Auditory-visual Tone Perception in Hearing Impaired Thai Listeners

Benjawan Kasisopa[1], *Nittayapa Klangpornkun*[2], *Denis Burnham*[1]

[1] MARCS Institute, University of Western Sydney, Australia
[2] Department of Electrical and Computer Engineering, Faculty of Engineering, Thammasat University, Thailand

b.kasisopa@uws.edu.au, 5710300160@student.tu.ac.th, d.burnham@uws.edu.au

Abstract

This study investigated the effects of hearing impairment and auditory vs. auditory-visual perception of lexical tone by native Thai hearing impaired listeners: Hearing Impaired with Hearing Aids (HI+HA), Hearing Impaired without Hearing Aids (HI-HA), and Normal Hearing (NH). Adults' discrimination of the 5 Thai tones was investigated in auditory-visual (AV), auditory-only (AO), and visual-only (VO) conditions. Generally, NH participants performed better than the two HI groups with hearing aids facilitating tone perception (HI+HA > HI-HA). The Falling-Rising (FR) pair of tones was the easiest to discriminate for all three groups and there was a similar ranking of the relative discriminability of all 10 tone contrasts across groups. There was better tone discrimination in AV than in AO and both were much better than VO; and this was equally the case for all groups. The results show that Hearing Impaired individuals either with or without hearing aids can and do use visual speech information to augment auditory perception of tone, but do so in a similar, *not a significantly more* enhanced manner as the Normal Hearing individuals.

Index Terms: tone perception, auditory-visual, hearing impaired, discrimination

10.21437/Interspeech.2015-613

1. Introduction

Lexical tones are linguistically contrastive in 70% of the world's languages [1]. While duration, F2 values, voice quality, and amplitude (perceived as vowel length, height, and quality; and loudness respectively) all play some part in lexical tone [2]-[4], the principal feature of lexical tone is F0 (perceived as pitch). In tone languages, particular changes in F0 height or contour result in a change of word meaning. For example, Bangkok Thai has 5 tones (see Figure 1), 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 [5]). For example, [maa33] (มา)-mid tone means 'to come', [maa21] (หม่า)-low means 'brew', [maa45] (ม้า)-high means 'horse', [maa241] (ม่า)-falling means 'grandmother', and [maa315] (หมา)-rising means 'dog'.

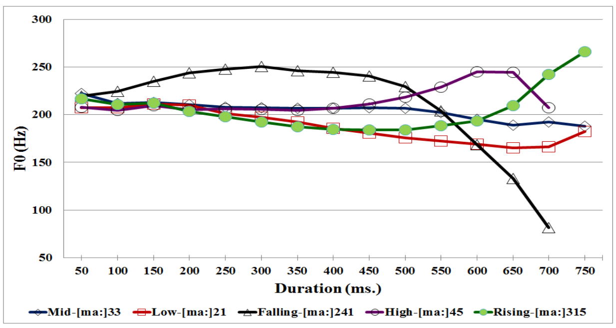

Figure 1: *F0 contours of 5 Bangkok Thai tones*

There has been extensive investigation of the auditory perception of consonants and vowels [6][7], and a number of factors have been found to influence speech discrimination. Of interest here, it is now very clear that visual information is used in speech perception when it is available [8], even in undegraded listening conditions [9].

While auditory [6] and auditory-visual speech perception [7][8][10] research has predominantly focused on consonants and vowels, of late there has been a welcome increase in perception studies involving lexical tone. However, these have mostly focused on NH adult populations; while we know that speech perception declines 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 hearing impaired native Thai listeners with and without hearing aids.

1.1. Hearing impaired and speech perception

It is well established that people with hearing loss have difficulty of hearing speech especially in noisy environments. The severity and nature of speech perception impairment depends on the frequency as well as the degree of hearing loss [11][12][13]. State-of-the-art hearing aids generally have limited effect in addressing the different signal-to-noise ratio (SNR) requirement issue in speech perception because they normally amplify everything, including noise, rather than enhancing only those speech sounds. The result is that they can hear the speech, but have difficulty understanding it [12][13].

One specific aspect of speech that may well differ from others is lexical tone which is based mainly on F0 height and contour. In this regard, HI listeners' tone perception performance is generally poorer than that of NH listeners and the use of hearing aids has not been found to result in any significant improvement to the speech intelligibility of tone perception in HI listeners [14][15][16]. A study on tone identification of Thai HI listeners using hearing aids in aided and unaided conditions [16] found that speech intelligibility for HI listeners is generally low even with hearing aids (percent correct was only ~67%) and hearing aids also did not improve the intelligibilities of Thai HI listeners in their study.

