Envelope-based inter-aural time difference localization training to improve speech-in-noise perception in the elderly

Maryam Delphi1, Yones Lotfi2*, Abdollah Moossavi3, Enayatollah Bakhshi4, Maryam Banimostafa5

Received: 17 Sep 2016 Published 3 Jul 2017

Abstract

Background: Many elderly individuals complain of difficulty in understanding speech in noise despite having normal hearing thresholds. According to previous studies, auditory training leads to improvement in speech-in-noise perception, but these studies did not consider the etiology, so their results cannot be generalized. The present study aimed at investigating the effectiveness of envelope-based interaural time difference (ITD ENV) localization training on improving ITD threshold and speech-in-noise perception.

Methods: Thirty-two elderly males aged 55 to 65 years with clinically diagnosed normal hearing at 250-2000 Hertz, who suffered from speech-in-noise perception difficulty participated in this study. These individuals were randomly divided into training and control groups: 16 elderlies in the experimental group received envelope-based interaural time difference localization training in 9 sessions, but 16 matched elderlies in the control group did not receive any training. The ITD ENV threshold and spatial word recognition score (WRS) in noise were analyzed before and after the localization training.

Results: Findings demonstrated that following the training program, the interaural time difference envelope threshold and spatial word recognition score (WRS) in noise were improved significantly in the experimental group (p≤ 0.001). Moreover, a significant difference was detected in interaural time difference envelope threshold and spatial word recognition score (WRS) in noise (p≤ 0.001) before and after the training in the experimental group.

Conclusion: The results of the present study revealed the effectiveness of envelope-based interaural time difference localization training in localization ability and speech in noise perception in the elders with normal hearing up to 2000 Hz who suffered from speech-in-noise perception difficulty.

Keywords: Interaural Time Difference, Training, Speech-in-Noise Perception, Localization

Introduction

Many elderly individuals, despite having normal hearing thresholds, complain of difficulty understanding speech in noise. Speech-in-noise perception difficulty is one of the disabling consequences of hearing impairment with severe impact on communication (1). One of the reasons for speech perception difficulty, especially in environments with multiple sources of sound, is localization difficulties or spatial processing disorders (SPDs) (2, 3). Localization of sound is the first cue for segregating the target auditory data from the non-target ones, and it is one of the important auditory functions for perception following target speech in noisy situations (4, 5). Researchers have emphasized that in normal individuals, sound localization improves the signal to noise ratio for about 2-3 dBs.
and therefore, contribute to better comprehension of conversations in noisy environments (4, 6, 7).

Sound localization is frequency specific, ie, sounds with lower frequencies than 1,500 Hz are localized by interaural time difference (ITD), while interaural level difference (ILD) is used for higher frequency sounds; and spectral cues contribute to sound localization (8). Speech is a modulated signal that contains 2 different types of ITDs: ITD fine structure and ITD envelope (9). The latter means that at higher frequencies, the envelope, a slow modulation (low frequency) of the carrier frequency, can transmit the ITD information, which is known as ITD in the envelope (ITD ENV) (10). Interaction of ITD ENV and ITD FS is still a controversial issue (9, 10). Previous studies on localization by ITD ENV have used stimuli that were limited to frequencies above the phase-locking range, leading to poor localization. Thus, such studies concluded that ITD ENV was not an important cue in localization (11). Recent studies have shown that ITD ENV information begins to overwhelm in binaural performance near and below the frequency at which ITD for fine structure becomes an ambiguous cue to spatial hearing (800 Hz for humans) (9, 11). Moreover, ITD ENV plays a more essential role in spatial hearing and speech-in-noise perception (10, 11). ITD ENV processing includes the interaction of excitation and inhibition in lateral superior olive (LSO) (12). Wright et al. proposed that this balancing in LSO is more plastic than coincidence detection in medial superior olive (MSO) (12).

