Canadian Audiologist
The Official Publication of the Canadian Academy of Audiology

Using Speech-to-Text Automatic Speech Recognition Software to Improve Accessibility in Audiology Practice

Published February 27th, 2021

Janine Verge, AuD, Aud(C) - Reg
Nikita Crowell, BA, MSc



Communication Barriers in Audiology Clinics

As audiologists, it is essential to make our services as accessible as possible to the clients we serve. In-person communication barriers can exist during all points of contact during a hearing test appointment, including registration, case history taking, testing, and counseling. Examples of at-risk populations that may require additional communication strategies might include:

- People living with hearing loss with no hearing aids.
- People with or without hearing aids living with moderate to poor word discrimination.
- People living with auditory neuropathy or auditory processing disorders.
- People who have poor speech reading skills.
- People with sudden hearing losses and/or cochlear implant candidates.

Communicating with clients on the other side of the audiometric test booth has always been a

unique challenge for audiologists; the setup for audiometric booths often involves suboptimal lighting, and double-glazed windows can reduce or completely block lip or speech reading making clients rely solely on auditory information for communication.

Other more recent communication barriers are face masks, plexiglass, and social distancing that have been necessary due to the COVID-19 pandemic. These have created communication barriers for everyone, especially for people living with hearing loss.[1–3] A recent article, written by hearing health advocate Shari Eberts, in the *Canadian Audiologist* provides an excellent review of the current day-to-day challenges of living with hearing loss in the time of COVID-19: https://canadianaudiologist.ca/issue/volume-8-issue-1-2021/column/acessibility-issues/

Due to these recent and on-going barriers, both audiologists and their clients would benefit from a means of supplementing verbal communication during hearing assessments. One solution is using automatic speech recognition software, such as Otter ai, to provide accurate speech-to-text in real-time during a hearing assessment to reduce barriers.

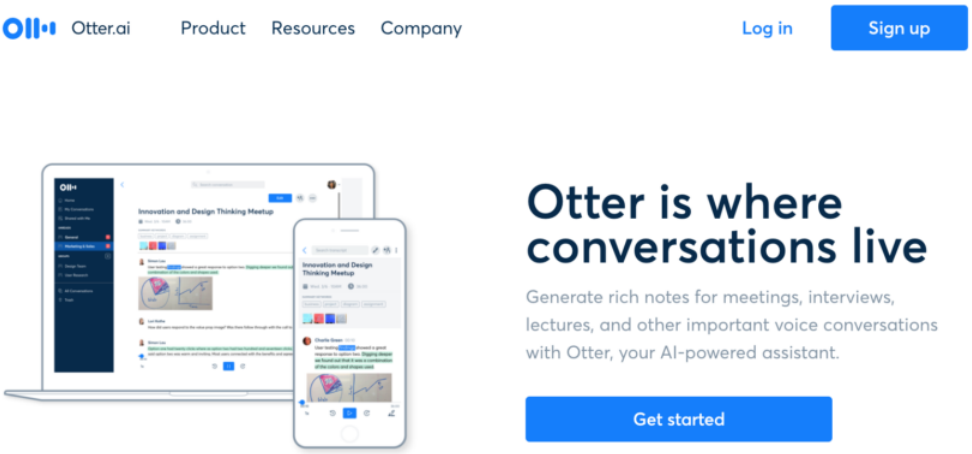
Automatic Speech Recognition Tools

Automatic speech recognition (ASR) is by no means new technology, but historically, it has not been accurate or efficient enough to become a staple in healthcare practices. Thanks to the exponential technological improvement led by tools like Siri and Cortana, voice recognition has become far more effective in the past decade and has become a commonly used tool for everything from virtual meetings to controlling "Smart Home" software through voice commands (e.g., Google Home, Alexa).

In the context of an audiological assessment, ASR software can be used synchronously to transcribe what the audiologist is saying in real-time. In doing so, the client will be able to augment the auditory conversation with written text. Real-time speech-to-text offers the chance to have longer, more naturally flowing conversations in a timelier fashion than writing with pen and paper.

In addition to facilitating conversation and instruction over the course of the appointment, research suggests that supplementing the spoken information delivered by a health professional with synchronous captioning can help the patient better recall the details of the information and improve understanding in noise.[4,5]

Using Otter.ai as a Speech-to-Text Tool



There are several speech-to-text tools available, each with its own strengths and weaknesses. One such tool is *Otter.ai* (https://otter.ai/), which is available as a smartphone/tablet app as well as online via desktop computer. The authors of this paper chose to focus on *Otter.ai* for two main reasons: (1) *Otter.ai*'s transcriptions are well known for accuracy and automatically inserts punctuation as well as breaks between phrases and different speakers, and (2) once you have created an account, you can record 600 minutes of transcriptions per month for free, all of which are stored separately on your account and accessible afterward.

Limitations to ASR Speech-to-text Software Applications

The main limitations of using automatic speech recognition tools are the need to use an internet connection and a software pack or a monthly subscription. There are nonetheless, many apps and programs available for free, and while no automatic speech recognition tool should be expected to have perfect accuracy, the technology will continue improving as the popularity of such devices grows. Another issue that arises when using specific automatic speech recognition tools is the risk of breaching privacy and confidentiality by allowing third-party programs to record the conversation when conversing with your clients. Thus, the audiologist needs to familiarize themselves with the privacy policy of whichever tool they intend to use, limit the use of identifying information during the conversation, and crucially seek informed consent from the client before beginning the recorded conversation. Lastly, it should also be recognized that ASR tools can be limiting for clients with low literacy skills or vision impairment.

