Evaluating the Accuracy of Speech to Text Applications for Cochlear Implant Candidates

Prithwijit Roychowdhury, BS; Melissa Castillo-Bustamante, MD; Dhrumi Gandhi, MS; Renata Knoll, MD; Matthew J. Wu, BS; Elliott Kozin, MD; Aaron Remenschneider, MD MPH
Disclosures

• We have no conflicts or disclosures
Background

Impact of Hearing Loss and Universal Face Masking in the COVID-19 Era

Richard D. Ten Hulzen, MD, and David A. Fabry, PhD
Background

Face masks degrade high-frequency speech sounds

![Graph showing the degradation of high-frequency speech sounds with different types of masks]

- Surgical mask
- Cotton jersey
- N95 respirator
- Cotton/spandex blend jersey
- PVC window
- Plastic shield

 go.illinois.edu/MaskAcoustics

(Corey 2020)
Effective Communication Recommendations

Computer-Assisted Real-Time Transcription (CART)

Pandemic Recommendations

“We urge medical professionals to familiarize themselves with backup communications methods in the event traditionally accessible services are unavailable or delayed during the COVID-19 crisis.”

Speech to Text (STT) Applications
Background

How accurate are some common STT applications for iOS when compared to CART?

Does wearing an N95, using medical terminology, or being a non-native English speaker make a difference in accuracy?
Methods: Selected Passages

Normal Speech Patterns

THE RAINBOW PASSAGE

When the sunlight strikes raindrops in the air, they act as a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

Medical Terminology

Cochlear Implant Surgery Consent

A cochlear implant is a device that restores partial hearing to patients with severe to profound hearing loss who do not observe a benefit with a hearing aid. The device is surgically implanted in the inner ear and activated by a receiver worn outside the ear. It does not make sound louder or clearer. Instead, the device bypasses damaged parts of the auditory system and directly stimulates the nerve of hearing allowing individuals who are profoundly hearing impaired to receive sound.

https://oregonent.com/audiology/hearing-aids/cochlear-implants-mapping/

https://medium.com/@fofaflowers02/listening-saves-lives-how-acoustic-speech-analysis-can-indicate-heart-failure-sofia-flores-b4be6adcc919
Methods: Transcript Recording
Methods: Example Accuracy Assessment

Levenshtein Distance used to assess the difference between the transcript and the original text

Cl Consent transcript via CART service by native English speaker without N95: 2.30% error rate

A COCHLEAR IMPLANT IS A DEVICE THAT RESTORES PARTIAL HEARING TO PATIENTS WITH SEVERE TO PROFOUND HEARING LOSS WHO DO NOT OBSERVE A BENEFIT WITH A HEARING AID. THE DEVICE IS SURGICALLY IMPLANTED IN THE INNER EAR AND ACTIVATED BY A RECEIVER WORN OUTSIDE THE EARS. IT DOES NOT MAKE SOUND LOUDER OR CLEARER, INSTEAD THE DEVICE Bypasses DAMAGED PARTS OF THE AUDITORY SYSTEM AND DIRECTLY STIMULATES THE NERVE OF HEARING ALLOWING INDIVIDUALS WHO ARE PROFUNDLY HEARING IMPAIRED TO RECEIVE SOUND. TO HRETURN NORMAL HEARING AND BENEFITS VARY FROM ONE INDIVIDUAL TO ANOTHER MOST ADULT COCHLEAR IMPLANT PATIENTS NOTICE AN IMMEDIATE IMPROVEMENT IN THEIR COMMUNICATION SKILLS. OTOLARYNGOLOGISTS ARE EAR AND THROAT SPECIALISTS WHO PERFORM IMPLANT SURGERY. YOUR PRIMARY CARE PHYSICIAN CAN REFER YOU TO AN IMPLANT CLINIC FOR AN EVALUATION. THE IMPLANT TEAM INCLUDES AN OTOLARYNGOLOGIST, AUDIOLoGIST NURSE AND OTHERS WHO WILL DETERMINE YOUR CANDIDACY FOR A COCHLEAR IMPLANT AND REVIEW THE APPROPRIATE EXPECTATIONS AS A RESULT OF THE COCHLEAR IMPLANT FOR YOUR SPECIFIC CASE. COCHLEAR IMPLANT SURGERY IS USUALLY PERFORMED AS AN OUTPATIENT PROCEDURE UNDER GENERAL ANESTHESIA. AN INCISION IS MADE BEHIND THE EAR TO OPEN THE MASTOID BONE LEADING TO THE MIDDLE EAR SPACE. ONCE THE MIDDLE EAR SPACE IS EXPOSED AN OPENING IS MADE IN THE COCHLEA AND THE IMPLANT ELECTRODES ARE INSERTED. THE ELECTRONIC DEVICE AT THE BASE OF THE ELECTRODE ARRAY IS THEN PLACED BEHIND THE EAR UNDER THE SKIN. THE RISKS OF COCHLEAR IMPLANT SURGERY INCLUDE FAILURE TO IMPROVE HEARING INFECTION, TYPANIC MEMBRANE PERFORATION DAMAGE TO THE FACIAL NERVE, WORSENING OF TINNITUS OR BALANCE CHANGES IN TASTE BLEEDING INFECTION NEED FOR FURTHER PROCEDURES Meningitis and Spinal Fluid LEak. THESE RISKS CAN OCCUR AT THE TIME OF THE SURGERY OR AT ANY TIME IN THE FUTURE. BE SURE TO ASK YOUR OTOLARYNGOLOGIST FOR ADDITIONAL WRITTEN INFORMATION INCLUDING BROCHURES PROVIDED BY THE IMPLANT MANUFACTURERS. YOU NEED TO BE FULLY INFORMED ABOUT THE BENEFITS AND RISKS OF COCHLEAR IMPLANT SURGERY AND FOLLOW UP CARE.