Localization, like other auditory skills, is affected by aging. Aging process can cause an increment in temporal jitter, resulting in binaural masking level difference (BMLDs) (13) reduction, elevated ITD and ILD threshold, and difficulty understanding speech in noise (2, 14-16).

Localization training is caused by brain plasticity and leads to improving speech perception in noise (17, 18). Moreover, some studies have confirmed the role of ITD ENV in sound localization (10, 11), but no study has focused on ITD ENV-based localization training. We focused on localization training based on ITD ENV in the elderly because LSO that tuned to ITD ENV is more plastic than MSO, and also the ITD discrimination ability, localization, and temporal envelope information were reduced in the elderly compared to normal individuals (1, 2). It is hypothesized that ITD ENV-based localization training may change the elderly’s ability to use spatial clues for segregating target speech from competing signals/noise and improve their speech in noise perception. Therefore, the present study aimed at investigating the effect of ITD ENV-based localization training in speech in noise perception in the elderly.

**Methods**

**Participants**

The study population was selected consecutively among the elderly males presented to Ahwaz audiology clinics over a 9-month-period from June 2015 to February 2015. In this study, 46 elderly males with speech in noise perception compliant were visited; 32 elderly males aged 55 to 65 years were selected based on inclusion criteria. All participants complained of difficulty for perception of speech in noisy conditions, and using the Persian version of the temporal jitter test, we found that their word discrimination score was lower than 56% (19). According to the result of temporal jitter test in 2011, the mean±SD of word discrimination score in 0dB signal to noise ratios was 65.93±9.60 in young adults (19). The participants’ auditory thresholds were below 20 dB at 250–2000Hz frequency range and less than 40 dB at 4–8KHz. Pure tone average difference was less than 5 dBHL between the 2 ears. All participants had normal IQs and were right-handed. Their otoscopic and tympanometric results were also normal, and none of them had a history of neurologic disease or injury, or did not participate in any other training programs during the ITD ENV-based localization training. Fourteen elderly males were excluded because they had hearing loss or did not meet other inclusion criteria. Moreover, 32 elderly who met the inclusion criteria were randomly divided into 2 groups according to the random number table produced by SPSS 21 software. Sixteen elderly were placed in the experimental group (They received training for sound localization based on ITD ENV) (mean±SD age of 60.50±2.52 years), and 16 elderly in the control group that received no training (mean±SD age of 60.18±1.51 years). All the above inclusion criteria were the same for the 2 groups.

The present study had two parts. First, it evaluated the ability to localize ITD ENV and spatial word recognition score (WRS) in noise in the elderly before the training; and second, it investigated these two after the training.

**Part 1: Evaluation**

**ITD ENV Threshold**

Measurement of the ITD ENV threshold was based on the identification of the auditory stimulus direction from the midline to the left and to the right with a change in the ITD (10, 20). The Persian nonword “bamâsh” was selected as the oral stimulus from the list of disyllabic Persian nonword (8). The advantage of nonwords are their similarity to natural signal, but they are meaningless and reduce the effect of meaning in localization (21). Duration of the stimulus was 1.5 seconds and had a rise-decay time of 25 milliseconds. The overall level of presentation for both ears was 70 decibels (20).

A constant-stimulus two-interval forced-choice (2 IFC) paradigm was utilized to measure ITD ENV threshold (22-24). In this method, the participants are requested to indicate whether the second interval was displaced from the midline by pressing the button. The percentage of correct responses was calculated for each ITD ENV value and displayed as psychometric function. 2IFC directly estimates the psychometric function. A 75% correct point criterion was used as the threshold (23-27).

The first interval contained a reference stimulus with 0µs ITD ENV, while the second interval contained the target stimulus with different ITD ENV values (10, 24). The percentage of correct left/right discriminations was calculated as a function of ITD ENV (10, 24). The range of ITD ENV was 10 µs to 100 µs in 10 µs steps within the range of 100 µs to 350 µs, and the steps were 50 µs. Two
blocks were presented, consisting of 30 repetitions for each ITD ENV value. Each block consisted of 15 repetitions of each ITD ENV value that were organized in 3 subblocks contenting 5 randomize sequence (28).

