ARIG lightning up-date number 2

Speech to text transcription apps - what do we know and what relevance do they have for those with hearing loss?

Welcome to the second in our lightning up-date series on interesting topics in adult rehabilitation. This time we review some of the evidence about the effectiveness of transcription apps and hopefully spark some interest and debate about their value for people with hearing loss.

Speech to text apps - What are they and why are they important?

For many people with hearing loss, relying on sound alone makes communicating problematic (effortful, frustrating, tiring) or even impossible. Now that mobile phone use is ubiquitous and the accuracy of automatic speech recognition is improving, the use of speech to text transcription (STT) apps that facilitate instantaneous text support of speech are becoming a viable option for routine use by people with hearing loss.

Are STT apps effective?

Several studies have looked at app transcription accuracy as compared to the gold standard of stenography. Stenography proved the most accurate (5-7% error rates), although STT apps performed pretty well, with error rates of between 15 and 37% depending on the app used (Roychowdhury et al 2023). These findings remained true even when N95 masks were being worn and medical terminology was used. Unsurprisingly, the accuracy of STT apps varies under different acoustic conditions. They perform better with slower speech (Lyall et al 2016) and in quiet (Pragt et al 2022).

STT transcription is faster than writing or typing, and according to the findings of Lyall et al. (2016), just as accurate. This is worth considering in healthcare settings, where information on care is of high importance and time is often pressured.

Whilst many apps perform well with spoken English, no formal studies have tested how well they perform with speakers who have strong accents or with non-native speakers. It is likely that accuracy will be reduced, and that some apps will cope better than others with particular accents or dialects – although again this has not been tested. Ideally, apps should have the capacity to 'learn' an accent with time. Otter Ai claims to handle a variety of accents ([Otter.ai - Voice Meeting Notes & Real-time Transcription](#)) but this does not appear to have been robustly compared against other apps.

It seems that currently apps are primarily trained with the English Language as seen from the work of Pragt (2022) who assessed word error rate in transcription in both English and Dutch for a variety of commonly used apps and found that there were fewer errors with the English dialogue than the Dutch dialogue. The authors concluded this was due to the way the language was implemented within the apps and reflected the training data used in its production.

 A limited number of STT apps have been formally evaluated. Among those that have, there were varying levels of accuracy. Roychowdhury et al (2023) examined 3 apps, two of which were developed specifically for the hearing impaired (Ava 24/7 and Live Transcribe), and one that was developed for dictated note taking (iPhone Dictation). The former two are paid apps, while the latter is available by default on all iPhones. They found that Ava 24/7 and iPhone Dictation made fewer transcription errors than Live Transcribe.

Do STT apps provide benefit to hearing impaired patients?

Pragt et al (2022) assessed the performance of four STT apps (AVA, Earfy, Live Transcribe, and Speechy) using standard clinical speech tests in both quiet and noise. Their conclusion was that performance in quiet was equivalent to the average expected results of an individual with moderate hearing loss. In noise, the apps performed at a level comparable to CI recipients, and, Live Transcribe performed best when performance across all measures and conditions was considered. However none of the apps performed as well as a normal hearing listener, suggesting that currently STT alone is insufficient for easy communication. However, it would be a helpful support tool.

No studies have looked at the indirect benefits of STT, such as increased confidence in communication or greater willingness to engage in activities whether social, educational or health-related. This may be particularly relevant in important conversations such as medical appointments where clear understanding is vital and people with hearing loss are often anxious about missing or misunderstanding information (Saunders & Oliver, 2022).

Are there downsides to STT apps?

A possible downside to using a STT app in a social situation is that the user might miss out on lipreading, facial cues and body language because they are focused on the screen of their mobile phone. No studies to date have examined this. Relative to stenography, a further disadvantage of STT apps is that they are currently cannot differentiate between speakers, which a stenographer can of course do very well.

Are STT apps cost effective?

In quiet situations, performance approaches the gold standard of stenography - at significantly reduced cost and without the need for pre-scheduling of a stenographer. Many apps are free or have a small monthly charge (Roychowdhury 2023), making them highly cost-effective relative to stenography. However, for important situations or more complex listening environments, the accuracy of stenography may over-ride considerations of cost.

Summary and conclusions

Whilst the accuracy of STT apps varies, the best performing apps approached the accuracy of stenography in quiet environments, whether or not a mask was being worn. Their performance in noise, however, diminished. Factors such as accents, specialised vocabulary, and the implications of their use for psychosocial and non-auditory aspects of communication have not been investigated. Nonetheless, in light of the fact that many are free to use, it seems appropriate to make sure that people with hearing loss are aware of their existence.

References

Lyall F C, Clamp P J, Hajioff D. Smartphone speech-to-text applications for communication with profoundly deaf patients J Laryngol Otol. 2016 Jan;130(1):104-6. doi: 10.1017/S0022215115003151. Epub 2015 Nov 27.

Pragt L, van Hengel P, Grob D, Wasmann J A. Preliminary Evaluation of Automated Speech Recognition Apps for the Hearing Impaired and Deaf. Front Digit Health. 2022 Feb 16;4:806076. doi: 10.3389/fdgth.2022.806076. eCollection 2022.

Roychowdhury P., Dhrumi GandhI M, Knoll R M, Wu M J, Kozin E D, Remenschneider A K. Evaluating the accuracy of speech to text applications for cochlear implant candidates during COVID-19 . Cochlear Implants Int. 2023 Jan;24(1):1-5. doi: 10.1080/14670100.2022.2120450. Epub 2022 Sep 23.

Saunders GH, Oliver F. Impact of Hearing Loss on Communication During Remote Health Care Encounters. Telemed J E Health. 2022 Sep;28(9):1350-1358. doi: 10.1089/tmj.2021.0490. Epub 2022 Jan 11. PMID: 35020484.