Phonation and Aging in White Hmong

Christina M. Esposito[a] , Kevin Schäfer[b] & Sameer ud Dowla Khan[c]
[a]Macalester College, USA [b]UCSB, USA [c]Reed College, USA
esposito@macalester.edu, kevinjschaef@gmail.com, skhan@reed.edu

ABSTRACT

Age-related changes in physiology such as the ossification of the ribcage cartilages [19, 25] and the thickening of the vocal folds [15] may change speakers' voices, such that older voices have greater variability [21, 22, 23] and phonatory instability [13]. Unlike previous linguistic work on aging, the current study examines the effects of aging in a language with contrastive phonation: White Hmong. Forty-five native speakers ranging from 23–71 years of age were recorded producing words with contrastive breathy, modal, and creaky phonation. H1*−H2* and cepstral peak prominence (CPP) were measured based on previous studies [6, 7, 11]. Results showed that phonation measures were surprisingly stable across age groups, while variability (dispersion) substantially increased with age. This demonstrates that speakers of a language with contrastive phonation can maintain production distributed around stable targets, even as control of articulation is reduced.

Keywords: phonation; White Hmong; aging

1. INTRODUCTION

The aging process produces a number of changes to the muscular system, the larynx, the vocal folds, and the respiratory system, all of which can affect speech. These include physical changes such as the ossification and calcification of the ribcage cartilages [19, 25], and the thickening of the vocal folds [15], as well as neurological changes such as decreased synaptic activity [27].

Studies have examined the effects of aging on speech, finding age-based differences in 1) f0 [4, 5, 16], 2) intonational patterns [1], 3) spectral mean for /s/ [28] and 4) overall variability [3, 19, 21, 22, 23, 27]. Most relevant for the current study, voice quality differences have also been observed: older voices are often perceived as hoarse or breathy [13, 26]. This might be due to an increase in vocal fold mass, especially for post-menopausal women [15, 16] and/or incomplete glottal closure in both older women and men [3, 16]. (Younger women also possess incomplete glottal closure, but in a different location than older women [24]). These age-related changes to the vocal folds and to the degree of glottal closure can result in an increase in noise and/or breathiness. [8] reported that elderly speakers' voices are characterized by lower harmonic-to-noise ratio, an indicator of increased noise, when compared to middle-aged and younger voices. In addition, [12] found that older speakers have lower cepstral peak prominence (CPP) values; as CPP is a measure of noise and/or aperiodicity, this result suggests more aperiodic/non-modal phonation. Furthermore, elderly speakers display increased phonatory instability, changing glottal configuration more frequently than young speakers [13].

All the aforementioned studies have focused on languages that do not contrast phonation type, raising the question of how aging might affect phonation when phonemic contrast is at play. In such a language, can we expect changes similar to those reported for languages like English? Or will the production of phonation types as indicated in the acoustic signal remain consistent across age? To answer these questions, we examined three phonemic phonation types (breathy, modal, creaky) produced by speakers of White Hmong ranging from 23–71 years of age, resulting in the first study of its kind to our knowledge.

2. THE LANGUAGE

White Hmong (*Hmoob Dawb*) is a Hmong-Mien language spoken in Laos and Thailand and by large diasporic communities in the US, Australia, and France. The phonetic properties of phonation (and by extension tone, as the two features are linked) in Hmong have been the subject of a number of studies [6, 7, 9, 10, 11, 17, 20]. White Hmong is described as contrasting seven tones: rising (45), mid (33), low (22), mid-rising (24), high-falling (52), high-falling breathy (42), and low-falling creaky (21). However, recent research on connected speech [11] proposed revisions to this description, finding that the high-falling breathy tone is more accurately 54 or 53.

Several acoustic measures of spectral tilt and noise have successfully distinguished phonation types in White Hmong, including: the amplitude of the first harmonic (H1*) [6], and the amplitude of the first harmonic minus the amplitude of the second harmonic (H1*−H2*) [6, 7, 20], where the asterisk indicates that the harmonic amplitudes were corrected for the effects of adjacent formants using [18], an extension of [14]. [11] showed H1*−H2* did not adequately distinguish modal vs. non-modal phonation in connected speech and found CPP to be a better measure.

