The Role of Pitch, Duration, and Lexical Tone in the Production of Voiced and Voiceless Burmese Nasals

Ludwig Adisiswoyo

Linguistics, University of California, Santa Barbara

Abstract

Burmese is a language of South-East Asia featuring a contrast between voiced and voiceless nasals. Voicing is an articulatory phenomenon involving the vibration of vocal folds and is the mechanism behind contrastive sounds in English such as /p/-/b/ and /t/-/d/. This contrast pertains to nasals—a typically voiced category including English consonants such as /m/ and /n/—in Burmese. I conducted a production study examining acoustic properties associated with the voicing contrast in Burmese nasals. The results confirmed well attestedpatterns found in the literature and includes a novel finding regarding an interaction between three factors and its correlation with voicing.

Introduction

In the world's languages, sounds may be categorized into perceptually distinct units called phonemes that are used to differentiate words. Phonemes may be as different as /l/ and /t/, used to distinguish the English words lackand tack, or as similar as /t/ and /d/, which distinguish pairs of words like toand do. The consonants /t/ and /d/ draw on a distinction calledvoicing, which results from the timing of vocal fold vibration relative to the production of a consonant. The vocal folds are located in the larynx andmodulate airflow from the lungs during speech production. Consonants are voiced if the vocal foldsvibrate before the release of the consonant, as in /d/, and are voiceless if vocal fold vibration occurs after its release, as in /t/.Burmese is a language of South-East Asia (spoken primarily in Myanmar) featuring a typologically unusual voicing contrast between its nasals. Nasals are a group of consonants including the English /m/, /n/, and /ŋ/ (orthographically represented as ng). They are produced by a lowering of the velum, a sluggish articulator located at the back of the mouth, which when lowered allows for airflow to pass through the nasal cavity. This is in combination with gestures of other articulators involved in producing consonants, such as the lips, tongue, teeth, and palate.

In Burmese, nasals are found at four places of articulation, the location within the oral cavity where articulators form a constriction for a consonant: 1. Bilabials, which are produced by movement ofthe lips, resulting in consonants such as /m/ and /b/. 2. Alveolars, which involve contact between the tongue tip and the alveolar ridge, a protruded area at the roof of the mouth behind the teeth, and include /t/, /s/, and /n/. 3. Palatals, which involvecontact between the tongue body and the hard palate, and include /ɲ/, the consonant in the middle of Spanish words such as baño. 4. Velars, which are produced by contact between the back of the tongue and the velum, and include /k/ and /ŋ/. Voicing contrast occurs at all these locations in Burmese nasals. Although relatively rare, voiceless nasals have been observed in various regions of the world, with concentrations in South-East Asia, Central America, the Pacific Northwest, and even in Icelandic. Typically, they can be divided into two categories based on their timing of voicing (i.e. vocal fold vibration): one consisting of two components, a voiceless and voiced phase, and the other consisting of only a voiceless phase (Bhaskararao & Ladefoged 1991). Languages differ in how their voiceless nasals are produced, with Burmese falling under the initial category, which are also known as pre-aspirated nasals, since they are characterized by a noisy voiceless period followed by a voiced murmur (Chirkova et al.2019, Dantsuji 1984, Ladefoged & Maddieson 1996, Maddieson 1984b, Ohala 1975, Ohala & Ohala 1993)1.Differing articulatory constrictions and the resulting airflow from oral and nasal cavities mark the contrast between the two nasal categories. Most descriptions of the voiceless period within pre-aspirated nasals only mention the presence of nasal airflow (Bhaskararao & Ladefoged 1991, Ladefoged & Maddieson 1996, Ohala 1975, Ohala & Ohala 1993). However, the voiceless phase consists of both nasal and oral airflow. The articulators remain open for the duration of the voiceless phase, only closing for the brief voiced phase preceding the vowel (Blankenship et al. 1993, Chirkova et al. 2019). Following this is the fully voiced segment consisting of only nasalairflow. This stands in contrast with fully voiceless nasals (found in languages, such as Angami, a language of South-East Asia) also known as voiceless aspirated nasals, which actually consist of two phases as well (Blankenship et al. 1993, Chirkova et al. 2019). While the oral cavity remains closed throughout most of the nasal—thus only allowing for nasal airflow—there is an opening of the articulatory stricture at the end of the nasal resulting in a brief partially voiced period with simultaneous oral and nasal airflow. The voiceless portions of these nasals do not help convey information about place of articulation since the nasal cavity, from which the turbulent airflow is released, lacks the capacity for varying location or degree of constriction (Bhaskararao & Ladefoged 1991, Chirkova et al. 2019, Dantsuji 1986, Ohala 1975). Noise spectrums cannot be generated by cavities leading up to the nostril, and the resulting nasal airflow is low in intensity (Ohala 1975, Ohala & Ohala 1993). Therefore, listeners cannot reliably perceive differences between places of articulation based on the voiceless phase alone. The voiceless phase thus mainly serves to signal a contrast with voiced nasals. On the other hand, the voiced phase serves as a transition to the following vowel and signals place of articulation (Bhaskararao & Ladefoged 1991, Dantsuji 1984, 1986, Ladefoged & Maddieson 1996, Ohala 1975, Ohala & Ohala 1993). The offglide in the transition to the following nasal displays cues characteristic of different points of articulation. Dantsuji (1986) segmented the voiced phase into three portions—onset, murmur, and transition—and found that the murmured and transition phases carry formant values indicative of manner and place of articulation The present study examines acoustic properties involved in signaling the contrast between voiced and voiceless nasals in Burmese, found for example in the minimal pair

