Perceptual Separation of Sensorineural Hearing Loss and Auditory Neuropathy Spectrum Disorder

Shuo Wang, PhD; Ruijuan Dong, MD; Dongxin Liu, BS; Yuan Wang, BS; Yitao Mao, MD, PhD; Hua Zhang, MD; Luo Zhang, MD, PhD; Li Xu, MD, PhD

Objectives/Hypothesis: The present study aimed to examine whether the response patterns to the chimeric lexical tone tokens, combined with their pure tone audiometry (PTA) results, could separate listeners with sensorineural hearing loss (SNHL) from listeners with auditory neuropathy spectrum disorder (ANSD).

Study Design: Case-control study.

Methods: Forty-three SNHL subjects and 46 ANSD subjects participated in a Mandarin lexical tone perception test using original and chimeric tone tokens. Ten sets of monosyllables, with four tone patterns for each, were processed through a 16-channel chimeric synthesizer in which a temporal envelope (E) from a monosyllabic word of one tone was paired with a fine structure (FS) from the same monosyllable of other tones.

Results: Significantly negative correlations were present between tone perception scores and PTA0.5–4 kHz for both SNHL (P < 0.001) and ANSD (P < 0.001) subjects. Overall, 72.4%, 66.4%, and 46.3% of the tone responses were consistent with FS for the SNHL subjects with mild, moderate, and severe degree of hearing loss, respectively; and 28.4%, 23.1%, and 22.7% were consistent with FS for the ANSD subjects, with the equivalent degree of hearing loss. Similarly, 17.6%, 24.2%, and 37.7% of the tone responses were consistent with E for the SNHL subjects with mild, moderate, and severe degree of hearing loss, respectively; and 45.5%, 44.3%, and 36.5% were consistent with E for the ANSD subjects.

Conclusion: Subjects with SNHL and ANSD may be separated by representing their FS- and E-consistent tone responses as a function of their pure-tone hearing thresholds.

Key Words: Lexical tone recognition, auditory chimera, sensorineural hearing loss, auditory neuropathy spectrum disorder.

Level of Evidence: 3b.

Laryngoscope, 126:1420–1425, 2016

INTRODUCTION

Temporal information in a signal can be partitioned into temporal envelopes (E), which are defined as amplitude contours of the signal, and fine structure (FS), which is defined as the instantaneous phase information in the signals, based on the Hilbert transform. Employing synthesized acoustic stimuli called auditory chimeras, Smith et al. demonstrated that E is most important for speech perception and FS is most important for pitch perception.

From the Otolaryngology–Head and Neck Surgery, Beijing Tongren Hospital, Beijing Institute of Otolaryngology, Capital Medical University (S.W., D.R., L.X., Y.W., H.Z., L.Z.), Beijing, People’s Republic of China; and Communication Sciences and Disorders, Ohio University (Y.M., L.X.), Athens, Ohio, U.S.A.

Editor’s Note: This Manuscript was accepted for publication July 24, 2015.

This work was funded in part by grants from the National Natural Science Foundation of China (81200754), the 2012 Beijing Nova Program (Z121107002512033), Beijing Natural Science Foundation (7122034, 7154190), the Capital Health Research and Development of Special from the Beijing Municipal Health Bureau (2011-1017-04), and the Basic- Clinical Cooperative Research Foundation of Capital Medical University (12YL46).

The authors have no other funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Luo Zhang, MD, PhD, Beijing Institute of Otolaryngology, 17 Hougu Hutong, Beijing 100005, China. E-mail: dr.luozhang@gmail.com

DOI: 10.1002/lary.25595

Laryngoscope 126: June 2016

Lexical tones in a tonal language are an integral part of the speech and manifest themselves in the pitch contours or the tone patterns. In this regard, Mandarin Chinese is a tonal language in which up to four tone patterns can be associated with any monosyllabic words. Xu and Pfingst used specifically synthesized chimeric tone tokens to investigate the relative contributions of E and FS in Mandarin tone recognition in subjects with normal hearing, and demonstrated that the FS played a dominant role in lexical tone recognition in these subjects. Similarly, Wang et al. used chimeric tone tokens to examine the roles of E and FS in Mandarin tone recognition in listeners with conventional sensorineural hearing loss (SNHL). These authors reported that hearing-impaired listeners relied increasingly on E to perceive tones and progressively less on FS as hearing loss became more severe, indicating that SNHL listeners have progressively reduced ability to use FS for lexical tone recognition, whereas the ability to use E for tone recognition may not be degraded. Other studies have demonstrated that, although listeners with SNHL have an impaired ability to use FS in speech and pitch perception, they might still have the ability to use E equivalent to that of subjects with normal hearing.

