

# Impact of Language Barriers and Age on English Hearing Test App Accuracy for Polish Users

Małgorzata Pastucha, MSc<sup>1</sup> , Elżbieta Gos, PhD<sup>2</sup> ,  
Henryk Skarżyński<sup>3</sup> , and W. Wiktor Jedrzejczak<sup>2</sup> 

## Abstract

**Objective.** The development of health applications (apps) includes those for testing hearing, although most of them are available only in English. This study investigates whether poor English language proficiency creates a barrier for Polish users in the accuracy of such an app in measuring self-determined hearing thresholds.

**Study Design.** The study compared hearing thresholds measured by an English-language app and a professionally conducted reference test, with attention to participants' English proficiency and age.

**Setting.** The English-language app "Hearing test, Audiogram," was used to determine hearing thresholds. A reference test was performed by an audiologist using specialized equipment.

**Methods.** Participants were 87 nonnative English language speakers aged 16 to 88. They were divided into 3 groups based on their proficiency in English: no knowledge (Group 1), basic (Group 2), and advanced (Group 3). The mean differences between hearing thresholds determined using the app and the reference tests were measured for each group.

**Results.** The accuracy of the results varied according to the level of English proficiency. A statistically significant difference was found between Group 1 (no knowledge) and Group 3 (advanced), with mean differences of 13.6, 9.3, and 6.7 dB for Groups 1, 2, and 3, respectively, meaning that discrepancies were larger in the less proficient groups. However, when participant age was considered, language proficiency was less important and was no longer a significant factor.

**Conclusion.** This study revealed that English language proficiency does affect the accuracy of mobile app-based hearing tests, but age of the user is also important.

## Keywords

English, hearing test, language skills, mobile app, reliability

Received September 13, 2024; accepted November 2, 2024.

OTO Open  
2024, Vol. 8(4):e70047  
© 2024 The Author(s). OTO Open  
published by Wiley Periodicals LLC  
on behalf of American Academy of  
Otolaryngology-Head and Neck  
Surgery Foundation.  
DOI: 10.1002/oto.2.70047  
<http://oto-open.org>

WILEY

The increasing prevalence of mobile applications (apps) has meant they are now extensively used in both everyday life and medical research.<sup>1</sup> Some apps are designed to test hearing, acting like an audiometric examination based on pure tones or speech.<sup>2</sup> Examples include "uHear," "Hearing Test," and "Hearing Test, Audiogram".<sup>3,4</sup> These apps allow users to monitor their own hearing status, which is useful in home settings or where access to specialist services is limited.<sup>5</sup>

One challenge with using apps to test hearing is a possible linguistic barrier. The majority of hearing test apps have been developed in English,<sup>2</sup> which can present difficulties for those who are not fluent in this language. Another factor that can limit the accuracy of app-based tests is the age of the user. Older people tend to have more problems with vision, motor coordination, and facility with mobile devices, and these may impair the accuracy of the results.<sup>6,7</sup>

Although mobile apps have great potential for self-monitoring hearing, it is still not clear whether non-English speaking users can operate them effectively. Studies to date have largely considered users who are fluent in English.<sup>8-10</sup> Consequently, further investigation is required to ascertain the impact of linguistic barriers on the accuracy of the results obtained.

The objective of this study was to ascertain whether Polish-speaking individuals can effectively use an English-language hearing assessment app, considering that Polish-language apps are rarely available. Additionally, the study aimed to determine whether the level of English proficiency affects the accuracy of results obtained using

<sup>1</sup>Department of Experimental Audiology, Institute of Physiology and Pathology of Hearing, World Hearing Center, Warsaw, Poland

<sup>2</sup>Department of Otorhinolaryngo surgery, Institute of Physiology and Pathology of Hearing, World Hearing Center, Warsaw, Poland

<sup>3</sup>Department of Teleaudiology and Screening, Institute of Physiology and Pathology of Hearing, World Hearing Center, Warsaw, Poland

## Corresponding Author:

Małgorzata Pastucha, MSc, Department of Experimental Audiology, Institute of Physiology and Pathology of Hearing, Mochnackiego 10 Street, Warsaw, Poland.

