INTERSPEECH 2016
September 8–12, 2016, San Francisco, USA

Auditory-Visual Lexical Tone Perception in Thai Elderly Listeners with and without Hearing Impairment

Benjawan Kasisopa[1,2], *Chutamanee Onsuwan*[3,4], *Charturong Tantibundhit*[3,5], *Nittayapa Klangpornkun*[5], *Suparak Techacharoenrungrueang*[6], *Sudaporn Luksaneeyanawin*[6], *Denis Burnham*[1,2]

[1] MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Australia
[2] HEARing Cooperative Research Centre (HEARing CRC), Australia
[3] Center of Excellence in Intelligent Informatics, Speech and Language Technology, and Service Innovation (CILS), Thammasat University, Thailand
[4] Department of Linguistics, Faculty of Liberal Arts, Thammasat University, Thailand
[5] Department of Electrical and Computer Engineering, Thammasat University, Thailand
[6] Center for Research in Speech and Language Processing, Chulalongkorn University, Thailand

b.kasisopa@westernsydney.edu.au, consuwan@tu.ac.th, tchartur@engr.tu.ac.th, 5710300160@student.tu.ac.th, suparak.t@student.chula.ac.th, sudaporn.l@chula.ac.th, denis.burnham@westernsydney.edu.au

Abstract

Lexical tone perception was investigated in elderly Thais with Normal Hearing (NH), or Hearing Impairment (HI), the latter with and without Hearing Aids. Auditory-visual (AV), auditory-only (AO), and visual-only (VO) discrimination of Thai tones was investigated. Both groups performed poorly in VO. In AV and AO, the NH performed better than the HI group, and Hearing Aids facilitated tone discrimination. There was slightly more visual augmentation (AV>AO) for the HI group, but not the NH group. The Falling-Rising (FR) pair of tones was easiest to discriminate for both groups and there was a similar ranking of relative discriminability of all 10 tone contrasts for the HI group with and without hearing aids, but this differed from the ranking in the NH group. These results show that the Hearing Impaired elderly with and without hearing aids can, and do, use visual speech information to augment tone perception, but do so in a similar, *not a significantly more* enhanced manner than the Normal Hearing elderly. Thus hearing loss in the Thai elderly does not result in *greater* use of visual information for discrimination of lexical tone; rather, *all* Thai elderly use visual information to augment their auditory perception of tone.
Index Terms: tone perception, auditory-visual, hearing impaired, discrimination

1. Introduction

In lexical tone languages, changes in fundamental frequency (F0) height or contour result in a change of word meaning. While duration, F2 values (vowel length and height), voice quality, and amplitude (loudness) contribute to tone [1],[2], and [3], F0 is the principle determinant of tone perception. Thai is an example of a tone language. Bangkok Thai has five tones, three level or static tones (Low-21, Mid-33, and High-45) and two contour or dynamic tones (Falling-241 and Rising-315) (numbers are Chao values [4]). For example, [paa33] (ปา)-mid tone means 'to throw', [paa21] (ป่า)-low, 'forest', [paa45] (ป้า)-high, 'aunt', [paa241] (ป๊า)-falling, 'father (Chinese origin)', and [paa315] (ป๋า)-rising, 'father; sugar daddy'.

Over and above extensive investigations of the auditory perception of consonants and vowels [5] and [6], it has also been found that visual information from the lips and face augments auditory speech discrimination [7], even in undegraded listening conditions [8].

Lately, both auditory [5] and auditory-visual [6], [7], and [9] perception of lexical tones have received attention. Such studies have focused mainly on normal hearing (NH) adult populations and, while we know that speech perception of consonants and vowels deteriorates with hearing loss, little is known about the specific effects of hearing impairment on the auditory, let alone the auditory-visual perception of tone. The study reported here extends this research to auditory and auditory-visual perception of Thai lexical tone in native Thai elderly without and with hearing impairment, the latter with and without hearing aids.