Auditory neuropathy spectrum disorder (ANSD) is a disorder characterized by dysynchrony of the auditory nerve firing but with normal cochlear amplification function. Studies have demonstrated that listeners with ANSD...
have a significantly impaired ability for temporal processing.\textsuperscript{11–13} Clinically, ANSD is diagnosed by the presence of otoacoustic emissions (OAEs) and/or cochlear microphonics in combination with absent or abnormal auditory brainstem responses (ABRs).\textsuperscript{14,15} Difficulty in speech perception, especially in noisy conditions, is also a prominent feature in listeners with ANSD. Although some studies have demonstrated the existence of a negative correlation between the pure-tone thresholds and the open-set speech perception ability for listeners with SNHL,\textsuperscript{16,17} there is no correlation between speech perception ability and pure-tone thresholds in listeners with ANSD.\textsuperscript{18} Furthermore, speech perception ability in the majority of listeners with ANSD is significantly poorer than that of listeners with SNHL having the equivalent degree of hearing loss, although some listeners with ANSD have comparable speech perception ability to listeners with SNHL,\textsuperscript{15,19} suggesting that it is not possible to distinguish between these groups of listeners purely on the basis of their speech perception ability.

Wang et al.\textsuperscript{20} recently investigated the role of E and FS in Mandarin lexical tone recognition in listeners with ANSD and demonstrated that these individuals had markedly compromised temporal processing ability but normal spectral resolution. In contrast, listeners with SNHL had poor frequency resolution but normal temporal gap detection. Furthermore, FS processing was deteriorated and E processing was somewhat enhanced in listeners with ANSD as compared to listeners with SNHL, as indicated by tone recognition responses to chimeric tone tokens. In view of these findings, it is hypothesized that it may be possible to distinguish between listeners with SNHL and ANSD based on their ability to use FS or E to perceive lexical tones. The aim of the present study was thus to examine the perceptual response patterns to chimeric tone tokens combined with their pure tone audiometry (PTA) results in subjects with SNHL or ANSD.

MATERIALS AND METHODS

Subjects

Overall, 89 native Mandarin Chinese-speaking subjects, comprising one group of 43 subjects with SNHL (28 females and 15 males) and one group of 46 subjects with ANSD (19 females and 27 males), were recruited from the Clinical Audiology Center in Beijing Tongren Hospital, China, from 2012 to 2014. The study protocol was reviewed and approved by the Institutional Review Board of Beijing Tongren Hospital, and all subjects provided informed written consent to participate in the study.

Subjects with SNHL were aged from 15 to 52 years old (mean ± standard deviation [SD] = 22.5 ± 10.6 years), and none presented a history that was suggestive of the auditory processing disorder. Subjects with ANSD were aged from 15 to 46 years old (mean ± SD = 25.9 ± 7.2). Both groups of subjects had bilateral hearing loss ranging from mild up to severe based on the World Health Organization criteria for classifying the degree of hearing loss.\textsuperscript{21} Figure 1 shows the mean pure-tone hearing thresholds at frequencies of 500; 1,000; 2,000; and 4,000 Hz (PTA\textsubscript{0.54 kHz}) across the two ears for both groups of subjects. Note that the listeners with SNHL tended to show a sloping high-frequency hearing loss, whereas those of ANSD tended to show a low-frequency hearing loss. All subjects in both groups had positive distortion product OAE at the frequency range of F2 from 0.7 to 6 kHz and the F2/ F1 ratio of 1.22. Their acoustic reflex thresholds for pure tone from 0.5 to 4 kHz were either absent or elevated. The ABRs were recorded using the Smart Evoked Potentials (EP) module in the intelligent hearing system (IHS) System. The click stimuli were presented monaurally at an intensity starting at 80 dB nHL, with a rate of 11.1/second. A total of 2,048 sweeps were recorded using alternating polarities, rarefaction, and condensation. An absence of responses was considered when no wave V was identified at the maximum presentation level (100 dB nHL). The ABRs were absent for all subjects with ANSD but were recorded for all subjects, with SNHL with the ABR thresholds ranging from 50 to 100 dB nHL.
Test Materials

