**Market Survey of Public Speaking Assistive Interfaces and Other Related Systems**

**1) VoiceVibes (Extension for Zoom)**

**Website:** <https://www.bigtincan.com/features/conversation-intelligence/>

**Zoom marketplace link:** <https://marketplace.zoom.us/apps/TX2v6vjxRfqP41AaTFUjIA>

**Description** **(Taken from the zoom app market place)**: “AI-powered coaching tool analyzes your recordings to show exactly how you sound, and where you can improve, so you can make the best impression every time. “

**How does it work:**

“Record or upload any speech or practice session, or easily select and analyze segments of your zoom recordings (from 15 seconds to 15 minutes) to receive personalized, automated feedback in minutes. AI algorithms trained on millions of human perception ratings provide helpful feedback on your oral delivery while human collaboration features enable additional feedback through coach and team sharing, scorecards and reporting. What's more - you can even see exactly where you sounded most confident, boring or captivating - 20 different "vibe" areas in all!”

**Automated feedback includes:**

* Pace
* Pausing
* Strength of Opening
* Clarity
* Upspeak
* Um/Uh filler counter
* programmable word and phrase detection
* Confident
* Boring
* Authentic
* Captivating
* Pushy
* Timid
* 14 additional vibes....

**See (very cheesy) Video on Zoom Marketplace:**

<https://marketplace.zoom.us/apps/TX2v6vjxRfqP41AaTFUjIA>

**Some screen shots of the extension (Although I believe the GUI may have had an update since):**

Graphical user interface

Description automatically generated with medium confidence

**App also appears to allow other users to comment:**

Graphical user interface

Description automatically generated with low confidence

Graphical user interface

Description automatically generated

**Technologies used:** N/A

**Key take aways:**

* Integration with **zoom** as a downloadable extension.
* Zoom calls can be **imported** to be analysed OR Audio or Video files, or publicly available URLs
* Does **not** provide feedback in real time, it provides **delayed** feedback.
* Provides feedback **only** on audio behaviour.
* Processing appears to be done on the **cloud** (Although I am not certain)**.**
* **Assignments** can be set for other users. (For example Receive an overall rating of ‘X’)
* States it can track and analyse **multiple speakers**.

***Other Notes:***

* *I haven’t tried it out so I don’t know how well any of this actually works.*
* *Although this is an extension for zoom, it appears to have been purchased by another company and is included in their collection of Training programs. As a result is seems to operate more as part of their own app and the zoom extension simply allows you to import zoom calls/recordings rather than that its been specifically designed for zoom.*

**2) PowerPoint’s ‘Speaker coach’**

* Will be available in Microsoft teams in ‘early 2022’ (Although It doesn’t appear to be a feature just yet).
* The following videos from Microsoft give a good idea of how this works:
  + <https://www.youtube.com/watch?v=FuIgDcUl4lc>
  + <https://www.microsoft.com/en-us/videoplayer/embed/RWKfoc>

**Speaker Coach provides feedback on:** pacing, pitch, your use of filler words, informal speech, euphemisms, and culturally sensitive terms, and it detects when you're being overly wordy or are simply reading the text on a slide.

After each rehearsal, you get a report that includes statistics and suggestions for improvements.

Graphical user interface, application

Description automatically generated

**Speaker Coach also provides (very basic) feedback on visual behaviour:** “For example, Speaker Coach pops up a warning if you are too far back from the camera or too close. Additionally, notifications appear if you obstruct your face with your hand or don't make eye contact with the audience.”

**Key take-aways:**

* Integration with **Microsoft teams**
* Provides feedback on both **audio behaviour** and very basic **visual body behaviour**.
* Provides **real-time** feedback with on-screen guidanceand **delayed feedback** after a presentation is finished with a ‘rehearsal report’.
* Can only analyse one person speaking at a time.
* Appears that its intended use is to give feedback to users rehearsing for a presentation on PowerPoint.
* Only understands English.
* It doesn’t look like video can be uploaded, it has to be recorded straight from PowerPoint.

