Meeting 1: Corti

**Time and data:** 01/03/2021 (9.00-17.00)

**Participants:** All group members

**Questions:**

1. Skal vi træne både ASR model og WaveNet model på det samme datasæt (eller 2 forskellige)? ASR-modellen bruger LibriSpeech validation- og testsættet (dette kan dog blot udskiftes). Tacotron 2 (ved sammenligning med state-of-the-art) er pretrained på LJ Speech. Hvad er din holdning. Vi evaluerer stadig på Libri Speech.
   1. Træn TTS: LJ Speech (generet)
   2. Træn ASR: Libri Speech (baseline)
   3. Test ASR: Libri Speech (baseline)
   4. Træn ASR på synthesized data (generet)
   5. Test ASR op Libri Speech og sammenlign med baseline
   6. Evt. gør samme med Tacotron 2 (hvis der er tid til det)
2. Kan du give os nok dansk trænings data til at træne TTS modellen fra bunden eller ville dette være nødvendigt at lave transfer learning. Hvilken slags dansk data er det?
   1. Formentlig ikke tid til det (muligt bachelor).
3. Hvad anbefaler du ift. at lave transcript embeddings (metoder til at lave dem selv vs. pre-trained model).
   1. Kan vi få en opsummering på mail at lave vores transcript embeddings. Kunne vi bruge f.eks. BIRD til en global embedding?
4. Data augmentation (bedste metoder)
   1. Data augmentation kan laves på output synthesized data for at tilføje ”noise” før der testes på ASR (i slutningen). Vi kan ikke se de giver nogen mening hvis vi ikke skal teste på alarmopkald (at teste på lydbøger har ingen støj og er rimelig monotone). Lars sagde vores output fra TTS modellen er meget clean og man derfor gør det. Vi ser på det senere.
5. Tidsplanen (Gannt)
   1. Passer fint. Dog behøver vi nødvendigvis ikke at sætte tid af til ASR (det ville være fedt – men vil hellere have bedre evaluering af metoder uden at kigge på ASR – ASR er ekstra stof til rapporten der kunne være fedt).

Ekstra:

LB ibrary til at få phoenmener, CCT og (spacy) global conditioning + AN4 (start med at overfitted for at se modellen virker)

Støj at se på støjkilder for ASR.

Træn WaveNet på LJ Speech Træn ASR på Librispeech Light TRAIN Test/Træn WaveNet/ASR på Librispeech Light Evaluation Træn ASR model på hele det syntetiserede Librispeech Light TRAIN data. Evaluér forskellen på de to ASR modeller