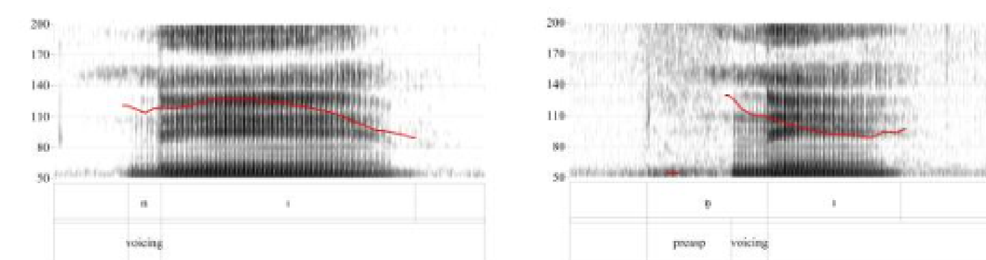
'tread/step on' vs. /n̥ĩ/ 'drive away/banish'. A pilot study suggested that the presence of the voiceless phase alone is not the only cue on which listeners rely to perceive voicing categories. Cues such as pitch, the perceptual correlate of fundamental frequency (the rate of vocal fold vibration), and duration seemed to be promising variables which are equally important: preliminary data suggested that voiceless nasals are produced with higher pitch and are longer than their voiced counterparts. Therefore, I also examined interactions between pitch and duration to assess the relative contribution of the two cues to signaling the contrast in voicing. An added focus of interest is the fact that pitch is also used to signal tone contrasts between words in Burmese. For example, the word □□ □□/m̥á/ 'wrong/incorrect' is produced with a high toneand the word □□ /m̥à/ 'locative; order (sth)' is produced with a low tone. The use of pitch both to help convey contrasts in voicing of nasals and also to mark tonal contrasts in vowels raises questions about its capacity (or "functional load") to signalboth types of contrasts.

Autonomy

This production experiment focuses on only individual words to limit the number of variables possibly influencing the results. The first step was procuring a list of minimal pairs to present to speakers. I mainly sourced words from Judson's English and Burmese dictionary(Judson & Stevenson 1953) and the Burmese index on wiktionary (Index:Burmese in Wiktionary). The results were 67 tokens containing nasals, with 21 minimal pairs. Of the 67 tokens, 53 occurred word-initially and 14 were in word-medial position. Words were varied in terms of the place of articulation of the nasal (bilabial, alveolar, palatal, velar) and the tone category (high, low, checked, and creaky). No voiced equivalents were collected for the word-medial tokens and so they were discarded from the analysis. Twenty-three more multisyllabic words with voiced and voiceless tokens in similar medial environments were also used as fillers. Following this, I consulted a native Burmese speaker to verify that all the tokens were likely to be recognized by Burmese speakers as words, resulting in 1 medial voiceless nasal and 8 filler nasal tokens being discarded. Overall, 81 tokens were randomly sorted over four iterations, resulting in 324 total tokens. I recorded four speakers reading from this list on a Marantz pmd 660 solid-state recorder and an AKG 51/SHER c520 headworn condenser microphone. In total, there were 1296 data points. Of these, 416 tokens were left out of the analysis due to them being fillers, occurring in medial position (since there is no comparison with voiced equivalents), or being errors, leaving a total of 880 data points. Nasals occurring in words targeted for analysis were segmented using Praat (Boersma and Weenink 2020) into three relevant phases: voiceless segment, voiced segment, and following vowel. The segmentation is illustrated in Figures 1a, which contains a spectrogram and an overlaid pitch trace (the red line) for a voiced nasal,and 1b, which depicts a voiceless nasal. I ran a Praat script gathering duration measurements for each phase alongside pitch measurements at the beginning, medial, and ending points of each phase. Pitch measurements were taken at the three points within each phase to provide a representation of the time course of pitch throughout each phase. These measurements, as well as speaker, gender, lexical item, place of articulation, the