**Other Notes/Thoughts:**

* *This is one of the very few systems that automatically provides visual feedback to the user and even then it provides it in a very basic manner.*
* *Would we be able to develop a more complex system that would be able to provide more detailed feedback on visual behaviour? Why has this not been touched upon particularly, especially by a large company like Microsoft, is it feasible with current technology?*

**3) Zoom Nonverbal feedback and meeting reactions:**

**Link for more information:** <https://support.zoom.us/hc/en-us/articles/115001286183-Nonverbal-feedback-and-meeting-reactions->

**How it works:**

“If the meeting organizer enables the Nonverbal feedback and Meeting reactions features, [**meeting participants**](https://support.zoom.us/hc/en-us/articles/360040324512) can place an icon in their video panel and beside their name in the participants panel to communicate with the [**host**](https://support.zoom.us/hc/en-us/articles/360040324512) and other participants without disrupting the flow of the meeting. For example, selecting the**Slow down** icon places the icon in your video panel and beside your name to indicate you would like the host or presenter to go slower. “

**Here are some examples of the emojis and nonverbal feedback that can be provided:**

Graphical user interface, text, application, chat or text message

Description automatically generated

**Key take-aways:**

* This is a very basic real-time form of public speaking feedback integrated into zoom.
* No form of AI used to automate this process, feedback is given by other users.

**Other Notes/Thoughts:**

* By using other users to provide feedback, more consistent and more relevant feedback to the situation can be given

**4) Rhema: Real-time In-Situ Intelligent Interface to Help People with Public Speaking**

**Link to article:** <https://dl.acm.org/doi/10.1145/2678025.2701386>

**Description:**

* Intelligent user interface for google glass.
* Operates by automatically detecting the speaker’s volume and speaking rate in real time and provides feedback during the actual delivery of speech.
* Created as part of research, published in 2015

**How does it work:**

* Provides feedback on just speech so only analyses audio behaviour.
* Records the live speech of the wearer of the google glass and **transmits the audio to a server (processing is done on a server rather than within the device)** for automated processing of **volume and speaking rate** which is then presented back to the user. This **allows for analysis to occur in real time** and allows for the interface to function during an actual speech delivery.

**How does it provide feedback? :**

* Experimented with 2 different strategies of information delivery: 1) Continuous stream of information and 2) Sparse Delivery of recommendation.
* Participants were significantly more pleased with their speech while using the sparse over the continuous.

**Key Take-aways:**

* Provides **real-time** feedback with in-device guidance.
* Integrated with google glass.
* Provides feedback on only **audio behaviour.**
* Can only analyse one person speaking at a time.

**Other Notes/Thoughts:**

* *With the rising popularity of wearable devices (Smart watches, smart glasses, smart jewellery, virtual reality) perhaps it might be an interesting idea to develop a system that can be integrated with one of these.*
* *Are we looking for the analysis to occur in real time? Or would we like to develop a system where you submit an audio or video recording, the data is processed, and then it is returned to the user. As discussed in Monday’s meeting, analysis of visual behaviour is much harder than just audio, this may be a challenge to give feedback based on visual behaviour in real time and if we would want to provide real time feedback to our user then analysis audio behaviour may be much more realistic.*
* *The study mentions that providing feedback to users during a live speech has a number of inherent design challenges. “One challenge is to keep the speakers informed about their speaking performance without distracting them from their speech. A significant enough distraction can introduce unnatural behaviors, such as stuttering or awkward pausing … Human attention is a limited resource, and the user might completely ignore the interface if the interaction requires too much attention.”*
* *The experiment with 30 participants showed that the real time interface did* ***add value*** *during the actual speech delivery*

**5) RocSpeak:**

**Website:** https://machinteraction.ur.rochester.edu/rocspeak/

**Both of you are probably aware of this system but if not, link to video:** <https://youtu.be/0fo4x5SuhfU>

**Key Take-aways:**

* Provides very detailed automated feedback based on both visual and audio behaviour – this is the only system I have seen that is able to do this with just a camera.
* Can only analyse one person speaking at a time.
* Provides only **delayed feedback** after a presentation is finished, **no real-time feedback**.
* It’s a web application where processing appears to be done on the cloud.