Email: [m.pastucha@ifps.org.pl](mailto:m.pastucha@ifps.org.pl)

mobile apps. The study did so by determining the differences between tests using the app and the results from a reference test conducted by an audiologist. It also looked at another factor, age, that might affect the accuracy of results.

## Methods

### Study Procedure

The study used 2 methods to assess hearing: first, a self-administered test via a mobile app and then a reference test conducted by an audiologist. The tests were always done in this order to mirror real-world conditions and circumvent any potential “learning effect”. Both tests were performed during a single session. Before the self-administered test, each participant received written instructions, in Polish, on how to complete the hearing test using the mobile app. The app was installed on the participant's mobile device. If they did not have an Android smartphone or headphones, the researcher provided the necessary equipment.

After completing the tests, participants filled out a survey that collected demographic data, including age and sex. The survey also included a closed-ended question about their English language proficiency. This question required participants to provide a subjective assessment of their own language skills, and, based on the responses, participants were categorized into 3 groups: no knowledge, basic knowledge, and advanced knowledge of English.

### Hearing Tests

#### Mobile App

The app “Hearing Test, Audiogram” was used for the study. It was chosen based on the following criteria: free access, available on Android, and ability to determine the hearing threshold separately for each ear. The app uses tonal stimuli for frequencies of 125, 250, 500, 1000, 2000, 3000, 4000, and 8000 Hz over a sound intensity range of 0 to 90 dB. During the test, the subject is presented with a continuous tone of increasing intensity, and the result is recorded upon pressing the smartphone screen.

The results are presented in the form of an audiogram for each ear separately. In accordance with the obtained results, the app automatically categorizes the degree of hearing loss and provides a concise description of it. The app enables the user to compare the results with the age standard.

In accordance with the instructions provided by the manufacturer, calibration of the app is not required. Testing was carried out on Android mobile devices using the associated headphones.

#### Reference Test

The reference test was conducted by an audiologist using the Sense Examination Platform and Sennheiser HDA200

headphones.<sup>11</sup> Hearing thresholds for 250, 500, 750, 1000, 1500, 2000, 3000, 4000, 6000, and 8000 Hz were determined using the modified Hughson-Westlake procedure.<sup>12</sup> Sounds were presented at levels from 0 to 80 dB HL, and the hearing threshold was determined when 2 of 3 correct responses were obtained at a given intensity.

### Study Sample

We performed a power analysis to determine the necessary sample size for our study, specifically for the analysis of variance (ANOVA) analysis. We assumed a large effect size (0.4), reflecting a strong relationship between the variables being studied. Based on this requirement, we calculated that a total sample size of  $N = 66$  participants was needed to adequately power the study. To ensure the robustness and reliability of the results, we recruited a larger sample of participants, exceeding the minimum requirement of 66. This allowed us to take into account possible exclusion of subjects while still retaining adequate power.

### Participants

The study was comprised of Polish volunteers from the community aged 16 years and older. All were non-native English language speakers. We excluded those who didn't have enough free time, didn't want to install the app on their mobile device, or were uncooperative. A total of 91 individuals were subjected to testing, of whom 4 were later excluded, resulting in a final group of 87 participants. The study group comprised 58 women and 29 men, with ages ranging from 16 to 88 years (mean age 48.7 years,  $SD = 20.2$ ). Participants were divided into 3 groups based on their proficiency in English: Group 1, no knowledge (43 participants); Group 2, basic (24 participants); and Group 3, advanced (20 participants).

### Ethical Considerations

Prior to the study, all subjects were informed about its nature. Informed consent was obtained from all participants. The study was conducted in accordance with the guidelines of the Declaration of Helsinki and was approved by the Ethics Committee of the Institute of Physiology and Pathology of Hearing (consent no. KB.IFPS 3/2023).

