

Exploring the feasibility of remote administration of speech audiometry: A comparative study of conventional and digital methods

DIGITAL HEALTH
Volume 10: 1–8
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076241271834
journals.sagepub.com/home/dhj



Chen Yuan 

Abstract

Objective: This study investigated the effectiveness of remote administration of speech audiometry, an essential tool for diagnosing hearing loss and determining its severity. Utilizing two software tools for remote testing, the research aimed to compare these digital methods with traditional, in-person speech audiometry to evaluate their feasibility and accuracy.

Design: Participants underwent the Cantonese Hearing in Noise Test (CHINT) under three listening conditions—quiet, noise from the front, and noise from the right side—using three different administration methods: the conventional in-person approach, video conferencing software, and remote access software.

Study Sample: Fifty-six Cantonese-speaking adults residing in Hong Kong participated in this study.

Results: Analysis revealed no significant differences in CHINT scores among the three administration methods, indicating the potential for remote administration to yield results comparable to those of conventional methods.

Conclusions: The findings supported the feasibility of remote speech audiometry using the investigated digital tools. This study paved the way for the wider adoption of tele-audiology practices, particularly in situations where in-person assessments are not possible.

Keywords

Telehealth, tele-audiology, speech audiometry, speech perception in noise, hearing in noise test

Submission date: 14 January 2024; Acceptance date: 2 July 2024

Introduction

The World Health Organization (WHO) estimates that approximately 900 million people worldwide suffer from disabling hearing loss, leading to potential costs of around USD 750 billion annually if unaddressed.¹ This presents a substantial barrier to accessing comprehensive hearing care services, which include, but are not limited to, hearing aid (HA) provision. Tele-audiology emerges as a component of the broader solution, providing access to audiological care for individuals in remote or underserved areas.² The COVID-19 pandemic has notably accelerated the shift from traditional in-person audiological services to remote delivery, markedly changing perceptions and usage of tele-audiology. Before the pandemic, the value

of tele-audiology was recognized by 44.3% of respondents in a survey, a figure that rose dramatically to 87.1% during the pandemic.³ The adoption of tele-audiology services similarly increased from 41.3% pre-pandemic to 61.9% during it, indicating a robust appreciation for its benefits.³ The US Department of Health and Human Services observed steady telehealth usage rates from April 2021 to

Department of Special Education and Counselling, The Education University of Hong Kong, Tai Po, Hong Kong SAR, China

Corresponding author:

Chen Yuan, Department of Special Education and Counselling, The Education University of Hong Kong, Tai Po, Hong Kong SAR, China.
Email: chen@eduhk.hk



August 2022, noting that, despite a peak during the pandemic, telehealth use has remained significantly higher than pre-pandemic levels. This trend indicates a sustained adoption of telehealth by both patients and providers.⁴

Tele-audiology encompasses a broad spectrum of services, including screening (e.g. pure-tone audiometry, otoacoustic emission, tympanometry), diagnosis (e.g. auditory brainstem response, video-otoscopy, tympanometry, pure-tone audiometry), and intervention (e.g. cochlear implant mapping, tinnitus management, education, counselling, and hearing aid programming), as detailed in previous studies.² Despite the comprehensive capabilities that tele-audiology offers, there has been limited focus on tele-speech-audiometry, a component of audiological assessments. Speech audiometry plays a role in the evaluation of hearing loss, providing essential information on an individual's ability to understand speech under various conditions. This testing is indispensable for diagnosing the degree of hearing loss and for guiding the appropriate choice of interventions, such as hearing aids, cochlear implants, or other therapeutic measures.⁵⁻⁸

When speech audiometry is carried out, it is not limited to a single environmental condition; rather, it encompasses a variety of scenarios to mimic real-life situations as closely as possible. Testing speech perception in quiet conditions is foundational, serving as a benchmark for an individual's ability to process speech without background noise.⁹ This scenario is crucial for determining the most basic level of speech understanding and for identifying any primary difficulties with speech recognition.

