**Abstract**

**Introduction.** Hearing is essential for our interaction and communication, and hearing loss can

negatively affect social life. Speech audiometry is essential for assessing speech perception and

there are various methods including speech in noise, but traditional methods have limitations.

Automated speech recognition (ASR) systems have the potential to streamline these tests.

Some studies show that ASR technology is as reliable as traditional methods and can improve

the accessibility of hearing loss diagnostics. Challenges for audiologists include language

barriers and the ability to accurately assess test results. A speech-to-text system can potentially

support audiologists by providing more accurate and consistent results in speech audiometry.

**Purpose.** This study was conducted to evaluate the agreement between a transcription

application and the audiologist's assessment of speech perception tests in noise and to

investigate whether the transcription application can be an aid to the audiologist.

**Method.** The study contained 17 participants, of which 16 had normal hearing and 1 had

moderate to severe hearing loss. All spoke Swedish as their first language. Subjects performed

bilateral tone audiometry to determine the appropriate sound level for the speech perception

test. The speech perception test was performed with background noise and used six different

lists of numbers that were randomly selected for each participant. In speech perception tests,

we used a transcription application that converts speech to text to assess speech perception.

The results from the application were compared statistically with the assessments performed

by an audiologist.

**Results.** The results of descriptive analysis showed that the mean scores for audiology

assessment were higher compared to application assessment, but this difference is not a

clinically relevant difference. Although the mean was higher for the audiology results, the

standard deviation for both tests was similar, indicating comparable variability in the results

for the two tests. The scatterplot showed a positive correlation between the audiologist's and

the application's assessment. Furthermore, the paired t-test showed a significant difference

between the audiologist's and the application's results, which was supported by the t-value (t

(33) = 4.8) and the strong p-value (p < 0.001), but this difference was not a clinically relevant

difference. The mean difference between the audiologist's and the application's results was

2.3 %-points, with a confidence interval on 95 %, between 1.4 and 3.3 %-points.

**Conclusion.** The study showed that the audiologist's assessment was slightly better than the

transcription application's assessment. Despite a small difference in mean and a strong positive

correlation, the statistical paired t-test showed that the difference was significant. This

indicated that the audiologist's assessment was preferred for assessing speech perception

compared to the transcription application.

**Keywords**

Automatic Speech Recognition (ASR), Interrater reliability, Phonemic balance (PB) lists,

Speech in noise, Speech perception, Speech-to-text (STT)