

The OptiCue Speech Tool

(Working Title)

JUSTIN LEE – INTERVIEW, DESIGN & ANALYSIS;

ALIREZA HASHEMLOO – INTERVIEW, DESIGN & ANALYSIS;

YING QING SET – INTERVIEW, DESIGN & ANALYSIS

Problem and Solution Overview

Communication is the cornerstone of successful interpersonal relationships, but it is often the case we are unaware of what we say and how we say it. Most people don't have planned scripts for what they want to say or analyze what they've already said, instead they just *respond*. The fact is many people are not wholly effective at speaking with other people. They may not have the right things to say, or they may not know how to say it, or knowing both of these things still botch their delivery.

Our proposed solution is a device in the form of eyeglasses that will allow for a user to both record and actively monitor their speech and have a display that can provide notes, tips, alerts, and notifications regarding what they are saying and how they are saying it.

Contextual Inquiry Target, Stakeholders, and Participants

The contextual inquiries were focused on determining what undesirable speech patterns exist and what steps people currently take in order to fix them. At the time of the inquiry we imagined that the primary design users would be individuals looking to improve their speaking ability, other stakeholders would include public speakers, speech therapists, speech clinics, and any other professional with an interest in helping people speak well.

Our participants were chosen to include normal people who can improve their speaking ability, and professionals that seek to help those that wish to improve. This resulted in speaking with two students, and two professionals.

SID | MAJOR: PHARMACY | STUDENT AT THE UNIVERSITY OF WASHINGTON

INTERVIEW ENVIRONMENT: CASUAL CONVERSATION IN NORDHEIM COURT

This first interview is unique in that they were not told until after speaking for some time that the interviewer was actually taking note of their use of filler words, and then discussed these habits. This allowed us to determine that people are typically unaware of their habits.

JESSICA | MAJOR: PSYCHOLOGY | STUDENT AT THE UNIVERSITY OF WASHINGTON

INTERVIEW ENVIRONMENT: CASUAL INTERVIEW SESSION IN PACCAR HALL AT UW CAMPUS.

This interview was unique in that it was the first conversational interview regarding factors that detract from effective communication. This interview and the one before are described in the Contextual Inquiry Check-In from assignment 2c. This brought our attention to the situational aspect of speech.

HANNAH | MAJOR: COMMUNICATIONS | TUTOR AT THE UW PUBLIC SPEAKING CENTER

INTERVIEW ENVIRONMENT: INTERVIEW AT THE PUBLIC SPEAKING CENTER

The tutor, Hannah, described some of the habits they try to solve in preparing the students for public speaking. The demographics of the students who come to the center she noted to be primarily international students, with nearly all of the students who came to the center being shy. When asked to describe what is done in order to help these students, a lot of practice, repetition, and teaching of organizational tactics are used. This was our first expert related to fixing public speaking problems, and allowed us to think about how we might aim to solve these problems.

PROFESSOR MATT MCGARRITY | PHD COMMUNICATION AND CULTURE, INDIANA UNIVERSITY, 2005; UW SENIOR LECTURER, INTRODUCTION TO PUBLIC SPEAKING; FOUNDER AND DIRECTOR OF THE UW PUBLIC SPEAKING CENTER

INTERVIEW ENVIRONMENT: INTERVIEW AT HIS OFFICE

This interview was unique in that in the realm of public speaking he was the most valuable expert that we hoped to reach. In contrast to our previous thoughts, in addressing vocal segregates (uh, um, ohh, etc) he states emphatically that the presence or absence of these segregates does not matter, and that instead the contexts in which they are used is more valuable. His approach focuses on teaching good judgment on recognizing and acting *in context*.

Contextual Inquiry Results and Themes

The common theme we have seen is that speech habits are universal; while most people are unaware of them in practice most people want to improve their public speaking ability regardless. It is also increasingly apparent that the social context in which someone is speaking is *the* most important factor. All these inquiries agree that problems exist in people's ability to communicate with each other, as evidenced by messages getting lost, misinterpreted, or worse. But the problems aren't as simple as the mere presence or absence of filler words and nervous tics, and would require large amounts of data and human analysis to determine what makes someone an effective speaker. The most common practice seen in improving speaking ability is practice. While organizational techniques and forced awareness of certain tics can be seen as helpful, what's most important is the development of good judgment. Next is a description of the contextual inquiry held with Professor McGarrity, which is the most valuable and credible inquiry thus far.

