



Automatic Speech Recognition (ASR): The History, Current Technologies, and Practical Tips

#	Question	Answers from Linda Kozma-Spytek and HLAA Staff
1	I love "live transcribing app" for android phone. Is there any app that is equal or good as "live transcribing app" for iPhone?	While there are iOS ASR apps (found by searching Apple's App Store), I don't know of any research that does a direct comparison of Google Live Transcribe, specifically, with these other ASR apps.
2	I wonder if Ms. Kozma could comment on latency and group delay in speech recognition systems. I think that problem is common to both speech recognition and hearing aid design. It shows up in both - especially with speakers who have high syllabic rates in speech. It is the one weak point in my otherwise useful hearing aids; with such speakers, volume is fine, but understanding is diminished.	Delay is inherent in any system that requires processing of the speech signal. Delay that introduces asynchrony between the audio signal and a visual signal (lip movements and/or text) can impact understanding - just as you've noted. Research has provided recommendations for how long these delays can be before impacting speech understanding. Good design will take delay into account but there can be technical challenges to implementing the recommendations, and the recommendations are typically based on average results. Any given individual may be more sensitive to delay than the average would suggest.
3	Has anyone come up with a way to use ASR with Zoom?	Zoom does permit the integration of ASR captions via Otter at certain paid service levels. This link discusses that integration: https://blog.otter.ai/zoom/
4	How do we add app with ASR to cell phone?	As with any app, all you need to do is download the app via the developer's website or the Google play store or the Apple app store, depending on which mobile device you have. Be sure to read or view all the information provided by the app developer to understand the app's features and functions and how to use them.
5	Just yesterday, Shari Eberts said that Zoom is finally considering making captions free for all, just as Google Meet finally agreed to do, a few weeks ago. Do you have any news regarding this new development? Is there hope that Zoom will make captions free for all, especially since the Covid19 situation developed, a few months ago? I have loved the Google Meet caption platform. It's made a huge difference in video conferencing with my family and with my support group, for HOH people, in Israel.	Generally, companies do not make these kinds of plans known outside the company in advance of any announcements they might make, and I have not heard of any. However, Zoom and other services have adjusted their service plans for particular cases due to the pandemic so it may be possible.



Automatic Speech Recognition (ASR): The History, Current Technologies, and Practical Tips

6	Why are the captions on TV (both cable and broadcast channels) so poor when technologies such as Otter are so much faster and more accurate? Why don't the TV broadcasters simply adopt Otter or Google transcribe?	We definitely observe a wide range of caption quality on TV. There are actually a variety of methods utilized to provide captions for live broadcasts and some include ASR. But I'd caution against assuming Otter or Google's Live Transcribe would be a solution. ASR caption quality can also vary widely for a variety of reasons, some of which were discussed during the webinar.
7	Why has it taken Android so long to implement Android 10 for the MF Apple/Android. I have it on my Pixel 4 but can't find it on computers or tablets. I have put an order for a Walmart new low-end tab only because says uses Android 10.	I don't know about the decision-making process used to roll out services or updates to services. However, it's not uncommon to see companies do this kind of step-wise process with new software releases.
8	I will be buying a portable device -- i.e. an iPad or a tablet -- in order to utilize ASR when I am at work. Do you recommend iPad over an Android device?	The type of tablet that you get, if it is specifically for ASR, will depend on the application that suits your needs best and whether it is compatible with iOS or Android or both. There are any number of considerations that go into a technology recommendation, including, most importantly, the needs of the user.
9	When people are using masks (less and less), it's impossible to understand what they are saying, without ASR enabled, and/or clear face masks. I've noticed that, sadly, people have stopped wearing face masks...especially sad, because, there is the threat of more waves of the virus, to occur, within the next few months.	
10	So far, I am only finding ASR programs and apps in English. I need Spanish ASR. At work is where I will need to be using ASR. At work, we use Spanish +/- 90% of the time. Do you know of any Spanish ASR?	There are ASR programs/apps that enable either translation from one spoken language to text in another language or a transcription of Spanish. Both Live Transcribe and Microsoft Translator have one or both of these functions. There may, of course, be other programs/apps that do the same thing.
11	THANK YOU!	Thanks.
12	There are MANY other transcription apps, for iOS.... also, the live transcribe for iOS has a charge associate with it.	You're right, as I mention in the webinar, there are many alternatives to the examples I provided. I would like to note that Live Transcribe for iOS is made by a developer other than Google. So, Google Live Transcribe for Android OS and Live Transcribe for Apple iOS are two different programs by two different developers.