**Other Notes/Thoughts:**

* *It has quite a basic GUI and as far as I can tell was created purely for research purposes. – perhaps this could be an interesting thing to develop into a mobile app. It has already been done but might be popular in the market if repackaged for mobile app or zoom extension. Is there any way in working in collaboration with Rochester HCI to build upon their system, improve the GUI and develop it into an app for the IOS marketplace or even an extension for zoom?*

**6) Cicero:**

* Will skip past as I am sure you are familiar.

**7) Logue:**

**Website:** <https://dl.acm.org/doi/10.1145/2702123.2702314>

**Description:** “With wearable technology, such as Google Glass, we now have the opportunity to augment social interactions and provide realtime feedback on one’s behaviour in an unobtrusive way. In this paper we present Logue, a system that provides realtime feedback on the presenters’ openness, body energy and speech rate during public speaking. The system analyses the user’s nonverbal behaviour using social signal processing techniques and gives visual feedback on a head-mounted display”

**Visual example of how it works:**

Diagram

Description automatically generated

**Key Take-aways:**

* Provides **very detailed** automated feedback based on both **visual and audio behaviour** – this system does this by making use of google glasses and an external camera.
* Provides high quality real-time to the user.

**Other Notes/Thoughts:**

* *I really like the way this system provides the real-time feedback to the user through functional icons. If we are looking for our system to provide real time feedback, I think taking inspiration from the way it is presented to the user in Logue is really well done.*

***Example of the Icon Set featured in Logue:***

A picture containing website

Description automatically generated

Graphical user interface

Description automatically generated

**Virtual Speech: Professional Development Training for the Modern Workplace (Mobile App and Web App):**

**Website:** <https://virtualspeech.com/>

**Description:**

* “Self-paced training courses designed to improve your skills in the most effective way.”
* Provides Online training exercises and Virtual Reality training exercises.
* Provides courses on Presentation Skills, leadership, sales and more.
* Provides practice exercises that allows the user to apply what they learned in example scenarios. This can be done online or through using VR.

**How it works?**

**See Video:** <https://virtualspeech.com/product>

**How is it monetized:**

* Provides paid personal and business subscriptions which gives the user access to all of the materials on their website. $45 per month or $34 yearly – Perhaps a quite expensive?

**Key Take-Aways:**

* It’s a web application where processing is done on the cloud.
* Provides both real time feedback and delayed results.
* Isn’t just for practising presenting, it allows the user to train for a variety of different scenarios (Presentations, Collaborative Remote Training, Leadership communication, Media training, Training trainers, Persuasive Business Storytelling, interview Practice e.t.c).
* Integrates with Virtual Reality.
* Seems to have been quite popular with individuals and businesses.

**Other Notes/Thoughts:**

* *The public speaking interface that we develop doesn’t have to be just for presenting like a lot of the current systems are, it could include a variety of different scenarios, we don’t have to just develop a System for feedback on presentations, there are lots of other scenarios, we can even think of a specific University related scenario that might be useful for University students that has not been covered yet?*
* *Is there any cutting edge technology that would be very beneficial if utilized for the development of a public speaking interface (VR, AR, MR, Smart Watches and other wearable devices)? Even though VR has already been utilized by virtual speech, is there different ways in which VR can be utilized to develop a public speaking interface? Something to think about.*
* *Using something like VR allows for head movements and hand gestures to be tracked quite easily and accurately, may be a more realistic alternative to using AI to analyse visual behaviour via a camera?*

**Orai for Public Speaking (Mobile App and Web App):**

**Website:** <https://www.orai.com/>

**Monetization:** Offers in app purchases in the form of a monthly subscription plan for $10 a month per user. It also offers enterprise options but does not quote prices for those.

**KEY FEATURES LISTED ON THE APP STORE:**

• Instant speech feedback using AI:

* Filler words (um, you know, basically)
* How fast you speak
* You energy level
* Vocal clarity

• Micro-lessons on public speaking

• Detailed performance tracking

• Audio recording and playback

• Accurate transcript of what you say

• Save your previous recordings

• Ability to share your recording and transcript with friends/colleagues

• Freestyle mode to practice anytime, anywhere

• Script mode to practice a prepared speech; simply copy/paste it into the app and tap record

• And more!