Professor McGarrity states that good speaking equals good judgment. The judgment to in the moment, naturally decide speed, cadence, intonation, pitch, tone, cues, and more in order to speak appropriately and effectively for the context at hand. So in order to teach someone to be a good speaker, requires a humanistic approach and much practice in order to develop that judgment. He uses the analogy that while coaching a soccer player, while you can show them the analyses of how to kick a ball, and describe what formations lead to the most effective strategy, they won't get better unless they actually play.

He then describes the Toastmasters organization, where in each meeting one person is designated as a "buzzer" to sound off whenever someone uses one of the aforementioned vocal segregates. In his professional opinion, while they *think* this makes them better speakers, and while people *think* speaking statistics such as words per minute, word counts, and the like are important, these are *not* important in the context of effective speaking. When then asked if there are any tools he'd like to see in this space, he responded that it is difficult to replace human feedback on a person's speaking.

As the team moved forward in the design process, it has become more evident that the sort of tool we previously considered for constantly recording and monitoring one's speech data throughout the day has extremely limited use and the amount of data gathered that would also have to be analyzed in order to provide any meaningful benefit would dissuade most potential users. However it remains that a sort of tool that will allow a user to see some sort of feedback and information on their speaking would be wanted for certain situations.

Task Analysis Questions

1. Who is going to use the design?

Anyone looking to improve their speaking skills is a potential user, whether it's in preparation for a speech, presentation, or day to day communication, we could provide the tools to improve. As the design acts as an information display on both what to say and how to say it, the team is looking at professions that require a lot of speaking and information retrieval, such as reporters, politicians, lecturers, and more.

2. What tasks do they now perform?

People still have to communicate on a day-to-day basis, but often will make no attempt to change their habits unless confronted about them, or in preparation for certain contexts. So they might practice in front of a mirror, record themselves on video, or practice with friends. The few who are very motivated to improve public speaking ability will find lessons, tutors, or clinics in order to improve. They might be using notecards or Teleprompters to guide content, and rely on visual cues from staff members to adjust how they are speaking, such as hand gesture to indicate the speaker needs to be more forceful, or quiet.

3. What tasks are desired?

The primary desired task is that customers want a way to receive feedback on their communicative ability. Customers also want statistics on their speaking, although as stated previously by Professor McGarrity, this data isn't necessarily helpful but is wanted nonetheless. They want to monitor their speech, and see information on what they still need to talk about.

4. How are the tasks learned?

Most users learn by speaking with other people to practice how they speak, but learning speaking content depends on the context. Such as a politician reviewing notes on prewritten answers to debate questions, or a sports reporter having player statistics read to him over a headset.

5. Where are the tasks performed?

The tasks that revolve around speech can be performed in any environment or context, but the team's design will want to focus on professional settings—speeches, presentations, interviews, debates, press conferences, news stations.

6. What is the relationship between the person and data?

The data we would observe falls under two categories. The first is data on the context in which the speaking is taking place, and the second is data on content and delivery of that speech. The contextual data is then primarily observed by the person, and should be used by them in order to determine how to shape their method and delivery of communication. The speech data is produced by the person, and would require some method of recording that data for later analysis, and this includes what they say in addition to how they say it.

7. What other tools does the person have?

They have internet access to try and find tips and lessons to improve, the people around them to speak and practice with, and professionals they can seek out to work on their speaking ability. They

might have notecards and Teleprompters to guide content, or in the cases of some reporters, headsets and cue cards.

8. How do people communicate with each other?

People communicate in a variety of ways, but our focus is on the verbal and nonverbal aspects of speech, as opposed to written word or other mediums. In professional settings they may hear feedback from supervisors and peers and coworkers.

9. How often are the tasks performed?

People speak nearly every day of their lives. The frequency of something like public speaking however, is more dependent on the individual, but good speaking habits carry over into every day of our lives. For our probable customer however, it depends on their profession. A reporter may want to use this daily, a politician only during press conferences.

10. What are the time constraints on the tasks?

This can be a continual process of constant self-improvement, or constrained by the urgency of their need to be prepared, such as practicing for a presentation later in the week. The ability to perform a speech can be greatly improved by practice over a relatively short period of time, but the judgment needed to be an effective speaker and communicator on a day-to-day basis is learned and absorbed over a much greater period of time. There would also be the time required to determine and then transmit the data that might go on the display, such as having another person look up a statistic they then send to the user's display that they can then use.