### Statistical Analysis

In order to evaluate accuracy of the results, the absolute discrepancies between the hearing thresholds determined with the app and the results of the reference test were calculated for each of the 7 common frequencies: 250, 500, 1000, 2000, 3000, 4000, and 8000 Hz. From these differences, an overall discrepancy index was calculated for both ears combined, representing the average difference in decibels. Smaller values between the results

obtained from the app and the reference test indicate a more accurate test.

A one-way ANOVA was used to assess differences between groups of participants, by level of English proficiency. This was followed by an analysis of covariance (ANCOVA) in which the age of participants was included as a covariate. A significance level of  $P < .05$  was adopted. Analyses were performed using IBM SPSS Statistics software (version 24).

## Results

The largest discrepancy between the results of the self-administered hearing test and the reference hearing test was observed in the group with no knowledge of English. Discrepancies ranged from 4.3 to 61.1 dB, with an average of 13.6 dB (SD = 11.0). In contrast, the group with basic proficiency in English showed smaller discrepancies, ranging from 4.3 to 18.1 dB, with an average of 9.3 dB (SD = 3.8). The most consistent results were found in the group with advanced proficiency in English, where discrepancies ranged from 1.8 to 21.1 dB, with an average of 6.7 dB (SD = 4.4).

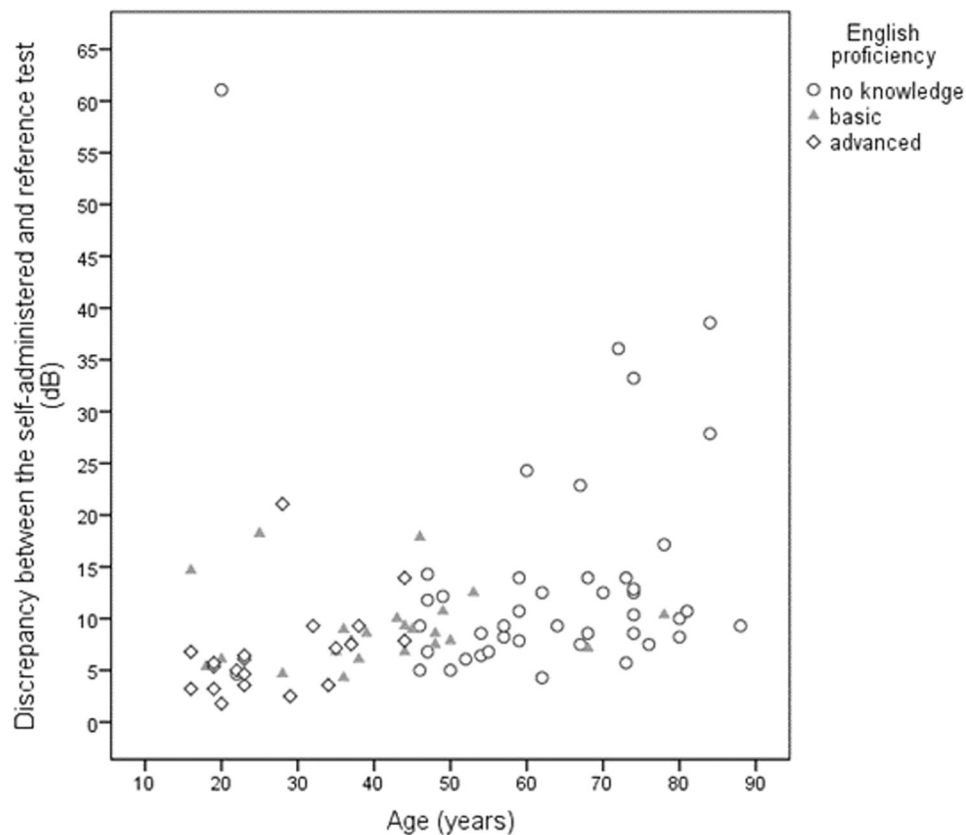
The discrepancy between the results of the self-administered and reference hearing tests across the 3 groups was compared using an ANOVA test. The effect of English proficiency was statistically significant:  $F(2, 84) = 5.31$ ;  $P = .007$ ;  $\eta^2 = 0.11$ , and a post hoc test revealed that the

difference between the first and third groups was statistically significant ( $P = .008$ ). This indicates that the results obtained by subjects with no knowledge of English were significantly less consistent than those obtained by subjects with advanced proficiency in English.