**Example of the apps delayed feedback:**

Graphical user interface, text, application, chat or text message

Description automatically generated

**Example of the apps interactive lesson:**

Graphical user interface, application, chat or text message

Description automatically generated

Graphical user interface, text, application

Description automatically generated

**It also includes enterprise features:**

Graphical user interface, application

Description automatically generated

**Key Take-aways:**

* This was the **first suggestion** on the IOS store when I typed in “Public Speaking”. It also has a very high 4 and a half star rating with 250 reviews - very popular with users. This looks to be one of the most popular public speaking assistive interfaces apps on the IOS market place – perhaps this gives us a good indication of what consumers might be looking for.
* Not clear if processing is done on-device or on the cloud but my guess would be that it is a combination of the two.
* The app has a very clean and useable interface. If we do decide to develop an app, perhaps we can have a look at this app and understand WHY we like the interface and take inspiration and incorporate some of this into our system.
* Provides both real time feedback and delayed results.
* It doesn’t just provide feedback, it provides lessons for a variety of different lessons.

**Other Notes/Thoughts:**

* *I downloaded this application on my mobile device and it has a great GUI, runs smoothly and looks clean.*
* *I do feel that looking at which systems are more popular is a hugely beneficial indicator for identifying the best systems out there. This one appears to have very solid functionality and is also one of the most popular apps of this type.*
* *The progress tracking featured in the app is an interesting idea, perhaps something we should think about, if users can see tangible progress it may make them more likely to continue using the app and reinforce to the user the effectiveness of the app.*
* *Having lessons/courses included as features in public speaking assistive applications seems to be quite a common occurrence. It may be because these two features when used in conjunction with one another can be more effective than simply just providing feedback. It really depends what we want our system to do, but its something to think about…*

**TED Masterclass (Mobile App and Web App):**

**Website:** <https://masterclass.ted.com/>

**Skim through this video to show the key features that the app offers:** <https://youtu.be/b7id4rzgKIM>

**Key Features:**

Provides a series of lessons to teach users how to identify, develop and share their best ideas with the world.

**Monetization:** Offers in app purchases in the form of a monthly subscription plan for $50 a month per user (Quite a bit more than Orai which appears to be a key competitor). It also offers enterprise options but does not quote prices for those.

**Key Take-aways:**

* This was the second suggestion on the IOS store when I typed in “Public Speaking”. It also has a very high 4 and a half star rating with but with just 28 reviews. This tells me it is very popular with users.
* It is still one of the most popular public speaking interfaces on the market yet it **doesn’t provide feedback to users**, it only provides lessons – this tells me that **consumers heavily value the ‘lesson’ features** when trying to approve their public speaking ability.

**Speeko (Mobile app):**

**Website:** <https://www.speeko.co/>

**Skim through this video to see how it works:** <https://youtu.be/c-WI35WK6EM>

**Monetization:** Offers in app purchases in the form of ‘Speeko PLUS’ which is a subscription plan that gives the user access to more features and resources (similar to other apps of this type).

**Key Take-aways:**

* The app has a very clean and useable GUI.
* Provides both real time feedback and delayed feedback using AI.
* Provides only audio behaviour based feedback

**Ummo (Mobile app) (No Longer available):**

**This product doesn’t appear to be available any more but still might be useful to give an idea of what can be done.**

**Website Link:** <https://ummoapp.herokuapp.com/>

**Description:**

It takes in input by the user pressing the record button integrated within the app. It then analyses the user speaking and will highlight their use of ‘filler words’ in real time. It will also give delayed feedback in the form of a report after (See images on their site).

Graphical user interface, application

Description automatically generated

**Key Take-Aways:**

* Has a nice clean interface, looks like one of the more user-friendly applications.
* Is available on mobile devices on the app store
* It also provides delayed feedback on pace, volume and pronunciation.

**Other Notes/Thoughts:**

* This mobile app was quite 1 dimensional. Placing a large focus on identifying filler words that the user says. Perhaps this is why it has not found success in the longer term.