However, real-world communication rarely occurs in perfectly quiet environments. Consequently, assessing speech perception in noise becomes essential. This aspect of testing reflects more realistic conditions where background sounds are present, which is a common challenge for individuals with hearing loss.^{6,10} Understanding how well a person can comprehend speech with noise from the front is particularly telling, as this situation is typical in many day-to-day interactions, such as conversations in busy settings where the speaker and the noise source are in the same direction.^{6,10,11}

Equally important is the evaluation of speech perception with noise coming from the side. This condition tests the ability of individuals to focus on speech while ignoring lateral noise, a skill necessary for effective communication in environments.^{12,13} Testing in these varied conditions—quiet, noise from the front, and noise from the side—provides a comprehensive picture of an individual's speech understanding capabilities. It offers critical insights into the complex nature of hearing loss and its impact on everyday communication. Through such thorough assessments, audiologists can tailor interventions more precisely, enhancing the individual's ability to engage in meaningful interactions across a broad spectrum of environments.

Despite speech audiometry's acknowledged importance in audiology, its integration into tele-audiology has been limited. An international survey of audiologists revealed that the greatest barriers to the use of tele-audiology were associated with clinical equipment, such as multiple technologies across manufacturers, which can be used to conduct measurements remotely.³ In addition, respondents suggested that research was needed to confirm the validity of telehealth services.^{3,14} In response to these challenges, the current study aims to examine the feasibility of administering speech perception remotely using two widely accessible software tools: video-conferencing and remote-control software. The hypothesis guiding this research posits that there will be no significant differences in the outcomes of speech perception tests whether administered in-person, via video-conferencing, or through remote-control software.

Materials and methods

Participants

A total of 56 participants aged 18 to 50 years (mean age = 26.23, SD = 7.85) were recruited from the Education University of Hong Kong using online advertisements, with 67.9% of them being female (n = 38). Eligibility for participation required native proficiency in Cantonese, with no specific criteria set for hearing status. The homogeneity of the participant group helps to reduce variability potentially stemming from factors not central to the study's objectives, such as differences in technological proficiency, cognitive abilities, and auditory function. The scenarios devised for this research aim to replicate the conditions of speech perception evaluations typically encountered in routine physical examinations and hearing function screenings. Demonstrating the feasibility of conducting speech perception assessments remotely in this study opens the door to future research into their applicability among more diverse populations, especially older adults with hearing impairments. Written consent was obtained from all participants prior to the start of the study.

Measures

The Cantonese Hearing in Noise Test (CHINT)¹⁵ is designed to assess an individual's ability to perceive sentences under various acoustic conditions, emphasizing the adaptive nature of this assessment. Central to the CHINT is the evaluation of how well participants can recognize sentences when presented amidst noise or in quiet environments. This ability is quantified through the measurement of Speech Reception Thresholds (SRTs), which determine the minimal level of presentation in quiet conditions or the required signal-to-noise ratio (S/N) in noisy conditions for accurate recognition of 50% of the sentences.

Incorporating principles from the English Hearing in Noise Test,¹⁶ the CHINT consists of 12 lists, each with 20 sentences. These sentences are used to evaluate speech perception capabilities by masking them with speech spectrum-shaped noise during the test. The assessment is conducted through insert earphones, delivering both the target sentences and masking noise under three distinct listening scenarios: in quiet, with noise coming from the front (NF), and with noise coming from the right side (NS).

Moreover, the CHINT leverages head-related transfer functions (HRTFs) to simulate sound-field listening conditions through insert earphones, enhancing the realism of the test environments.¹⁷ HRTFs describe the complex spatial filtering effects of the outer ear, including the torso, head, and pinnae, on sounds from different directions, thereby providing a comprehensive assessment of a participant's ability to perceive speech in varied acoustic scenarios using insert earphones.¹² The CHINT was conducted employing the three methodologies below.

Conventional method. An experimenter operated the computer and delivered instructions directly to the participants. The CHINT was administered face-to-face, following the guidelines outlined in the CHINT manual.

Using remote-access software. TeamViewer version 15.27.3, a remote-access software, was employed to facilitate the administration of the CHINT from a different location through the internet. The selection of TeamViewer was based on its capabilities for remote control and communication, which are essential for conducting tests without the physical presence of an experimenter.

