university of groningen / faculty of mathematics and natural sciences / artificial intelligence

Reconstructing non-fluent aphasic speech using large language models

Graduation Project Proposal
(Language Modelling and Natural Language Processing)

Ruhi Mahadeshwar (s4014456)

December 20, 2023

Internal Supervisor: dr Tsegaye Misikir Tashu (Faculty of Science and Engineering, University of Groningen
External Supervisor: dr Frank Tsiwah (Faculty of Arts, University of Groningen)

Artificial Intelligence
University of Groningen, The Netherlands

1 Introduction

Language is the main tool humans use for communication. Effective communication is important to build communities, convey emotions and communicate essential information. From previous research, we know that in the brain, there are two areas which are responsible language processing. These are Broca's and Wernicke's areas. Damage to these regions (due to stroke or injury) can lead to Broca's aphasia or Wernike's aphasia[1].

Broca's aphasia is a type of non-fluent aphasia in which there is a partial loss of the ability to produce language whether it be spoken, written or signed. However, comprehension of language is generally intact. Speech becomes agrammatic and severely reduced. Function words such as prepositions and articles are left out. There is also a struggle to retrieve words (word-finding difficulties). Patients who suffer from non-fluent aphasia are often aware of their condition and this leads to frustration and social isolation in daily life since effective communication does not take place.

Therefore, this project aims to reconstruct sentences of individuals[2] with non-fluent aphasia (characterised by agrammatism and incomplete sentences) through the use of large language models (LLMs), with a focus on decoder only transformer architecture. The findings of this project will contribute significantly to the developments of assistive technology for individuals with non-fluent aphasia as well as their caregivers. This project does not aim to develop assistive technology, rather it aims to investigate whether assistive technology can be done, and if yes then what is the best approach to it. Another important thing to note is that the project will not look at transcribing recorded speech to sentences.

2 Theoretical Framework

As mentioned earlier, one of the regions in the brain which is responsible for language processing and speech production is Broca's area (Friederici, 2011). Damage to this region, due to stroke, injury, brain tumors or brain infections leads to Broca's aphasia. Studies of interactions with people who suffer from Broca's aphasia display the following (Raymer & Rothi, 2018, Chapter 8): halting and reduced speech, commonly classified as apraxic; non-fluent language output; limited vocabulary with the usage of more nouns than verbs; usage of verbs with omission/misuse of inflections; articles, pronouns and adjectives might be eliminated (agrammatism); sentences produced might lack in syntactic structure and morphological detail.

Further, there are two main accounts of disordered language processing and language productions in non-fluent aphasic speech. These are the trace deletion theory and working memory theory. Trace deletion theory (Grodzinsky, 2000) proposes that Broca's area is where the manipulation of traces is localised to and that damage to this area leads to the loss of the ability to represent and track traces

[1] In my project, I will focus on non-fluent aphasia such as Broca's aphasia.

[2] I will only focus on native speakers of English.

of syntactic movement. Traces (Chomsky, Anderson, & Kiparsky, 1973) are how the brain encodes words with their correct thematic roles. The working memory theory suggests that individuals with aphasia have impaired working memory which contributes to impairments of sentence processing (Wright & Fergadiotis, 2012).

Rehabilitation and assistance of patients with aphasia can be used to target language processing impairments at the sentence level. As of now, there is no standard treatment for Broca's aphasia since treatment needs to be as per a patients needs. Speech and language therapy help patients to communicate their wants and needs. However, this can take long periods of time (months or years). Therefore, the use of assistive technology is needed for the immediate impact of loss of verbal communication in daily life.

Developments made in natural language processing (NLP) using machine learning methods allows computers to process human languages and perform tasks, such as text summarization, semantic analysis, translation and language generation. Deep learning architectures, such as transformers (Vaswani et al., 2017) allow for such tasks to happen efficiently. However, training such deep learning models on the task of completing non-fluent aphasic speech requires a large enough dataset of sentences which were uttered by non-fluent aphasics and corresponding corrected versions of the sentences. Constructing such a dataset would require an effort to recording speech, transcribing speech and correcting speech. Due to this, NLP research focused on aphasia is limited.

To overcome this, Misra et al. (2022) show us that an synthetic dataset of non-fluent aphasic speech can be created by augmenting the grammatically correct C4 dataset (Raffel et al., 2020) through the use of linguistic features extracted from small scale studies on non-fluent aphasic patients. They further show that the pretrained T5 model (Raffel et al., 2020) can be fine-tuned using the augmented dataset to suggest 5/3 most likely corrected sentences given an aphasic sentence as input. Thus, they demonstrate the augmented datasets can be used to develop assistive devices for non-fluent aphasic patients through transfer learning.

3 Research Question

Thus, the main research question of this project is "Can decoder-only transformers be used to reconstruct non-fluent aphasic speech?". I would be focusing on decoder-only transformers, namely LLaMA (Touvron et al., 2023). This project would go beyond the research of Misra et al. (2022) by using an LLM to reconstruct an original agrammatic aphasic sentence. This LLM would be trained only with synthetic aphasic.

Firstly, I will develop a rule based model to augment a dataset and create a new dataset which reflects the speech of non-fluent aphasics. This can be done using linguistic features from studies on non-fluent aphasics as described earlier, and through the results of verbal tests, such as the cookie theft picture. This is done to overcome the lack of large scale studies of linguistic deficits in aphasia.

Further comparisons of the augmented dataset to an original non-fluent aphasic dataset will be done to show if the augmented dataset is an accurate representation of non-fluent aphasic speech in terms of linguistic features commonly seen in non-fluent aphasics.