1.1. Aging, hearing impaired, and speech perception

It is well established that the elderly, especially those who are hearing impaired (HI), have more difficulty hearing and understanding speech [10] and [11]. The manifestation of the hearing impairment depends the degree of hearing loss but also on the frequency of the sounds in a particular context [10], [12], and [13]. In this regard, increased age is usually accompanied by frequency-range degradation that begins at the high end of the usual 20 Hz to 20 kHz range of human hearing. This gradual high-frequency loss eventually encroaches upon the frequency range necessary for speech perception, around 250 to 4,000 Hz [14]. State-of-the-art hearing aids (HAs) generally have limited effect in addressing the different signal-to-noise ratios (SNR) in speech perception because they normally amplify all sounds, including noise, rather than enhancing only speech sounds. The result is that hearing-impaired people can hear speech, but they have difficulty understanding it [12] and [13].

http://dx.doi.org/10.21437/Interspeech.2016-908

Since lexical tone perception is based primarily on F0 height and contour, and since in the elderly there is a progressive high frequency loss, there may be particular difficulties for elderly people perceiving lexical tones. As would be expected hearing impaired (HI) listeners' tone perception performance is generally poorer than that of NH listeners, however the use of HAs has not been found to result in significant improvement in speech intelligibility for HI listeners' tone-based distinctions [15], [16], and [17]. For example, a study of tone identification by Thai HI listeners using hearing aids in aided and unaided conditions by a subset of the current authors [17] found that identification of Thai lexical tones by Thai HI listeners is generally low and that percentage of correct identification did not improve significantly from when the listeners were without (64.6%) to with (66.5%) HAs.

Turning to auditory-*visual* speech perception by HI listeners, it has been found that auditory and visual information complement each other, and that AV speech is generally more intelligible than AO speech [18] and [19]. Visual speech information can provide up to 15dB improvement in SNR [20], which is significant given that a 1dB increase in SNR corresponds to a 5-10% improvement in intelligibility [21]. Indeed, visual speech information is particularly advantageous for HI listeners who tend to rely more heavily on the visual component of AV speech than do NH listeners [16] and [19].

Given the relative scarcity of research on tone perception, and in particular AV tone perception, it is of interest to investigate if and to what extent native Thai elderly especially those with HI make use of visual speech information in their perception of Thai tones.

1.2. Auditory perception of tone

Cross-language studies have shown that native Thai listeners perceive Thai tones better than do non-native listeners of other tone or pitch-accent languages, who in turn perceive tones better than non-native non-tone language listeners [22] and [23]. There are also consistent differences in discrimination of particular pairs of tone types, with better discrimination of Dynamic-Dynamic (DD) pairs (Falling-Rising/FR) than Static-Static (SS) pairs (Mid-Low/ML, Mid-High/MH, and Low-High/LH) and in turn better than for Static-Dynamic (SD) pairs (Low-Rising/LR, Mid-Rising/MR, High-Rising/HR, Low-Falling/LF, Mid-Falling/MF, and High-Falling/HF) [23]. In addition, in accord with the proposal of a physiological bias towards better perception of rising pitch contours based on studies of the frequency following response (FFR) in the brainstem [24], it has been found that within SD pairs, those involving rising tones are discriminated better than those involving falling tones [23].

1.3. Auditory-visual perception of tones

There is a 40%–80% augmentation of AO speech perception when speech in a noisy environment is accompanied by the speaker's face [25], and there is now evidence for visual influences in lexical tone perception in Cantonese [26], Mandarin [27], and Thai [23]. There is auditory-visual augmentation of lexical tone perception, which surprisingly occurs irrespective of language background – for native tone-, non-native tone-, and non-native-non-tone language listeners alike [23], [27], [28], and [29], but somewhat paradoxically, there is better use of visual information for tone in visual-only (VO) (lip-reading) conditions by *non-native*, *non-tone* language listeners than by native tone or non-native tone language listeners [23] and [27]. This is thought to be the result of tone language listeners not attending to the visual tone information in lip-reading conditions as they have learned to rely on the usually more powerful auditory information for tone, whereas non-native, non-tone language listeners are not attuned to tones and so are more open to *any* relevant information, including visual information to solve a difficult perceptual problem [23] and [27].

In this study, auditory and auditory-visual tone discrimination is investigated in native Thai NH and HI elderly in clear listening conditions. The aim is to establish the base discriminability of Thai tones by native Thai HI elderly, so only clear (no background noise) speech was used because hearing aids generally tend to amplify *all* noise making speech less intelligible [12]. Based on the above literature, the hypotheses are:

a) ***Auditory Tone Perception*:** NH elderly will perform better than HI elderly, and HI elderly will perform better with than without their HAs.
b) ***Auditory-Visual Tone Perception*:** Given that HI listeners may rely more upon visual input [19], they should make better use of visual information than the NH as indexed by (i) VO tone discrimination, and (ii) visual augmentation of tone discrimination (the degree of advantage for auditory-visual over auditory-only conditions).
c) ***Tone Contrasts*:** NH and the HI listeners should perform similarly on discrimination of different types of tone pairs, i.e., DD better than SS, better than SD pairs [23].
d) ***Rising vs Falling Tones*:** There should be, in accord with FFR brainstem studies, better discrimination of tone pairs involving rising tones than those involving falling tones [26] and [27].

