



Charlemont grant report

Recipient name:	Dr Sébastien Le Maguer
Discipline and subject area:	Sciences; Computer Science and Informatics
Amount and year awarded:	€2,500
Title of project:	An automatic auditory investigation of speech synthesis systems production

Summary of findings:

Currently, speech synthesis is evaluated by human listeners who are giving scores based on their perception. In addition to these costly methodologies, the community uses machine learning modelling to predict the human listeners score for a synthetic speech sample. However, these models are black boxes, and it is currently impossible to understand the reasons behind their predictions. On the other hand, psychoacoustics and hearing research can provide additional tools which can enable a better understanding concerning the perception of synthetic speech. The goal of the present project aims to use Auditory Nerve Models to simulate the reception of a synthetic speech signal and to investigate its specificities before its interpretation.

Preliminary research, conducted before receiving the grant, showed that the output (i.e., neurograms) of Auditory Nerve Models are a valid alternative representation of speech to the ones which are standardly used in the speech synthesis field. However, my knowledge of these models was too limited to keep on the investigation. Therefore, the grant was used to finance a visit to the Auditory Engineering Laboratory (AEL) to get the missing knowledge to enable further exploration of synthetic speech using hearing research. The other main goal of this visit was to initiate a collaboration with Prof. Ian Bruce, a recognized expert in auditory modelling. This collaboration aims to investigate the use of Auditory Nerve Models, developed in the AEL, to analyse synthetic speech. Due to the current pandemic, this travel only took place in the middle of October 2021. As this visit was envisioned to take place at the start of the project, the implementation of the project has been delayed as well.

Nonetheless, during the visit to the AEL, we started to conduct a set of experiments to investigate how Auditory Nerve Models can inform us about the intelligibility of synthetic speech. Preliminary results show that neurograms identify different artifacts than the usual representations of synthetic speech. From an application standpoint, if the present investigation leads to conclusive results, this could provide an alternative perspective to explain the behaviours of TTS (Text to Speech) models. Furthermore, it could also impact the Hearing Research community by providing insights into the advantages and drawbacks of using synthetic speech for their experiments.

Plans for continuing collaboration:

This grant was used to initiate a term collaboration between myself and the AEL. During the visit financed by the grant, I not only got the opportunity to be introduced to the auditory nerve model developed in the laboratory, but I also got a glimpse of the other activities conducted by the group. As a result of this visit, we identified multiple directions that we can collaborate on from a near to a further away future.

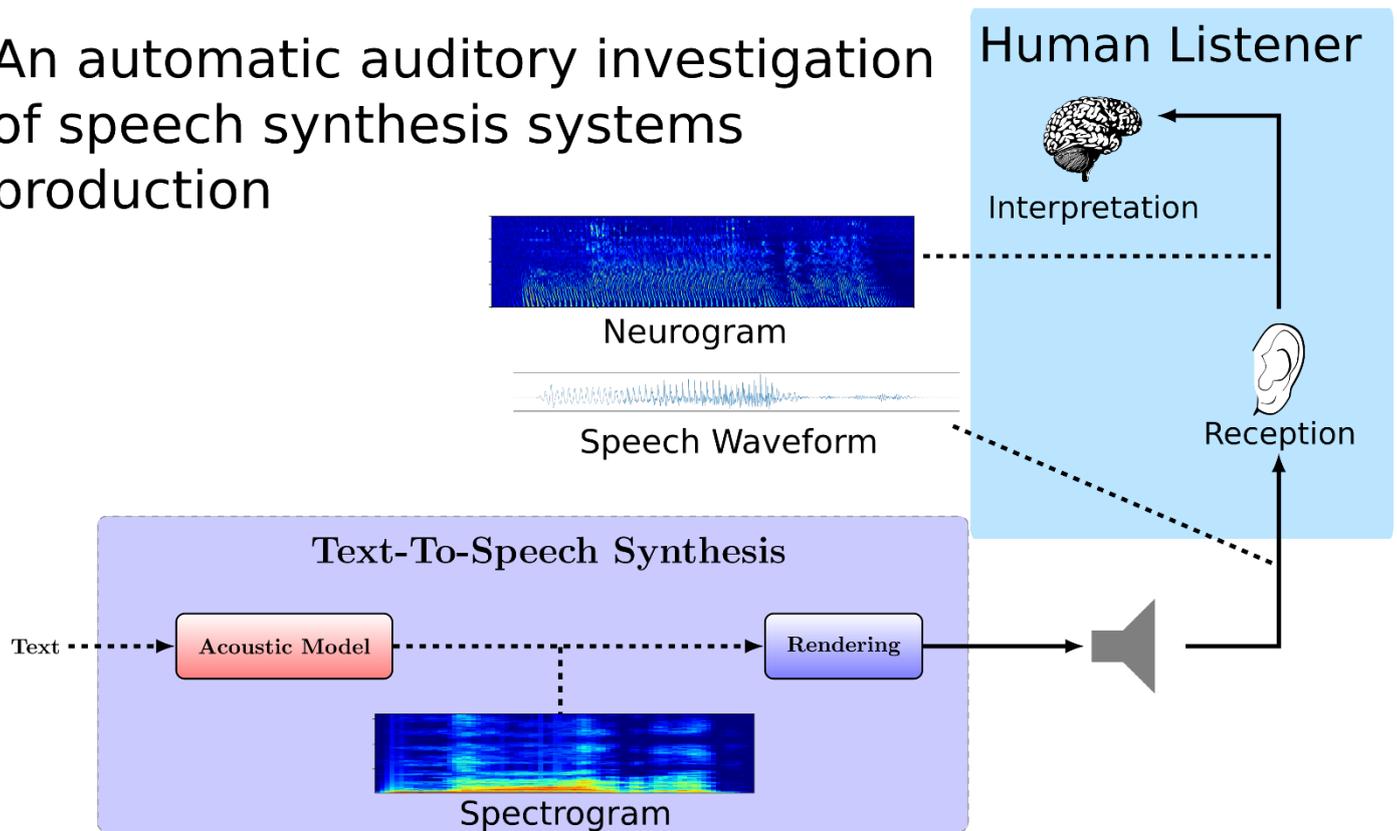
Firstly, we will keep on investigating the specificities of synthetic speech and the explainability of modern Text-To-Speech synthesis technologies using the Auditory Nerve Model. As the model developed in the AEL is in constant

Charlemont grant report

development, the impact of any implemented improvements will also be investigated for synthetic speech. My knowledge of synthetic speech and Pr. Bruce's knowledge of auditory modelling is the ideal combination to conduct such an investigation.

In addition to this primary research topic, we also plan to investigate the perception of synthetic speech by a hearing-impaired audience. This collaboration could inform our societies about the impact of synthetic speech as it becomes ubiquitous, and there is no investigation on the perception of such speech.

An automatic auditory investigation of speech synthesis systems production



Published work and publication plans:

During the visit, we initiated a set of experiments, and we expect to submit the results of these experiments next year as an academic paper. In addition, the exchange conducted during the visit identified the hearing-aid public as blind spot of the speech synthesis community. We expect to start experiments which are dedicated to this audience next year and to submit the results in a venue next year as well.

Dissemination and plans for future dissemination:

Due to the nature of the visit and the delay caused by the pandemic, no dissemination has yet been conducted. However, we expect to publish our work into conferences dedicated to speech science next year.



Charlemont grant report

Collaborations and planned collaborations:

The collaboration involves the Trinity College Dublin and the Auditory Engineering Laboratory of McMaster University.

Outreach and engagement activities:

To this date, we didn't conduct any engagement activities but we expect to do so in the future depending on the results we obtained.