Then, a pretrained LLaMA model will be fine-tuned using the synthetic agrammatic aphasic dataset to suggest 5/3 most likely corrected sentences given a non-fluent aphasic sentence as input. The fine-tuned model can then be evaluated on the validation set from the augmented dataset first and then the original non-fluent aphasic dataset. This will show if the fine-tuned model can be used in assistive devices even if it was trained on an augmented dataset. Further comparisons of the fine-tuned model can be done to an encoder-decoder model (such as T5 in Misra et al. (2022) trained on the same augmented dataset to see if both models are able to complete the given task or if they differ in some way.

4 Methods

To answer the research question, I would first look at whether it is possible to generate a synthetic dataset. Comparisons of the synthetic dataset to the real-life dataset can be done through measures of sentence complexity (such as mean length of utterance and noun/verb ratio as used by Misra et al. (2022). This will be done to see if the synthetic dataset is an accurate representation of non-fluent aphasic speech in terms of linguistic features (or lack of) commonly seen in people with non-fluent aphasia.

Comparison of the results of fine-tuned model on the synthetic data and on original non-fluent aphasic dataset will be done through manual human annotations. Native English speakers will be shown the input aphasic sentence and the top 5/top 3 reconstructed sentence. Then, they will be asked to rate the reconstructed sentences based on most likely reconstruction to least likely reconstruction. This will show if the model can be used in assistive devices even if it was trained on a synthetic dataset.

By using the metrics above, further comparisons of the model can also be done to an encoder-decoder model trained on the same synthetic dataset to see if both models are able to reconstruct aphasic speech. Other metrics such as training time, number of hyperparameters and number of parameters can be used to see which one of the models can be more effectively used in assistive technology.

5 Scientific Relevance for Artificial Intelligence

From this graduation project, I would like to show that augmenting datasets based on small scale studies of non-fluent aphasia can be used to develop real-time assistive technology to reconstruct non-fluent aphasic speech. This would show that the tremendous effort of gathering a dataset from transcribing the speech of non-fluent aphasics would not be required. Through the use of LLM, I can

show that real-time completion of speech can be possible successfully. By doing so, I will be able to show that it is possible to use artificial intelligence to help the daily struggle of aphasics to reduce their frustration and social isolation, and, to improve their quality of life and ability to effectively communicate.

6 Planning

- WP1 Conducting research into aphasia and loooking at studies on non-fluent aphasics to extract language features of aphasic speech and write background research.
- WP2 Make synthetic dataset.
- WP3 Compare synthetic dataset. Also write methods for making synthetic dataset and comparison results.
- WP4 Start fine-tuning LLaMa model. Write down methods for training LLaMa model
- WP5 Evaluate fine-tuned LLaMa model. Write results. Fine-tune T5 model.
- WP6 Evaluate T5 model. Write results. Compare two results.
- WP7 Buffer period, continue writing thesis.
- WP8 Continue writing thesis.

I have also planned my Christmas holidays from 25-30th but I will be working a bit during them.

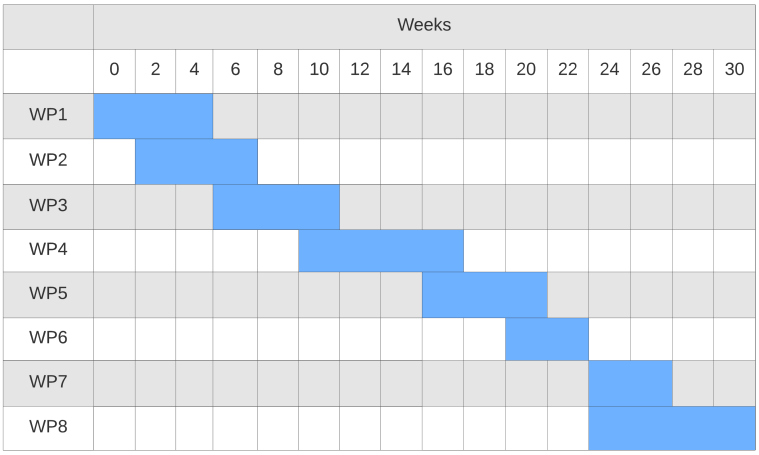

Figure 1: Duration of work parts (WP)

7 Resources and Support

For day-to-day supervision, I will go to Frank, as discussed with both supervisors. I would meet with Frank bi-weekly to ensure that the project is on track. I do not expect heavily detailed feedback for every supervision session. I would also like to have meetings with Frank and Tsegaye (maybe once every 5/6 weeks) to ensure that both supervisors are up to date with my project. If I need to meet with Tsegaye for help with LLMs, I would inform him at least 2-3 days in advance.

For most of the project, I will be able to use my laptop to work. For the fine-tuning of the models, I would be able to use the Hábrók cluster or any other computational resource provided by the Artificial Intelligence department. If I need to work at the university, I would be able to use room 295 in Bernoulliborg as my workspace.

References

Chomsky, N., Anderson, S., & Kiparsky, P. (1973). Conditions on transformations. *1973*, 232–286.

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological reviews*, *91*(4), 1357–1392.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without broca's area. *Behavioral and brain sciences*, *23*(1), 1–21.

Misra, R., Mishra, S. S., & Gandhi, T. K. (2022). Assistive completion of agrammatic aphasic sentences: A transfer learning approach using neurolinguistics-based synthetic dataset. *arXiv preprint arXiv:2211.05557*.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, *21*(1), 5485–5551.

Raymer, A. M., & Rothi, L. J. G. (2018). *The oxford handbook of aphasia and language disorders*. Oxford University Press.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., ... Lample, G. (2023). *Llama: Open and efficient foundation language models.*

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Wright, H. H., & Fergadiotis, G. (2012). Conceptualising and measuring working memory and its relationship to aphasia. *Aphasiology*, *26*(3-4), 258–278.