presence of a voiceless phase, tone, location (whether medial or initial), the preceding vowel, the following vowel, the presence vs. absence of a coda consonant (the consonant following the vowel), and the coda consonant (the phoneme) are the result-

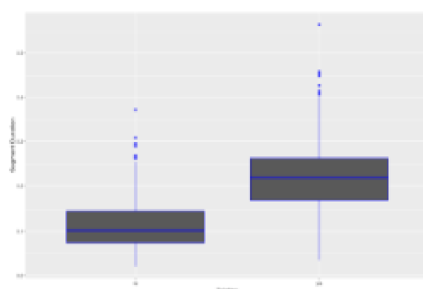
Figure 1a. နင်း /nĩ/ 'tread/step on' **1b.** နှင် /n̥ĩ/ 'drive away/banish'



Results & Discussion

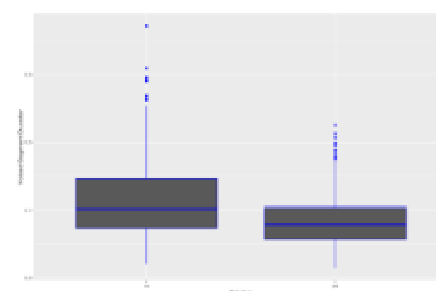
Variables were analyzed usinga mixed-effects model through the glmerfunctionin R (R core team 2017). I chose a mixed-effects modelto account for speaker variation, since it does not use fixed parameters. Therefore, if a speaker has a lower than average pitch reading in 1a.—but exhibits a pitch difference between 1a. and 1b. similar to other speakers—the outlying measurement may not be indicative of a trend, but instead is an effect of another variable. This was intended to illustrate the effectiveness of certain variables as predictors of whether a nasal is voiced or voiceless. In order to determine the most relevant variables, I used ggplot (Wickham 2016) to compare pitch and duration measurement—since they contain the only numeric values—in relation to other variables. Pitch and duration values were used as predictors of other dimensions, such as tone, voicing, place of articulation, and vowel quality, by comparing them across subcategories of these variables. These comparisons are also split by voicing category (for all comparisons except voicing) to examine whether voicing also influences measurement results. This allowed me to observe potentially two-or three-way interactions between any of these variables. One pattern can be observed in duration comparisons between different voicing categories. The duration of voiceless nasals, including both voiced and voiceless phases, are on average twice as long as voiced nasals. This is depicted in Figure 2a.Comparisons of only the voiced phases of both nasals reveal that the durations of voiced nasals are overall greater than the voiced phase of voiceless nasals. This is depicted

Figure 2a.



Vd.duration		Vl.duration	
Min.	:0.0210	Min.	:0.0350
1st Qu.	:0.0740	1st Qu.	:0.1680
Median	:0.1020	Median	:0.2190
Mean	:0.1136	Mean	:0.2224
3rd Qu.	:0.1460	3rd Qu.	:0.2630
Max.	:0.3720	Max.	:0.5640

2b.