3. METHODS

3.1. Participants

Forty-five native speakers (20 female, 25 male) of White Hmong were recorded, with ages ranging from 23–71. They were born in Laos, Thailand, or the US, but all resided in the Minneapolis/St. Paul area of Minnesota, US, at the time of recording. The number of participants per age by decade is provided in Table 1.

Number of Participants	Age (by decade)
8	20s
12	30s
9	40s
3	60s
3	70s

Table 1: Number of participants per age by decade.

Twenty speakers over 60 were originally recorded, but they struggled with the task, and their data was unusable. All spoke English in addition to White Hmong; the reported age of English onset ranged from 5–57 years of age. All reported daily use of and literacy in White Hmong.

All speakers indicated their exercise level using guidelines from the US Centers for Disease Control and Prevention; all reported some degree of regular exercise. Speakers reported no history of smoking and no history of vocal fold disorders or strokes.

3.2. Recording

Speakers were recorded producing the list of 70 monosyllabic words used in [6]. These included all six vowels [i, e, ɨ, a, u, ɔ] and all seven tones of the language: 38 words had modal phonation, 17 had creaky phonation, and 15 had breathy phonation. Onset consonants varied, but were restricted to oral obstruents. Aspirated consonants and nasalized vowels were avoided. Words were read in the frame *rov hais* ____ *dua* [tɔ24 hai22 ____dua33] ‘Say ____ again’.

3.3. Measurements

The measures used in the current study were CPP and H1*–H2*, as defined in section 2; these are the measures that successfully distinguished phonation in previous studies of the language [6, 7, 10, 11, 20].

3.4. Analysis

The data were modeled in terms of phonation and dispersion to determine if there were correlations between age and phonation, or between age and dispersion of phonation measures, calculated as the standard deviation. (Age was treated as a continuous variable.) Models were fit using one of the measures of voice quality (H1*−H2* and CPP) as the dependent variable; CPP was log-transformed for regularity. Based on exploratory statistics, the data from three of the 45 participants were removed, given insufficient evidence that they significantly distinguished more than two phonation types at any time point using either measure. The three unusable participants were all female and aged 23, 30, and 43.

The resulting linear mixed-effects regression models were fitted for each of the dependent phonemic phonation (breathy, modal, or creaky), variables (H1*−H2* or CPP) with fixed effects of time point of the measurement (beginning, middle, end), and age group, with random intercepts for individuals and age of English exposure.

4. RESULTS

4.1. Mixed effects models of phonation measures

Linear mixed effects regression models were fit to H1*−H2* and log CPP. A maximal model was reduced by eliminating interaction effects when doing so did not result in a significantly different model. To prevent overfitting, the random effects were subject to a Principal Components Analysis following the method outlined by [2]. This process eliminated the random effects of gender and exercise level, which at any rate were unbalanced, with reported exercise covariant with gender and age. Models without these random effects outperformed those that included them. Furthermore, adding gender as a random effect was deleterious to the predictive power of the models, which suggests that no effect of menopause was observable in the data. Age of English onset had little effect on the data. Thus, gender, exercise level, menopausal status, and age of English onset will not be discussed here.

The model for H1*−H2* (Fig. 1) was moderately successful in characterizing variance (conditional r^2=0.369). There were significant main effects of time point (χ^2=111.9, df=2, p<0.01) and phonation (χ^2=912.0, df=2, p<0.01). The main effect of age was not significant (χ^2=0.0223, df=1, p<1.0). All two-way interactions were also significant: age and time point (χ^2=10.7, df=2, p<0.01); age and phonation (χ^2=6.29, df=2, p<0.05); and time point and phonation (χ^2=138, df=4, p<0.001). The effect sizes (regression coefficients) are, however, miniscule. The greatest change is found in creaky vowels, but mean H1*−H2* does not exceed an increase of 4dB over the span of five decades.