**LikeSo:**

**Website:** <https://sayitlikeso.com/>

**Description:**

“LikeSo is your personal speech coach. LikeSo offers a fun and effective way to train against verbal habits and practice speaking articulately, confidently and without all of those “likes,” and “sos.”

**Key Features:**

**It has two modes of play:**

* **FreeStyle** - In FreeStyle, you can speak for up to 30 minutes. Select the words you want the app to train against. You can “pause” or “stop” at any time.  As in TalkAbout mode, at completion you will be presented with an analysis of your talk session and see your results.
* **Talk About** - conversation game to practice talking on the fly. You Choose a topic from a list of 12 (e.g. The Job Interview, The College Interview, Debate Team, Pop Culture Favorites, Speed Dating, etc.). Your session will include five prompts.  Select your Talk Time and the filler words you want to train against. Speak into your smartphone’s microphone and LikeSo will listen and capture your speech. At completion, you will be presented with an analysis of your talk session and see your results.

**Monetization:** A $5 one off fee for purchasing the app (This is quite a bit cheaper than the other apps of a similar type).

**Key Take-aways:**

* Provides both real time feedback and delayed feedback using AI.
* Perhaps slightly **less sleek GUI** in comparison to the competitor apps such as Orai.
* Not clear if processing is done on-device or on the cloud but my guess would be that it is a combination of the two.
* The ‘**Talk about’** is an interesting feature that I haven’t come across in other systems, perhaps we could think about implementing this into our own system

**Voice Analyst:**

**Website:** <https://speechtools.co/voice-analyst>

**Monetization:** $14 one off App purchase fee.

**Key Take-aways:**

* Provides delayed feedback using AI.
* This app places a particular focus on pitch and volume.
* Quite a basic app, not many features.

**Other Notes/Thoughts:**

* Voice analyst allows you to organise your recordings into folders on the cloud. I feel like this is quite a valuable feature that I haven’t particularly seen on many other apps. This is feature is definitely something we should consider doing if it would make sense for our system.

**Samsung BeFearless:**

**Website:** <https://www.oculus.com/experiences/gear-vr/942681562482500/>

**Description:**

* **VR App** designed by Samsung to help users overcome their fear of public speaking.
* Provide a **series of scenarios** (Job interview, business lunch, team meeting, management presentation, and job fair). The app will respond to the volume of your voice, speaking pace, eye contact and **heart rate**.

**Key Take-aways:**

* In this app, virtual humans speak to you in different scenarios. It provides quite **a real simulation** of the real life experience.
* Uses AI to provide **Delayed feedback**. Example (This is for their ‘Fear of heights’ program but it is similar for the public speaking program as well):

Graphical user interface, application

Description automatically generated

**Other Notes/Thoughts:**

* This system analyses the users **heart rate**. I think this is quite an interesting idea. Perhaps if it would make sense, our system could integrate with a **smart watch** (since they are so common), and make use of heart rate to give an even higher level of feedback to the user. For example, if they have a really high heart rate, they could be provided with real time feedback telling them to relax.

**Conclusion and Key take aways from the market survey:**

**General Overview:**

The Majority of public speaking assistive interface systems that I found came in the form of a mobile app or a web app.

**What key indicators do automated feedback systems commonly look for?**

* Indicators in which feedback is given on most commonly includes:
  + Volume of speech/energy
  + Speaking pace (most commonly words per minute)
  + Number and length of pauses
  + Number of ‘filler words’ (um, like, etc. )
  + Clarity of speech
  + Pitch modulation.
* Uniquely, The ‘Samsung BeFearless’ system analyses the users heart rate. I think this is quite an interesting idea. Perhaps, if it would make sense, our system could integrate with a smart watch device (which are very common nowadays), and make use of heart rate to give an even higher level of feedback to the user. For example, if they have a really high heart rate, they could be provided with real time feedback telling them to relax.