Vd.Vd.Duration		Vl.Vd.Duration	
Min.	:0.0210	Min.	:0.01400
1st Qu.	:0.0740	1st Qu.	:0.05675
Median	:0.1025	Median	:0.07900
Mean	:0.1136	Mean	:0.08369
3rd Qu.	:0.1460	3rd Qu.	:0.10400
Max.	:0.3720	Max.	:0.22500

Pitch results also showed some correlation with voicing. Overall pitch movement throughout the word, calculated across tone categories, is depicted in Figure 3a. Voiced nasals are on the left and voiceless nasals on the right in each plot comprising the figure. Voiceless nasals with a starting pitch of 200hz decrease to a pitch of 180hz at the point of transition to the following vowel. Voiced nasals typically have a roughly level pitch profile that starts lower than the voiceless nasal and ends on nearly the same level at the midpoint of thefollowing vowel. This pattern was consistent across all lexical tones, although, again, variations were observed, primarily in the pitch during the nasal. The difference between pitch was greatest in creaky and checked tones. For creaky tone, the pitch ofvoiceless nasals averaged between 246hz at the onset, while it was 173hz for voiced nasals. For checked tone, the difference is almost identical with voiceless nasals having an average pitch of 243hz, while voiced nasals have an average of 173hz. The gap decreased in high tones, with voiceless nasals averaging 221hz and voiced nasals 171hz. Differences in pitch are smallest in low tones, with an average pitch of 185hz for voiceless nasals and 157hz for voiced nasals. Onset pitch across lexical tones is illustrated in Figure 3b.

Figure 3a

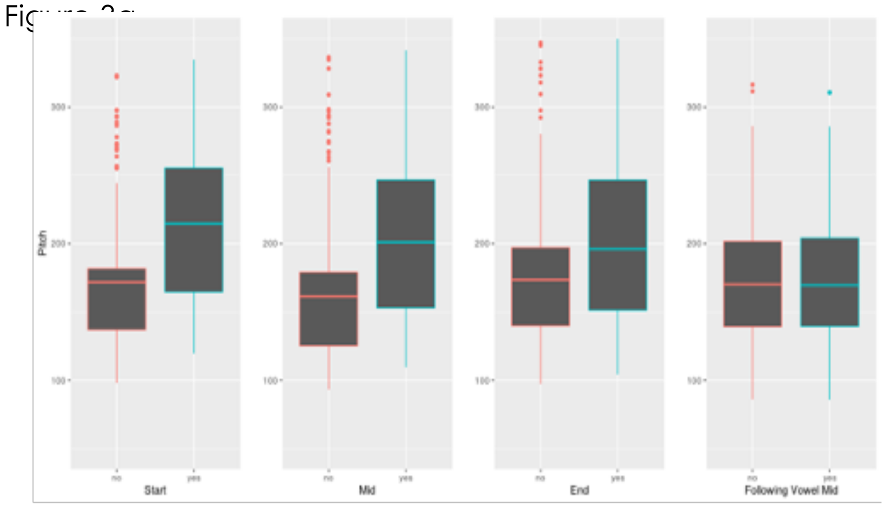
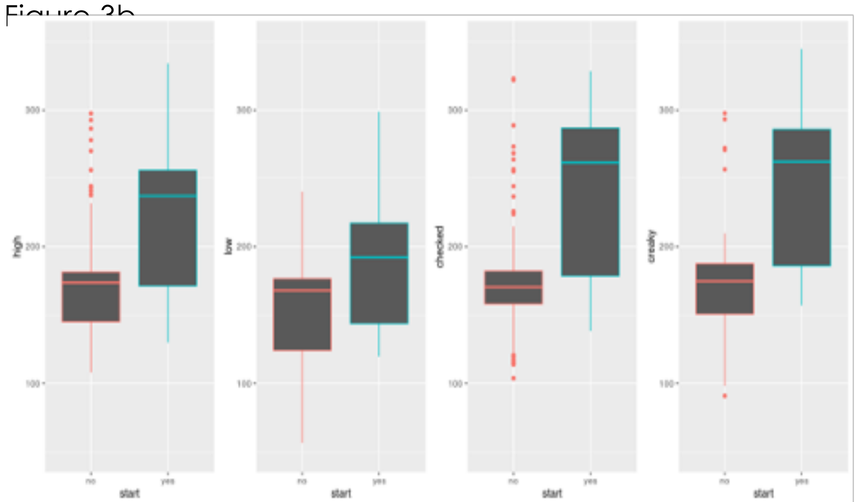