2. Method

2.1. Participants

Thai Hearing Impaired Elderly (HI groups): 37 native Thai HI listeners (mean age 66.35 years, *SD* = 11.5, 13 females) were recruited in Bangkok, Thailand. All had moderately severe to profound hearing impairment with pure-tone thresholds more than 68dB HL in both ears at octave frequencies from 250 Hz to 8 kHz (range: 44-115dB HL in left (*SD* = 17.36), and 48-198dB HL (*SD* = 16.06) in right ear). These 37 HI participants were tested both with (HI+Aid) and without (HI-Aid) hearing aids.

Thai Normal Hearing Elderly (NH group): 18 native Thai elderly listeners (mean age 63.19 years, *SD* = 3.12, 10 females) were recruited in Bangkok, Thailand. Most had NH with pure-tone thresholds lower than 25 dB HL; a few had mild hearing loss with thresholds lower than 40 dB HL, at octave frequencies from 250 Hz to 8 kHz in both ears.

2.2. Experimental design

Participants were tested on an AX same/different paradigm. The design was 2 [**groups**: NH/HI (with (HI+Aid/HI-Aid nested within the latter) x 3 [**modes**: AV/AO/VO] x 10 [**tone contrasts** – SS (ML/MH/LH); SD-falling (MF/HF/LF); SD-rising (MR/HR/FR); and DD (FR)] x 4 [same/ different pairing conditions: 2 different, AB or BA, trials, and 2 same, AA or BB, trials]. There was a total of 120 trials presented to each participant (240 for the HI, 120 in HI+Aid and 120 in HI-Aid).

2.3. Stimulus materials

As in Burnham *et al*. [23], stimuli consisted of six Thai CV syllables (C = [k/kʰ]; V = [a:/i:/u:]) each carrying all five of the Bangkok Thai tones. Of the 30 syllables 21 were words and nine non-words. Syllables were recorded audio-visually in citation form from a native Thai female speaker in a sound-treated booth with 25 fps, 720 x 576 pixels, and 48 kHz 16-bit audio. Three good quality exemplars of each syllable were selected. Sound level was normalised and all videos were compressed using the msmpeg4v2 codec. In the AO mode, a still image of the speaker was displayed.

2.4. Procedure

Participants were tested individually in a sound-attenuated room on individual laptop computers running DMDX software [30] with the face presented in the centre of the screen. The auditory stimuli were presented via Sennheiser HD 25-1 II headphones (for the NH) (and via Boss Companion 2 Series ii loud speakers for the HI groups) connected through an EDIROL/Cakewalk UA-25EX USB audio interface unit at a comfortable hearing level.

For each session, there were 120 trials (3 modes x 10 tone pairs x 4 same-different AB orders) presented in two blocks of 60 trials with a short break between. At the start of each block, four extra warm-up trials were presented: 1 AV, 1 AO, and 1 VO trial in the training session, then another AV trial.

Participants were asked to determine whether two tones played sequentially with an inter-stimulus interval of 500 msec. were the same or different (AX task) with a time-out limit for each trial of five seconds. If a participant failed to respond on a particular trial, one additional chance to respond was given in an immediate repetition of that trial.

2.5. Data processing and analysis

As in Burnham *et al.* [23] *d'* scores were calculated for each of the 10 tone pairs in each condition (*d'* = *Z*(hit rate) – *Z*(false positive rate) with adjustments made for probabilities of 0 (= .05) and 1 (= .95), where *hit* is a 'different' response on an AB or BA trial, and a *false positive* is a 'different' response on an AA or BB trial.

Separate Analyses of Variance (ANOVAs) were conducted to compare Hearing Groups and Conditions in AV, AO, and VO modes for the 10 tones contrasts, and for visual augmentation (AV *minus* AO). Alpha was set at 0.05, and effect sizes are given for significant differences.