Spatial Word Recognition Score (WRS) in Noise

Spatial WRS in noise included delivering monosyllabic Persian words (29) to ears through the headphones in 7 different spatial locations at 0, -30, +30, -60, +60, -90, and +90 azimuth degrees by changing interaural time difference in the presence of white noise with a signal to noise ratio of 0 dB (30). Five monosyllabic words were presented in each spatial location, and the participant’s task was to repeat them (30). The percentage of correctly localized spatial WRS in noise was calculated.

**Part 2: Training**

The experimental group was trained for formal localization skills in a course of 9 sessions using disyllabic non-words with different ITD ENV values (12). The participants were asked to respond by showing the direction of the sound in the right or left side (24). Each ITD was frequently presented to the individuals so that they could identify it properly. The ITD ENV training program began for each participant at 20 milliseconds above the ITD ENV threshold that was obtained in the evaluation part so that localization would be attainable for the participants (31). Each training step was delivered to the individual 10 times. If the presented ITD could correctly localize the source, it would then be reduced in 10 millisecond steps. In the case of no improvement in new ITD ENV, the previous ITD ENV was practiced again. To investigate the effectiveness of the training program, ITD ENV threshold and spatial WRS in noise was obtained post training.

This study was approved by the ethics committee of the University of Social Welfare and Rehabilitation Sciences (code IR.USWR.REC.1394.3).

**Statistical Analysis**

Statistical analysis was conducted using SPSS software (Version 21). The Shapiro-Wilk test was used to verify the normal distribution of the data, which showed the abnormal distribution of data in spatial WRS and the mean of ITD ENV thresholds (p ≤ 0.001) in both groups. The Wilcoxon test was conducted to compare spatial WRS and ITD ENV thresholds pre- and post-training. The 2 groups were compared using Mann-Whitney test to show the training effects. Significance level was set at 0.05 (5%), with confidence interval of 95%.

The participants’ consents were obtained before the tests, and they could quit the program at any time. The control group had the option to participate in the training program after the completion of this research.

**Results**

Tables 1 and 2 demonstrate the mean scores and the standard deviation for ITD ENV thresholds and spatial WRS in the 2 groups before and after the auditory localization training.

Wilcoxon test was conducted to compare the results within each group. A significant difference was found in ITD ENV threshold before and after training in the training group (p ≤ 0.001); no difference was observed in the control group. Also, a significant improvement was observed in Spatial WRS in noise in the experimental group (p ≤ 0.001), but no significant changes were observed in ITD ENV thresholds (p= 0.180), and spatial WRS (p=0.102 at 0° azimuth, 0.564 at +30° azimuth, 0.234 at -90° azimuth and 0.317 for other spatial locations) after the training in the control group.

Mann-Whitney test was used to investigate the effectiveness of the training on ITD ENV thresholds and spatial WRS in noise. The results revealed a significant difference in ITD ENV threshold between the 2 groups. In other words, the ITD ENV threshold performance in the experimental group was improved significantly (p≤ 0.001). Moreover, a significant difference was found between the 2 groups in spatial WRS in noise (p≤ 0.001).

Figure 1 displays the psychometric function for the

**Table 1.** Descriptive Statistics of ITD ENV Threshold (µs) in the Two Groups Before and After Training

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before Timing</td>
<td>After Training</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>130.00</td>
<td>51.12</td>
<td>92.50</td>
<td>44.49</td>
</tr>
<tr>
<td>Control</td>
<td>136.25</td>
<td>53.89</td>
<td>133.75</td>
<td>53.27</td>
</tr>
</tbody>
</table>

**Table 2.** Percentage of Spatial WRS (word Recognition Score) in Noise at 7 Spatial Locations in the Two Groups Before and After Training