Turning to auditory-visual information in speech perception by HI listeners, a number of studies have found that auditory and visual information complement each other, and that AV speech is generally more intelligible than AO speech [17][18]. It has been shown that visual speech information can provide up to a 15dB improvement in SNR [19], which is significant given that a 1dB increase in SNR corresponds to 5-10% intelligibility improvement [20]. Visual speech information is especially beneficial when the speech signal is degraded or impoverished and so especially enhances intelligibility in noisy environments [21][22][23]. Visual speech should be particularly advantageous for HI listeners who tend to rely more heavily on the visual component of AV speech than do NH listeners [15][18].

Given the relative paucity of research on tone perception, and in particular AV tone perception, it is of interest to investigate if and to what extent native Thai HI listeners 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 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 [24][25]. 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, 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, High-Falling/HF) [25]. In addition, in accord with the proposal of a physiological bias towards better perception of rising pitch contours based on frequency following response (FFR) studies in the brainstem [26], it has been found that within SD pairs, those involving rising tones are discriminated better than those involving falling tones [25].

1.3. Auditory-visual Perception of Tones

A 40%–80% augmentation of AO speech perception has been found when speech in a noisy environment is accompanied by the speaker's face [27], and there is now evidence for visual influences in lexical tone perception in Cantonese [28], Mandarin [29] and Thai [25]. While there is auditory-visual augmentation of lexical tones presented in noisy environments across native tone-, non-native tone- and non-native non-tone language listeners [25][29][30][31], there is, somewhat paradoxically, better use of visual information for tone in VO conditions by non-native, non-tone language listeners than by native tone or non-native tone language listeners [25][29]. This is thought to occur because there has been found to be information for tone in the face, but that as tone language listeners are attuned to tones 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 have not learned this and so are more open to any relevant information, including visual information to solve a difficult perceptual problem [25][29].

In this study auditory and auditory-visual tone perception of tone is investigated in native Thai HI listeners, and compared with a reference group of Thai NH listeners. Note that the test was conducted using only the original clean speech without background noise. This is because hearing aids generally tend to amplify *all* noise so making speech less intelligible [12]. In addition, previous studies have also found that HI listeners' performance is generally poorer in noisy conditions [12][13]. Thus, background noise was not included her as the aim was to establish the base discriminability of Thai tones by native Thai HI listeners. Based on the above literature, the hypotheses are:

a) NH will perform better than HI listeners in auditory tone perception, and HI individuals with hearing aids will perform better than HI without hearing aids.
b) Given that HI listeners may rely more upon visual input [18], they should be able to make better use of visual information than the NH (i) in VO presentations, and (ii) for visual augmentation (AUG) of tone perception, as indexed by the degree of advantage for auditory-visual of auditory-only conditions.
c) 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 [25].

2. Method

2.1. Participants

Thai Hearing Impaired Adults (HI groups): 25 native Thai hearing impaired listeners (mean age 63 years, *SD*=15.7, 9 females) were recruited in Bangkok, Thailand. These constituted two groups - hearing impaired with hearing aids (HI+HA) and hearing impaired without hearing aids (HI-HA). All had moderately severe to profound hearing impairment with pure-tone thresholds more than 68dB HL in both ears (range: 44-108dB HL in left (*SD*=15.25), and 51-108dB HL (*SD*=14.03) in right ear).

Thai Normal Hearing Adults (NH group): 36 native Thai adult listeners (mean age 29 years, *SD*=4.0, 21 females) were recruited in Sydney, Australia. Mean duration in Australia prior to testing was 2 years (*SD*=2.8). All had normal hearing with pure-tone thresholds lower than 25dB HL in both ears.

2.2. Experimental design

Participants were tested in an AX same/different paradigm. A 3 [Groups: NH/HI+HA/HI-HA] x 3 [Modes: AV/AO/VO] x 10 [Tone contrasts – SS (ML/MH/LH); SD (MF/HF/LF/MR/HR/FR); and DD (FR)] x 4 [AB control order with same/ different pairing conditions: 2 different, AB or BA, trials, and 2 same, AA or BB, trials] x 2 [repetitions for NH] or x 1 [repetition for HIs]) design was employed.