Figure 3b

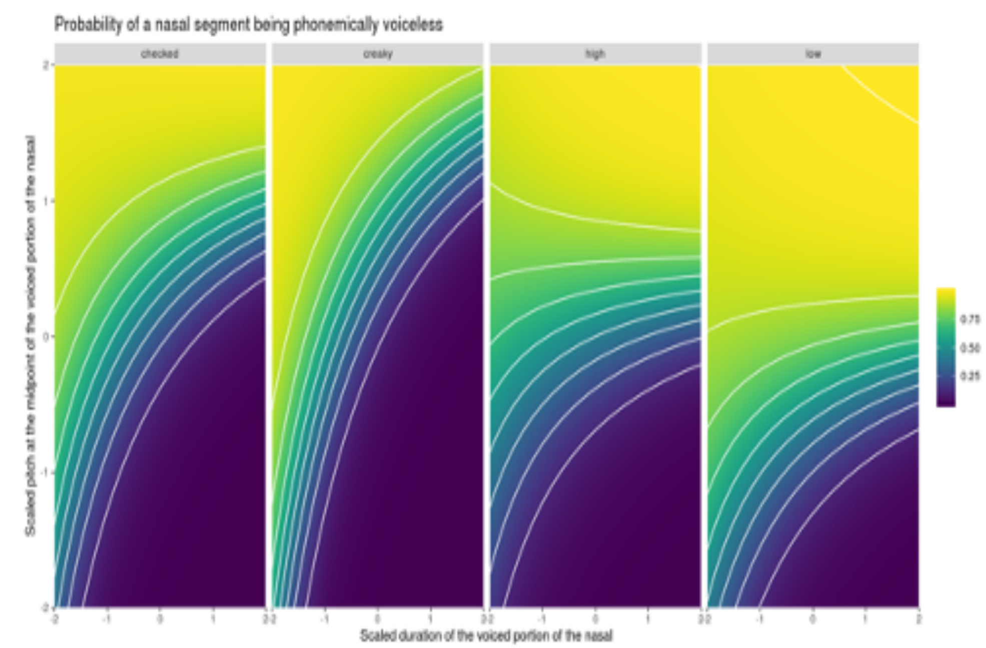


Other interactions between pitch or tone with specific lexical items and surrounding vowel quality exhibited no differences. Two variables, coda consonant and place of articulation, showed minor effects. Duration is shorter and pitch is higher for words ending in glottal stops, i.e. those with a checked tone. The duration of all words ending in a glottal stop is below 250 milliseconds, while most words without a coda have a duration greater than 250 milliseconds. Meanwhile, the pitch of words ending in a glottal stop averages between 175hz to 250hz, while words without a coda glottal stop only average between 150hz to 200hz.

Thus, glmer models were created including some or all of the following independent variables: pitch, duration, tone, place of articulation, and coda consonant. The dependent variable is the presence of a voiceless phase (which differentiates the voicing categories), and the random effect of speaker is included. Pitch measurements at the mid point of the consonant are the predictor since there is an effect on pitch at onset positions. ANOVA tests were used to compare these models. Models including coda consonants (AIC = 740.15) were less efficient than models without it (AIC = 735.80, p<.001). Furthermore, models including place (AIC = 750.2) were revealed to be less efficient than models excluding it (AIC = 736.8, p<.001). Thus, both variables were excluded.

Afterwards, interactions between all three variables were added into the predictions. Two more variables were added—one for the interaction between pitch and duration, and another for the interaction of tone category and duration. This was decided after more ANOVA tests with models containing differently ordered interactions between these three variables. None of them proved to be significantly better than the other (p = 1). However, interactions with duration as the common point resulted in the lowest AIC of 746.06. This is in comparison to an AIC of 765.29 when tone was the common point, and an AIC of 765.40 when it was pitch. Therefore, the final model has pitch, tone, duration of the voiced phase, the interaction between pitch and duration of the voiced phase, and the interaction between tone and duration of the voiced phase as the final independent variables.