Automatic Speech Recognition (ASR): The History, Current Technologies, and Practical Tips

13	Do you know of any Spanish ASR other than Watson Speech-to-Text?	There are ASR programs/apps that enable either translation from one spoken language to text in another language or a transcription of Spanish. Both Live Transcribe and Microsoft Translator have one or both of these functions. There may, of course, be other programs/apps that do the same thing.
14	Why the need of the intermediate person for phone conversations?	ASR can be more accurate when you control the speaker characteristics and the environment. Using a single, trained speaker in a quiet, controlled environment as is done by many current IPCTS providers has the potential to produce more accurate captions. The added step of revoicing does introduce another source of delay.
15	With the recent approvals by FCC for ASR for MachineGenius, does that mean that ASR is now mature enough to meet the accuracy the FCC requires? I know it's improving, but I thought we were a ways away from that yet. What about accents, dialects, tone deafness, etc?	The mandatory minimum standards set by the FCC for IPCTS do not include specific metrics for assessing the quality of transcription. The FCC Disability Advisory Committee has recommended that quality metrics be established to create minimum performance standards regardless of the platform or technology used. That said, one of the filings related to Machine Genius' application report the results of evaluations of accuracy and latency for this offering compared to other offerings.
16	From what I am hearing so far in this webinar, it sounds like I have to use a cell phone. Can I get ASR through an iPad or a tablet?	Yes, you can.
17	(In case this is not covered...) Where do things stand with free captions for all videoconferencing platforms for people with hearing loss?	I don't know where this stands at the moment. Generally, companies do not make these kinds of plans known outside the company in advance of any announcements they might make, and I have not heard of any. However, Zoom and other services have adjusted their service plans for particular cases due to the pandemic so it may be possible.
18	Is there something like Google's Live Caption for iPhones?	I don't know of a similar app to Google's Live Caption for iPhones.
19	Which of these products would be best for a Zoom or WebEx meeting?	I've used both web-based captioning on my laptop (e.g., Otter) and app-based captioning on my mobile device (e.g., Live Transcribe iOS) for videoconferencing meetings. The best product is the one that meets your caption quality requirements and feature and function needs.



Automatic Speech Recognition (ASR): The History, Current Technologies, and Practical Tips

20	CART at times is "financially challenging" - in a physical meeting, which of these products would work best as a "replacement" for a CART person at the meeting transcribing?	Whatever ASR product you might use as an alternative to CART depends very much on your meeting needs. I wish it were as easy as saying this is the best product but it's not. I'd suggest picking several based on your meeting needs, trying them each in turn and selecting the one with the best fit. There is no substitute for actual experience with a product to assist your decision-making process.
21	Is this meeting using Live Caption?	Yes, this webinar is being captioned by a CART provider.
22	Video is well-synched - audio is far behind.	This is hard to control because each person has different internet speeds and may be running different programs at the same time on their computer that could affect the experience.
23	In these days of social distancing, I am finding that ASR programs have trouble distinguishing voice from background noise if the speaker is six feet away. Is anything being done to improve ASR's recognition of human voice?	For ASR applications on a mobile device, you might consider using a directional microphone that plugs directly into your mobile device. You'll want to pay attention to the physical connection the mic uses to be sure it and the mic itself are compatible with your mobile device.
24	Does Linda use the free version of Otter? Or does she have a premium account?	I've used both.
25	As a hearing aid user, kudos to Ms. Kozma for clear diction!!!	Thank you.
26	Which of these options might be best for captioning purposes when in a Zoom or FaceTime call so that I know exactly what's being said?	I have been able to use Otter to generate captions for Zoom and FaceTime. There may be other options using other apps or a set up like the example I mentioned in the Webinar using Live Transcribe as an alternative to Live Caption on an Android phone.
27	Is there a monthly charge for Otter? or do you have to pay a user fee? Most people seem to prefer apps and programs that are free of charge (or minimal charges)	Otter has both a free plan and paid plans.