<table>
<thead>
<tr>
<th>Spatial Location</th>
<th>Group</th>
<th>Before Training</th>
<th>After Training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>At 0° Training</td>
<td>42.50</td>
<td>14.37</td>
<td>68.75</td>
</tr>
<tr>
<td>Control</td>
<td>37.50</td>
<td>12.38</td>
<td>32.50</td>
</tr>
<tr>
<td>At +30° Training</td>
<td>21.25</td>
<td>15.43</td>
<td>51.25</td>
</tr>
<tr>
<td>Control</td>
<td>20.00</td>
<td>10.32</td>
<td>21.25</td>
</tr>
<tr>
<td>At -30° Training</td>
<td>21.25</td>
<td>15.43</td>
<td>51.25</td>
</tr>
<tr>
<td>Control</td>
<td>20.00</td>
<td>10.32</td>
<td>18.75</td>
</tr>
<tr>
<td>At +60° Training</td>
<td>20.00</td>
<td>16.32</td>
<td>46.25</td>
</tr>
<tr>
<td>Control</td>
<td>21.25</td>
<td>13.60</td>
<td>22.50</td>
</tr>
<tr>
<td>At -60° Training</td>
<td>20.00</td>
<td>16.32</td>
<td>46.25</td>
</tr>
<tr>
<td>Control</td>
<td>21.25</td>
<td>13.60</td>
<td>22.50</td>
</tr>
<tr>
<td>At +90° Training</td>
<td>21.25</td>
<td>12.57</td>
<td>46.25</td>
</tr>
<tr>
<td>Control</td>
<td>27.50</td>
<td>14.37</td>
<td>26.25</td>
</tr>
<tr>
<td>At -90° Training</td>
<td>21.25</td>
<td>13.20</td>
<td>46.25</td>
</tr>
<tr>
<td>Control</td>
<td>26.25</td>
<td>12.30</td>
<td>28.25</td>
</tr>
</tbody>
</table>
Utilizing ITE ENV localization training to improve speech-in-noise perception

mean of ITD ENVs in the experimental (1a) and control (1b) groups before and after training. The psychometric function represents the percentage of correct responses as a function of ITD ENV localization. Figure 1a demonstrates the improvement of ITD ENV threshold in the experimental group, but figure 1b shows that the ITD ENV threshold did not change in the control group. In the psychometric function, the threshold was estimated to be at the 75% probability level for correct points (25-27). The percentage of the correct responses increased with the increment of ITD ENV, and all individuals had lower ITD ENV threshold after training.

Discussion

Elderlies with difficulty in speech perception in noise were recruited and trained with ITD ENV-based localization; and finally, the possible changes in ITD ENV threshold and spatial WRS in noise were tracked.

Prior to the training program, the mean of ITD ENV thresholds were 130 and 136 microseconds for the experimental and control groups, respectively. These values were higher than those of the Laback’s study (22) who found the mean of ITD ENV threshold of 56 microseconds for young participants with normal hearing. This threshold rise can be attributed to the binaural hearing disorder in the elderly. This is in agreement with findings by Grose, who believed that age-related factors contribute to a weaker binaural processing and a higher ITD threshold among the elderly compared to those of young individuals (14). In a similar study, Dobreva et al. indicated a deterioration of auditory spatial cues in sound localization with age advancement (16). Age-related changes, independent of hearing threshold, leads to a reduction in temporal coding and spatial separation cues such as ITD (32). It seems that these age-related changes reduced the effectiveness of temporal envelope information on spatial segregation and sound localization (33, 34). This reduction may be due to prolonged neural refractory times, loss of myelin integrity, decreased brain connectivity, and deficits in spectro-temporal processing (35).

The spatial WRS in noise in different locations was found to be from 20% to 42.50% before training. This finding was supported by Jaffary et al. who found that the mean score for speech-in-noise perception in elderly and young individuals was 45% and 65%, respectively (19). Many previous studies have confirmed speech in noise perception difficulty in elderlies (13, 19, 36, 37). It appears that reduction in localization and spatial processing skills is one of the leading causes of difficulty in speech in noise perception in the elderly (38, 39).