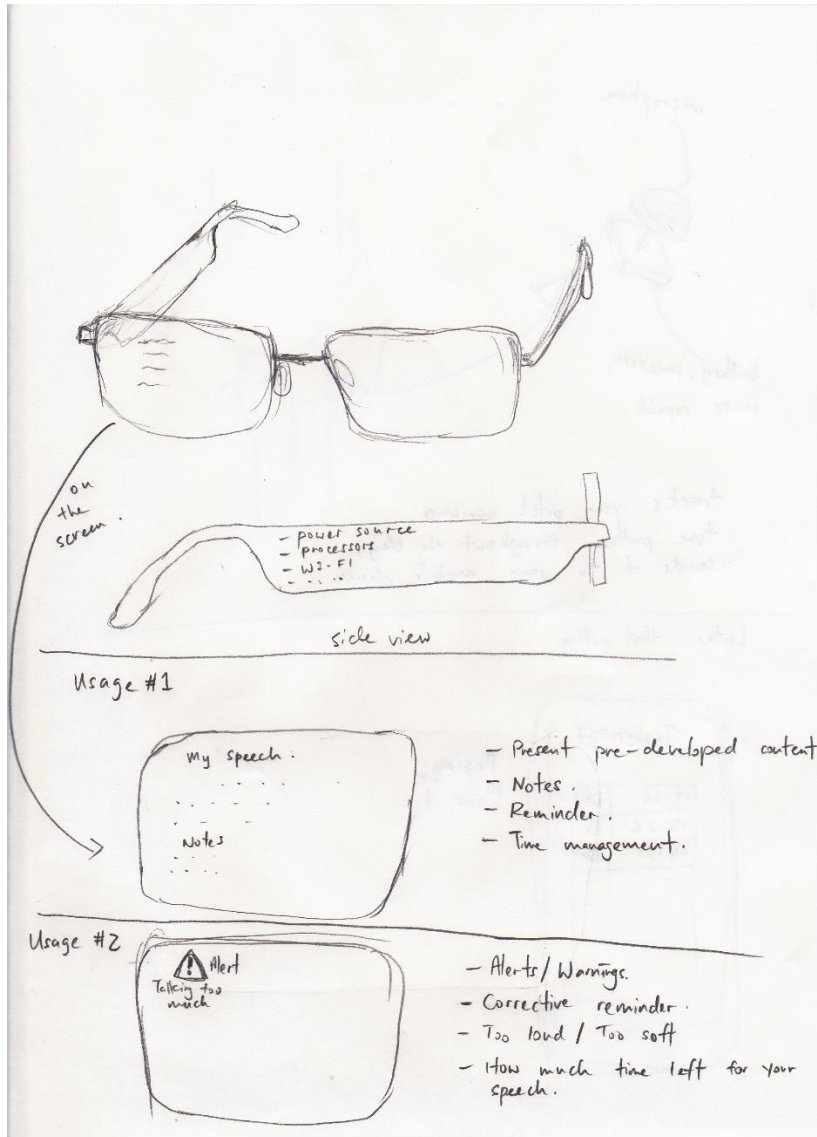
11. What happens when things go wrong?

People make mistakes, can't get their message or ideas across, and depending on the context this can be forgivable or disastrous. This could be flopping a major presentation or pitch, resulting in major business losses, or an argument caused by a misunderstanding. When people cannot speak well, they will often be associated with nervousness, insecurity, anxiety, and a whole list of negative traits. The information on the user's display may show up too late, or not be helpful, and this may also affect the user's ability to speak on some subject, or the notifications and tips it gives to improve the user's speaking may also be ineffective.