28	In looking into using Otter, it looks like you need Otter for Teams to be able to use it with virtual meetings like with Zoom. Otter for Teams is expensive. Do you know if they are considering making it available for the Premium level?	I don't know but you could look at this link: https://blog.otter.ai/how-to-transcribe-any-video-meetings-with-otter-ai/ It describes, as I did in the webinar, how to transcribe live, side by side with the meeting software using Otter Basic or Otter Premium.
29	re: Google meet...how come I can't see more than two people on the screen, simultaneously?	I don't know but would suggest you try adjusting the layout. Select "More Options" - you may need to click on the three vertical dots on the lower right-hand corner of your Google Meet screen. Then click on "Change Layout."
30	What technology is being used for this webinar?	We are using Zoom for the webinar platform and a live CART provider.
31	It's also worth mentioning that Google Meets prefaces the text with whomever is speaking. This is a HUGE advantage over others when reading captions.	
32	I've tried Google Meet. I find that I can only avoid echos by muting everyone who is not speaking. Is there a fix?	Unfortunately, I'd need to know more about the audio set up you and others are using to suggest an alternate fix.
33	Apparently Zoom is in the process of offering captioning embedded in the meetings. BUT only on paid accounts. BUT they did say will make it free to DHH upon verification of hearing loss.	
34	What is the best option to use to caption a Zoom meeting?	I have been able to use Otter to generate captions for Zoom and FaceTime. There may be other options using other apps or a set up like the example I mentioned in the Webinar using Live Transcribe as an alternative to Live Caption on an Android phone.



Automatic Speech Recognition (ASR): The History, Current Technologies, and Practical Tips

35	Can you speak to privacy issues? I like Otter’s accuracy (compared to other ASR tools), but don’t like the request to take over my microphone and save my conversations.	Privacy and security are issues anytime data is transferred over the Internet, as I’m sure you know. Almost all ASR is currently cloud-based, and encryption is commonplace. The best you can do is carefully read a company’s policy and decide if the benefits outweigh the risk for you. You can limit a program’s control of your microphone/video to only when you use an application but that adds another layer of complexity because in most cases you must do this manually. Once we get to on-device only transcription, this will help but you’ll still need to inform yourself as much as possible about how data gets stored and used.
36	I use the Live Caption App in church meetings - like the deacon board. However, the microphone on they iPhone or iPad will only pick up voices 4-5 feet away. Do you know of an external microphone that till pick up voices 10 or more feet away? The ones I have found are recording mics.	I don’t have a particular mic to suggest but there are microphones specifically designed for mobile devices that generally use a 3.5 mm audio plug. Some mics are made specifically for iPhones models that no longer have this audio jack. And, there are also some Bluetooth mics for mobile devices. If the mic plugs directly into the mobile device, you’ll want a directional microphone that has higher sensitivity. Due to the mic’s directionality, you’ll need to point the mic you’ve plugged into the phone at the speaker. Another option is to use an omni-directional Lavalier mic also plugged directly into the phone. These mics are on a cord that may be long enough to place at the center of a table around which your meeting is taking place while the phone remains close to you in order to view the captions. Like the Lavalier mic, a Bluetooth mic could also be put in a convenient place closer to the other people in the meeting to pick up their voices while the phone remains with you.
37	Do you see the quality, convenience, and cost of ASR ending business for CART providers?	There’s no way to know the impact of ASR on CART providers’ businesses over time. If only I had a crystal ball. I will say that I expect ASR to only get better with time. I think what’s true - even right now - is that ASR is filling a gap in caption provision where CART has never played a role. One example is captioning on-demand for in-person, face-to-face conversation.
38	Can you turn on captions during Face Time on iPhone and if you have an iPhone can you go to Google Meet with other iPhone users and use captions?	As far as I know FT does not have caption support. You can download the Google Meet app and start or join a meeting with others as long as everyone has internet access, a Gmail account and the meeting link.