3. Results

Figure 1 below shows the mean *d'* scores collapsed over tone contrasts for each group and condition for AV, AO and VO scores.

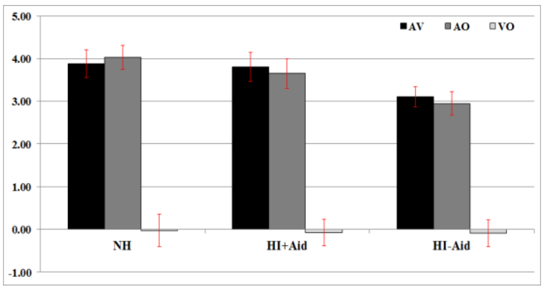

Figure 1: *Mean d' scores in AV, AO, and VO conditions.*

Between groups and conditions

***AV and AO together*:** ANOVA for AV and AO performance showed that overall the NH group performed better (Mean = 3.88 [AV] and 4.03 [AO]) than the HI participants in both their +Aid' (Mean = 3.80 [AV] and 3.65 [AO]) and –Aid (Mean = 3.11 [AV] and 2.94 [AO]) conditions, although the only significant differences were between the NH group and HI participants in the –Aid condition [$F_{(1,53)} = 5.96$, $p < .05$] and, contrary to previous *identification* results [17] here in a *discrimination* task HI participants were better with than without their Hearing Aids [$F_{(1,36)} = 10.20$, $p < .01$] (see Fig. 2A).

***AV and AO separately*:** When broken down into effects for AV and for AO, NH were significantly better than HI participants generally (across +Aid and –Aid conditions) [$F_{(1,53)} = 4.75$, $p < .05$] in the AO but not the AV mode. On the other hand within the HI group, HI participants were significantly better in the +Aid than –Aid condition for both AO [$F_{(1,36)} = 8.65$, $p < .01$] and AV [$F_{(1,36)} = 10.37$, $p < .01$]. Moreover, the HI+Aid > HI-Aid effect was greater for DD contrasts than SD contrasts for AO [$F_{(1,38)}$=6.30, p<.05], but not for AV.

***VO Performance*:** VO discrimination performance was analysed in separate ANOVAs, but no between-groups effects or between-HI condition effects were found (see Fig. 2B).

So in general NH were better than HI, and HI+Aid better than HI-Aid, but this was shown only in AV and AO modes not the VO mode.

***Effects of Tone Contrast*:** The most dominant effect of Tone Contrast across the two groups and the two HI conditions was that of DD tone pairs being easier to discriminate than SD pairs [$F_{(1,53)} = 10.95$, $p < .01$]. This difference was especially apparent in HI+Aid vs HI-Aid comparisons (see next section).

***Visual Augmentation (AV minus AO)*:** As can be seen in Fig. 1, there appears to be a significant augmentation of AO tone perception by the addition of visual information for the HI but not the NH participants, that is, as would be expected, visual information provides more augmentation for people with a hearing impairment. In comparison across groups this did not appear as a significant effect except that in the NH vs HI-Aid comparison there was significantly more visual augmentation for HI-Aid for DD than SD contrasts whereas for NH there was no visual augmentation for either [$F_{(1,53)} = 5.77$, $p < .05$] (see Fig. 2C).

Visual augmentation was most apparent for comparison of HI participants with and without their HAs. There were significant effects of AO/AV x HI+Aid/HI-Aid x SD/DD effect [$F_{(1,36)} = 9.39$, $p < .01$]. This shows that for HI people without their HAs, there is a large effect of visual information for the DD tone pair, and this visual information is not used when they have their hearing aid on (and neither is it used by elderly NH people) (see Fig. 2C).

***Rising versus Falling Tones*:** There was an overall effect of SD rising > SD falling for NH and HI+Aid combined, showing that, in line with the physiological bias found with FFR response measures [26] and [27], in optimal (NH) or near optimal (HI+Aid) *auditory* perception conditions rising tones are better perceived than falling tones. There was also an effect of AO/AV x HI+Aid/HI-Aid x SD-Rising vs SD-falling [$F_{(1,36)} = 6.03$, $p < .05$]. This shows (see Fig. 2C) that there is similar visual augmentation for falling tone contrasts with and without HAs, whereas for rising tone contrasts, there is greater visual augmentation when wearing than not wearing HAs, that is visual information is more useful when added to the inherent bias in the auditory system [27] for rising tones by adding one's hearing aid, than when that inherent bias is not present – in falling tone contrasts.

