Poster Abstract: Smartphone Support for Persons who Stutter

Thiemo Voigt
Uppsala University and SICS
Sweden
thiemo.voigt@it.uu.se

Kasun Hewage
Uppsala University
Sweden
kasun.hewage@it.uu.se

Per Alm
Uppsala University
Sweden
per.alm@neuro.uu.se

Abstract—Stuttering is a very complex speech disorder that affects around 0.7% of adults while around 5% of the population have stuttered at some point. A large percentage of the affected people tend to speak more fluently when their own speech is played back to their ear with some type of alteration. While this has been done with special devices, smartphones can be used for this purpose. We report on our initial experiences on building such an application and demonstrate problems with delay caused by the lack of real-time support for audio playback in the Android operating system. We also discuss ideas for future work to improve app support for people who stutter.

I. INTRODUCTION AND BACKGROUND

Stuttering is a complex speech disorder, resulting in motor symptoms of speech disruption and difficulties to initiate speech. One of the typical traits of stuttering is its variability, and that it may be manipulated and influenced by a range of different strategies. Two such strategies are reading in unison with someone else, or speaking to the pace of a rhythm, for example, a metronome. At least since the end of the 1950:s there has been an interest in manipulation of auditory feedback of stuttering persons. It has been shown that many different manipulations may reduce stuttering, temporarily or for longer periods:

- *Masking noise (Masking Auditory Feedback, MAF)*. MAF uses noise in headphones to mask the normal auditory feedback. The effect appears also if the feedback is not completely masked, though with somewhat smaller effect.
- *Delayed Auditory Feedback (DAF)*. Using DAF, the sound from the speech is transmitted to a headphone in one or two ears with a delay of about 50 to 250 ms. Short delays, about 50 ms, may reduce stuttering but at the same time allows speech with a normal rate. Longer delays require progressively slower speech, with stretched speech sounds: one has to wait for the sound to reach the ear before beginning the next one. Long delays may be used for training of a speech technique with soft and slower speech.
- *Frequency shifted Auditory Feedback (FAF)*. When FAF is used the pitch of the feedback is shifted up or down. The effects on speech of shifts up or down are quite similar, and it seems to be more a matter of comfort, and gender identity, which type of shift is preferred. The most common way to do this shift is "octave shift", so that all frequencies are multiplied

with a factor. For use with stuttering a shift of up to about ± 0.5 octaves may be used. The important aspect of FAF, however, does not seem to be that the new sound is as accurate as possible. Rather, the important aspects seems to be that it is different from the original sound but at the same time not perceived as unpleasant. In this way the distortion of the harmonics and the formant frequencies may be a good thing.

- *Enhanced Auditory Feedback (EAF)*. Based on the old finding of reduced stuttering when the auditory feedback is masked the general assumption seems to have been that making the feedback stronger would make stuttering worse. However, the contrary seems to be the case. We have tested this hypothesis in a student project in Uppsala, supporting this idea.

While earlier special devices were designed and sold as tools to assist stutterers, these functions can now be implemented as smartphone apps. Smartphone apps could also implement more than one functionality. For example, DAF and FAF can be combined. While there are several apps available, they are either quite costly or have limited functionality. Our short term goal is to develop an inexpensive app that contains a lot of insights we have gained in our research on stuttering during the last 10 years [2]. Developing such an app is, however, not straightforward due to the lack of real-time support in the Android OS. This does not allow the users to speak at normal rate which is of course unpleasant for many users. In this poster abstract, we report on our first insights and future plans of developing a sophisticated multi-functional app for people who stutter.

The research community has developed apps in similar contexts. For example, Hao et al.'s app monitors sleep quality [3]. Lu et al. have developed an Android app for stress detection [5] and a scalable framework for modeling sound events on mobile phones called SoundSense [6].

II. FIRST INSIGHTS

In order to gain more insights about the feasibility of providing an app that supports persons who stutter we implemented a DAF app on an Android smartphone.