Tasks

1. **Monitoring and optionally recording, replaying user's speech.** This feature allows the user to review the way that they speak. For example, like how you would practice for a speech in front of a mirror, or in front of a friend, a user could record themselves then review it to see how they sound. This task allows the users to record themselves speaking naturally without the need to hold a phone or be in front of a static microphone.
2. **Determining what makes someone an effective speaker, and what detracts from being an effective speaker.** Many people have no definite idea of what makes someone an effective or confident speaker, and after determining what kind of speaker they wish to be, they can learn more concretely what they should or should not do when speaking, and guiding a user to follow good habits.
3. **Keeping track of statistics on speech patterns and making the results accessible and human readable.** The user would expect to be able to see some sort of analytics on their speech habits. Although as explained by Professor Matt McGarrity, the sort of statistics such as words per minute, number of "umms", "uhhs", or "likes", are fairly irrelevant in the context of being an effective communicator, they are the sorts of things a user would want and expect to see nonetheless. In addition to examples mentioned, this could include contextual data, such as who you are speaking to, where, and why.
4. **Providing a platform for users to get feedback on their speech.** Inspired by a project where Brazilian students are connected with Americans living in retirement homes in order to create a safe and welcoming platform for language exposure and practice <https://www.youtube.com/watch?v=-S-5EfwpFOk>. This could be for getting feedback on a speech for a class, preparation for a business presentation, or improving language skills, but will allow users to get real feedback from people in order to improve by connecting people who need to talk with people who are willing to listen.
5. **Allowing users to practice and learn the native cadences, pitch variation and other patterns.** While this could be considered falling under the third category, as it is a type of analysis on speech data, specific focus is placed on intonation patterns and pitch variance. The importance of this specific data warrants it being categorized as its own task. In the context of learning native speech patterns or in the context of becoming a more expressive speaker and avoiding a monotone voice, knowing and being able to compare a user's speech to samples of expressive, native speakers can allow the user to learn how to speak more naturally.
6. **Improving users' speaking vocabulary.** People who are learning to speak another language will often state that their listening vocabulary is much larger than their speaking vocabulary. Or stated more simply, while they are able to understand much more of what they hear, they feel more limited in what they know to say. Users may start with a set of words or phrases they use most commonly as their go-to responses, but soon build up their repertoire of responses and no longer feel limited to a small set of sentences they are comfortable with.

Proposed Design Sketches



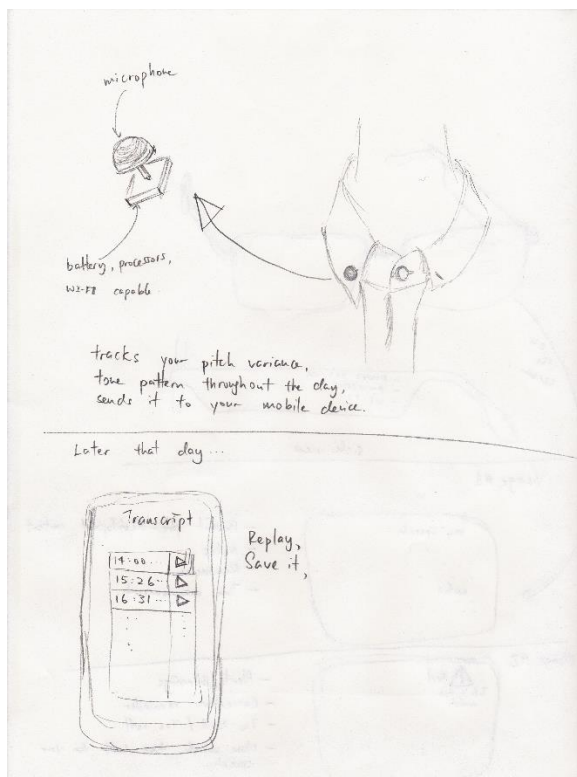
Design (A): This design is a platform for connecting people who need feedback on their talking, and those willing to provide that feedback. This design is merely a connecting tool giving people a place to get feedback from many other human users who will be able to provide more honest suggestions as they would not be biased compared to the speaker asking for advice from a friend.

It addresses **task one, monitoring and recording one's speech**; this is addressed by having the user speaking in front

of a camera, which can optionally record portions for them to play over, and in showing the user on the screen is similar to practicing with a mirror and they will be monitored by other users who are watching and providing feedback. This also addresses **task two, determine what makes someone an effective or ineffective speaker**; this can be addressed by hearing the feedback from other users and the advice they have. This is also the primary focus of **task four, connecting people in order to get human feedback on their speech**; this is the core idea of this design and is accomplished when the speaker has someone in his chatroom to speak with and get feedback from.