2.3. Stimulus materials

Stimuli consisted of 6 Thai CV syllables (C = [k/k^h]; V = [a:/i:/u:]) each carrying all 5 of the Bangkok Thai tones. These are either words (*n* = 21) or non-words (*n* = 9). The 30 syllables were recorded in citation form from a native Thai female speaker. Productions were recorded audio-visually 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. Note that 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 [32] with the visual component presented in the centre of the screen. The auditory stimuli were presented via Sennheiser HD 25-1 II headphones connected through an EDIROL /Cakewalk UA-25EX USB audio interface unit at a comfortable hearing level.

For the NH participants, there were 240 trials split into 2 blocks (120 trials in each). In each block, 40 trials in each mode made up of 10 tone pairs in 4 AB orders were presented randomly and across blocks different stimulus exemplar repetitions were used. Block order was counterbalanced between participants. At the start of each block, 4 extra trials were presented: 1 AV, 1 AO, and 1 VO trial in the training session, then another AV trial placed at the start of the test session as a warm-up trial. For the HI participants only half the number of trials were presented (120 trials only, as there was only 1 not 2 repetitions).

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 at 5 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. NH participants were given a break between the two blocks while the HI participants were given a break half-way through their one block of trials.

2.5. Data processing and analysis

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 'same' response on an AA or BB trial.
Separate Analyses of Variance (ANOVAs) were conducted for a) Hearing Groups x AV/AO/VO modes; b) Hearing Groups x tones contrasts for three modes, and AUG separately. Alpha was set at 0.05, and effect sizes are given for significant differences.

3. Results

The mean *d'* scores over individual tone contrasts for each group are shown for AV, AO and VO scores in Figure 2.

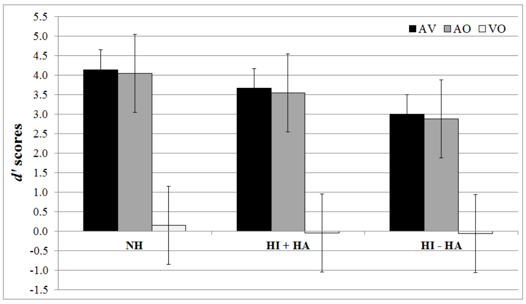

Figure 2: *Mean d' scores in AV, AO, and VO conditions*

3.1. AV, AO and VO tone perception

NH group performed significantly better than the two HI groups [$F_{(1,83)}$=11.45, p<.01]; and the HI+HA group were better than the HI-HA [$F_{(1,83)}$=4.07, p<.05]. Also significant was better tone perception for AV and AO than for VO [$F_{(1,83)}$=723.13, p<.001] and between AV and VO [$F_{(1,83)}$= 678.52, p<.001], but *not* AV and AO (p>.30). There were also significant interactions between Groups x AV*AO/VO [$F_{(1,83)}$=6.52, p<.05] and x AV/VO [$F_{(1,83)}$=4.42, p<.05]. All groups generally had poor discrimination of tone contrasts in VO when no auditory component was presented, although there appeared to be slightly better VO performance for the NH group. Regarding AUG, while there was a slight superiority of AV over AO, i.e., some visual-augmentation, this was not significant for any of the groups.

3.2. Relative discriminability of each tone pair

Mean *d'* scores for the 4 types of tone contrasts: SS, SD-falling, SD-rising and DD were examined in more details, in the AO; the VO; and the AV conditions, and in the Augmentation separately. The F-values for significant factors and their interactions (i.e., F >3.956) of the tone contrasts for each condition are shown in Table 1.

Table 1: *F-values for different tone contrasts*

	AO	VO	AV	AUG
Group (NH vs. HI groups)	**9.88**		***12.05***	
HI+HA vs. HI-HA			4.79	
SS*SD/DD	***17.19***		***12.19***	
SS vs. SD-rising	***14.62***			
Group x SD	6.01			4.60
HI groups x SS vs. SD	**9.15**			
HI groups x SD-rising			5.67	

Note: empty cell means not significant; *Italic figure* = p<.05; **Bold figure** = p<.01; and ***Bold Italic figure*** = p<.001.