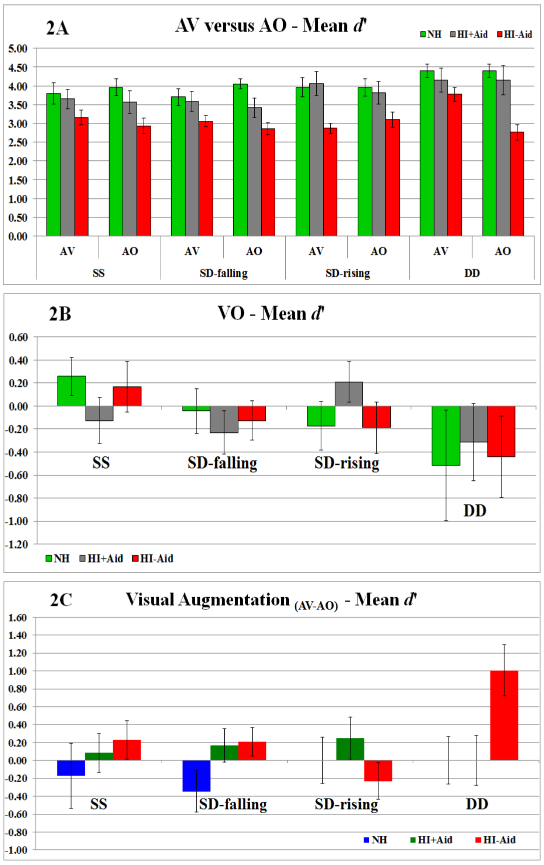

Figure 2: *Mean d' scores by tone contrasts.*

4. Discussion

This study clearly shows that Thai NH elderly perform better than the HI groups in their discrimination of native tones, and that hearing aids assist hearing impaired elderly to perform better in tone perception. These results contradict those found in Tantibundhit *et al*. [17], in which it was found that conventional HAs (with 250 Hz – 2 kHz frequency range; the same as those in the present study) failed to assist in Thai tone perception. It is possible that this is because Thai tones fall outside this range, but if so then this should also have applied in the present study. Thus it remains that there are different results for the two studies. The reasons for this difference could be due to 1) the nature and difficulty of the tasks (AX discrimination here versus tone identification task in [17], with the latter usually being considered more difficult); and/or 2) that the AV augmentation effect occurring mainly for tone contrasts involving DD versus SD pairs (with DD > SD) and for SD-falling and SD-rising pairs (with SD-falling > SD-rising). As the Tantibundhit *et al*. [17] study did not test tone *pairs*, no such contrast-specific effects could manifest in their results. Further research on the intricacies of the AV augmentation for particular tone contrasts is warranted.

The visual tone perception effects were only found in AV and AO but not in the VO condition. This is in line with previous studies that found better VO tone perception by non-tone language perceivers [23] and [27]. It is interesting that visual augmentation was stronger for tone pairs with rising tones than with falling tones, and that there was more visual augmentation for DD contrasts in the HI-Aid than the HI+Aid group, as showed in Fig. 2A. While this could be due to better use of visual information when not wearing a HA, this is more likely to be due to the HI-Aid discrimination scores being quite low in AO than in AV where the HI+Aid's scores were almost the same in the two conditions. Nevertheless, it is clear that HI participants without HAs made more use of visual information than those with hearing aids.

Finally, this study also found that the HI participants showed the expected bias for rising tones [24] only with their HAs, indicating the possibility that here both visual and auditory information were required for the rising tone bias to be evident.

5. Conclusion

In conclusion, the results of this study show that tone perception of the Thai elderly is similar to the hearing impaired elderly when they wear hearing aids. So, while HI causes an understandable deterioration of tone perception, HI Thai perceivers with Hearing Aids were able to use auditory and visual information effectively, in a manner similar to, but slightly worse than their NH counterparts. However, in general it seems that Thai HI perceivers ignore potentially beneficial visual information for tone perception except in particular cases, e.g., without their Hearing Aids for the DD tone contrasts. Thus it appears there may be the potential for Thai HI listeners to *learn* to use visual information for tone in other tone perception situations.

