Machine-Aided Keyword Suggestions for Aphasic Speech

15-300, Fall 2020

Amy Li

November 5, 2020

Webpage: `https://amy1351.github.io/15-400.html`

1 Project Description

1.1 Faculty Advisors

I will be working with Professor Graham Neubig and PhD student Xinyi Wang in the Language Technologies Institute. Their research focuses on the areas of machine translation, multi-lingual NLP, and natural language understanding. We will also be collaborating with Professor Bonnie Nozari in the Department of Psychology, who studies the cognitive and neural bases of language production.

1.2 Project Overview

In this project, we hope to apply neural machine translation techniques to help patients with aphasia. Aphasia is a condition in which someone has suffered an injury to a part of the brain that is responsible for language. This is most commonly caused by strokes, although it can also be caused by head injuries, neurological diseases, or brain tumors. Aphasia hinders a person's ability to effectively communicate using language, and aphasia patients must often go through speech-language therapy sessions to improve some of the language skills they have lost [6]. By applying language modelling and style transfer techniques to aphasic speech, we hope to make it easier for aphasia patients to communicate effectively with others.

One of the most common problems with aphasic speech is that aphasia patients cannot retrieve the keywords they want to use. Instead, the patient might use unrelated words or made-up

words in place of the ones they have forgotten, which makes their speech difficult for others to understand. To help address this issue, we will create a language model to perform the prediction of keywords in aphasic speech. By creating suggestions for the most likely keywords that an aphasia patient will want to say, we can help them remember words that they have forgotten.

To train our language model, we will use data from AphasiaBank, a shared database containing transcriptions of conversations with aphasia patients. This also includes a set of control data from conversations with neurotypical patients in which the same questions were asked. To generate additional training data, we will perform data augmentation on standard English text using style transfer techniques to make it look similar to aphasic speech. This can then be used to pre-train our language model. We will also experiment with working off of existing language models such as GPT-2 which can perform word prediction on neurotypical English speech. Afterwards, we will fine-tune the language models on the aphasic English. Lastly, we will experiment with different modifications to these language models to optimize their performance on the aphasic speech data.

One of the most significant challenges in this project will be evaluating how well the language model finds the correct keywords. For now, we plan to use the language model to predict the keywords that should be in the next part of the aphasic speech when there is a pause recorded in the transcription, since this is an indication that the patient might have forgotten the next word to use. The model will use the context from each question or topic to predict the next most likely keyword. Then, we will extract the keywords from the control data in AphasiaBank to see if the keywords were correctly predicted. We will use standard measures such as perplexity, recall@N, and mean average precision to evaluate how well the model performs. In the future, we would also like to work directly with aphasia patients to demonstrate the utility of the system to them and receive feedback on its effectiveness.

1.3 Project Impact

If we are able to create a successful model, we will show that neural language modelling can be used as a tool to help aphasia patients communicate more effectively. There are also possible applications of this project in speech-language therapy, since a keyword prediction system could help patients to practice improving their language skills over time. Since aphasia is such a common condition that impacts about 1 million people in the U.S. alone [6], the results of this

research could help to improve the quality of life for many people who are affected by it.

2 Project Goals

2.1 100% Goal

Through doing this research project, I hope to learn more about the area of natural language processing and language modelling in particular. My goal by the end of Spring 2021 is to have a working language model that can successfully predict keywords for aphasic speech. I would also like to have a clear set of evaluation criteria to prove that the language model works and produces useful results.

2.2 75% Goal

If things go slower than expected, I would still like to have a working language model that is able to process aphasic speech data from AphasiaBank. Even if there is not a clear set of evaluation criteria, the language model should at least be able to generate keyword predictions that make intuitive sense given the context of the speech.

2.3 125% Goal

If there is extra time, I would like to work directly with aphasia patients to show them the keyword prediction system and demonstrate its utility. This will also allow me to receive feedback on how it could be improved. I would also like to write a final report or paper on the results of the research project. In the future, I could also create an interface with speech processing capabilities for the language model so that the keyword predictions can be generated in real time as an aphasia patient is speaking to help suggest keywords they have forgotten.

