Assisting People with Public Speaking Using Smart Phone

Debarun Das
DED59@pitt.edu

Anuradha Kulkarni
ANK163@pitt.edu

ABSTRACT

Fear of public speech has led to the development of different speech feedback technologies. However, current feedback systems using a device like Google Glass is expensive, may cause headache and has the potential to distract both the audience and the speaker. To resolve this, we implemented a feedback system using different vibration themes of smartphones. We conducted a user study to analyze the distraction, pace, volume, engagement of speech and other related parameters using the feedback system, and compared them to the case when the speech was given without using the feedback system. The results show that while the clarity of speech has improved in speech with the feedback system, the distraction of the speaker is higher when compared to speech without the feedback system. However, the distraction of the audience in both the cases (with and without feedback) is almost similar.

Keywords

Public Speaking; Google Glass; Vibrational Feedback; User Study; Speaker Feedback; Audience Feedback; Video Feedback.

INTRODUCTION

It is said that 'an average person ranks the fear of public speaking more than he fears death' [1]. This fear can have a greater impact professionally and personally. There are techniques to overcome these fears which require rigorous commitment and time. In addition to presenting meaningful content, the speaker needs to modulate their volume and vary speaking rate to retain the audience's attention.

To address the above problems, there has been development of different technologies to provide feedback to the speaker during speech. One such technology is the use of Google Glass to provide feedback to the user during speech [2]. The feedback provided is visualized by the speaker in the Google Glass device that he is wearing, and he can adjust the volume and/or the rate of the speech accordingly. However, there are some drawbacks of using Google Glass for this purpose. *Firstly*, Google Glasses are expensive. *Secondly*, they have the potential to distract both the audience and the speaker during a speech. Distraction of audience may be caused because of the general tendency of them skeptical about being recorded by the device. Moreover, Google Glass, being a new and expensive device, most people are not accustomed or comfortable to see the speaker using the device. *Thirdly*, it has the potential to cause distraction to the speaker himself who wears it [3]. *Fourthly*, using Google Glass has some health concerns. Wearing the device may cause eye strain or headache [4].

To address these concerns, we implemented a simpler mechanism to provide user feedback during speech. We used different vibrational themes for to provide feedback to the speaker based on the speaker's volume and/or rate of speech. There are multiple reasons for choosing a smartphone for this purpose. Smartphones are less expensive when compared to Google Glass. Also, feedback using vibrational themes provides a simpler system design for use.

We conducted a within subject design for the two different cases, where the speakers gave a scripted speech with the feedback mechanism and without the feedback mechanism. The results were then analyzed after a threefold evaluation system (speaker feedback, audience feedback, video feedback). The main contributions of this paper are as follows:

- We present a simple speech feedback mechanism using vibrational patterns in smartphones. This information allows the user to vary their volume and pace of speech accordingly.
- We explore different vibration patterns and finalized 4 different patterns to help the user recognize the different feedbacks.
- Our experiment with 6 participants demonstrated that the feedback mechanism helps to improve the clarity of the speech and also does not cause any significant distraction to the audience.

RELATED WORKS

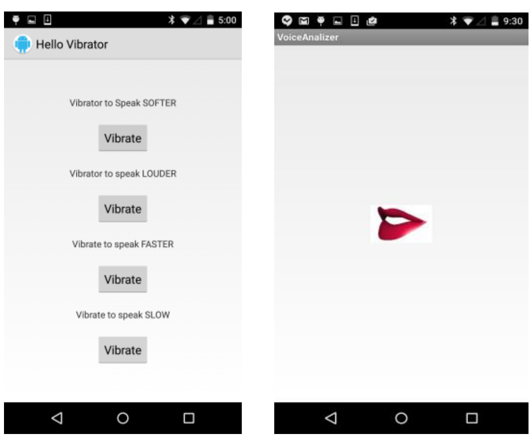
Different feedback strategies for improving public speaking has been explored by Chollet *et al.* in [5]. The authors mainly explore three study conditions: a non-interactive virtual audience, non-verbal virtual audience feedback and a direct visual feedback. The results show that interactive virtual audience feedback shows the best results [5].