These results are promising, but further research involving perception in noise, perhaps also including acoustic enhancements (hyperarticulation) of tones are required in Thai and in other tone languages before definitive conclusions can be drawn. For now it can be concluded that hearing aids indeed improve lexical tone perception in a simple same-different task, and that AV augmentation of AO perception occurs in unaided HI listeners under certain conditions.

6. Acknowledgements

This research was supported by an Australian Research Council (ARC) Discovery grant (DP0988201) to the last author. The authors acknowledge the financial support of the HEARing CRC, established and supported under the Australian Government's Cooperative Research Centres Program. The authors would like to thank Dr Pittayapon Pitathawatchai and Dr Nida Rueangwit from Thammasat University Hospital for supporting the testing in Thailand.

7. Appendix

Table of abbreviations

HI	Hearing Impaired
NH	Normal Hearing
HA(s)	Hearing Aid(s)
HI+Aid	Hearing Impaired wearing Hearing Aid(s)
HI-Aid	Hearing Impaired without Hearing Aid(s)
AV	Auditory-visual
AO	Auditory-only
VO	Visual-only
SS	Static-Static tone contrasts
SD-falling	Static-Dynamic (falling) tone contrasts
SD-rising	Static-Dynamic (rising) tone contrasts
DD	Dynamic-Dynamic tone contrast

8. References

[1] Abramson, A. S. "Static and dynamic acoustic cues in distinctive tones," *Language and Speech,* vol. 21, no. 4, pp.319-325, 1978.

[2] Henderson, E. J. A. "Tonogenesis: Some recent speculations on the development of tone," *Transactions of the Philological Society,* vol. 112, pp.1-24, 1981.

[3] Tseng, C.-Y., Massaro, D. W. and Cohen, M. M. "Lexical tone perception in Mandarin Chinese: Evaluation and integration of acoustic features," *J. of Chinese Linguistics,* vol. 13, pp. 267-289, 1985.

[4] Chao, Y.R. "A system of tone-letters", *Le Maitre Phonetique*, vol. 45, pp. 24-27, 1930.

[5] Werker, J. F., and Logan, J. S. "Cross-language evidence for three factors in speech perception," *Perception & Psychophysics*, vol. 37, pp. 35-44, 1985.

[6] Campbell, R., Dodd, B., and Burnham, D. (Eds.). *Hearing by Eye II: Advances in the psychology of speech reading and auditory–visual speech*. East Sussex: Psychology Press, 1998.

[7] Vatiktiotis-Bateson, E., Kuratate, T., Munhall, K. G., and Yehia, H. C. "The production and perception of a realistic talking face," In O. Fujimura, B. D. D. Joseph, and B. Palek (Eds.), *LP'98, Item order in language and speech*. Prague: Charles University, Karolinum Press. pp. 439-460, 2000.

[8] McGurk, H., and McDonald, J. "Hearing lips and seeing voices," *Nature,* vol. 264, pp. 746–748, 1976.

[9] Werker, J. F., and Tees, R. C. "Phonemic and phonetic factors in adult cross-language speech perception," *J. Acoust. Soc. Am.*, vol. 75, pp. 1866-1878, 1984b.

[10] Committee on Hearing, Bioacoustics and Biomechanics (CHABA), "Speech understanding and aging," *J. Acoust. Soc. Am.* Vol. 83, pp. 820–895, 1988.

[11] Li, F, and Allen, J. B. "Identification of Perceptual Cues for Consonant Sounds and the Influence of Sensorineural Hearing Loss on Speech Perception." In E. Lopez-Poveda, A., Palmer, R. Meddis (Eds.) *The Neurophysiological Bases of Auditory Perception*. New York: Springer, pp. 449-462, 2010. doi: 10.1007/978-1-4419-5686-6_42.

[12] Duquensnoy, A. J., and Plomp, R. "The effect of a hearing aid on the speech-reception threshold of hearing-impaired listeners in quiet and in noise," *J. Acoust. Soc. Am.,* vol. 73, no. 21, pp. 2166-2173, 1983.

[13] Verschuure, J., and van Benthem, P. P. G. "Effect of Hearing Aids on Speech Perception in Noisy Situations," *Audiology,* vol. 31, pp. 205-221, 1992.

[14] Strickland, E.A., Viemeister, N.F., Van Tasell, OJ. & Preminger, J.E. "Is useful speech information carried by fibers with high characteristic frequencies?," *J. Acoust. Soc. Am.* vol. 95, pp. 497-501, 1994.