Automatic Speech Recognition (ASR): The History, Current Technologies, and Practical Tips

39	What is future of ASR for TV?	If only I had a crystal ball. ASR is used for some live-captioned shows, along with human CART providers and ENT (Electronic Newsroom Technique - dialogue included on a teleprompter script is converted into captions). As I said above, ASR will only get better with time so I expect its role for TV will increase.
40	Which ASR would work well with internet video that do not include captions?	This depends on the platform you're using for viewing the internet video. If you are using a mobile device and happen to have a Pixel 2 or higher phone or Galaxy S20 phone, I would suggest trying Live Caption. If not, then it depends on whether you view the videos on your phone or on your computer. I discussed an alternative to Live Caption for captioning media on your computer during the webinar.
41	Linda, thank you, this is great. Will you have a "matrix" of characteristics that each of the technologies have? For example, which works with Android/Apple platforms or both and which of the characteristics it has of the many you listed?	I do not have nor do I know of anyone who has developed a matrix of characteristics.
42	The Speak See system and the Q system have been in the development stage for some time. Do you know if they will actually be available for public use? Both systems would be a great help to people with hearing loss.	I don't know if the SpeakSee system will become available for public use. By Q system, do you mean Android Q? Android Q was branded as Android 10, which was generally released September 2019.
43	How is the captioning for this webinar being performed?	It is being provided by a live captioner.
44	LOVE Shankar Verdatim on Hidden Brain!	
45	Do you also have error or accuracy rates compiled for each?	I do not have error rates compiled for each.
46	Have you tested accuracy when dealing with people who have accents or higher pitched voices?	I have not tested accuracy for people who have accents other than a primarily general American accent.



47	How do I include ASR in vimeo video?	This depends on the platform you're using for viewing the internet video. If you are using a mobile device and happen to have a Pixel 2 or higher phone or Galaxy S20 phone, I would suggest trying Live Caption. If not, then it depends on whether you view the videos on your phone or on your computer. I discussed an alternative to Live Caption for captioning media on your computer during the webinar.
48	Speech clarity and Rate of Speech important also to HA users.	
49	You mentioned ASR for people that have speech impediments. Can you tell me what I need to look for to try to stay up to date on this?	You can learn more about Google's Project Euphonia here: https://sites.google.com/view/project-euphonia/ And here: https://ai.googleblog.com/2019/08/project-euphonias-personalized-speech.html
50	Wow --- this is an awesome presentation! Great job Linda. I'm pleased that we'll be able to download some printed info afterwards. There is just so much to consider!	Thank you!
51	Are there any ASR services (and which ones) will work without Internet connections? Again, do you address in document you'll post? Thank you!	Google's Live Caption and Recorder work offline. There may be others, but these are the only applications I know of that do so.
52	Is it possible to get a chart / table, etc with information on various apps and programs, and the differences between each of them, such as for android / iOS devices, fee or no fee, purpose of app? What it offers? speech recognition? Translation? meeting captions, etc etc? It may help us to figure out what is most likely to meet our needs quicker.	I do not know that anyone has prepared such a chart.



53	You did not mention Google Duo which is helpful since it has video and is simple to use.	Google Duo is helpful. However, at the moment, to get captions using Google Duo, you must use Live Caption, which only a few phones support at the time of this webinar. It has been reported that Google may roll out captions for video calling in Duo.
54	Apologies if you already covered this, but is there a better alternative that Apple Visual Voice mail for transcribing phone voice mail?	There are third party apps for visual voicemail. I'd suggest doing an Internet search for this. A quick search I did came up with: https://www.igeeksblog.com/best-voicemail-apps-for-your-iphone/
55	Do you have any suggestions (what is currently available) to being able to see captioning on a cell phone or tablet, for students who may be in a classroom with a mask wearing teacher (muffled voice)?	This is a difficult question to answer because there are so many aspects to finding a workable solution that go beyond captioning alone.
56	How are the different ASR apps accounting for accents? Are any of them allowing the user to customize the "listening" of the app based on the accent of the speaker?	Some ASR applications account for not only language but also for dialects, which users can customize for. However, accents that are present for example in the speech of non-native speakers of a language can be challenging for ASR. I've never seen this as something users can customize.
57	Excellent presentation, Linda - exactly everything I didn't know, and needed to know, and more!!	Thanks.