Using Google Glass for the purpose of giving speech feedback was discussed by Tanveer *et al.* in [2]. The interface developed by them automatically detects the volume and rate of speech in real time to give feedback to the speaker. Two different feedback strategies were tested – continuous and sparse feedback strategies. The results showed that the audience was significantly more satisfied with the sparse feedback mechanism with respect to the continuous feedback strategy and no feedback strategy [2]. Based on this, we implemented the sparse feedback mechanism where the user is given feedback based on his/her average speed and loudness over a period of time. However, using Google Glass as the feedback mechanism has its share of disadvantages as explained in the previous section.

Exploring the design of efficient vibration based notifications was done by Saket *et al.* in [6]. The authors studied the effect of ten different vibration patterns to determine the urgency of the different vibrations as perceived by the users. The results showed that for effectively communicating urgency of vibration to the user, the vibration design pattern should be simple with the least number of on-off signal pairs. The discussion of different strategies for designing effective vibration patterns provided insight to develop our own custom vibration schemes.

TECHNICAL DETAILS

We implemented two Android applications (in Android Studio 2.2.2) for conducting the user study. The first app, named *"Hello Vibrator"*, has the option to play the four different vibration patterns [7]. There are different vibration patterns for speaking LOUDER, FASTER, SLOWER and/or SOFTER. The interface design of the app is simple (shown in Figure 1a.). The vibration pattern for speaking *louder* consisted of a single long vibration (2000 ms). For speaking *softer*, the vibration pattern consisted of five small vibrations (60 ms) with a gap of 150 ms in between the vibrations. For notifying the user to speak *faster*, the vibration pattern consisted of 3 long vibrations (500 ms) with a gap of 300 ms in between each of them. For speaking slower, the vibration pattern that was given as feedback included four small vibrations of 40 ms each with a long gap (1000 ms) between each of them.



a. b.

Figure 1. The two apps: a) "Hello Vibrator" and b) "Voice Analyzer"

The second app called the "Voice Analyzer" (shown in Figure 1b.) is used to give feedback to the user based on his/her speech [8]. The loudness of the speech is determined by the average amplitude of the sound signal over a fixed value of a counter. The average speed is determined by the average frequency of the speech over the fixed value of a different counter. In the implementation, we have considered a threshold value of volume over which it is considered a valid sound of speech. This allows us to filter most unwanted softer noises that may occur nearby.

METHOD

To measure the effectiveness of our feedback mechanism, we conducted a user study (within subject design). The experimental setup is described below.

PARTICIPANTS

We recruited 6 participants (4 males, 2 females) for the user study. Five of the participants were non-native English participants and one of the participants was a native English speaker. All the participants were from the Department of Computer Science, University of Pittsburgh.

APPARATUS

The device used for the purpose of giving vibrational feedback was Google Nexus 5 (Android 5.0.1, API 21).

PROCEDURE

The procedure of the experiment is explained in the following steps.

- The participants were initially briefed about the purpose of the experiment.
- The participants were then given 5-10 minutes to get accustomed to the four different vibration patterns using the *"Hello Vibrator"* app.
- When the participants felt comfortable with recognizing the different vibration patterns, the smartphone was tied to the upper arm, to maintain an almost constant distance between the mouth and the smartphone's microphone.
- The participants were then given a script to read, with the feedback and without the feedback system. The feedback was given using the *"Voice Analyzer"* app. Some participants spoke with the feedback system first and then without the feedback system. The rest of them gave the speech in the opposite order. This was done in order to account for the learning effect.
- After the two speeches, the speaker is required to answer some speaker survey questions (score in the range 1-5).
- Some of the participants also served as the audience when another participant was speaking. After the speech, all the members of the audience were required to answer some audience survey questions.
- Each of the speakers was video recorded (with the speaker's consent) when they gave the speech. This video was later analyzed by a select set of users and they were required to answer some video feedback questions based on the video analysis.

DESIGN

The experiment was a within subject design, where the user gave the speech with and without the feedback system. Following the Black Box metaphor, input is the user's speech and the output is the vibrational feedback pattern. The independent variables are the two cases (with and without feedback schemes). Some of the dependent variables are the volume of speech, the pace of the speech, distraction and engagement of speech, etc.

EVALUATION

The main goals of evaluation were:

- Effectiveness of the interface for user's performance.
- Evaluation of noticeable distractions.
- Ease of use of the interface.