3.2.1. AO condition (see Figure 3A)

NH group performed significantly better than the two HI groups. The DD pair (FR) was significantly more discriminable than all other pairs and the SD-rising were more discriminable than the SS pairs. There was also a significant interaction between the two HI groups and SS/SD tone contrasts showing that performance of the HI+HA group was better than HI-HA and that SD > SS. The results also showed that while the NH group found SD-falling ≈ SD-rising, the HI groups, especially HI+HA found SD-rising > SD-falling. These results show that the predicted SD-rising > SD-falling effect [26] was evident in Thai tone perception across hearing groups. Overall these results can be described as: **DD > SD (with SD-rising > SD-falling) ≈ SS.**

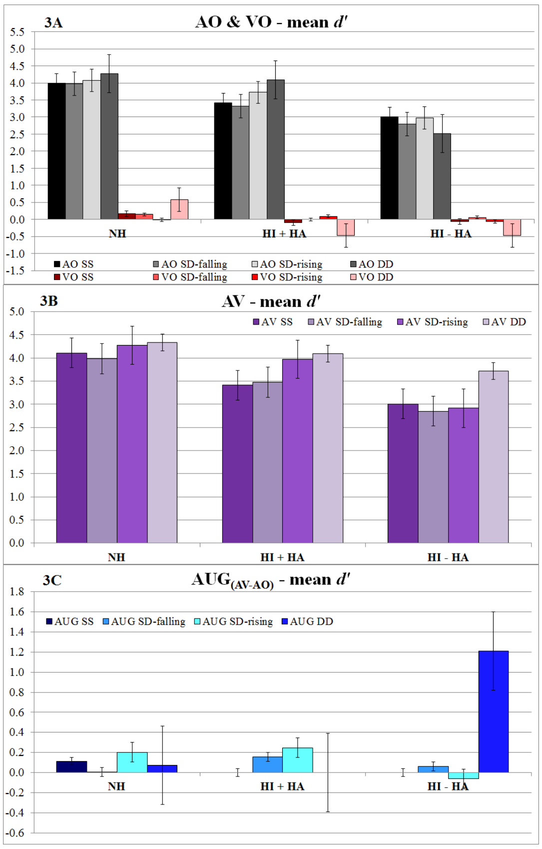

Figure 3: *Mean d' scores by tone contrasts*

3.2.2. VO condition (see Figure 3A)

There were no significant effects found in the VO condition even though all groups found DD slightly more discriminable than the other contrasts (see Figure 3A). Overall these results can be described as: **DD slightly > SD (with SD-rising ≈ SD-falling) ≈ SS**.

3.2.3. *AV condition (see Figure 3B)*

Similar to the results for AO, NH group's performance was significantly better than the two HI groups; the HI+HA was also better than the HI-HA group. All groups found the DD more discriminable than SS and SD pairs, especially the HI-HA group. All groups also found that SD contrasts were more discriminable than SS contrasts. There was also a significant interaction of the two HI groups with SD contrasts such that the HI+HA group found SD pairs, especially SD-rising more discriminable than did the HI-HA. Overall results can be described as: **DD > SD (with SD-rising > SD-falling) ≈ SS**.

3.2.4. *Augmentation (AV-AO) (see Figure 3C)*

There was no significant main effect found for AUG. However there was AUG under certain conditions. There was a significant interaction between Groups and SD-falling/SD-rising tone pairs. The HI-HA group found the DD pair more discriminable than did the other groups, but the HI+HA group did not show any visual benefit for AV vs. AO for the DD pair at all. Thus, there was an overall pattern of **DD > SD (with SD-rising > SD-falling) ≈ SS.**

4. Discussion

The results will be discussed in turn in terms of Hearing Impairment, Visual Augmentation, and Tone Contrast Effects.

4.1. Hearing Impairment

Generally this experiment shows that NH group performed better than HI groups in discriminating native lexical tone contrasts. However, the HI+HA performed significantly better than HI-HA group when auditory information was available, which shows that in contrast to the findings in [15][16] hearing aids do indeed assist with the perception of Thai lexical tones. The reason for differences between this and the previous studies could be that in this experiment a simple same-different task was used whereas previous studies used the more difficult tone identification task.

Overall discrimination in AV and AO conditions was much better than in VO, indicating that native Thai listeners tend to rely more on auditory more than visual information in speech perception. Even though all groups performed worse in VO condition, the effect was more pronounced (but not significantly so) in the HI groups, especially the HI-HA group. This is counterintuitive and unexpected; it would be expected that HI groups would rely more on lip-reading than would NH group. Further research on this is required.