The process began with the experimenter, situated in a separate room, taking control of the computer located within the sound booth. This remote setup allowed for the CHINT to be administered as if the experimenter was physically present, without actually being in the same room. The advantage of this approach was that the audio was produced directly from the participant's side, ensuring it was delivered without delay and without compression.

Furthermore, the use of TeamViewer's video conferencing feature was instrumental in delivering instructions to participants. This real-time communication ensured that participants fully understood the test procedures and what was expected of them, thereby facilitating a smoother testing process. Additionally, the demonstration of equipment usage, particularly the proper way to use insert earphones, was crucial for the accuracy of the CHINT. The experimenter demonstrated this using an additional pair of insert earphones, effectively showing participants the correct usage through the video conference.

Using video-conference software. In this condition, speech audiometry was administered remotely using ZOOM video conferencing software version 5.9.1. It does not

require participants to install CHINT administration software on their computers. Importantly, it also raises fewer concerns regarding privacy and data protection. For instance, the host organization—where the participants are tested—might be wary of allowing computers from other organizations to control their systems, a scenario observed in the study using remote-access software. However, while video conferencing software can streamline communication, it may compromise the audio quality due to the application of audio compression techniques.¹⁸ These techniques involve removing redundant or non-essential information from the audio signals to enhance the transmission efficiency.¹⁸

To closely mimic the audio quality of face-to-face testing conditions, we leveraged ZOOM's advanced audio settings to transmit high-quality stereo audio from the computer being administered. This was accomplished by selecting 'Share Screen,' proceeding to 'Advanced,' and choosing 'Computer Audio.' Such configuration was vital for conducting speech perception in noise tests, allowing for the delivery of sentences and noise across separate channels at varying intensities. Moreover, ZOOM's 'High-Fidelity Music Mode' was activated through the settings by navigating to audio—original sound for musicians. This specialized audio setting designed to optimize the platform for music and high-quality audio enhances audio quality via higher bitrate transmission, resulting in clearer, more detailed sound. It also reduces audio compression and disables post-processing to preserve the original sound's quality. Additionally, this mode lowers audio latency, which is crucial for synchronizing live performances or tests by minimizing the delay between sound production and reception. This setup ensures the audio fidelity in speech perception tests is on par with clinical settings.

Similar to the remote-access software administration (RASA) method, this procedure was conducted without facilitators present in the same room as the participants, with all operations managed by an experimenter in a separate room via ZOOM. The experimenter could neither see nor hear the participants without utilizing ZOOM. In contrast to the RASA, sentences and noise from the CHINT were generated on the experimenter's computer, and then transmitted to the participant's computer.

Procedures

Before the administration of the CHINT, participants underwent a pure-tone audiometry test in person. The CHINT was administered using three distinct approaches: conventional (CM), video-conference software (VCSA), and RASA using two ThinkPad E15 computers. To determine if variations in outcomes across different administration techniques (CM, VCSA, and RASA) were due to the use of different CHINT lists, the CHINT was administered

twice via the conventional method in 34 randomly selected participants out of the total 56 participants (designated as CM1 and CM2). The testing employed a pair of ER-3A insert earphones, which were calibrated beforehand by utilizing the ARSON-DAVIS SYSTEM 824 audiometer connected to an AEC201-A ear simulator and an AEC 304 occluded ear simulator. Specifically, the calibration involved displaying a calibration sound through the CHINT software and adjusting the computer volume until the audiometer indicated the sound level at around 80 dB SPL. For the duration of the study, all participants used the same set of insert earphones, with ear tips being replaced for each participant to maintain hygiene and consistency.

The study encompassed 12 test scenarios, integrating three listening settings (quiet, NF, and noise NS) with four administration strategies (CM1, CM2, VCSA, and RASA). In each scenario, one out of 12 CHINT lists, each consisting of 20 sentences, was selected at random for every test condition.¹⁷ The equivalency of these 12 lists has been previously established to ensure consistency in test difficulty and content.¹⁷ Moreover, the sequence of these test conditions was randomized using a Latin Square design to ensure a balanced distribution. Prior to employing each administration method, a trial test was conducted to confirm participants' comprehension of the instructions.