3 Project Milestones

3.1 1st Technical Milestone for 15-300

By the end of this semester, I should have read all the papers below, brainstormed possible language models and evaluation criteria to use, and pre-processed the aphasic speech data from AphasiaBank so that it is in a form that can be inputted into a neural language model.

3.2 Bi-weekly Milestones for 15-400

February 15th

I plan to work on this project over winter break. By February 15th, I hope to have finished training the standard GPT-2 model on the aphasic speech from AphasiaBank to gather preliminary results. I should also have finished the step to extract keywords from the control data in AphasiaBank so that they can be used to evaluate the performance of the language model's keyword predictions.

March 1st

By the second biweekly milestone, I should have the data augmentation step completed to convert standard English text into text that resembles aphasic speech.

March 15th

By the third biweekly milestone, I should have completed building the initial version of the custom LSTM-based and/or transformer-based language model to predict keywords in aphasic speech.

March 29th

By the fourth biweekly milestone, I should have pre-trained the custom model on the augmented neurotypical English text. I will then work on training this model with the aphasic speech data to fine-tune it.

April 12th

By the fifth weekly milestone, I should have gathered testing results for the custom keyword prediction language model on aphasic speech. I should also have results from trying out different hyperparameters and modifications to the model architectures, for both existing language models such as GPT-2 as well as the custom-built model.

April 26th

By the sixth biweekly milestone, I hope to have analyzed the results of the experiments and identified the language model that performs best on keyword prediction for aphasic speech. I also should have begun writing my final report.

May 10th

By the seventh and final biweekly milestone, I should have completed documenting the results of this research project in a final report or paper. I should also have identified possible applications or areas in which this project could potentially go next.

4 Literature Search

4.1 Language Modelling

Papers on language modelling will be useful to determine which existing language models we can build off of and fine-tune with the aphasic speech data. These papers will also be helpful for building our own custom model from scratch. These are some papers I have found so far, but I will continue reading more as I gather more information.

Papers: [3], [7], [9]

4.2 Style Transfer

Papers on text style transfer will be useful for performing data augmentation to transform neurotypical speech into text that resembles aphasic speech. They could also be useful for an extension project to help fill in the grammatical structure in aphasic speech to make it more fluent and understandable.

Papers: [2], [8]

4.3 Aphasia

Papers and resources on aphasia will be useful for helping us understand how aphasic speech is generated and how we can maximize the positive impact that our project can potentially have on aphasia patients.

Papers and websites: [1], [4], [5], [6]

5 Resources Needed

We will use Python and the deep learning framework PyTorch to create the language models. We will also use aphasic speech data and neurotypical speech data from AphasiaBank, which Professor Noxari has shared with me. However, we still need to find another dataset containing more neurotypical English text which can be used for data augmentation to boost the amount of training data we have.

To complete this research project, we will also need GPU machines for computing power. Professor Neubig has granted me access to some of the LTI servers which I will be using for training models and running evaluations.

References

[1] National Aphasia Association. Aphasia definitions, Jun 2018.

[2] Junxian He, Xinyi Wang, Graham Neubig, and Taylor Berg-Kirkpatrick. A probabilistic formulation of unsupervised text style transfer, 2020.

[3] Linmei hu, Juanzi Li, Liqiang Nie, Xiaoli li, and Chao Shao. What happens next? future subevent prediction using contextual hierarchical lstmt. 02 2017.

[4] Duc Le, Keli Licata, and Emily Mower Provost. Automatic quantitative analysis of spontaneous aphasic speech. *Speech Communication*, 100:1 – 12, 2018.

[5] Duc Le and Emily Mower Provost. Improving automatic recognition of aphasic speech with aphasiabank. In *Interspeech*, pages 2681–2685, 2016.

[6] National Institute of Deafness and Other Communication Disorders. Aphasia, Nov 2020.

[7] A. Radford, Jeffrey Wu, R. Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. 2019.

[8] Sandeep Subramanian, Guillaume Lample, Eric Michael Smith, Ludovic Denoyer, Marc'Aurelio Ranzato, and Y-Lan Boureau. Multiple-attribute text style transfer. *CoRR*, abs/1811.00552, 2018.

[9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017.