In addition, we decided to test whether or not the significant effect of English proficiency was confounded by participant age. The results of the ANCOVA showed that this was indeed the case. The effect of English proficiency was not statistically significant after controlling for the effect of participant age,  $F(2, 83) = 2.06$ ;  $P = .133$ . Previously found differences between groups became statistically nonsignificant. This means that when participants with different levels of English proficiency are of similar age, they have a similar discrepancy between the results of the self-administered and reference hearing tests.

**Figure 1** shows the relationship between age, English proficiency and discrepancy between the results of the self-administered and reference hearing tests.

The correlation between English proficiency and discrepancy between the results of the self-administered and reference hearing tests English proficiency was positive and moderate,  $\rho = 0.41$  ( $P < .001$ ). The correlation between age and discrepancy between the results of the self-administered and reference hearing tests was also positive but slightly stronger,  $\rho = 0.49$  ( $P < .001$ ). It



**Figure 1.** Relationship between age, English proficiency and discrepancy between the results of the self-administered and reference hearing tests.

shows that as age increases, the discrepancy between the results of both tests also tends to increase.

## Discussion

The findings of the present study suggest that while English proficiency does have some effect on the accuracy of mobile hearing test apps, this factor is not as crucial as initially assumed. The results indicate that participants with advanced English proficiency showed slightly more consistent results (about 5 dB compared to those with no knowledge of English); however, this effect diminished when the age of the participants was taken into account. Thus, age also appears to be a significant factor in determining the accuracy of these measurements. Although English proficiency may contribute to more accurate self-administered hearing test results, it is not the primary determinant when age is considered. This highlights the importance of accounting for age when interpreting the outcomes of mobile hearing tests, rather than emphasizing the role of language proficiency alone.

There are many studies available in the literature that focus on evaluating the sensitivity and specificity of hearing screening tests using mobile apps. However, the majority of these studies were conducted in controlled settings, with the tests being administered by the researcher or participants after prior training<sup>13–17</sup> or comprehensive familiarization with the instructions.<sup>9,18–20</sup> Such approaches do not necessarily reflect the realities of using mobile apps in everyday life. For example, Whitton et al<sup>21</sup> applied a different methodology, which led to different results, finding that neither participant age nor ownership of a tablet significantly impacted the accuracy of hearing tests performed using a mobile app in a home setting.

Moreover, in a number of studies on mobile hearing test apps, lack of English language skills was an exclusion criterion<sup>8–10</sup> or the tests were conducted with the help of a translator.<sup>18</sup> It is possible that both these approaches may have affected the accuracy of the results obtained, although this has not been sufficiently addressed in the literature. To the best of our knowledge, the present study is one of the first to investigate the effect of language competence and age on the accuracy of results obtained during self-testing of hearing using a mobile app.

The impact of linguistic competence has only been considered in studies that have used verbal tests. The results show that non-native speakers perform worse compared to speakers that are fluent in English.<sup>22,23</sup> When considering English proficiency in terms of the ability to independently determine hearing thresholds, it is relevant to examine the work of Rourke and colleagues,<sup>24</sup> who showed that children were able to intuitively understand and perform tests despite a lack of language proficiency. However, Rourke and colleagues did not assess the extent to which this ability translated into test accuracy. Likewise, a limitation of our study is that we

did not objectively verify participants' English proficiency, relying instead on self-reported language skills, and this could have introduced some bias.

Our findings indicate that difficulties in understanding commands from apps, and consequently the accuracy of the results, are more closely associated with the age of users than with their linguistic proficiency. Older users often face manual difficulties, such as impaired precision of movements, reduced motor skills, or impaired vision, which may hinder the use of touch interfaces and affect the correct use of mobile devices.<sup>25–27</sup> Such issues may lead to greater discrepancies between app-based scores and reference scores, which could ultimately result in lower accuracy.