The duration of the testing procedure was approximately 1 hour. The participant was granted a 10-minute intermission roughly halfway through the session. Additional breaks were available upon request to accommodate participant needs. All tests were performed in a certified sound booth by an audiologist (i.e. the experimenter) at the Integrated Centre for Wellbeing (I-WELL) of the Education University of Hong Kong. A Wi-Fi connection characterized by a latency of 5.2 ms, a jitter of 1.6 ms, a download speed of 108 Mbps, and an upload speed of 93 Mbps was established for the tests. These connection metrics were verified through the OFCA Broadband Performance Test (<https://speedtest.ofca.gov.hk/>).

Statistical analysis

To evaluate differences in SRTs across various administration methods (CM, VCSA, RASA) and listening conditions (quiet, NF, NS), a two-way repeated-measures analysis of variance (ANOVA) was employed. Additionally, to investigate the disparities in SRTs among different administration methods, we calculated the differences in SRTs from the first (CM1) and second (CM2) conventional method testing denoted as Δ SRT (CM1-CM2); the differences between the first conventional method testing and the ZOOM administration is marked as Δ SRT (CM1-VCSA); and the differences between the first conventional method testing and the TeamViewer administration is indicated as

Δ SRT (CM1-RASA). Given that the reported list equivalence for the CHINT is within ± 1 dB,¹⁷ it was expected that Δ SRT (CM1-CM2), Δ SRT (CM1-VCSA), and Δ SRT (CM1-RASA) would also fall within this range. This hypothesis would be corroborated by a two-way repeated ANOVA, which should reveal no significant differences among Δ SRT (CM1-CM2), Δ SRT (CM1-VCSA), and Δ SRT (CM1-RASA), confirming the consistency and reliability of the test results across different administration methods. The significance level was set at 5%, and SPSS (v29.0.2) was used to perform all the statistical analyses.

Results

Pure-tone audiometry results revealed that 54 of the 56 participants demonstrated normal hearing, with their hearing thresholds reaching 20 decibels hearing level (dB HL) or lower at frequencies of 500, 1000, 2000, and 4000 Hz in both ears. Meanwhile, two participants showed signs of mild hearing impairment characterized by a decrease in hearing sensitivity to a range between 21 and 40 dB HL at these frequencies in the better-performing ear. The SRTs using different administration methods are shown in Table 1.

Repeated ANOVA showed that there was a significant main effect of the listening condition (quiet, NF, NS), $F(2, 110) = 3891.74$, $p < 0.05$, $\eta_p^2 = 0.99$. However, there was no significant main effect of the administration methods, $F(2, 110) = 3.09$, $p > 0.05$, $\eta_p^2 = 0.05$, and interaction effects between listening condition and administration methods, $F(4, 220) = 2.02$, $p > 0.05$, $\eta_p^2 = 0.04$.

The analysis of SRT differences between various administration methods, as highlighted in Table 2, underscores the consistency in SRTs across conventional, ZOOM, and TeamViewer modalities. This consistency is encapsulated by the mean difference scores across all test conditions, which remained within the ± 1 dB range.¹⁷ Such a range aligns with the list equivalence previously established for the CHINT,¹⁷ affirming the reliability of remote audiometry methods compared to traditional in-person assessments. Further confirmation of these findings is provided by the repeated-measures ANOVA, which reveals no significant differences in the Δ SRT (CM1-CM2), Δ SRT (CM1-VCSA), and Δ SRT (CM1-RASA), $F(2, 66) = 0.07$, $p > 0.05$, $\eta_p^2 = 0.00$.

Discussion

The present study investigated the feasibility of remote administration of speech audiometry using two different digital tools and found no significant differences in SRTs across different administration methods. The results of this study closely align with those of prior research comparing pure-tone audiometry thresholds obtained through

Table 1. Speech reception thresholds (SRTs) across different listening conditions and administration methods [range] (standard deviation).