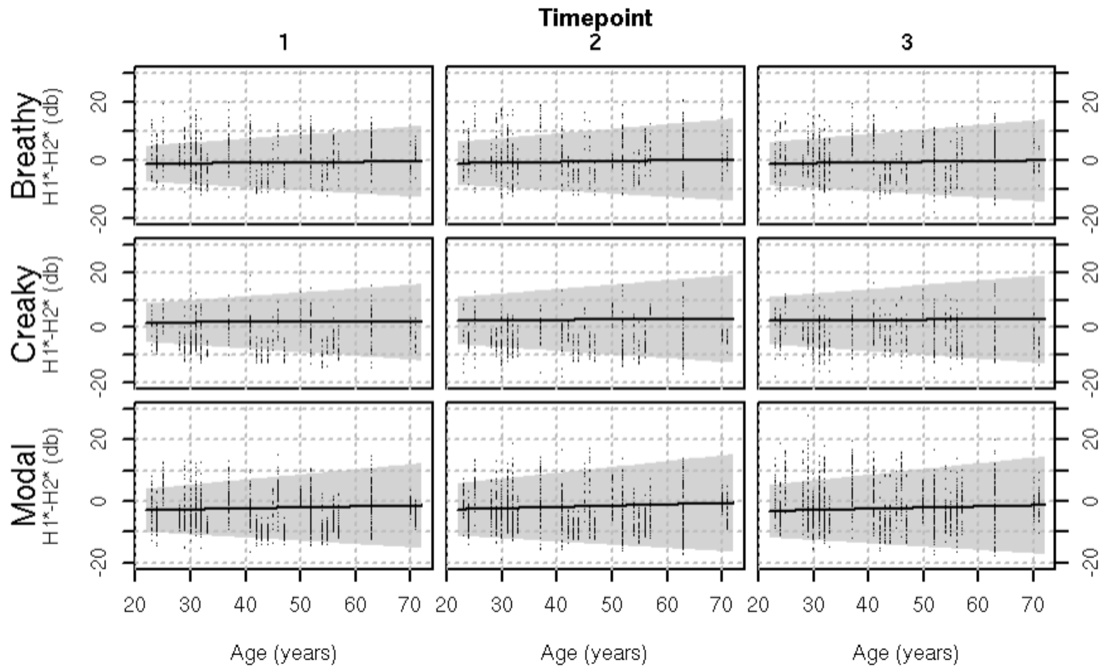

Figure 1: Token H1*−H2* (dB) values by participant age at each time point (1 = beginning, 2 = middle, 3= end of vowel) for each phonation type with a regression line and 95% confidence intervals (light gray).