What you need to get started to Use Otter ai in Audiology Practice

First, you will need a *computer screen* visible to the client (inside or outside of the audio booth, or both, depending on your needs). If you are seated next to the client, a smartphone or tablet/iPad could be used.

Second, you will need a *microphone connected directly to the computer*(the program may not recognize a mic plugged only into the audiometer). Other aspects to consider are:

- Microphones connected via AUX (the standard stereo input/output; often labelled with a graphic of headphones or a microphone) or USB should both work.
- Although not ideal for speech-to-text, Table microphones will work and may not achieve the same level of accuracy as a headset or clip-on microphone.

- If using a smartphone or tablet/iPad, the built-in mic will likely suffice (although a headset microphone may still provide better accuracy).

Third, when using Otter ai:

- It requires an internet connection.
- An account creation & sign-in is required.
- It is free to use for 600min per month (Otter Pro = $100 USD yearly; Otter Business = $720 USD yearly at the time of this publication).

Step-by-step Instructions for using Otter ai:

1. Navigate to Otter ai (https://otter.ai/).
2. Sign-in (or create an account if you have not yet done so).
3. Ensure you have a microphone connected to your computer. You can click the record button on Otter ai to verify. A pop-up reading "could not start audio source" will appear if there is no microphone connected.
4. To begin transcribing your conversation, hit the record button in Otter ai's *Home* tab. This will open a new tab that will show your speech-to-text transcription. Near the bottom of the screen, you will see a pause button and a stop button:
    - Hitting *pause* will stop recording your conversation until you hit the *resume recording* button that replaces the pause. When you resume recording, the transcription will begin in a new paragraph.
    - Hitting stop will end that transcription session and store the conversation.
    - If you find the transcription's text too small to read easily, you can zoom in by pressing and holding "*ctrl"* and either scrolling using the mouse wheel or hitting the "+" on the keyboard.
5. You can access your transcriptions later by navigating to the *My Conversations* tab in the Otter ai sidebar. Your conversations are arranged by time of recording (most recent at the top) and show the length of the recording and a list of key terms used. Click on the desired "note" to open that transcription. You can then change the title (click on "note" and enter your title), as well as edit the transcript itself. You can insert paragraph breaks, type in additional lines, remove unnecessary text, and change any errors. You can also copy/paste or share your transcription. This can be useful if you wish to provide the client with notes on the session regarding what you addressed in the session before they leave by printing them out.
6. Another useful feature is the "*Search conversations"* bar found at the top right of Otter ai. This will search all of your conversations for whatever term or phrase you enter and show all possible locations it occurred with the surrounding text.
7. Past conversations can be deleted or moved to different folders by clicking on the three dots next to each note in the *My Conversations* tab. Folders can be accessed and created in the sidebar (the 6th tab down).

General Tips for using Speech-to-Text Applications in Audiology Practice with your Clients

- Always obtain informed consent to use speech-to-text applications.

- To avoid privacy issues, avoid using identifying names/info during the conversation.
- Keep up to date with your chosen speech-to-text program's privacy policy.
- Many people could benefit from speech-to-text. Consider using it as a standard instead of only when a client is in crisis.
- Speak at a normal pace with clear enunciation.
- Ensure that the font size is large enough for the client to read easily.
- Monitor the transcription while you are speaking to check for errors, so you can clarify as needed.
- Pay attention to the client's facial expressions/body language to gauge their understanding.
- You may need to slow down or stop talking to allow the client to catch up in reading what has been said or review and consider the recent information more deeply.
- Ask the client directly for feedback as you go (e.g., is the transcription working well, would slowing down help, would they prefer an alternative means of communication).
- Be sure to turn off your transcription during speech testing (SRT, WRS) when the computer is in the audiometric booth.

Conclusion

Automatic Speech Recognition Software applications, such as Otter ai, should be considered as another communication tool audiologists can use during clinical practice. It will reduce communication barriers due to facial masks, social distancing, physical barriers (such as audiometric booths), and the personal communication barriers our clients are receiving our services for in the first place.

References

1. Chodosh J, Weinstein B, and Blustein, J. Face masks can be devastating for people with hearing loss. BMJ : British Medical Journal (Online) 2020;370:M2683.
2. Goldin A, Weinstein BE, Shiman N. How do medical masks degrade speech reception? Hearing Review. May 2020. Available at: https://www.hearingreview.com/hearing-loss/health-wellness/how-do-medical-masks-degrade-speech-reception
3. Ten Hulzen R and Fabry D. Impact of Hearing Loss and Universal Face Masking in the COVID-19 Era. Mayo Clin Proc 2020;95(10):2069–72.
4. Krull VE and Humes L. Text as a supplement to speech in young and older adults. Ear Hear 2016;37(2):164–76.
5. Spehar BJ, Tye-Murray N, Myerson J, and Murray D. Real-time captioning for improving informed consent: patient and physician benefits. Region Anesth Pain Med 2016;41(1):65–68.