The evaluation process was threefold. Firstly, we evaluate the feedback provided by the audience, then we evaluate the feedback provided by the speaker and finally, we evaluate the video feedback

AUDIENCE FEEDBACK

The audience filled in the audience survey questions after each speech. These questions focused mainly on comparing the performance of the speaker for the two cases (with and without feedback). Some of the questions included "How engaging the speech was (with and without feedback)?" and "How distracted the audience was (with and without feedback)?". The audience was required to give a score between 1 and 5.

SPEAKER FEEDBACK

The speaker answered the survey questions related to the feedback system after the speech. The questions designed for this feedback focused only on the different parameters of the vibrational feedback system ("how easy it was to remember the vibrations?", "how effective was the feedback system?"). The speakers were required to give a score between 1 and 5.

VIDEO FEEDBACK

The video of each speaker was analyzed by a select set of users. The main aim of this evaluation was to determine how distracted the speaker was (based on the number of pauses in speech, etc). Also, it helped us study the comparison of the quality of the speech for the two cases.

RESULTS AND DISCUSSION

We analyze the results obtained in the three feedback schemes, Audience Feedback, Speaker Feedback and Video Feedback.

AUDIENCE FEEDBACK ANALYSIS

Figure 2 shows the plot for the ratings provided by the audience in the Audience Feedback Survey form. We also conducted 'F test Two-Sample for Variances' and 't-test: Paired Two Sample for Means' for analysis [9, 10].

We found that the speech engagement with feedback is significantly more than without the feedback. (t(5) =3.06, p=0.028). This confirmed by the plot in Figure 2, where we see that the speech engagement is higher with the feedback system than without the feedback system.

There is no significant effect on the audience distraction during the speaker's speech. ($F_{1,5}$ =1.7, p>0.05) We found that there is no significant difference, i.e. the difference in the mean rating for distraction from audience perception with and without feedback is not convincing enough to state that there is a significant difference between them (t (5) =-0.2, p=0.84). This is further confirmed by the plot shown in Figure 2, where we see that the audience distraction rating is almost similar in both the cases.

Further, we found that there is a significant effect on the pace of speech on the completion of the speech. ($F_{1,5}$ =0.429, p=0.0024) We found that the pace of speaker's speech without feedback is significantly more than the pace of the speaker's speech with feedback. (t(5) = 3.16, p=0.025). This is confirmed by the plot in Figure 2.

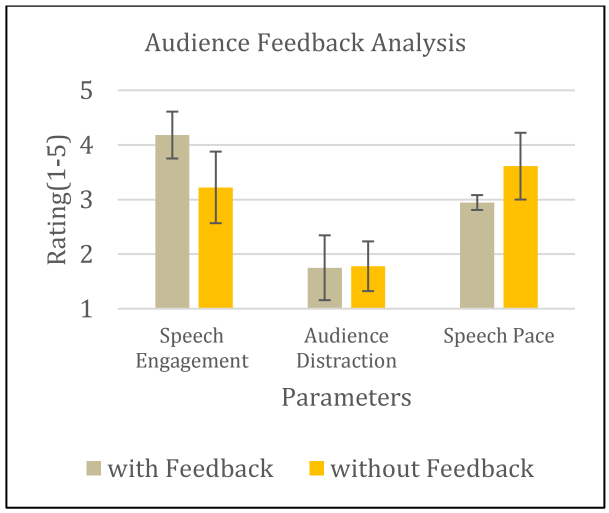


Figure 2. Audience Feedback Analysis Plot

SPEAKER FEEDBACK ANALYSIS

Based on the feedback provided by the speaker after the speech (shown in Figure 3), we can determine that it is neither too easy nor too difficult (moderate rating) to remember the vibrations. Interpretation of vibration to perform the required action (like speaking louder or softer based on their respective vibration patterns) is not too easy (low to moderate rating) for the speaker. Rating for ease of use of the feedback system is moderate to high. This is attributed to the simple design of the feedback mechanism. Also, based on the plot in Figure 3, we find that the users were mostly anxious to use the vibrational feedback system (moderate to high rating). This can be attributed to the fact that the users needed more time to get habituated to using the system. Also, comparison of this parameter with respect to that of another new feedback system would have helped us in analyzing the effectiveness of our system with respect to

