Speak

Speak! - speech synthesis meeting.

The purpose of these regular informal meetings is to discuss and share progress relating to speech synthesis (audio and visual) research - within CSTR specifically as well as in the field generally. Talks are intended to be short and informal, with an emphasis on discussion, interaction and feedback. Relevant references should be sent round in advance to encourage everyone to contribute.

Everybody with an interest in speech synthesis (audio and visual) research is welcome.

- Meetings will typically be held in the Instrumented Meeting Room ("IMR" - room 3.07), on Level 3 of the Informatics Forum building (though this may vary on odd occasions.)
- At present, the standard time for these meetings is **Thursdays at 2-3pm.**

**Note:** To edit the current schedule, please go to [this page](#).

Speak! - speech synthesis meetings schedule for 2017/2018

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.09.17</td>
<td>-</td>
<td>Language at Edinburgh Welcome Reception - 4pm in IF 4.40 (Mini-Forum 2)</td>
</tr>
<tr>
<td>21.09.17</td>
<td>Srikanth</td>
<td>Semester 1 Planning Meeting</td>
</tr>
<tr>
<td>28.09.17</td>
<td>Felipe</td>
<td>Generative Adversarial Network-based Postfilter for STFT Spectrograms</td>
</tr>
<tr>
<td>05.10.17</td>
<td>Catherine</td>
<td>Postponed (due to illness)</td>
</tr>
<tr>
<td>12.10.17</td>
<td>Catherine</td>
<td>Neural style transfer</td>
</tr>
<tr>
<td>19.10.17</td>
<td>Carol</td>
<td>Inaudible Voice Commands</td>
</tr>
<tr>
<td>26.10.17</td>
<td></td>
<td>ICASSP deadline (few people away)</td>
</tr>
<tr>
<td>09.11.17</td>
<td>Oliver Watts</td>
<td>2 prosody papers from Interspeech: Rendel et al., Weakly-Supervised Phrase Assignment from Text in a Speech-Synthesis System Using Noisy Labels and Klimkov et al., Phrase break prediction for long-form reading TTS: exploiting text structure information.</td>
</tr>
<tr>
<td>Date</td>
<td>Name</td>
<td>Event Description</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>16.11.17</td>
<td>Martti Vainio</td>
<td>Hierarchical representation of speech prosody using Continuous Wavelet Transform (CSTR Talk)</td>
</tr>
<tr>
<td>30.11.17</td>
<td></td>
<td>CSTR Christmas Lunch</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Samples: <a href="https://avdnoord.github.io/homepage/vqvae/">https://avdnoord.github.io/homepage/vqvae/</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Semester 2</td>
</tr>
<tr>
<td>25.01.18</td>
<td></td>
<td>Most people away</td>
</tr>
<tr>
<td>01.02.18</td>
<td>Srikanth</td>
<td>Planning Session</td>
</tr>
<tr>
<td>08.02.18</td>
<td>Oliver</td>
<td>Tacotron 2: Natural TTS Synthesis by Conditioning WaveNet on Mel Spectrogram Predictions</td>
</tr>
<tr>
<td>15.02.18</td>
<td>Srikanth</td>
<td>visual to sound: generating natural sound for videos in the wild</td>
</tr>
<tr>
<td>22.02.18</td>
<td>Cassia</td>
<td><a href="https://www.spin-lab.org">Summary of SPIN workshop</a></td>
</tr>
<tr>
<td>01.03.18</td>
<td>Simon</td>
<td>spoofing / anti-spoofing / generative models / adversarial techniques</td>
</tr>
<tr>
<td>08.03.18</td>
<td>Felipe</td>
<td>Speaker-dependent wavenet vocoder and parallel wavenet</td>
</tr>
<tr>
<td>15.03.18</td>
<td>Zack</td>
<td>Presentation on his Interspeech work?</td>
</tr>
<tr>
<td>22.03.18</td>
<td></td>
<td>Interspeech deadline</td>
</tr>
<tr>
<td>29.03.18</td>
<td>Zack</td>
<td>ICASSP papers?</td>
</tr>
<tr>
<td>05.04.18</td>
<td>Simon</td>
<td>Speaker Odyssey Keynote</td>
</tr>
<tr>
<td>12.04.18</td>
<td>Sam</td>
<td>Prosody?</td>
</tr>
<tr>
<td>19.04.18</td>
<td></td>
<td>Invited speakers?</td>
</tr>
<tr>
<td>26.04.18</td>
<td>Korin</td>
<td><a href="https://www.spin-lab.org">Crowd-sourcing prosodic annotation</a></td>
</tr>
<tr>
<td>03.05.18</td>
<td>Avashna</td>
<td>DNN-adaptation session</td>
</tr>
<tr>
<td>10.05.18</td>
<td>Carol</td>
<td>Intelligibility measures / source separation ?</td>
</tr>
<tr>
<td>17.05.18</td>
<td>Felipe</td>
<td>Tutorial on signal processing</td>
</tr>
<tr>
<td>24.05.18</td>
<td>Jason</td>
<td>Literature review? ICASSP papers?</td>
</tr>
<tr>
<td>31.05.18</td>
<td>Jennifer</td>
<td>Multimodal (speech, video, text) emotion</td>
</tr>
</tbody>
</table>

**Suggestions for future presentations**

- Sessions on Merlin, wavenet, anti-spoofing, DNN-adaptation?
- Tutorials - One tutorial for statistical analyses
- Talk from Junichi about his work at NII?
- Comparison of all waveform-based approaches? latest work on GANs?

**Older Speak!**

The Speak! wiki was recently moved to this location.

For older schedules, please check the old wiki: [https://wiki.inf.ed.ac.uk/CSTR/Speak](https://wiki.inf.ed.ac.uk/CSTR/Speak)