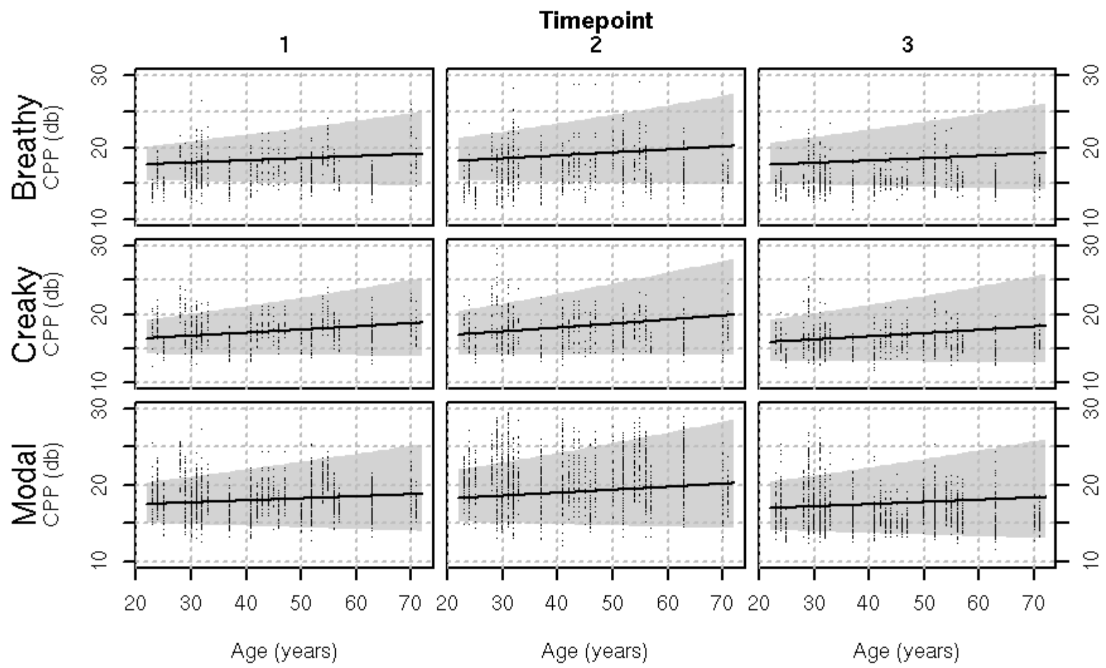

Figure 2: Token CPP (dB) values by participant age at each time point (1 = beginning, 2 = middle, 3= end) for each phonation type with a regression line and 95% confidence intervals (light gray).

The model for CPP (Fig. 2) characterized comparatively more of the variance (conditional r^2=0.475). As with the model of H1*−H2*, age was not a significant main effect (χ^2=0.0989, df=1, p<1.0), but there were significant main effects of time point (χ^2=2410, df=2, p<0.001) and phonation (χ^2=1308, df=2, p<0.001). All two-way interactions were also significant: age and time point (χ^2=36.21, df=2, p<0.001); age and phonation (χ^2=70.70, df=2, p<0.001); and time point and phonation (χ^2=551.8, df=4, p<0.001). The effects were similarly small, but minute increases in CPP (less than 3dB over five decades) were seen across phonation types and time points with age, suggesting slightly *less* noise in older speakers' voices. Generally, CPP was highest across phonation types at the midpoint of the vowel. The interaction with the greatest effect size involving age resulted only in a ~2.88dB increase in CPP at the midpoint of breathy vowels (i.e. slightly less breathy at that time point).

4.2. Mixed effects models of dispersion

Finally, dispersion was quantified as the standard deviation for *each participant's* (untransformed) CPP measures at each timepoint within each phonation category. Another pair of linear mixed effects models was fitted to dispersion measures as dependent variables, with the same fixed effects of phonation type, time point, and age, with random intercepts for individuals and age of English onset. The model with H1*−H2* as dependent variable did not have any significant main effect of age or interactions with age and is not reported here.

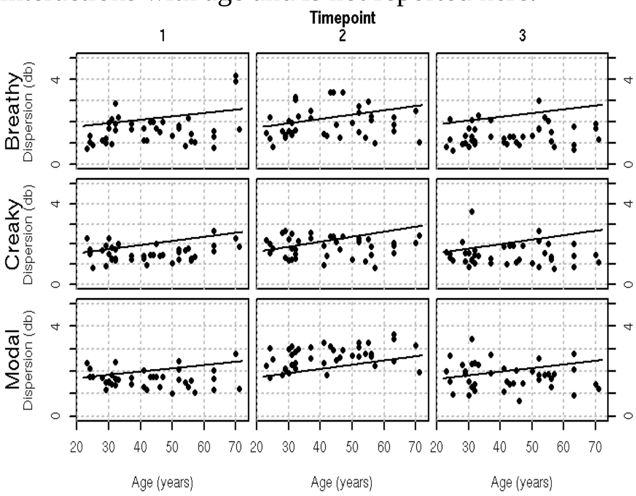

Figure 3: Dispersion of individual participants' CPP for each time point (1=beginning, 2=middle, 3= end); each phonation type is plotted with a regression line.

