consonants in Japanese. *Journal of the International Phonetic Association*, *44*(3), 237–260.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford, UK: Blackwell.
Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
Liljencrants, J. & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast, *Language*, *48*, 839–862.
Local, J., & Simpson, A. P. (1999). Phonetic implementation of geminates in Malayalam nouns. In J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 595–598). San Francisco, CA, USA: University of California.
Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neuroscience letters*, *98*(2), 217-221.
Newton, B. (1991). The Cebuano language and generative phonology. *Philippine Quarterly of Culture and Society*, *19*(4), 253–263.
Prince, A. S. (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry*, *11*(3), 511–562.
Ridouane, R. (2010). Gemination at the junction of phonetics and phonology. In C. Fougeron, B. Kuhnert, M. D'Imperio, & N. Valle (Eds.), *Papers in laboratory phonology X* (pp. 61–90). Berlin, Germany: Mouton de Gruyter.
Shryock, A. (1993). A metrical analysis of stress in Cebuano. *Lingua*, *91*(2–3), 103–148.
Tsujimura, N. (2007). *An introduction to Japanese linguistics* (2nd ed.). Cambridge, MA: Blackwell.
Urbaniak, G. C., & Plous, S. (2018). Research Randomizer [Computer software]. Version 4.0, retrieved 23 May 2018 from http://www.randomizer.org/
Wolff, J. (1972). *A dictionary of Cebuano-Visayan* (Vols. 1-2). Ithaca, New York: Cornell University.
Zorc, R. (1993). Overview of Austronesian and Philippine Accent Patterns. *Oceanic Linguistics Special Publications, 24*, 17–24.

Appendix A. Grammar notation

1SG	–	first person singular
ERG	–	(postponed) ergative case
DPS	–	direct passive punctual subjunctive
INT	–	intensifier

Appendix B. Phonetic transcription notation

ˈ (apostrophe)	–	stress
ː (colon)	–	length
- (dash)	–	morpheme boundary
. (dot)	–	syllable boundary

Appendix C. Independent *t*-test results

Cebuano Speaker	Comparison	mean diff. (ms)	t(df)	p
1	level 0 vs. level 1	154.35	$t(80.00) = 12.93$	$p < .001$
	level 1 vs. level 2	21.86	$t(79.00) = 1.60$	n.s.
	level 2 vs. level 3	8.07	$t(80.00) = 0.53$	n.s.
	level 3 vs. level 4	33.45	$t(79.00) = 1.29$	n.s.
	level 4 vs. level 5	70.00	$t(78.00) = 3.06$	$p < .01$
2	level 0 vs. level 1	61.69	$t(70.70) = 9.01$	$p < .001$
	level 1 vs. level 2	2.54	$t(79.00) = 0.35$	n.s.
	level 2 vs. level 3	7.91	$t(80.00) = 1.09$	n.s.
	level 3 vs. level 4	14.40	$t(81.00) = 1.67$	n.s.
	level 4 vs. level 5	5.71	$t(80.00) = 0.57$	n.s.
3	level 0 vs. level 1	107.91	$t(80.00) = 7.22$	$p < .001$
	level 1 vs. level 2	31.13	$t(76.00) = 1.79$	n.s.
	level 2 vs. level 3	1.61	$t(77.00) = 0.08$	n.s.
	level 3 vs. level 4	7.83	$t(78.05) = 0.40$	n.s.
	level 4 vs. level 5	52.104	$t(61.07) = 1.96$	n.s.
4	level 0 vs. level 1	51.00	$t(51.77) = 5.01$	$p < .001$
	level 1 vs. level 2	30.27	$t(74.00) = 1.90$	n.s.
	level 2 vs. level 3	22.99	$t(74.00) = 1.36$	n.s.
	level 3 vs. level 4	57.33	$t(73.00) = 3.24$	n.s.
	level 4 vs. level 5	3.93	$t(74.00) = 0.19$	n.s.
5	level 0 vs. level 1	100.48	$t(80.00) = 7.08$	$p < .001$
	level 1 vs. level 2	0.20	$t(78.00) = 0.01$	n.s.
	level 2 vs. level 3	7.62	$t(80.00) = 0.52$	n.s.
	level 3 vs. level 4	34.03	$t(81.00) = 1.95$	n.s.
	level 4 vs. level 5	30.17	$t(80.00) = 1.51$	n.s.
6	level 0 vs. level 1	97.90	$t(80.00) = 5.72$	$p < .001$
	level 1 vs. level 2	8.15	$t(78.00) = 0.41$	n.s.
	level 2 vs. level 3	21.93	$t(80.00) = 1.08$	n.s.
	level 3 vs. level 4	26.37	$t(79.00) = 1.29$	n.s.
	level 4 vs. level 5	10.014	$t(78.00) = 0.42$	n.s.
7	level 0 vs. level 1	65.02	$t(80.00) = 6.15$	$p < .001$
	level 1 vs. level 2	4.06	$t(78.00) = 0.33$	n.s.
	level 2 vs. level 3	12.63	$t(80.00) = 0.83$	n.s.
	level 3 vs. level 4	10.48	$t(80.00) = 0.63$	n.s.
	level 4 vs. level 5	3.83	$t(79.00) = 0.25$	n.s.