## Conclusion

Although English proficiency affects the accuracy of results in mobile hearing tests, our findings indicate that age is also a decisive factor, diminishing the importance of language skills. Therefore, when selecting a hearing test app, it is essential to consider the specific characteristics of the user. Particularly, the interface should be adapted to meet the needs of older individuals by ensuring it is intuitive and easy to use. This adaptation may be more crucial than ensuring the app's language compatibility with the user's native language.

## Acknowledgments

The authors thank Andrew Bell for comments on an earlier version of this article.

## Author Contributions


**Małgorzata Pastucha**, conceptualization, data curation, investigation, methodology, writing; **Elżbieta Gos**, conceptualization, formal analysis, investigation, methodology, writing; **Henryk Skarżyński**, conceptualization, investigation, supervision; **W. Wiktor Jedrzejczak**, conceptualization, investigation, methodology, supervision, writing.


## Disclosures


**Competing interests:** None.


**Funding source:** None.

## ORCID iD

Małgorzata Pastucha  <http://orcid.org/0000-0002-4866-1866>

Elżbieta Gos  <https://orcid.org/0000-0003-3173-3867>

Henryk Skarżyński  <https://orcid.org/0000-0001-7141-9851>

W. Wiktor Jedrzejczak  <https://orcid.org/0000-0001-8404-0672>

## References

- Sezgin E. Can We use commercial mobile apps instead of research mobile apps in healthcare research? *Front Public Health*. 2021;9. doi:10.3389/fpubh.2021.685439
- Almufarrij I, Dillon H, Dawes P, et al. Web- and app-based tools for remote hearing assessment: a scoping review. *Int J Audiol*. 2023;62(8):699-712. doi:10.1080/14992027.2022.2075798