After ITD-ENV based localization training, the ITD ENV threshold and spatial WRS in noise was significantly improved in the experimental group. The control group did not show any significant changes. Because we only used lateralization based on ITD-ENV training, these improvements could be attributed to the training method. The observed influence of ITD ENV training in improving localization and spatial WRS in noise is consistent with previous behavioral studies, which commonly showed the effectiveness of localization training (12, 17, 18, 40, 41). Kawashima et al. found that ITD ENV has a positive influence on localization and confirmed its contribution to the plasticity of auditory localization (40). Other previous studies also investigated the effectiveness of localization rehabilitation programs. In a study by Wright et al. the ITD differentiation threshold of 32 listeners was assessed before and after the completion of ITD training. They confirmed that ITD training leads to an improvement in ITD discrimination (12). Our results are in accordance with theirs as they indicated that ITD ENV training leads to improving ITD ENV threshold. Study of Kuk et al. indicated that the laboratory-based and home-based localization training programs had positive effects on localization ability (18). Also, Tyler et al. found that spatial auditory rehabilitation can improve localization and speech in noise perception (17). The study by Cameron et al. showed an improvement of 10 decibels in speech perception threshold in listening in spatialized noise (LISN) after the training program (41). Findings of the current study were in line with all the mentioned studies. According to Kumpik’s study, spatial hearing plasticity is considerable during the developmental period. This plasticity remains in adulthood (42). Our result also hints to probable presence of localization plasticity in an old age. It appears relearning is caused reweighting in localization circuits and new spatial maps (42).
In the present study, no difference was observed between ITD ENV thresholds before and after training at 10, 20, 300, and 350 microseconds because none of the participants could differentiate 10 and 20 microsecond ITD ENV before and after the training. However, all participants could identify ITD ENV at 300 and 350 microseconds prior to and following the training. The percentage of correct identification of ITD ENV increased with the increase in ITD ENV (Fig. 1).

Our findings showed the effectiveness of ITD ENV-based localization on spatial WRS in noise after 9 sessions of training. Moreover, because the ITD ENV-based localization training program resulted in improving speech-in-noise perception, its effectiveness might have been limited only to the training stimulus and thus could not be generalized to other stimuli.

The constant-stimulus two-interval forced-choice (2 IFC) paradigm was used to estimate the ITD ENV threshold. An advantage of this method is an estimation of the participants guessing behavior by inserting blank trials in pretest schedule. The disadvantage of constant-stimulus 2 IFC is the significant amount of time required to obtain sufficient data points to display the psychometric function (43).

Finally, this study was conducted with a small sample size and in a short period of 9 sessions. Thus, further investigation with larger sample size and longer duration will help generalize the findings of this study. We did not use an age-matched valid memory test in our inclusion criteria, which made it a scientific limitation for our study. Furthermore, to evaluate the long-term effects of the training, follow-up studies in intervals of several months after training should be planned.

Conclusion

Our findings revealed that elderly with normal hearing sensitivity up to 2,000 Hz and a speech-in-noise difficulty could benefit from the ITD ENV training program. This training improved localization ability and speech in noise perception. According to the current study, it seems that plasticity of localization and spatial hearing is still maintained in old age. This study may indicate a potential for future training programs for the elderly to help them overcome speech in noise difficulties.

Acknowledgments

The authors of this paper are thankful to Dr. Mehrkian, S and Bakhit, M for their useful comments. In addition, they gratefully acknowledge the cooperation of all participants in this study. This study was supported by University of Social Welfare and Rehabilitation Sciences.

Financial Support

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflict of Interests

The authors declare that they have no competing interests.

References

1. Murphy DR, Daneman M, Schneider BA. Why do older adults have difficulty following conversations? Psychol Aging. 2006;21(1):49.

http://mjiri.iums.ac.ir

Med J Islam Repub Iran. 2017 (3 Jul); 31.36.
Utilizing ITE ENV localization training to improve speech-in-noise perception