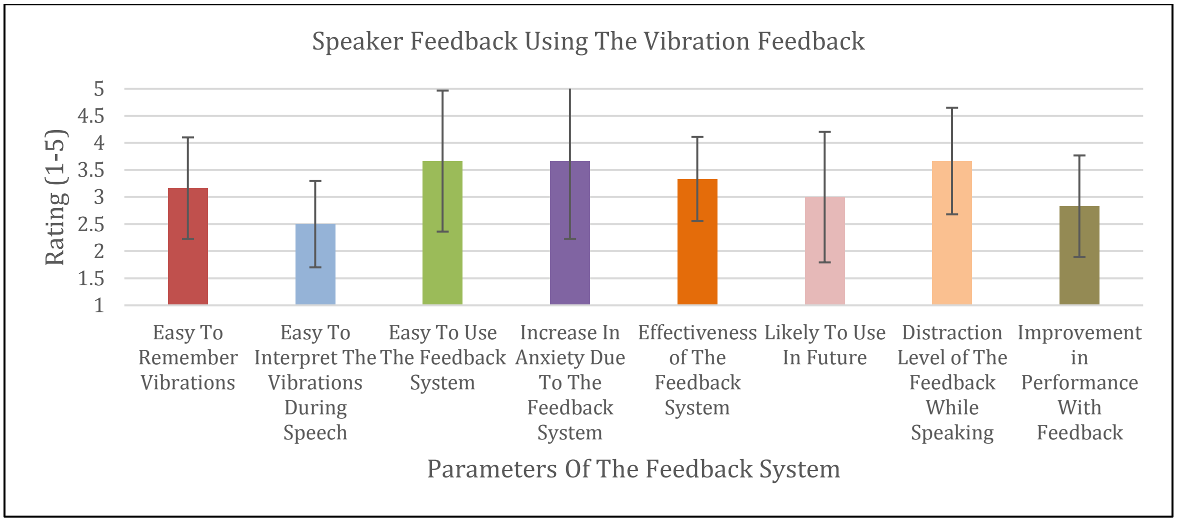


Figure 3. Speaker Feedback Analysis

this parameter better. Further, we see that the effectiveness of the feedback system is moderate. Most users were undecided whether they were likely to use the feedback system in the future (moderate rating). The rating for feeling distracted while speaking is moderate to high. This implies that in the view of the speaker, most of them felt that they were distracted while giving speech using the vibrational feedback system. Finally, the user ratings indicated moderate score for the parameter that indicates whether their performance improved with the use of the feedback system. This implies that most users were probably unsure about their performance improvement with the use of the feedback system.

VIDEO FEEDBACK ANALYSIS

Figure 4 shows the plot for video feedback analysis. We also conducted F test and paired two-tailed t test for analysis [9, 10].

There is no significant effect on the pause in the speech during the speech. ($F_{1,5}$ =0, p > 0.05). The observed difference in both the mean value is not convincing enough to say that the pause effect on speech with feedback and without feedback differ significantly (t(5) = 2.42, p=0.059). This is also seen in the plot (shown in Figure 4).

There is a significant effect on the speaker's distraction during the speech. ($F_{1,5}$ =7.809, p=0.02) between the two cases (with and without feedback). However, based on the video feedback's t test values, we found that we cannot confirm if there is a significantly huge difference on the mean rating values between the two cases (t(5) = 1.557, p = 0.18).

There is no significant effect on the speaker's stiffness on completion of the speech using both the mechanism. ($F_{1,5}$ =4.2, p>0.05). Based on t test values, we found that the observed difference in the means is not convincing enough to say that speaker's stiffness is found to be more with feedback than without feedback mechanism. (t(5) = 0.79, p=0.465). We can confirm the same from the respective plot in Figure 4, which does not show much difference between the ratings in the parameters of the two cases.

There is a significant effect on the volume of the speaker on the completion of speech using both the mechanisms. ($F_{1,5}$ =0.08, p=0.009). However, based on the t test, the difference in values is not fully convincing enough to significantly say one feedback mechanism is 'highly' better than the other. (t (5) = 1.938, p = 0.11).

Based on the t-test for the speed of the speech, we found that the speed of speaker is significantly more when the speaker is speaking without any feedback mechanism than using the feedback mechanism. (t(5) =8.944, p < 0.05). This is confirmed by the respective plot in Figure 4.