58	<p>Some examples of iOS speech to text apps:</p> <ul style="list-style-type: none">iHearU2Say & go audio notesVoice IttChatableListenAllHearing HelperAppmyear (\$9.99 one-time price)EarlyMicrosoft Translator <p>Not ASR, but helpful for communication with people who wear masks:</p> <ul style="list-style-type: none">CardzillaBuzz Cards <p>Cardmedic is an app developed by a doctor, who works in ICU....it enables patients who are on ventilators, be able to communicate with medical staff who wear masks and the HOH person on the vent, needs to communicate with that person.</p>	Thanks.
59	<p>Do you think we are close to being able to provide our own CART when we go to meetings? Do we have to warn the speakers that we are recording them?</p>	<p>I know some people are using ASR for captioning at meetings right now. I make it a practice to always let everyone know when they are being recorded. It is not only a courtesy but there may also be a legal requirement to do so.</p>
60	<p>I was wondering if you have any experience with InnoCaption+ I have downloaded it to both my work and personal cell phones and supports me in understanding my callers.</p>	<p>I do not have any user experience with InnoCaption. I cannot use relay services because I am hearing.</p>
61	<p>The other day I viewed a webinar, but it had no provision for captioning. I used Text to Hear but it was expensive/. Were there alternatives?</p>	<p>I discussed a couple of options for using ASR to caption media - like a webinar - during my webinar.</p>



62	Has anyone done an accuracy study of all these ASR systems?	No, not to my knowledge.
63	Interact Streamer is coming out with caption for all tool for Zoom soon.	
64	The Chromebook will run Android apps, including GLT. Also has an HDMI output, easy to connect to a display.	
65	What version of Microsoft Power Point provides ASR?	Please see: https://support.microsoft.com/en-us/office/present-with-real-time-automatic-captions-or-subtitles-in-powerpoint-68d20e49-aec3-456a-939d-34a79e8ddd5f#OfficeVersion=Web
66	Regarding health care providers and maintaining Health Information Privacy. Which way would be the best for privacy?	Privacy issues, both the technological and legal aspects, are outside my area of expertise.
67	Zoom recently integrated Otter for their meetings, but not webinars. Is there a work around to have Otter captions appear on the screen live?	Otter captions would need to be shown in a separate browser window. On their website, Otter mentions that support for Zoom webinars is coming soon.
68	Can I close caption Zoom meetings? How can I advocate for this?	Zoom meetings can be captioned. I'd suggest you talk to HLAA about how to best advocate for this.
69	Do you know if Google is ever bringing its actual Live Transcribe for iOS? Nothing on iOS compares to the Live Transcribe and its abilities.	Google is aware that people want Live Transcribe for iOS, but I don't know whether they have any plans for this.
70	Hi, I would like more information about products that would recognize difficult speech like deaf speech, so persons can be understood by others in a meeting.	I don't know of any products that currently support this. Please see the links above to Google's Project Euphonia.
71	Note that Google Slides only works with Chrome. I found this by googling how to use google slides with captions. So, googling how to use products often helps.	Thank you - so true.



Automatic Speech Recognition (ASR): The History, Current Technologies, and Practical Tips

72	Do you know if there is a way for when I host work zoom meetings for me to get a transcript from my recording? It would be ok afterwards, doesn't need to be in real time.	Otter's paid plans allow you to upload audio files in order to generate transcripts from them.
73	Google Pixel phones has a recorder app that provides captioning. Is that app available for non-Pixel android phones?	The app seems to be available on at least some Android smartphones running Android 9 Pie or Android 10. The best way to find out if the app can be installed on your particular phone is to go to the Google Play store and search for it. There are many recorder apps out there so make sure you select the Recorder app where Google LLC is the developer.
74	In what situation would you use a live captioner as opposed to ASR. Today's event is a live captioner.	This is a complex question that requires careful consideration of many factors for the given situation and input from the individual(s) who require the captions.
75	What's the best way for the hearing-impaired that will allow them to go to a public speech (i.e. graduation) and read on their cell phone what is being said?	I would suggest contacting your regional ADA center (https://adata.org/) to determine if the individual with hearing loss has protections under Title II of the ADA in this circumstance. If the person does, I would suggest requesting CART generated by a human captioner.
76	Is there a "pen" microphone to use in public stores to pick up ASR on your phone? Only I will hold the "pen" due to COVID.	There are microphones specifically designed for mobile devices. Some are directional and plug directly into the phone so you would simply point your phone toward the person speaking. Others are Lavalier mics that also plug directly into the phone and are on a cord or wireless Bluetooth mics so that you would need to hold the mic and point it toward the person speaking while also holding the phone.
77	Do you recommend ASR to be used in legal or medical settings where accuracy is crucial and could impact health or safety?	This is a complex question, and I do not provide recommendations of this sort. However, I would recommend that anyone needing caption support in these settings consult their regional ADA center and state agency for people with hearing loss to understand their rights under the American with Disabilities Act.
78	Are you using a special microphone? You sound very clear.	I was using the microphone built into my MacBook Pro laptop.
79	Regarding machine genius, do you know which relay provider is using them?	Machine Genius is the relay provider.