The original test materials consisted of 10 sets of Chinese monosyllables (bai, di, tan, fei, guo, hu, liu, ma, qi, and she) with four tone patterns each, resulting in a total of 40 commonly used Chinese words. These original words were recorded in an acoustically treated booth using one adult male and one adult female native Beijing Mandarin speaker. Because there is some variation in the durations of each syllable for different tones during natural speech production, the speakers were asked to record the 40 monosyllabic words multiple times. In order to eliminate the effect of duration cues on tone recognition, tokens in which the durations of four tones in each monosyllabic word were within 5-milliseconds precision were chosen as the original tone tokens. Speech signals were captured through an M-Audio, 200 Scenic View Drive, Cumberland, RI 02864, Unites States digital recording interface attached to a computer, and recordings were made at a 44.1-kHz sampling rate and 16-bit quantization.

The auditory chimera technique was used in the present study to create the chimeric tone tokens, which contained FS from one tone pattern and the E from a different tone pattern. The chimeric tone tokens were generated in a 16-channel condition only because this not only reduce test time but also avoided the potential effects of reconstructed E from the FS when fewer channels were used. Because the auditory chimera algorithm developed by Smith et al. utilizes the Hilbert transform that rigorously computes FS and E based on the mathematical formula, without any free parameters to limit the E extraction, this may induce the blurriness between E and FS in the chimeric stimuli. Therefore, in order to prevent this blurriness, the temporal envelope extracted in each filter band using the Hilbert transform was low-pass filtered with a cutoff frequency of 64 Hz. Overall, 240 chimeric tone tokens were generated (i.e., 12 chimeric combinations × 10 sets of syllables × 2 voices) using the auditory chimera technique and were used in conjunction with the 80 original tone tokens to assess the tone recognition performance in the SNHL and ANSD groups of listeners.

Procedures

The tone recognition task was conducted in an acoustically treated booth. A Grason-Stadler A Welch Allyn Company 1 Westchester Drive Milford, NH 03055-3056 clinical audiometer connected to a computer was used to adjust the intensity of the tone tokens, and a four-alternative, forced-choice procedure was employed for the tone recognition test. A custom-made graphical user interface written in MATLAB (MathWorks, Natick, MA) was used to present the stimuli and to record the responses during the experiment. Subjects listened to the chimeric speech materials presented bilaterally through headphones, with the intensity of the stimuli set at the most comfortable loudness level, and tone tokens were presented in a randomized order. A set of four Chinese characters with four tone patterns and their pinyin were displayed on the screen (e.g., bai ma; ma bai; ma ma; ma ma) on presentation of the stimuli, and the subjects were asked to select the tone pattern (or Chinese word) in the token that they had heard using the computer. Prior to the formal test, all participants were given instructions and undertook a practice session listening to the tokens and using the computer to familiarize themselves with the test, which took approximately half an hour for each subject.

Statistical Analysis

The results were analyzed statistically using Statistics Package for Social Science 16.0 (SPSS Inc., Chicago, IL). A Pearson correlation was used to analyze the correlation between pure-tone threshold and tone recognition performance, as well as the correlation between tone responses that were consistent with E and FS for each group of subjects. An independent-samples t test was used to compare the mean differences in tone responses consistent with FS or E in subjects SNHL and ANSD. An ellipsoid model was used to plot the assessment of tone responses consistent with both FS and E as a function of pure-tone threshold of each group of subjects.

RESULTS

Correlation Between Pure Tone Threshold and Tone Recognition Performance

Tone recognition performance employing the original tone tokens depended on the degree of hearing loss. For mild, moderate, and severe degrees of hearing loss, the SNHL group scored 99.4%, 95.0%, and 86.9% correct on tone recognition, respectively, whereas the ANSD group scored 68.7%, 54.3%, and 34.0% correct, respectively. Figure 2 shows the correlation between the tone recognition scores and pure-tone hearing thresholds for both subjects with SNHL and ANSD. There was a significant correlation between tone recognition scores and pure-tone hearing thresholds for both groups of subjects with SNHL and ANSD. ANSD. An ellipsoid model was used to plot the assessment of tone responses consistent with both FS and E as a function of pure-tone threshold of each group of subjects.