[15] Lee, K.Y., van Hasselt, C.A, and Tong, M.C. "Tone perception in Cantonese-speaking children with hearing aids," *Ann. Otol. Rhinol. Laryngol.*, vol. 117, no. 4, pp. 313-316, 2008.

[16] Gandour, J., Carney A., Nimitbunnasarn C., and Amatyakul, P., "Tonal confusions in Thai patients with sensorineural hearing loss," *J. Speech Lang. Hear. Res.*, vol. 27, no.1, pp. 89–97, 1984.

[17] Tantibundhit, C., Onsuwan, C., Klangpornkun, N., Phienphanich, P., Saimai, T., Saimai, N., Pitathawatchai, P., and Wutiwiwatchai, C. "Lexical tone perception in Thai normal-hearing adults and those using hearing aids: a case study," In *INTERSPEECH 2013 – 14th Annual Conference of the International Speech Communication Association, August 25-29, Lyon, France, Proceedings*, 2013, pp. 2262-2266.

[18] Heifer. K. S. "Auditory-visual perception of clear and conversational speech," *J. Speech Lang. Hear. Res.,* vol. 40. pp. 432-443, 1997.

[19] Most, T., Rothem, H., and Luntz, M. "Auditory, *Visual,* and Auditory-Visual Speech Perception by Individuals with Cochlear Implants versus Individuals with Hearing Aids," *Amer. Ann. the Deaf*, vol. 154, no. 3, pp. 284-292, 2009.

[20] Sumby, W. H., and Pollack, I. "Visual contribution to speech intelligibility in noise," *J. Acoust. Soc. Am.,* vol. 26, pp. 212–215, 1954. doi:10.1121/1.1907309

[21] Grant, K. W., and Braida, L. D. Evaluating the articulation index for auditory-visual input. *J. Acoust. Soc. Am.,* vol. 89, pp. 2952–2960, 1991. doi:10.1121/1.400733

[22] Wayland, R. P., and Guion, S. G. "Training English and Chinese listeners to perceive Thai tones: A preliminary report," *Language Learning*, vol. 54, no. 4, pp. 681-712, 2004.

[23] Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., Xu Rattanasone, N., Schwarz, I.-C., and Webster, D. "Universality and language-specific experience in the perception of lexical tone and pitch," *Appl. Psycholinguistics*, vol. 36, pp. 1459-1591, 2015. doi:10.1017/S0142716414000496, 2014.

[24] Krishnan, A., Gandour, J. T., and Bidelman, G. M. "The effects of tone language experience on pitch processing in the brain stem," *J. Neurolinguistics*, vol. 23, pp. 81–95, 2010.

[25] Sumby, W. H., and Pollack, I. "Visual contribution to speech intelligibility in noise," *J. Acoust. Soc. Am*. Vol. 26, pp. 212–215, 1954.

[26] Burnham, D., Ciocca, V., and Stokes, S. "Auditory–visual perception of lexical tone," In P. Dalsgaard, B. Lindberg,H. Benner, &Z.-H Tan (Eds.), *Proc. the 7th Conf. on Speech Commun. and Technology, EUROSPEECH 2001 Scandinavia,* pp. 395–398, 2001. Retrieved from http://www. isca-speech.org/archive/eurospeech_2001

[27] Smith, D, and Burnham, D. "Facilitation of Mandarin tone perception by visual speech in clear and degraded audio: Implications for cochlear implants," *J. Acoust. Soc. Am.*, vol. 131, no. 2, pp. 1480-1489, 2012.

[28] Mixdorff, H., Hu, Y., and Burnham, D. "Visual cues in Mandarin tone perception," In I. Trancoso, L. Oliviera, and N. Mamede (Eds.), *Proc. the 9th European Conf. on Speech Commun.and Technology*. Bonn, Germany: International Speech. Communication Association, pp. 405-408, 2005.

[29] Mixdorff, H., Charnvivit, P., and Burnham, D. K. "Auditory–visual perception of syllabic tones in Thai," In E. Vatikiotis-Bateson, D. Burnham, and S. Fels (Eds.), *Proc. Auditory–Visual Speech Process. Int. Conf.* Adelaide, Canada: Causal Productions, pp. 3-8, 2005.

[30] Forster, K. I., and Forster, J. C. "DMDX: A windows display program with millisecond accuracy," *Behavior Research Methods, Instruments, and Computers,* vol. 35, pp. 116-124, 2003