The model with CPP as the dependent variable (Fig. 3) was moderately successful at characterizing the variance (conditional r^2=0.402). All of the main effects were significant: age (χ^2=5.80, df=1, p<0.05), time point (χ^2=102, df=2, p<0.001), and phonation (χ^2=42.0, df=2, p<0.001). There was a significant interaction of time point and phonation (χ^2=23.9, df=4, p<0.001) and an interaction of age and time point (χ^2=5.77, df=2, p<0.1) approaches significance.

In this model, breathy vowels have the lowest dispersion for younger speakers at the midpoint, and other phonation types at the end of the vowel (and secondarily the midpoint). There is a trend for all phonation types throughout the vowel toward greater dispersion with greater age. This effect is most limited for modal vowels and strongest for creaky vowels. Dispersion at the midpoint is most affected, such that the model predicts that dispersion should be greater at the midpoints for participants in their 50s and older.

5. DISCUSSION AND CONCLUSION

When compared to research on languages without phonemic phonation contrasts, the results of the current study are mixed. Counter to expectation, this study finds no evidence that the voices of older White Hmong speakers are substantially more non-modal than those of younger speakers by either measure of phonation (H1*−H2* or CPP). In fact, the H1*−H2* models strongly suggest stable means across age groups. If anything, CPP measures suggest (negligible) *increase* with age, i.e. marginally *less* aperiodic noise in older voices.

However, in line with previous studies, older speakers' voices were found to be more variable than those of younger speakers. That is, the primary effect of age appears to involve increased variability (dispersion) rather than an overall change to or a failure to approximate a target level of phonation. This variability is strongest for non-modal voice qualities, in particular creaky tone.

It is possible that creak and breathiness in White Hmong are more affected by variability because they are already distinguished from modal phonation by tonal correlates. Indeed, previous research on tone identification in White Hmong has shown that listeners do not rely on creaky voice to perceive the low-falling creaky *-m* (21) tone, but instead, rely on f0 and duration [10]; this might explain why the variability effect in the current study is strongest for creaky vowels. As speakers of White Hmong age, they may allocate their limited capacity to coordinate and modulate articulation between different correlates of tone. Next steps in this research would be looking at variability in the tone, to determine if tonal patterns remain consistent as speakers age.

Replicating these methods across a diverse range of languages, particularly those that have different types of phonemic phonation contrasts, would shed light on the generalizability of these results. In future work, the inclusion of a non-tonal language with phonemic phonation contrasts, such as Gujarati, would provide an indication of the role of functional load, if any, in aging-related changes.

6. REFERENCES

[1] Barnes, D. R. 2013. Age-related changes in the production of linguistic prosody [Master's thesis, Purdue University].

[2] Bates, D., Kliegl, R., Vasishth, S., & Baayen H. 2015. Parsimonious Mixed Models. arXiv:1506.04967 [stat], arXiv: 1506.04967.

[3] Biever, D. M., & Bless, D. M. 1989. Vibratory characteristics of the vocal folds in young adult and geriatric women. *Journal of Voice* 3, 120–131.

[4] Brown, W. S., Jr., Morris, R. J., Hollien, H., & Howell, E. 1991. Speaking fundamental frequency characteristics as a function of age and professional singing. *Journal of Voice* 5, 310–315.

[5] Eichhorn, J. T., Kent, R. D., Austin, D., & Vorperian, H. K. 2018. Effects of aging on vocal fundamental frequency and vowel formants in men and women. *Journal of Voice 32*, 644.e1-644.e9.

[6] Esposito, C. M. 2012. An acoustic and electroglottographic study of White Hmong tone and phonation. *Journal of Phonetics* 40, 466–476.

[7] Esposito, C.M., & Khan, S.D. 2012. Contrastive breathiness across consonants and vowels: A comparative study of Gujarati and White Hmong. *Journal of the International Phonetic Association* 42, 123-143.