Fig. 2. Correlation between pure tone hearing thresholds (PTA0.5–4 kHz) and tone-perception scores for subjects with SNHL (blue triangles) and ANSD (red squares), respectively. ANSD = auditory neuropathy spectrum disorder; PTA = pure tone audiometry; SNHL = sensorineural hearing loss.
Correlation Between Tone Responses Consistent With E and FS

Tone recognition performance to the chimeric tone tokens was classified based on whether the responses were consistent with either the E or the FS. Both subjects with SNHL and ANSD demonstrated a negative correlation between tone responses that were consistent with E and FS [SNHL: \( r (43) = -0.962, P < 0.001 \); ANSD: \( r (46) = -0.393, P < 0.01 \)] (Fig. 3), indicating that as the ability to use FS cues was decreased in these subjects, their ability to use E cues to compensate for this deficit in tone recognition was increased.

Effects of Pure Tone Thresholds on Tone Responses Consistent With FS and E

Figure 4 shows the responses consistent with FS and E, depending on the degree of hearing loss in both SNHL and ANSD subjects. Overall, 72.4%, 66.4%, and 46.3% of the tone responses were consistent with FS; and 17.6%, 24.2%, and 37.7% were consistent with E for the SNHL subjects with mild, moderate, and severe degrees of hearing loss, respectively. Similarly, 28.4%, 23.1%, and 22.7% of the tone responses were consistent with FS; and 45.5%, 44.3%, and 36.5% were consistent with E for the ANSD subjects with mild, moderate, and severe degrees of hearing loss, respectively.

Comparison of the mean differences in tone responses consistent with FS or E in subjects SNHL and ANSD demonstrated that, despite substantial overlaps between the two groups of subjects, all differences were statistically significant (\( t \) test, all \( P < 0.01 \)). Assessment of tone responses consistent with both FS and E as a function of PTA\(_{0.5-4\text{kHz}}\), using an ellipsoid model that represented the best fit to at least 95% of the data, demonstrated that the tone recognition performance of the subjects with SNHL was clearly different from that of subjects with ANSD (Fig. 5). The three-dimensional (3D) plot of the model was constructed, presuming that the data were normally distributed, and using the mean of the data to represent the center of the ellipsoid and the lengths of the three semi-axis to represent twice the SD of the data along each corresponding axis.
The present study demonstrated that averaged audiometric hearing thresholds (i.e., PTA0.5–4 kHz) were observed between tone recognition performance and the hearing thresholds of SNHL listeners scoring more than 90% correct with tone recognition tests. The tone recognition scores ranged from 100% to 70%, with the majority of SNHL listeners with severe SNHL ranging from 100% to 70%, with the majority of SNHL listeners scoring more than 90% correct with tone recognition tests. A significant negative correlation was observed between tone recognition performance and the averaged audiometric hearing thresholds (i.e., PTA0.5–4 kHz). In contrast, listeners with ANSD performed much worse in tone recognition tests. The tone recognition scores ranged from the chance level (i.e., 25% correct) to nearly 100% correct, and only three of the 46 patients scored higher than 90% correct. Furthermore, in the ANSD listeners with PTA0.5–4 kHz values comparable to those SNHL listeners, tone recognition performance was approximately 50 percentage points lower than in listeners with SNHL. Moreover, there was a negative correlation between tone recognition performance and the hearing thresholds of subjects with ANSD. This suggests that the ability to perceive lexical tone in these individuals generally degraded progressively with increasing severity of hearing loss. Nevertheless, there were large individual differences in tone recognition performance, especially for those with mild to moderate hearing loss, as might be expected when assessing any perceptual measures. Note, however, that if we pooled together the data of the two groups of subjects, the correlation between tone recognition performance and PTA0.5–4 kHz disappeared. No correlation existed between tone recognition performance and pure-tone threshold at any individual frequencies if the data were pooled together for the two groups of subjects, suggesting that neither the pure-tone threshold alone nor the hearing loss configuration could account for the tone perception ability of all the subjects.

**Comparison of Tone Responses Using Chimeric Tone Tokens Between SNHL and ANSD Subjects**

We examined the relative contributions of FS and E cues in tone recognition by both SNHL and ANSD listeners using auditory chimeric tone tokens. In accordance with the findings by Wang et al., the present study demonstrated that in SNHL subjects, as the severity of hearing impairment increased, tone recognition responses consistent with FS progressively decreased, whereas tone recognition responses based on E cues progressively increased. Listeners with SNHL usually have the impaired frequency selectivity because of the enlarged bandwidths of the auditory filters. Bernstein and Oxenham observed that the transition between good and poor F0 difference limens in listeners with SNHL corresponded well with the measurement of frequency selectivity, suggesting that reduced frequency selectivity may lead to the poorer F0 discrimination. Thus, spectral resolvability in the auditory system may play an important role in processing each frequency component more or less independently, and the degraded frequency selectivity in listeners with SNHL may result in the impaired ability to use FS in tone recognition. Other studies have demonstrated that listeners with SNHL may have a preserved ability to use E cues comparable to normal-hearing listeners, thus suggesting that the ability to process E cues may compensate the deficit of FS processing in tone recognition so that even listeners with severe SNHL could still achieve a reasonably good tone recognition performance, as has been shown in the present study.