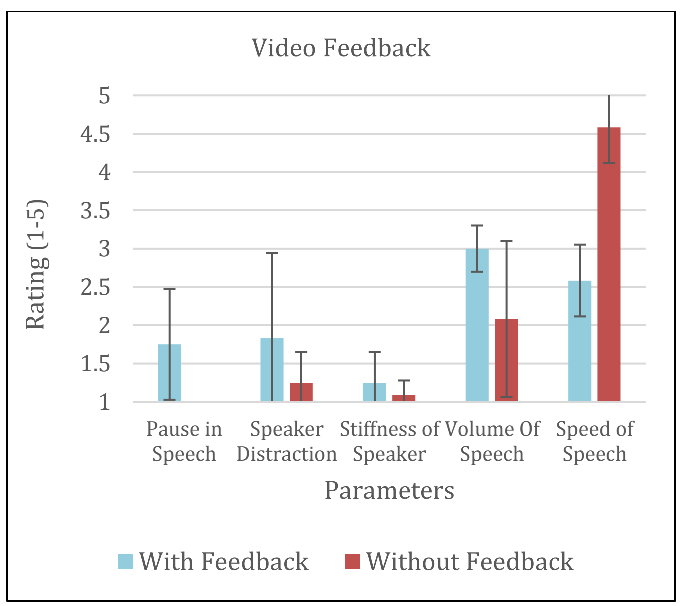


Figure 4. Video Feedback Analysis

CONCLUSION AND FUTURE WORK

In this project, we implemented a speech feedback system using vibrational patterns with a smartphone. The major motivation of the project was to counter the drawbacks of using modern technologies, like Google Glass for this purpose.

Based on the Audience Feedback analysis, we find that using the feedback mechanism during speech does not cause any significant distraction to the audience, when compared to the case with no feedback. Based on both the Audience and Video Feedback analysis, we observe that the pace of the speech is moderate when using the feedback mechanism (compared to a higher speech pace when speakers used no feedback). Also, from the video feedback analysis, we find that the volume of the speech is moderate with feedback mechanism, compared to lower volume without feedback. So, this implies that the clarity of the speech is higher with the feedback mechanism (moderate volume and pace). This is also confirmed by the results obtained in Audience feedback analysis for speech engagement.

Some of the possible future works include testing the system with the participants giving an actual unscripted speech. This would help to better analyze the effectiveness of the system for a giving an actual public speech. Another extension of this work is to implement the feedback system using a smartwatch. Using a smartwatch for this feedback mechanism will be an interesting choice because it would help the user to better detect the vibration patterns and has the potential to cause lesser distractions. However, a possible challenge in using a smartwatch is to determine an appropriate way to make sure that the speaker's voice reaches the smartwatch's microphone from a constant distance. A possible solution is to use a dedicated microphone (having a connection to the smartwatch). Furthermore, we can compare the performance of our feedback mechanism with the feedback mechanism using Google Glass [2]. This would help to better estimate the performance differences between the two systems. Finally, we can improve our feedback mechanism scheme by making the range of cutoff values for loudness and rate dynamic (varies with the size of the audience and the place).

ACKNOWLEDGMENTS

We thank all the participants in the research study. We also thank Professor Jingtao Wang for giving us constant guidance to complete the project work.

REFERENCES

1. http://www.briantracy.com/blog/public-speaking/27-useful-tips-to-overcome-yourfear-of-public-speaking/

2. M. Iftekhar Tanveer, Emy Lin, Mohammed (Ehsan) Hoque. "Rhema: A Real-Time In-Situ Intelligent Interface to Help People with Public Speaking", IUI 2015, Atlanta, GA, USA.

3. http://www.philforhumanity.com/ Google_Glass.html/

4. https://bl.newegg.com/the-pros-and-cons-of-google-glass/

5. Mathieu Chollet, Torsten Wortwein, Louis-Philippe Morency, Ari Shapiro, Stefan Scherer, "Exploring Feedback Strategies to Improve Public Speaking: An Interactive Virtual Audience Framework", UBICOMP '15, SEPTEMBER 7–11, 2015, OSAKA, JAPAN.

6. Bahador Saket, Chrisnawan Prasojo, Yongfeng Huang, Shengdong Zhao, "Designing an Effective Vibration-Based Notification Interface for Mobile Phones", February 23–27, 2013, San Antonio, TX, USA.

7. http://programmerguru.com/android-tutorial/android-vibrate-example/

8. http://www.doepiccoding.com/blog/?p=195

9. http://www.yorku.ca/mack/RN-HowToReportAnFStatistic.html

10. http://hci.cs.wisc.edu/courses/hci/lectures/fall2011/HCI-Week11-Lecture12.pdf