	Conventional method one (CM1)	Conventional method two (CM2)	Using the video-conference software (VCSA)	Using a remote-access software (RASA)
Quiet	18.82 dB	18.53 dB	17.89 dB	18.25 dB
	[13–36 dB]	[11–38 dB]	[15–25 dB]	[11–34 dB]
	(4.2 dB)	(5.45 dB)	(2.40 dB)	(4.64 dB)
Noise_Front (NF)	−4.25 dB	−4.56 dB	−4.09 dB	−4.55 dB
	[−7.00 dB to −1.10 dB]	[−6.90 dB to −3.10 dB]	[−6.70 dB to −1.10 dB]	[−6.60 dB to −2.70 dB]
	(1.09 dB)	(0.84 dB)	(0.96 dB)	(0.95 dB)
Noise_Side (NS)	−9.55 dB	−10.01 dB	−10.11 dB	−9.98 dB
	[−12.90 dB to −5.6 dB]	[−15.40 dB to −4.80 dB]	[−14.10 dB to −7.40 dB]	[−13.60 dB to −5.50 dB]
	(2.24 dB)	(2.08 dB)	(1.44 dB)	(1.54 dB)

Table 2. Difference in SRTs obtained using different administration methods: mean (standard deviation).

	Δ SRT (CM1–CM2)	Δ SRT (CM1–VCSA)	Δ SRT (CM1–RASA)
Quiet	0.65 dB (3.16 dB)	0.92 dB (3.30 dB)	0.57 dB (3.09 dB)
Noise front	0.56 dB (1.30 dB)	−0.17 dB (1.21 dB)	0.30 dB (1.11 dB)
Noise side	0.36 dB (2.04 dB)	0.56 dB (2.05 dB)	0.44 dB (2.13 dB)

remote testing with those from conventional face-to-face assessments. For instance, Swanepoel et al.¹⁹ utilized the RASA method to evaluate remote pure-tone audiometric testing and found no clinically significant differences between the remote and conventional face-to-face audiometry outcomes. Similarly, Botasso assessed the efficacy and feasibility of remotely administering pure-tone audiometry screening among elementary school children, confirming the reliability and feasibility of tele-audiometry for hearing screening in this demographic. These findings are corroborated by additional studies, such as Choi’s PC-based tele-audiometry²⁰ and the research of Krumm et al.²¹

However, these prior studies focused solely on pure-tone audiometry. Speech audiometry, in contrast, not only requires the ability to hear sounds, but also to understand speech, which may necessitate a higher-quality acoustic environment. Hughes et al.¹⁴ compared tele-speech-audiometry with traditional methods among cochlear implant users and reported significantly poorer speech perception in remote settings. This disparity was attributed to the absence of soundproof booths at the remote locations exacerbated by higher background noise levels and longer

reverberation times—factors known to adversely affect speech perception in individuals with hearing loss.^{22,23}

In this study, all the speech audiometry was completed in a sound booth. The audio was obtained from the same computer as in the conventional and RASA methods (i.e. TeamViewer). Therefore, the quality of the audio did not change when speech audiometry was administered using the RASA methods. The variations in SRTs between the conventional method and the RASA method may mainly come from the interactions between participants and administrators (e.g. how the instructions were delivered online) and the use of different CHINT sentences. However, these would not significantly affect the CHINT results, as no significant differences in SRTs were found between these two administration methods.

During video calls in the VCSA, audio data, alongside video, are transmitted over the internet. Initially, this involves converting and compressing the analog audio signal into a digital format using audio codecs.¹⁸ This helps minimize the data size while preserving the quality. Once encoded, the data are packaged and sent to the recipient, where they are decompressed and reverted to their analog form. However, variable network conditions can

cause packet delays or losses, potentially impacting audio clarity and leading to distortions.¹⁸ The codec selected is vital as it adjusts the acoustic bandwidth and bitrate in response to network changes, significantly influencing the audio transmission quality.²⁴