As shown in Figure 1, the main stages of an app that implements the techniques described in the first section are audio sampling, digital signal processing and audio playback. In a

Fig. 1: The main stages of a smartphone app that implements stutter reduction techniques.

typical mobile phone operating system such as Android, audio sampling and playback are done via Application Programming Interfaces (APIs) that are provided by the operating system. The techniques applied in the digital signal processing stage vary based on the application requirements, i.e., MAF, DAF, FAF, EAF or a combination.

While implementing a DAF application on Android we notice that a delay is being added even though the app is configured to add no delay. In other words, the app reads audio data from the microphone of the mobile device and writes them to the audio playback device without adding any delay. As an initial debugging step, we mechanically measure the delay being added by using an audio recording application that runs on an external computer. We use a metronome application that runs on the same external computer to generate periodic tick sounds and record the output of the mobile phone at the same time.

Fig. 2: Audio latency for DAF technique for Android apps including ours on a Nexus 4 device. The delay of all apps is too high to allow speech at a normal rate.

As a comparison to other DAF implementations, we also measure the latency for the DAF technique without any delay for several Android apps on a Nexus 4 device. The results are shown in Figure 2. The figure shows that the minimum delay that can be achieved is closer to the upper bound of the recommended delay range (50 - 250 ms) for the DAF technique. This implies that the person who uses these apps is forced to speak quite slowly which is a reason for some of the potential users to not use these apps.

Related to the main stages of typical app design, the latency of DAF relies on the delay of audio sampling, digital signal processing and audio playback. Besides hardware capabilities of the mobile devices, the support from the operating system

also plays a critical role in reducing the latency. The latency in audio playback on Android is a well known issue and got the attention of many audio application developers [1].

III. CONCLUSIONS AND FUTURE WORK

While there exists some apps for persons who stutter, the development on Android is currently hampered by the lack of existing real-time support which forces the users of the apps to speak slower than at normal rate. Furthermore, there are more directions to exploit that would give the users better support. In particular, these are in the separation of speech and environmental background noise as well as support for more sophisticated speech analysis.

A limitation of all available systems is that also environmental sound is altered and reaches the earphone. For the user, this can make it more difficult to hear what other persons are saying, and surrounding sounds may become unpleasant since they are also manipulated. We will investigate possible signal processing methods and evaluate if it is possible to apply these with sufficient low and constant delay that does not negatively impact the DAF or FAF functionality.

Current systems include little or no feedback loops for automatic calibration of the parameter settings for increasing or decreasing the delay for DAF, the frequency shift for FAF or the volume. Required for this are analysis methods that can identify when the stutterer talks with little or no problems. There exists some methods for the assessment including phonation interval analysis [4] that are suitable for assessing the current “performance” of the person who stutters and adapt the settings. It is of course also important to study the impact of using such a feedback loop on the person who stutters to ensure that it does not have a negative impact on the speech.

REFERENCES

- [1] Issue: Add APIs for low-latency audio. Web page: <https://code.google.com/p/android/issues/detail?id=3434>. Visited 2014-01-28.
- [2] P. Alm. *On the causal mechanisms of stuttering*. PhD dissertation, Lund University, 2005.
- [3] T. Hao, G. Xing, and G. Zhou. isleep: Unobtrusive sleep quality monitoring using smartphones. In *Proceedings of 11th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Rome, Italy, Nov. 2013.
- [4] R. J. Ingham, M. Kilgo, J. C. Ingham, R. Moglia, H. Belknap, and T. Sanchez. Evaluation of a stuttering treatment based on reduction of short phonation intervals. *Journal of Speech, Language, and Hearing Research*, 44:1229–1244, 2001.
- [5] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury. Stresssense: Detecting stress in unconstrained acoustic environments using smartphones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012.
- [6] H. Lu, W. Pan, N. D. Lane, T. Choudhury, and A. T. Campbell. Soundsense: Scalable sound sensing for people-centric applications on mobile phones. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services, MobiSys '09*, Krakow, Poland, June 2009.
- [7] R. Webster. A few observations on the manipulation of speech response characteristics in stutterers. *Journal of Communication Disorders*, 10:73–76, 1977.