*I do feel that looking at which systems are the most popular is a hugely beneficial indicator for identifying the best systems out there and what consumers value/ are looking for out of a system of this type. So…*

**What do the most of the successful public speaking apps and webapps all have in common?**

* They all have **good functionality, run smoothly and have a very clean, user friendly GUI**. If we do decide to develop an app, perhaps we can have a look at some of the apps I have discussed and understand **WHY** we like the interface and take inspiration and incorporate some of this into our system.
* Nearly all of them include some form of ‘**progress tracker’**, whether that is of the form of ‘daily warm ups’, ‘number of recordings’, ‘completed activities’, e.t.c. I believe this is quite key for users as if users can see **tangible progress**. It may make them more likely to continue using the app and reinforce to them the effectiveness of the app. The progress tracking feature is an interesting idea and is something I would feel would be a great thing to include if we decide to develop an app.
* Almost all are **all free to download** and include a paid **‘premium subscription plan**’ to access a greater number of features/ materials within the app. If an individual user who is looking to download a speaking assistive app they will most likely try a free one over one that requires an upfront fee, I believe that is a big reason as to why this business model proves more effective than others.
* Having lessons/courses included as a feature in public speaking feedback assistive applications seems to be quite a common occurrence. It may be because these two features when used in conjunction with one another can be more effective than simply just providing feedback. Given the popularity of the ‘TED Masterclass’ app which is **purely lesson based** and provides no feedback it may mean that it is a feature that consumers also really value. It really depends what we want our system to do, but its something to think about…
* Of the most popular systems that provide automated feedback on the market (not including academic projects such as ROC Speak, Cicero, e.t.c), the majority of them only provide feedback based on audio behaviour. This may leave a gap in the market for an automated public speaking feedback system that gives feedback based on visual behaviour via a camera.

**What Technologies did these systems use?**

Unfortunately it was unclear – Would you have any advice on how to go about researching this?

**What are the most common forms of monetization?**

* (Especially for systems that are in the form of an app), monetization is by far most commonly found in the form of a ‘premium subscription format’, where users pay monthly for access to extra features, materials e.t.c that are not available in the free/standard version of the system.
* The paid systems with a monthly fee ranged anywhere from $10 a month to $50 dollars a month for personal use, with the majority of these paid systems sitting at the upper end of this range.
* There was also a handful of paid apps that required a one-off initial fee to access all of the functionality of the system, this was in the range of $3 - $15.

**Systems I have not come across:**

*I am yet to find a non research system that analyses facial behaviour and body language to provide public speaking feedback.*

**For systems that analyse visual behaviour using a form of camera do they look at full body or just the face?**

*Of the systems that analyse visual behaviour I have come across a variety of different ways of doing it. For example, RocSpeak uses a camera to record the upper half of the users body, predominately looking at the face, whereas Logue makes use of a Kinect sensor and google glasses to record the users full body to provide* realtime feedback on the presenters’ openness, body energy and speech rate during public speaking.

**What systems that I have explored that are able to analyse an interaction between two people speaking (for interviews e.t.c)?**

*The ‘analysing and providing feedback for two speakers at once’ feature does not seem to be common at all. VoiceVibes is the only one I have found that claims is able to track, analyse and provide feedback on two speakers at once, although I have not tried it for myself.*

**Other App/ System Ideas to think about:**

* ROC Speak has quite a basic GUI and as far as I can tell was created purely for research purposes. – perhaps this could be an interesting thing to develop into a mobile app. It has already been done but might be popular if it was repackaged for market in the form of a mobile app or extension for zoom. Would there be any way we could work in collaboration with Rochester HCI to build upon their system, improve the GUI and develop it into an app for the IOS marketplace or even an extension for zoom?
* There doesn’t appear to be a Public speaking assistive interface that gives feedback on both visual and audio behaviour on the zoom market place – But I do not know this for sure though.
* Perhaps an app where users submit their presentations and provides features for real life professionals (could be public speaking coaches, the users’ Boss e.t.c) to provide feedback on their performance.
* Perhaps we could include a ‘teleprompt feature’ (See Prompt Smart: <https://promptsmart.com/> )
* Perhaps a system that provides automated feedback based on visual behaviour could go through a “quality control” process where individuals review the results before they are returned to the user – most likely not feasible though.

Link to

<https://docs.google.com/forms/d/1-Ie2eWxUJHYX6EGbXiveBAAoLCb1eDYN3695ujor0yg/edit>