In contrast to the SNHL listeners, the ability to use the FS cues in tone recognition by listeners with ANSD was severely degraded irrespective of the severity of the pure tone hearing thresholds. Their ability to process the E cues remained at a reasonable level with a large individual variability, and a weak correlation was observed between tone responses consistent with E and FS. More recently, it has been demonstrated that poor temporal resolution rather than the frequency resolution exerts the major detrimental effects on the FS cue processing for pitch perception for subjects with ANSD. The findings of the present study are in accordance with those of Narne, who demonstrated that fine structure processing was severely degraded in listeners with ANSD, and that the dramatically decreased performance in lexical tone recognition for the listeners with ANSD was due to the abolished ability to process the FS cues. Furthermore, the assessment of temporal modulation detection in listeners with ANSD showed greatly elevated thresholds in the temporal modulation transfer functions (TMTF) and a large individual variability in the TMTF, suggesting that listeners with ANSD may suffer from an abnormal processing of the E cues. The present study has indicated that the large variability in tone recognition performance in the subjects with ANSD may be due to the variable disruption of the processing.
of the temporal amplitude cues (i.e., the E cues) for tone recognition.

**Perceptual Separation of SNHL and ANSD Subjects**

Although speech perception difficulty is an important characteristic in identifying subjects with ANSD, evidence suggests that regular speech audiometry assessment may not be able to distinguish the perceptual differences between listeners with SNHL and with ANSD.\(^{15,19}\) Speech perception testing under noisy conditions may possibly distinguish subjects with ANSD from those with SNHL; some studies have shown that subjects with ANSD generally had greater difficulties in speech perception in background noise compared to subjects with SNHL.\(^{11,28}\) However, complete separation of the two groups of hearing-impaired listeners was not evident using this test. The present study provides an alternative method to separate listeners with ANSD from listeners with SNHL based on their ability to use FS and E cues in lexical tone recognition. In SNHL and ANSD subjects with equivalent degree of hearing loss, the subjects with ANSD showed a significant degradation in coding FS information relative to subjects with SNHL, especially for less than severe degree of hearing loss. Although some subjects with severe SNHL had similar ability to use FS information as subjects with ANSD, their preserved preserved ability to use E information was reasonably good, whereas the ability to use E cues in subjects with ANSD was dramatically degraded. Indeed, a 3D representation of the data incorporating the tone recognition responses consistent with FS and E and the PTA\(_{0.5–4k}\) Hz clearly separated the two groups of hearing-impaired listeners. In this respect, it is proposed that the chimeric tone recognition test might be used as an additional means to identify listeners with ANSD, particularly when the clinical audiological results are sometimes ambiguous in distinguishing ANSD from SNHL. A limitation of the chimeric tone recognition test, however, is that it can only be used in the tone-language-speaking populations; thus, the use of a simplified version of the chimeric tone recognition test using the upward and downward frequency-modulation sweeps, as described by Xu et al., is further proposed.\(^{29}\) This technique will eliminate the need for linguistic knowledge as a prerequisite and should be efficient in separating listeners with ANSD from listeners with SNHL from any linguistic background.

**CONCLUSION**

The present study aimed to examine whether the response patterns to chimeric lexical tone tokens, combined with their PTA results, could separate listeners with SNHL from listeners with ANSD. The study demonstrated that lexical tone recognition performance with the original tone tokens depended on the degree of hearing loss for both SNHL and ANSD listeners. Furthermore, a clear perceptual separation could be made between the SNHL and ANSD listeners by representing their tone responses, consistent with both FS and E as a function of pure tone hearing thresholds.

**Acknowledgments**

We thank all the subjects participating in this study and our colleagues at the Clinical Audiology Center in Beijing Tongren Hospital for recruiting the hearing-impaired subjects. We also thank Dr. J. L. Devalia for his help with editing an earlier version of the article.

**BIBLIOGRAPHY**


Wang et al.: Perceptual Separation of SNHL and ANSD