Zoom uses the Opus codec, which adapts to various audio types, from narrow-band speech to full-band stereo music, employing two main techniques: linear prediction (LP) and modified discrete cosine transform (MDCT).²⁵ The LP technique reduces bitrate and enhances compression effectiveness, while MDCT optimizes for higher frequencies or music by segmenting the audio spectrum into bands that align with the Bark scale, which mimics the human ear's frequency resolution.²⁵ Opus can dynamically adjust its mode based on network conditions, supporting everything from low bitrate speech to high-quality music transmission.^{18,25} We found that the audio processing in Zoom, as described above, does not significantly affect Speech Reception Thresholds (SRTs), consistent with the findings of Perepelytsia and Dellwo,¹⁸ who reported that acoustic compression in Zoom audio does not compromise the voice recognition performance.

Additionally, the dynamics of interactions between participants and administrators during testing can influence the SRT results. Specifically, remote administration involves communication through digital platforms, which can alter the rapport and clarity of instructions compared to in-person interactions.²⁶ For instance, technical issues, such as latency or audio quality, can affect the participant's understanding of test instructions, potentially impacting their performance.¹⁸ Moreover, the absence of physical presence might influence the participant's engagement and concentration during the test. In conventional settings, administrators can adapt their communication strategies based on real-time feedback from participants, such as body language or facial expressions, to ensure comprehension and comfort. This dynamic adjustment is more challenging to replicate with remote administration.²⁶ Nonetheless, the observed variations in SRTs are within a 1 dB range, aligning with the SRT fluctuations encountered when using different CHINT sentences (i.e. Δ SRT (CM1–CM2)). This indicates that the nuanced dynamics of participant-administrator interaction in remote assessments did not significantly affect SRT outcomes.

Limitations and suggestions

This study demonstrated that there were no significant differences among the three administration methods, indicating that the two non-physical-contact methods are viable for conducting speech audiometry. However, all participants in this study were relatively young and exhibited only normal-to-mild hearing impairment. Older adults and those with more severe hearing impairments might encounter difficulties understanding instructions online due to

unfamiliarity with the technology. Nevertheless, the instructions for CHINT administration were straightforward, requiring no participant operation of the software, and a trial test was conducted to ensure comprehension of the instructions. To alleviate potential communication barriers, especially among older adults, future studies could benefit from integrating facilitators or visual aids, such as detailed pictorial instructions.

In addition, insert earphones were used in the current study, which are rarely utilized as an outcome measure for hearing aid benefits. Sound field testing using speakers is more commonly employed when evaluating HA fitting outcomes. Compared to insert earphones, sound field speech audiometry is more easily affected by ambient noise. Therefore, further studies are warranted. It is important to note, however, that HA fitting encompasses much more than just speech audiometry. Speech audiometry has other significant clinical applications: it can better predict real-life communication skills; it can assist in diagnosing conditions, such as auditory neuropathy and central auditory processing disorder; it can determine candidacy for cochlear implants or HAs; and it can monitor treatment progress, such as in cases of sudden deafness.^{27,28} We hope the results of this study will contribute to these fields.

Furthermore, this study employed Zoom's video-conference function to represent VCSA and TeamViewer's remote-access function to represent RASA. While a variety of similar remote-access software options exist, such as Zoom's own remote-access capabilities, the choice of software appears to have a minimal impact on the results. However, due to potential differences in audio data processing and transmission, such as compression and noise reduction by different video-conferencing platforms, conducting additional validity tests is advised. These tests are essential to verify the efficacy of using alternative video-conferencing software for the administration of the CHINT test.

Implications

Despite its limitations, this study highlights several critical implications. First, the two remote administration software tools used are freely available and user-friendly, requiring no specialized equipment or additional personnel. These attributes are particularly beneficial in situations necessitating social distancing, such as evaluating speech perception among individuals with infectious diseases.

Second, tele-speech-audiometry proves vital for residents in remote or underserved areas, where access to audiological services is often scarce. For instance, in comparison to the large number of individuals with hearing loss, the availability of audiologists is quite restricted in China.²⁹ Speech audiometry is not routinely provided by some audiological service providers, requiring patients to travel across provinces to receive care.^{29,30} However, finding a

sound booth in hospitals or at hearing aid distributors is typically feasible. Therefore, developing region-specific tele-speech-audiometry protocols could significantly enhance accessibility and reduce the need for long-distance travel for audiological evaluations.