Results from this mixed effects model demonstrate that pitch and duration measurements produce different results depending on lexical tone. Figure 4 depicts four plots demonstrating interactions between tone, pitch, and duration. Each plot illustrates the effect of pitch and duration on voicing for a specific tone. Duration and pitch are scaled to account for the diversity of values. The gradient represents the probability of a voiced or voiceless outcome, with yellow associated with predicted voicelessness and blue with predicted voicing. Some consistent observations can be made for all lexical tones. At any given pitch, increased duration of the voiced phase leads to a greater probability of a nasal being voiced while decreased duration of the voiced phase leads to a greater probability of a nasal being voiceless. Conversely, at any given duration, there is a positive correlation between high pitch and a voiceless prediction and between low pitch and a voiced prediction. Furthermore, once pitch values reach a certain threshold, any increase in duration does not affect the outcome of voicing; there is a very high likelihood that these nasals are voiced. This threshold is highest for the checked tone, and then for high and low tones in decreasing order. Creaky tone stands out in showing a greater reliance on duration of the voiced phase. Thus, given sufficient duration of the voiced phase, the probability that, regardless of pitch level, the nasal is voiced is stronger than for the other three tones. This may be due to the fact that creaky and checked tones share similar pitch contours, but are differentiated by the fact that creaky tone possesses a greater duration (Watkins 2003). Shorter length also appears to predict voicelessness at lower pitch levels in creaky tone in comparison to other tones. For high tone, extremely short duration of the voiced phase with high pitch actually increases the likelihood of a voiced prediction unlike the pattern observed for the other three tones.



Conclusion

Initial results regarding pitch and duration confirmed previous findings in the literature. Mainly, voiceless nasals have greater duration and higher pitch than voiced ones. The current study demonstrates that this effect on pitch is localized to the nasal itself and does not perseverate through the following vowel. Rather, voicing of the consonant appears to have no effect on the pitch of the following vowel, aside from the period of transition from the nasal to the vowel.

An important novel finding of this study involves the interaction between pitch, duration of the voicing phase, and tone. Increased pitch in general increases the likelihood of a nasal being voiceless whereas increased duration of the voiced phase increases the probability of a nasal being a member of the voiced category. However, the interaction between pitch and duration of the voiced phase differs between tones. Results from the present study shed light on how a typologically unusual type of linguistic contrast is realized phonetically along different acoustic dimensions.

Acknowledgments

I would like to thank my mentors and participants who helped make this study possible. I am grateful to my mentor, Matthew K. Gordon, who guided me from the beginning stages of this project. I am grateful for his time, patience, and support, for arranging meetings, and reviewing drafts. Special thanks to Phillip G. Rogers, for tutoring me in R and helping me flesh out the methodology to produce these results. I also thank Caroline Crouch for helping inspire this project by making me aware of this topic's research potential. I want to express my sincerest gratitude to all the participants of this project, for helping me and allowing me to work on their language: San San Myint, Cynthie Tin Oo, Aung Tun Myint, and Dino Myint.

Works Cited

Bhaskararao, Peri & Peter Ladefoged. 1991. Two types of voiceless nasals. Journal of the International Phonetic Association 21, 80-6.

Blankenship, Barbara, Peter Ladefoged, Peri Bhaskararao & Nichumeno Chase. 1993. Phonetic structures of Khonama Angami. Fieldwork studies of targeted languages (UCLA Working Papers in Phonetics) 84, 127–141.

Boersma, Paul & Weenink, David. 2020. Praat: doing phonetics by computer [Computer program]. Version 6.1.09, retrieved 26 January 2020 from http://www.praat.org/

Chirkova, Katia & Zev Handel. 2013. Diachronic developments of voiceless nasals in Ersu, Lizu,

and related languages. Presented at the 46th International Conference on Sino-Tibetan

Languages and Linguistics, Dartmouth College, New Hampshire, August 7–10.

Chirkova, Katia., Patricia Basset & Angélique Amelot. 2019. Voiceless nasal sounds in three

Tibeto-Burman languages. Journal of the International Phonetic Association, 49(1), 1-32.

Dantsuji, Masatake. 1984. A study on voiceless nasals in Burmese. Studia Phonologica XVIII, 1-14.

Dantsuji, Masatake. 1986. An acoustic study on the distinction of place of articulation for voiceless nasals in Burmese. Studia Phonologica 20, 1–11.

Hill, Nathan W. 2011. An Inventory of Tibetan Sound Laws. Journal of the Royal Asiatic Society, Third Series 21, 441-457.

Hombert, Jean-Marie., John J. Ohala & William G. Ewan. 1979. Phonetic Explanations for the Development of Tones. Language 55(1), 37-58.