80	Are there any landline phones that offer ASR instead of CART?	I do not know of any landline phone that offers ASR instead of CART. Hamilton Relay offers software for using Cisco phones with their Relay Service but they don't use CART or ASR. They use revoicers plus ASR. Additionally, this only works for phones in large companies that can use Cisco phones.
81	Why do you have to download your video on which you want captions for Facebook.	I do not know why you have to download videos from FB to get captions. Not all video will have captions. If the video does have captions, here's a link that describes how to turn captions on for videos on FB: https://www.facebook.com/help/427723640933279?helpref=related&source_cms_id=261764017354370
82	Newer hearing aids and CI has bluetooth connectivity to cell phones and audio players. Is there a tool to interject ASR to capture the bluetooth audio?	No, not to my knowledge.
83	Do you think at live political speeches they could have a big screen and live caption in place of or in addition to the signed language?	Could they? YES!
84	I came in late -I'm wondering if this session will be posted online later? It's very informative.	It will be posted to the HLAA website.
85	Any recommendations about wireless connection of few devices (e.g. iPad and iPhones) so speakers see how their speech is being transcribe so they speak slower or clearer if needed. Wireless devices?	Yes, there are ways that speakers other than the person with hearing loss can see the captions generated from their speech. For example, if you are using Live Transcribe on your phone for an in-person meeting with someone, you can simply place the phone so that both you and the other person can see the text being generated. I would need to know about the technologies you have available in order to make specific suggestions. I'm uncertain though how diligently the other person would monitor the captions being generated to affect changes in their speech patterns or correct errors.



HLAA Webinar – 5/27/2020

Automatic Speech Recognition (ASR): The History, Current Technologies, and Practical Tips

86	Not a question but I'd like to put this on your radar. I did speak with an Otter representative and they said that Zoom has the code to make live captioning work for webinars (not just meeting a) but the integration has not been completed and activated by Zoom. Maybe an opportunity to encourage Zoom to roll this feature out as soon as possible for webinars.	Thanks. In this blog post from Otter: (https://blog.otter.ai/zoom/), they mention that support for Zoom webinars is coming soon.
87	You made mention of a training this Friday - can you please share details/links to the full info?	It was not training but a set of documents that were made available. Please see this link for documents that provide guidance on making remote workplaces accessible: https://www.deafhhtech.org/erc/2020/05/29/dhh-erc-and-deaf-hard-of-hearing-consumer-groups-provide-remote-workplace-guidance/
88	When I am in a Zoom meeting, I generally just put my phone next to the computer and read captions on Otter. When you say that Otter can work with Zoom, I assume you mean that it will somehow be integrated. Is that true if you are not the host or must the initiator of the meeting generate the captions?	In my experience, Otter can work with Zoom as an integrated feature under certain pricing plans. I also use Otter running in a separate browser window from Zoom, when I do not have access to the integrated feature. This link discusses Otter integration with Zoom: https://blog.otter.ai/zoom/
89	Thank-you so much!	You're welcome.
90	Can HLAA PUT THE CAPTIONS UNDER YOUR PICTURE?	No, we cannot. We are limited to the zoom webinar changes video format when the screen is shared for the PowerPoint presentation.
91	I like watching non-captioned older movies, any suggestions of ASR to use	There are some possibilities for using ASR to get captions for older movies, but I would need to know about the technologies you have available in order to make suggestions.