Lastly, contemporary studies typically require multi-center collaboration. Ensuring sufficient participant enrollment and maintaining compliance are crucial for the success of these studies. For example, our ongoing multicenter study investigates the effectiveness of online auditory training in enhancing speech perception in noise for adults with untreated mild hearing loss. Administering the speech perception test remotely has significantly increased the number of participants required and reduced travel demands on researchers.

Acknowledgments: I am grateful to all the participants who took the time to participate in this study.

Contributorship: CY solely undertook the conceptualization of the research topic, collection and analysis of data, and the composition of the written manuscript.

Data availability statement: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declaration of conflicting interests: The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the General Research Fund (GRF 18608423), University Grants Committee, Hong Kong SAR, China, and the Knowledge Transfer Seed Fund for 2023–2024 (DTSF03109), Department of Special Education and Counselling, The Education University of Hong Kong, Hong Kong SAR, China.

Guarantor: CY.

Informed consent from participants: The study has been approved by the Faculty Human Research Ethics Committee (HREC), the Education University of Hong Kong (FHREC:23/24-ER030). Written consent was obtained from participants prior to the study.

ORCID iDs: Chen Yuan  <https://orcid.org/0000-0001-6706-0962>

References

- Haile LM, Kamenov K, Briant PS, et al. Hearing loss prevalence and years lived with disability, 1990–2019: findings from the Global Burden of Disease Study 2019. *Lancet* 2021; 397: 996–1009. doi:10.1016/S0140-6736(21)00516-X
- Muñoz K, Nagaraj NK and Nichols N. Applied tele-audiology research in clinical practice during the past decade: a scoping review. *Int J Audiol* 2021; 60: S4–S12.
- Eikelboom RH, Bennett RJ, Manchaiah V, et al. International survey of audiologists during the COVID-19 pandemic: use of and attitudes to telehealth. *Int J Audiol* 2022; 64: 283–292.
- Lee EC, Grigorescu V, Enogieru I, et al. Updated national survey trends in telehealth utilization and modality (2021–2022). US Department of Health and Human Services 2023.
- Wong LL, Zhu S, Chen Y, et al. Discrimination of consonants in quiet and in noise in Mandarin-speaking children with normal hearing. *Plos one* 2023; 18: e0283198.
- Chen Y, Wong LLN, Chan SS, et al. Speech perception in noise is associated with different cognitive abilities in Chinese-speaking older adults with and without hearing aids. *Front Psychol* 2021; 12: 640300. DOI: 10.3389/fpsyg.2021.640300
- Chen Y, Wong LL, Kuehnel V, et al. Can dual compression offer better Mandarin speech intelligibility and sound quality than fast-acting compression? *Trends Hear* 2021; 25: 2331216521997610. Doi: 10.1177/2331216521997610
- Wong LLN, Chen Y, Wang Q, et al. Efficacy of a hearing aid noise reduction function. *Trends Hear* 2018; 22: 2331216518782839. DOI: 10.1177/2331216518782839
- Chen Y. Is Cantonese lexical tone information important for sentence recognition accuracy in quiet and in noise? *Plos one* 2022; 17: e0276254.
- Chen Y, Wong LL, Qian, et al. The Role of Lexical Tone Information in the Recognition of Mandarin Sentences in Listeners With Hearing Aids. *Ear Hear*. 2020; 41: 532–538.
- Mattys SL, Davis MH, Bradlow AR, et al. Speech recognition in adverse conditions. *Lang Cognitive Proc* 2012; 27: 953–978.
- Wong LLN, Chen Y and Leung KP. The Cantonese Hearing in Noise Test for children. *Trends Hear* 2019; 23: 2331216519837128. doi:10.1177/2331216519837128
- Chen Y and Wong LLN. Speech perception in Mandarin-speaking children with cochlear implants: a systematic review. *Int J Audiol* 2017; 56: S7–S16. DOI: 10.1080/14992027.2017.1300694
- Hughes ML, Goehring JL, Baudhuin JL, et al. Use of telehealth for research and clinical measures in cochlear implant recipients: a validation study. *J Speech Lang Hear Res* 2012; 55: 1112–1127. DOI: 10.1044/1092-4388(2011/11-0237)
- Wong L and Soli S. Development of the Cantonese hearing in noise test (CHINT). *Ear Hear* 2005; 26: 276–289. DOI: 10.1097/00003446-200506000-00004
- Nilsson M, Soli SD and Sullivan JA. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am* 1994; 95: 1085–1099. DOI: 10.1121/1.408469
- Wong LL and Soli SD. Development of the Cantonese Hearing in Noise Test (CHINT). *Ear Hear*. 2005; 26: 276–289. DOI: 10.1097/00003446-200506000-00004
- Perepelytsia V and Dellwo V. Acoustic compression in zoom audio does not compromise voice recognition