Index: Burmese in Wiktionary. From https://en.wiktionary.org/wiki/Index:Burmese

Judson, Stevenson, & Robert C. Stevenson. 1953. Burmese-English Dictionary. Unabridged Centenary Ed., as Rev. and Enl. by Robert C. Stevenson and F. H. Eveleth. ed. Rangoon: Baptist Board of Publications.

Ladefoged, Peter & Ian Maddieson. 1996. The sounds of the world's languages. Oxford & Malden, MA: Blackwell.

Maddieson, Ian. 1984. The effects on Fo of a voicing distinction in sonorants and their implications for a theory of tonogenesis'. Journal of Phonetics 12, 9-15.

Matisoff, James. 2001. Prosodic diffusibility in South-East Asia. In Aleksandra Aikhenvald and Robert Dixon (eds.), Areal diffusion and genetic inheritance, 291-327. Oxford: Oxford University Press.

Nishida, Tatsuo. 1975, On the Development of Tones in Tibetan. Acta Asiatica 25, 43-55.

Ohala, John J. 1975. Phonetic explanations of nasal sound patterns. In C.A. Ferguson, L.M. Hyman & J.J. Ohala (eds.), Nasalfest: Papers from a symposium on nasals and nasalization (eds.), 289–316. Stanford, CA: Stanford University, Linguistics Department.

Ohala, John J. & Manjari Ohala. 1993. The phonetics of nasal phonology: Theorems and data. In

Marie K. Huffman & Rena A. Krakow (eds.), Nasals, nasalization, and the velum, 225–249. San Diego, CA: Academic Press.

Okell, John. 1969. A reference grammar of colloquial Burmese. Oxford: Oxford University Press.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Watkins, Justin. 2003. Tone and intonation in Burmese. Proceedings of the International Congress of Phonetic Sciences 15, 1289-1292.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Bruner, J. S. (1986). Actual minds, possible worlds. Cambridge, Mass: Harvard University Press. DePietro, P. (2013). Social Media And Collaborative Learning. Counterpoints, 435, 47-62. Retrieved from http://www.jstor.org/stable/42982123

Domine, V. (2007). Chapter Seven: "Doing Technology" in the Classroom: Media Literacy as Critical Pedagogy. Counterpoints, 235, 131-147. Retrieved from http://www.jstor.org/stable/42979004

Kiernan, A. (2017). Futurebook Critics and Cultural Curators in a Socially Networked Age. In Barekat H., Barry R., & Winters D. (Eds.), The Digital Critic: Literary Culture Online (pp. 114-123). New York; London: OR Books. doi:10.2307/j.ctt207g899.14

Mackey, M. (2014). Learning to Choose: The Hidden Art of the Enthusiastic Reader. Journal of Adolescent & Adult Literacy, 57(7), 521-526. Retrieved from http:// www.jstor.org/stable/24034360

Mejias, Ulises (2006) "Teaching Social Software with Social Software," Innovate: Journal of Online Education: Vol. 2: Iss. 5, Article 2. Available at: https://nsuworks.nova.edu/ innovate/vol2/iss5/2

Naik, Y., & Trott, B. (2012). Finding Good Reads on Goodreads: Readers Take RA into Their Own Hands. Reference & User Services Quarterly, 51(4), 319-323. Retrieved from http:// www.jstor.org/stable/refuseserq.51.4.319

Nakamura, L. (2013). "Words with Friends": Socially Networked Reading on "Goodreads" PMLA, 128(1), 238-243. Retrieved from http://www.jstor.org/stable/23489284

Nathan, L., MacGougan, A., & Shaffer, E. (2014). If Not Us, Who? Social Media Policy and the iSchool Classroom. Journal of Education for Library and Information Science, 55(2), 112-132. Retrieved from http://www.jstor.org/stable/43686975

Sweeny, S. (2010). Writing for the instant messaging and text messaging generation: using new literacies to support writing instruction. Journal of Adolescent & Adult Literacy, 54(2), 121-130.

White, J., & Hungerford-Kresser, H. (2014). Character Journaling Through Social Networks: Exemplifying Tenets of the New Literacy Studies. Journal of Adolescent & Adult Literacy, 57(8), 642-654. Retrieved from http://www.jstor.org/stable/24034751

About the Author

Ludwig Adisiswoyo is a 4thyear linguistics major and TESOL minor. He is graduating this spring with an honors thesis. Afterwards, he intends to teach English before pursuing a PhD in linguistics with a focus on language documentation.