3. Irace AL, Sharma RK, Reed NS, Golub JS. Smartphone-based applications to detect hearing loss: a review of current technology. *J Am Geriatr Soc.* 2021;69(2):307-316. doi:10.1111/jgs.16985
4. Oremule B, Abbas J, Saunders G, et al. Mobile audiometry for hearing threshold assessment: a systematic review and meta-analysis. *Clin Otolaryngol.* 2024;49(1):74-86. doi:10.1111/coa.14107
5. Berampu RW, Adriztina I, Sofyan F, Machrina Y, Adenin I. Accuracy and pitfalls in the smartphone-based audiometry examination. *Iran J Otorhinolaryngol.* 2024;36(2):421-431. doi:10.22038/IJORL.2024.71187.3462
6. Chien CH, Tu TY, Chien SF, et al. Relationship between mandarin speech reception thresholds and pure-tone thresholds in the geriatric population. *J Formos Med Assoc.* 2006;105(10):832-838. doi:10.1016/S0929-6646(09)60270-9
7. Kobewka D, Ruller S, Daly B, et al. Usability, acceptability and clinical utility of a mobile app to screen for hearing loss in older adults in a geriatric rehabilitation hospital. *J Eval Clin Pract.* 2023;29(2):392-396. doi:10.1111/jep.13795
8. Abu-Ghanem S, Handzel O, Ness L, Ben-Artzi-Blima M, Fait-Ghelbendorf K, Himmelfarb M. Smartphone-based audiometric test for screening hearing loss in the elderly. *Eur Arch Otrhinolaryngol.* 2016;273(2):333-339. doi:10.1007/s00405-015-3533-9
9. Margolis RH, Frisina R, Walton JP. AMTAS®: automated method for testing auditory sensitivity: II. Air conduction audiograms in children and adults. *Int J Audiol.* 2011;50(7):434-439. doi:10.3109/14992027.2011.553206
10. Szudek J, Ostevik A, Dziegielewski P, et al. Can Uhear me now? Validation of an iPod-based hearing loss screening test. *J Otolaryngol Head Neck Surg.* 2012;41(Suppl 1):S78-S84.
11. Skarzynski PH, Kochanek K, Skarzynski H, et al. Hearing screening program in school-age children in Western Poland. *J Int Adv Otol.* 2011;7(2):194-200.
12. Carhart R, Jerger JF. Preferred method for clinical determination of pure-tone thresholds. *J Speech Hear Disord.* 1959;24(4):330-345. doi:10.1044/jshd.2404.330
13. Masalski M, Kręcicki T. Self-test web-based pure-tone audiometry: validity evaluation and measurement error analysis. *J Med Internet Res.* 2013;15(4):e71. doi:10.2196/jmir.2222
14. Larrosa F, Rama-Lopez J, Benitez J, et al. Development and evaluation of an audiology app for iPhone/iPad mobile devices. *Acta Otolaryngol.* 2015;135(11):1119-1127. doi:10.3109/00016489.2015.1063786
15. Yousuf Hussein S, Wet Swanepoel D, Biagio de Jager L, Myburgh HC, Eikelboom RH, Hugo J. Smartphone hearing screening in mHealth assisted community-based primary care. *J Telemed Telecare.* 2016;22(7):405-412. doi:10.1177/1357633X15610721
16. Corry M, Sanders M, Searchfield GD. The accuracy and reliability of an app-based audiometer using consumer headphones: pure tone audiometry in a normal hearing group. *Int J Audiol.* 2017;56(9):706-710. doi:10.1080/14992027.2017.1321791
17. Mahomed-Asmail F, Swanepoel DW, Eikelboom RH, Myburgh HC, Hall J. Clinical validity of hearScreen™ smartphone hearing screening for school children. *Ear Hear.* 2016;37(1):e11-e17. doi:10.1097/AUD.0000000000000223
18. Sandström J, Swanepoel D, Laurent C, Umeffjord G, Lundberg T. Accuracy and reliability of smartphone self-test audiometry in community clinics in low income settings: a comparative study. *Ann Otol, Rhinol, Laryngol.* 2020;129(6):578-584. doi:10.1177/0003489420902162
19. Manganella JL, Stiles DJ, Kawai K, Barrett DL, O'Brien LB, Kenna MA. Validation of a portable hearing assessment tool: Agilis Health Mobile Audiogram. *Int J Pediatr Otorhinolaryngol.* 2018;113:94-98. doi:10.1016/j.ijporl.2018.04.010
20. Lin HYH, Chu YC, Lai YH, et al. A Smartphone-based approach to screening for sudden sensorineural hearing loss: cross-sectional validity study. *JMIR Mhealth Uhealth.* 2020;8(11):e23047. doi:10.2196/23047
21. Whitton JP, Hancock KE, Shannon JM, Polley DB. Validation of a self-administered audiometry application: an equivalence study. *Laryngoscope.* 2016;126(10):2382-2388. doi:10.1002/lary.25988
22. De Sousa KC, Swanepoel DW, Moore DR, Smits C. A Smartphone national hearing test: performance and characteristics of users. *Am J Audiol.* 2018;27(3 Suppl):448-454. doi:10.1044/2018\_AJA-IMIA3-18-0016
23. Potgieter JM, Swanepoel DW, Myburgh HC, Smits C. The South African English smartphone digits-in-noise hearing test: effect of age, hearing loss, and speaking competence. *Ear Hear.* 2018;39(4):656-663. doi:10.1097/AUD.0000000000000522
24. Rourke R, Kong DCC, Bromwich M. Tablet audiometry in Canada's North: a portable and efficient method for hearing screening. *Otolaryngol Head Neck Surg.* 2016;155(3):473-478. doi:10.1177/0194599816644407
25. Cheong Y, Shehab RL, Ling C. Effects of age and psychomotor ability on kinematics of mouse-mediated aiming movement. *Ergonomics.* 2013;56(6):1006-1020. doi:10.1080/00140139.2013.781682
26. Wu HC. Electronic paper display preferred viewing distance and character size for different age groups. *Ergonomics.* 2011;54(9):806-814. doi:10.1080/00140139.2011.600775
27. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). *Ergonomics.* 2014;57(5):635-652. doi:10.1080/00140139.2014.895855