- performance. *Sci Rep.* 2023; 13: 18742. DOI: 10.1038/s41598-023-45971-x
19. Swanepoel DW, Mngemane S, Molemong S, et al. Hearing assessment—reliability, accuracy, and efficiency of automated audiometry. *Telemed J E Health.* 2010; 16: 557–563. DOI: 10.1089/tmj.2009.0143
 20. Choi JM, Lee HB, Park CS, et al. PC-based tele-audiometry. *Telemed J E Health* 2007; 13: 501–508. DOI: 10.1089/tmj.2007.0085
 21. Krumm M, Ribera J and Klich R. Providing basic hearing tests using remote computing technology. *J Telemed Telecare* 2007; 13: 406–410. DOI: 10.1258/135763307783064395
 22. Beutelmann R and Brand T. Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 2006; 120: 331–342. DOI: 10.1121/1.2202888
 23. Poissant SF, Whitmal III NA and Freyman RL. Effects of reverberation and masking on speech intelligibility in cochlear implant simulations. *J Acoust Soc Am* 2006; 119: 1606–1615. DOI: 10.1121/1.2168428
 24. Alzqhoul EA, Nair BB and Guillemin BJ. Impact of dynamic rate coding aspects of mobile phone networks on forensic voice comparison. *Sci Justice* 2015; 55: 363–374. DOI: 10.1016/j.scijus.2015.04.006
 25. Valin J, Vos K and Terriberry T. Definition of the opus audio codec. *RFC* 2012; 6716: 1–326.
 26. Martin N, Lalalde OM, Stokes C, et al. An evaluation of remote communication versus face-to-face in clinical dental education. *Br Dent J* 2012; 212: 277–282. DOI: 10.1038/sj.bdj.2012.226
 27. Chen Y. Speech audiometry. In: Hu XJ (eds) *Clinical audiology*. Hangzhou, China: Zhejiang University Press, 2024, pp.61.
 28. Chen Y and Wong LLN. Development of the Mandarin hearing in noise test for children. *Int J Audiol* 2020; 59: 707–712. DOI: 10.1080/14992027.2020.1750717
 29. Chen Y, Wong LLN, Zhu S, et al. Vocabulary development in Mandarin-speaking children with cochlear implants and its relationship with speech perception abilities. *Res Dev Disabil* 2017; 60: 243–255. DOI: 10.1016/j.ridd.2016.10.010
 30. Chen Y, Lena LNW, Zhu S, et al. A structural equation modeling approach to examining factors influencing outcomes with cochlear implant in Mandarin-speaking children. *PLoS One* 2015; 10: e0136576. DOI: 10.1371/journal.pone.0136576

List of abbreviations

ANOVA	Analysis of variance
CHINT	Cantonese Hearing in Noise Test
CM	Conventional method
CM1	First conventional method
CM2	Second conventional method
HA	Hearing aids
HRTFs	Head-related transfer functions
LP	Linear prediction
MDCT	Modified discrete cosine transform
NF	Noise from the front
NS	Noise from the side
RASA	Remote-access software administration
S/N	Signal-to-noise ratio
SRTs	Speech reception thresholds
VCSA	Video-conference software administration