4.2. Visual Augmentation (AUG)

AUG of AO tone perception by the addition of visual information (AV > AO) was noticeable for all groups, but not significant as a main effect. This may be because AUG is usually shown in conditions where auditory white or babble noise is added to the stimuli [25], but for reasons given above this was not feasible or advisable here. It may be the case that there would be AV-AO AUG in noisy condition, but this is yet to be tested. Additionally there was no significant group effect for AUG, i.e., HI groups did not show any better AUG than did the NH group. Again this could be due to the testing in quiet condition.

None of the groups used visual information *alone* (VO) very much in perception of their native tones, even though HI perceivers would be expected to do so more. This is consistent with the previous findings that native Thai listeners and other tone language listeners do not perceive tone in VO conditions as well as non-tone language listeners. [25][29]. However, the reason why the HI groups did not use visual information in tone perception, given their lack of auditory experience is still unclear. One reason could be that these participants did not have training that would allow them to use visual speech information efficiently, such as lip reading. The average duration of hearing aids use by these HI participants was 5 years (range: 0.5 – 20 years, *SD*= 5.64). Moreover, interviews with the participants after the test revealed that most said that they found the VO condition too difficult, so they guessed most of the time. This supports the notion that these participants did not have any training either implicit or explicit in lip reading.

Nevertheless, there *was* AUG under some conditions. HI-HA group did show visual AUG for the easiest tone pair, the DD pair, whereas the HI+HA group did not show any benefit. Thus AUG occurred for the easiest tone pair for the participants with uncorrected hearing impairment under some conditions. Again, further research is required to investigate this further.

4.3. Tone contrast effects

The above results showed that there were subtle effects due to AUG across groups for a particular tone contrast. There were also a number of other effects specific to the particular tone combinations. Overall the DD pair was the most discriminable compared to the other pairs, while the SS and SD pairs were more or less equally discriminable, with a slight edge for SD. Among the SD pairs, SD-rising pairs were better discriminated than SD-falling pairs in all conditions except in VO where both were equally badly discriminated. Thus when auditory information was available the results support the previous finding that rising tones are physiologically easier to discriminate than falling tones [25].

5. Conclusion

In conclusion, the results of this study show that while HI involves an understandable deterioration of tone perception, HI Thai listeners are 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 listeners ignored potentially beneficial visual information for tone perception except in particular cases - HI-HA group showed better AV than AO perception for the FR contrast than did the HI+HA group. The facial locus of the tone information is still unclear and awaits further research. These results are promising, but further research involving noisy condition, perhaps also including acoustic enhancements (hyperarticulation) of tones is 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 would like to thank Assoc. Prof. Charturong Tantibundhit and Dr Chutamanee Onsuwan from Thammasat University for supporting the testing in Thailand.

7. References

[1] Fromkin, V. *Tone: A linguistic survey*. New York: Academic Press. 1978.

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

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

[4] 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.

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

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

[7] 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.

[8] 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.

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

[10] 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.

[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] 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.Dekle, F. J., Fowler, C. A., and Funnell, M. G. "Audiovisual integration in perception of real words." *Perception* & *Psychophysics,* vol. 51, pp. 355-362, 1992.

[15] 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.Grant, K. W., and Seitz, P. F. "The use of visible speech cues for improving auditory detection of spoken sentences." *J. Acoust. Soc. Am.,* vol. 108, no. 3, pp. 1197-1208, 2000.

[16] 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.

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

[18] 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.

[19] 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

[20] 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

[21] Bernstein, L. E, Demorest, M. E., and Tucker, P. E. "Speech perception without hearing." *Perception and Psychophysics,* vol.62, no. 2, pp. 233-252, 2000.

[22] Lachs, L., Pisoni, D. B., and Kirk, K. I. "Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants; A first report." *Ear and Hear.,* vol. 33, pp. 236-251, 2001.

[23] Massaro, D. W., and Cohen, M. M. "Speech perception in perceivers with hearing loss: Synergy of multiple modalities." *J. Speech Lang. Hear. Res.,* vol. 42, pp. 21-41, 1999.

[24] 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.

[25] 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*. doi:10.1017/S0142716414000496, 2014.

[26] 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.

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

[28] 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

[29] 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.

[30] 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.

[31] 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.

[32] 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.