Design (B): Glasses are used in this design to allow a user to have a private, more innocuous display in order for the user to see notes on what they plan to say, other tips, alerts, and notifications they wish to see in order to become a better speaker. The glasses also integrate a microphone, battery system, and connectability with the user's phone. By having the glasses display notes on their speech, they can speak smoothly without the interruption of looking down at notes, alerts could indicate whether the speaker is talking too quickly, and display a timer to help improve pacing.

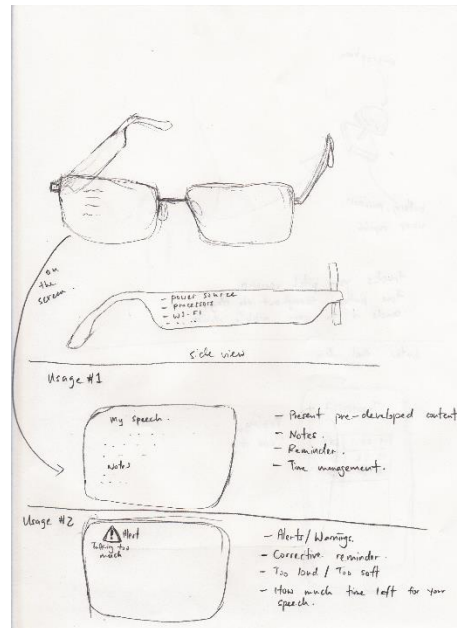
This addresses **task one, monitoring and recording one's speech**; this task is addressed by the microphone integrated in the glasses which will be able to listen to what is being said, and display notices or information as necessary. This also addresses **task two, determine what makes someone an effective or ineffective speaker**; as the notifications and information displayed will be able to guide the user towards better speaking, whether it's ensuring they won't forget the content of their speech, or giving indicators when they are speaking too quickly.



Design (C): This final design uses a type of collar-pin or clip that can be unobtrusively attached to the user's clothing. This device contains a microphone, can be synced with a phone, and contains a power source. This device is intended to passively record the user's speech data to be reviewed at a later time. This is what differentiates it from design 2, which acts as a head's-up display for active awareness and response to one's own speech, this design focuses on the idea of record throughout the day, review later at night. This then allows the user to behave and interact with others normally, and in this case gather selfinformatics over a period of time on the entirety of one's speaking habits, rather than just the preparation for speeches.

This design focuses on **task one, monitoring and recording one's speech**; this task is the core of this design, which records all the speech data throughout the day of the user. The user can later in review, replay any of this speech and hear how they speak. The second focus of this design is **task three, collecting speech analytics**; after taking this large amount of speech data throughout the day, various statistics and patterns can be pulled out and analyzed by the user. This could also flow into a cloud based database for researchers looking into human speech patterns and habits and provide them a massive treasure trove of speech samples from various users on various days. This design can also address **task five, analyzing pitch variance, and learning native speech patterns**; by recording, hearing, and analyzing one's speech throughout a day, the user can review portions of their speech and practice and determine areas to focus on. Through this awareness, they can also work on improving their fluency and natural speaking.

Selected Design.



The design chosen is design (B). The reason for choosing this design is based primarily on ensuring our design has a clear user base and is solving a specific problem. While we were initially focusing on something based on design (C) to record and review speech data throughout a day, issues brought to our attention during the critiques made it more evident that the user base for such a device would be extremely small and focused on “life-loggers”. The amount of data it would create and subsequently have to analyze would also be immense, and the benefits gained may not be worth the hurdles that would need to be faced. Design (A) would also require a lot of work in find ways to incentivize users to use the service, and to screen them. A primary reason for the team to choose this design is that it gives the user the ability to find and see information in real-time on a head’s-up display, which also allowed our user group to become more clearly defined. This design allows us to focus on professions such as reporters, sportscasters, politicians, lecturers, and others who may need to bring up information on the fly and also be able to record their speaking for later review and see information on their speaking that may allow them to improve, as speaking is a very large part of their jobs.

This design focuses on **task one, monitoring and recording one's speech** and **task two, determine what makes someone an effective or ineffective speaker**. These tasks have been deemed the most important as task one captures the core idea of personal informatics in speech; it is about gathering information on how the user talks, and allowing them to view that information, and also guide and improve their talking. The second task goes in tandem with the first, as by allowing the gathering of such information and also being able to use this as a tool for viewing content to speak on, the user can improve their speaking ability and be better prepared, and this tackles the major want and need of becoming a better speaker.

