Using Machine Learning to Detect and Predict Word Level Errors in Narratives of People with Aphasia

Rosa Zavaleta, M.S., CF-SLP
Jacob Brue, M.S., CS
Laura Wilson, Ph.D., CCC-SLP, CBIST
Sandip Sen, Ph.D.
DISCLOSURES

● Financial Relationships

- Jacob Brue receives a full scholarship and stipend from Tulsa CyberFellows for PhD research at the University of Tulsa
- Rosa Zavaleta is employed at Kindred Hospital; former student at The University of Tulsa
- Laura Wilson and Sandip Sen are full-time, salaried employees at The University of Tulsa

● Nonfinancial Relationships

- We used the AphasiaBank as the source of the data
- AphasiaBank is supported by NIH-NIDCD grant R01-DC008524 for 2022-2027
CONTENTS

1. Introduction
2. Errors in Aphasia
3. Paraphasias and Past Studies
4. Research Question
5. Big Picture Model
6. Specific Modules and Results
7. Future Directions
8. Questions and Comments
9. References
INTRODUCTION

aphasia research
APHASIA
An acquired communication disorder caused by neurological damage and marked by language deficits across all modalities

NON-FLUENT
Effortful production of language with varying comprehension skills

FLUENT
Fluid production of language with varying comprehension skills

ERRORS
Breakdown in communication on a spectrum of unintelligible speech to single word errors

PARAPHRASIAS: Substitutions of an intended word with a sound or unintended word (Dalton et al., 2018).
PARAPHASIAS TYPES

EXAMPLE: I HAVE A CAT.

SEMANTIC: substitutions are related to intended meanings

I have a dog.

PHONEMIC: substitutions are related to intended sounds

I have a hat.

NEOLOGISTIC: substitutions are non-words

I have a sark.

WHY PARAPHASIAS?

2018
Le and colleagues
Created a model to automatically detect paraphasias in the narratives of PWA.

2003
Bird and colleagues
Explored impact of imageability and part of speech in speech of PWA.

1995
Nickels and colleagues
Confirmed relationship with psycholinguistic factors. Namely imageability and age of acquisition.

1984
Butterworth and colleagues
Found relationship between paraphasia production and external factors like frequency.
RESEARCH QUESTION

DETECTION OF ERRORS
- the accuracy with which the model is able to detect the presence of a paraphasic errors

PERSONAL FEATURES
- age
- years of education
- sex
- race

CLINICAL FEATURES
- aphasia type
- aphasia severity
- time post onset
- comorbidity with apraxia
- comorbidity with dysarthria

WORD LEVEL FEATURES
- imageability
- frequency
- word length
- part of speech
- position in a sentence
- age of acquisition
DATA PROCESSING STEPS

1. Converted transcripts from AphasiaBank to XML format
2. Manually imputed missing intended words
3. Included word level features - imputed from a model when necessary
4. Filtered unusable utterances
5. Developed interpretable models to predict paraphasias
Aphasiabank Transcription: Cindessa uh was a poor boy.

Target sentence: Cinderella was a poor girl.

<table>
<thead>
<tr>
<th>features from aphasiabank</th>
<th>features from psycholinguistic database</th>
</tr>
</thead>
<tbody>
<tr>
<td>error tag</td>
<td>error tag</td>
</tr>
<tr>
<td>intended word</td>
<td>imageability</td>
</tr>
<tr>
<td>part of speech</td>
<td>frequency</td>
</tr>
<tr>
<td>location</td>
<td>age of acquisition</td>
</tr>
<tr>
<td></td>
<td>length of word (grapheme number)</td>
</tr>
<tr>
<td>error</td>
<td>100</td>
</tr>
<tr>
<td>girl</td>
<td>6.53</td>
</tr>
<tr>
<td>noun</td>
<td>320</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
DECISION TREES

Target sentence: Cinderella was a poor girl.

Is the part of speech a noun?

YES

... YES

ERROR

... NO

NOT AN ERROR

NO

... NO

... NO

ERROR

NOT AN ERROR
HISTORY OF MODELS FOR CURRENT PROJECT

Preliminary detection model at word level

Exclusions: neologisms, semantic errors without replacements, and clinical features

Preliminary detection model at sentence level

Included aggregate features of utterance and individual word features

Final detection model at word level

Included previously excluded features
WORD LEVEL DETECTION MODEL

- Included previously excluded neologistic and semantic errors with and without replacements
- Used a manual imputation strategy for words without replacement
- Included clinical features
Receiver Operating Characteristic (ROC) Curve

AUC (Area Under Curve) = 0.85

Test-train split (cross-validation)

- Found salient features supporting the detection of word level errors with high accuracy
  - Part of speech
  - Word frequency
  - Participant age
  - Word imageability
  - Aphasia duration
  - Severity status
SENTENCE LEVEL DETECTION MODEL

- **Main difference**: Include context through the use of features of surrounding words.

- **E.g.** I haves a cat.
  - The model would incorporate features of the intended word (have) as well as of the surrounding words within a small window.
  - The model will additionally use aggregate information from the whole utterance.

- We will continue to focus on interpretable models while expanding the input of the model.

- We anticipate that providing more context will create a more accurate detection model than our previous word level model.
Future Directions

Classification model at the word level
This model could not only detect an error but determine its type
Could incorporate video/audio language processing features

Classification model at the sentence level
Could include non core word features such as gestures, pauses, filler words, etc.

Prediction model for types of error
Given a speech sample, this model could predict the most likely errors a PWA may produce

Real time analysis of nuanced speech characteristics
Could serve as a tool for evaluations and prognosis
DISCUSSIONS AND CONCLUSION

- **Interdisciplinary work**: Innovative, interdisciplinary work between computer science and speech language pathology

- **Salient features**: Paraphasic errors in PWA are related to personal, word, and clinical level features  
  - Salient features from this project include part of speech, frequency, imageability, participant’s age, aphasia duration, and severity of aphasia

- **Clinical implications**  
  - Diagnosis: integrated model that can filter and identify various communication patterns  
  - Treatment: selection of target words, treatment approaches, predictive output within AAC, etc.  
  - Prognosis: using accurate medical data, models increase predictive clinical outcomes
ADDITIONAL TAKEAWAYS

The combination of machine learning applications and speech pathology can facilitate:

**Challenging** and unique research questions within each field

**Fine tuning** AI language models for speech and communication applications

**New** interactive and accessible technologies for improved care of people with aphasia
Questions/Comments?

Rosa.zv04@gmail.com
jacob-brue@utulsa.edu
References