[8] Ferrand, C. T. 2002. Harmonics-to-noise ratio: An index of vocal aging. *Journal of Voice* 16, 480–487.

[9] Garellek, M. 2012. The timing and sequencing of coarticulated non-modal phonation in English and White Hmong. *Journal of Phonetics* 40, 152–161.

[10] Garellek, M., Keating, P., Esposito, C., & Kreiman, J. 2013. Voice quality and tone identification in White Hmong. *The Journal of the Acoustical Society of America* 133, 1078–1089.

[11] Garellek, M., & Esposito, C. 2021. Phonetics of White Hmong vowel and tonal contrasts. *Journal of the International Phonetic Association*, 1-20.

[12] Garrett, R.K. 2013. *Cepstral-and spectral-based acoustic measures of normal voices*. Doctoral dissertation, The University of Wisconsin-Milwaukee.

[13] Gorham-Rowan, M. M., & Laures-Gore, J. 2006. Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders* 39, 171-184.

[14] Hanson, H.M. 1995. Glottal characteristics of female speakers. Ph.D. dissertation, Harvard.

[15] Hirano, M., Kurita, S., & Nakashima, T. 1983. Growth, development, and aging of the human vocal folds. In D. B. J. Abbs (Ed.), Vocal fold physiology: Contemporary research and clinical issues (pp. 22–43). San Diego, CA: College-Hill Press.

[16] Honjo, I., & Isshiki, N. 1980. Laryngoscopic and voice characteristics of aged persons. *Archives of Otolaryngology* 106, 149–150.

[17] Huffman, M. 1987. Measures of phonation type in Hmong. *The Journal of the Acoustical Society of America* 81, 495–504.

[18] Iseli, M., Shue, Y.-L., Alwan, A. 2007. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America* 121, 2283-2295.

[19] Kahane, J. 1980. Age related histological changes in the human male and female laryngeal cartilages: Biological and functional implications. In V. Lawrence (Ed.), Transcripts of the ninth symposium care of the professional voice (pp. 11-20). New York: The Voice Foundation.

[20] Keating, P., Kuang J., Garellek, M., Esposito, C.M., & Khan, S.D. to appear. A cross-language acoustic space for vocalic phonation distinctions. *Language.*

[21] Linville, S. E., & Fisher, H. B. 1985. Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. *The Journal of the Acoustical Society of America* 78, 40–48.

[22] Linville, S. E. 1988. Intraspeaker variability in fundamental frequency stability: An age-related phenomenon? *The Journal of the Acoustical Society of America* 83, 741–745.

[23] Linville, S. E., Skarin, B. D., & Fornatto, E. 1989. The interrelationship of measures related to vocal function, speech rate, and laryngeal appearance in elderly women. *Journal of Speech and Hearing Research* 32, 323–330.

[24] Linville, S. E. 2002. Source characteristics of aged voice assessed from long-term average spectra. *Journal of Voice* 16, 472–479.

[25] Malinowski, A. 1967. The shape, dimensions, and process of calcification on the cartilaginous framework of the larynx in relation to age and sex in the Polish population. *Folia Morphologica* 26, 118-128.

[26] Ptacek, P. H., & Sander, E. K. 1966. Age recognition from voice. *Journal of Speech & Hearing Research 9*(2), 273–277.

[27] Ramig, L. A., & Ringel, R. L. 1983. Effects of physiological aging on selected acoustic characteristics of voice. *Journal of Speech and Hearing Research* 26, 22–30.

[28] Taylor S., Dromey C., Nissen S.L., Tanner K., Eggett D., Corbin-Lewis K. 2020. Age-Related Changes in Speech and Voice: Spectral and Cepstral Measures.

[29] *Journal of Speech Language Hearing Research* 63(3), 647-660.

[30] Shue, Y.-L., Keating, P., Vicenik, C. 2009. VoiceSauce: A program for voice analysis. *The Journal of the Acoustical Society of America* 124(4), 2221.