Storyboards and Scenarios

This design focuses on **task one, monitoring and recording one's speech** and **task two, determine what makes someone an effective or ineffective speaker**.

A possible scenario:

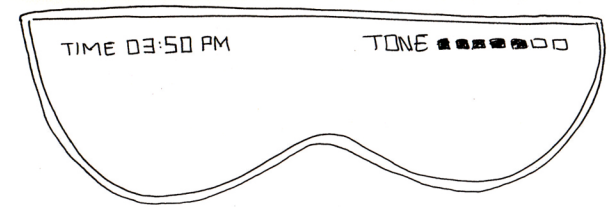
The user wants to know his/her speech pattern and is keen to make some kind of improvement. The user would then activate the recording and monitoring features of this glass and later that day, or later that week, the user could open up the log and check out the analyzed data and based on these data, the user could adjust what kind of improvement he/she would like to apply, and the glass would remind the user about his/her goal from time to time.

During the user's presentation, instead of using a cue Card, the user could have the information needed for the presentation displayed on the glass, that way, the user could increase the effectiveness of their presentation as they could keep eye contact with the audiences while looking at the information at the same.

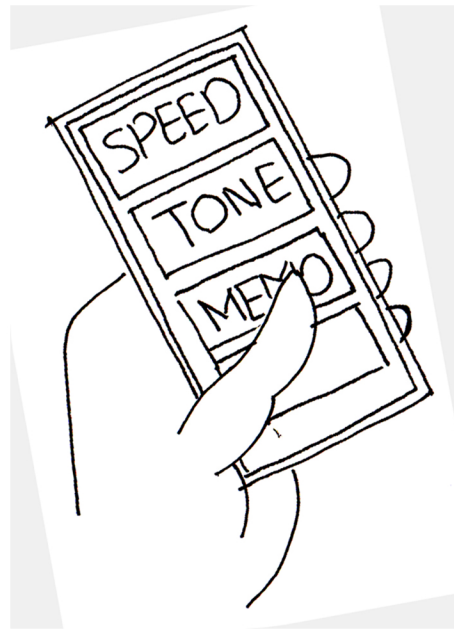
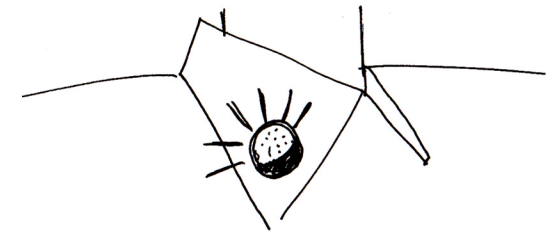
Storyboards and Scenarios



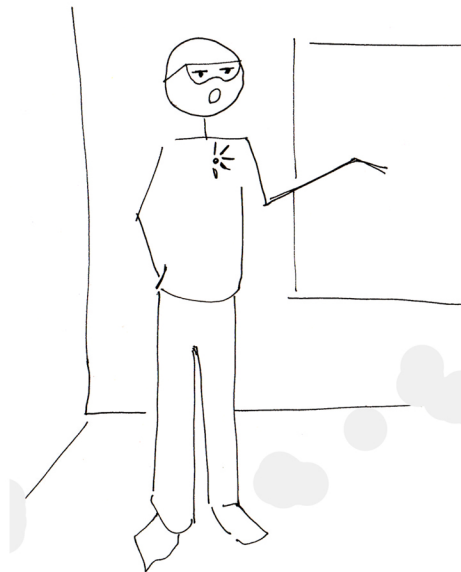
1. Daily conversations are recorded and evaluated for both real-time and post-speech analysis



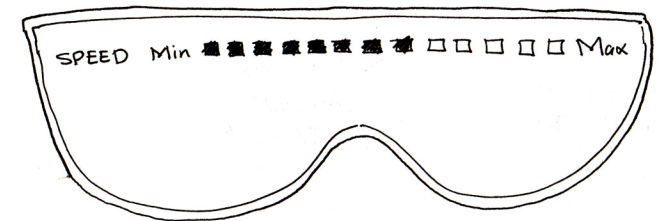
2. Real-time speech data received by the microphone and displayed via the glasses.



3. Speech analysis options for recorded speech



4. Formal speech data is recorded for both real-time and post-speech analysis



5. Display of data based on exercised speech qualities