Cl Consent transcript via CART service by native English speaker with N95: 5.73% error rate

A COCHLEAR IMPLANT IS A DEVICE THAT RESTORES PARTIAL HEARING TO PATIENTS WITH SEVERE TO PROFOUND HEARING LOSS WHO DO NOT OBSERVE A BENEFIT WITH A HEARING AID. THE DEVICE IS SURGICALLY IMPLANTED IN THE INNER EAR AND ACTIVATED BY A RECEIVER WORN OUTSIDE THE EARS. IT DOES NOT MAKE SOUND LOUDER OR CLEARER, INSTEAD THE DEVICE Bypasses DAMAGED PARTS OF THE AUDITORY SYSTEM AND DIRECTLY STIMULATES THE NERVE OF HEARING ALLOWING INDIVIDUALS WHO ARE PROFUNDLY HEARING IMPAIRED TO RECEIVE SOUND. TO HRETURN NORMAL HEARING AND BENEFITS VARY FROM ONE INDIVIDUAL TO ANOTHER MOST ADULT COCHLEAR IMPLANT PATIENTS NOTICE AN IMMEDIATE IMPROVEMENT IN THEIR COMMUNICATION SKILLS. OTOLARYNGOLOGISTS ARE EAR AND THROAT SPECIALISTS WHO PERFORM IMPLANT SURGERY. YOUR PRIMARY CARE PHYSICIAN CAN REFER YOU TO AN IMPLANT CLINIC FOR AN EVALUATION. THE IMPLANT TEAM INCLUDES AN OTOLARYNGOLOGIST, AUDIOLoGIST NURSE AND OTHERS WHO WILL DETERMINE YOUR CANDIDACY FOR A COCHLEAR IMPLANT AND REVIEW THE APPROPRIATE EXPECTATIONS AS A RESULT OF THE COCHLEAR IMPLANT FOR YOUR SPECIFIC CASE. COCHLEAR IMPLANT SURGERY IS USUALLY PERFORMED AS AN OUTPATIENT PROCEDURE UNDER GENERAL ANESTHESIA. AN INCISION IS MADE BEHIND THE EAR TO OPEN THE MASTOID BONE LEADING TO THE MIDDLE EAR SPACE. ONCE THE MIDDLE EAR SPACE IS EXPOSED AN OPENING IS MADE IN THE COCHLEA AND THE IMPLANT ELECTRODES ARE INSERTED. THE ELECTRONIC DEVICE AT THE BASE OF THE ELECTRODE ARRAY IS THEN PLACED BEHIND THE EAR UNDER THE SKIN. THE RISKS OF COCHLEAR IMPLANT SURGERY INCLUDE FAILURE TO IMPROVE HEARING INFECTION, TYPANIC MEMBRANE PERFORATION DAMAGE TO THE FACIAL NERVE, WORSENING OF TINNITUS OR BALANCE CHANGES IN TASTE BLEEDING INFECTION NEED FOR FURTHER PROCEDURES Meningitis and Spinal Fluid LEak. THESE RISKS CAN OCCUR AT THE TIME OF THE SURGERY OR AT ANY TIME IN THE FUTURE. BE SURE TO ASK YOUR OTOLARYNGOLOGIST FOR ADDITIONAL WRITTEN INFORMATION INCLUDING BROCHURES PROVIDED BY THE IMPLANT MANUFACTURERS. YOU NEED TO BE FULLY INFORMED ABOUT THE BENEFITS AND RISKS OF COCHLEAR IMPLANT SURGERY AND FOLLOW UP CARE.

Accuracy assessment provided by countwordsfree.com/comparetext

Red = Words Removed
Green = Words Added
Results: % Word error rate (Mean and 95% CI) by testing scenario

- Rainbow Passage
  - N95: 6.74%
  - No Mask: 5.08%

- Cochlear Implant Consent
  - N95: 7.14%
  - No Mask: 4.79%
Results: Change in accuracy with N95, CI passage, or NNE speaker for each STT App. Mean and 95% CI shown.
Discussion

- CART is the gold-standard

- Ava 24/7 and the native iPhone dictation application have relatively low error rates (< 20% each)

- Wearing an **N95** mask or using medical terminology did not make a difference

(Vos 2021)
Limitations

How a Speech Application Learns

Step 1. Record voices
- the quick brown fox jumps over the lazy dog

Step 2. Input voice data
- STT Engine Algorithm
- da ✗
- kick ✗
- bound ✗
- the ✅
- quick ✅
- brown ✅

Step 3. Train the speech algorithm
- Deep Learning Architecture
- https://research.mozilla.org/machine-learning/

Speech algorithms are constantly improving

Wired or Bluetooth enabled microphones were not assessed

Common Voice Project
Open Source STT Engine
Instamic.com
Conclusion

CART or real-time stenography is the most accurate solution to facilitate communication with patients who are deaf or hard of hearing and do not use sign language.

If CART is not available, certain iOS speech to text applications are reasonable alternatives even if speakers wear an N95 or use medical terminology.
Thank you to our Team!

Thank you to the Interpreter Services Office!

• Amy Sands
• Tanya Rossi
• Jessica Winiecki

Please feel free to reach out to me with any questions!

Prithwijit_Roychowdhury@meei.harvard.edu
References


"When I saw the doctor, I didn’t understand what the doctor was saying, but was scared to ask for repetitions because it’s a waste of doctor’s time and makes the waiting list longer."

“I wish medical personnel would take hearing loss more seriously. . . . To them hearing loss is easy to deal with. . . . They just talk louder at you. They don’t truly have sensitivity or compassion for it.”

(Reeves 2005, lezzi 2004)