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.

THE UNIVERSITY of TULSA

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

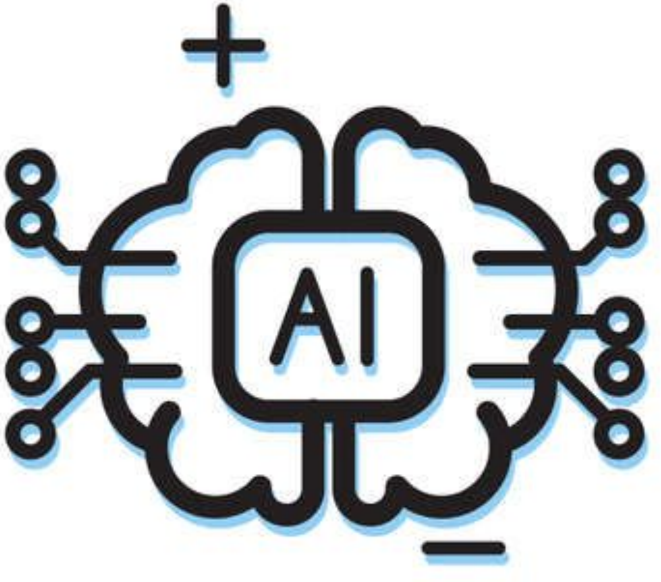
- Nonfinancial Relationships

-We used the AphasiaBank as the source of the data

- AphasiaBank is supported by NIH-NIDCD grant R01-DC008524 for 2022-2027

CONTENTS

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

PARAPHASIAS: 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 = **PERSONAL FEATURES** + **CLINICAL FEATURES** + **WORD LEVEL FEATURES**

DETECTION OF ERRORS	PERSONAL FEATURES	CLINICAL FEATURES	WORD LEVEL FEATURES
-the accuracy with which the model is able to detect the presence of a paraphasic errors	-age -years of education -sex -race	-aphasia type -aphasia severity -time post onset -comorbidity with apraxia -comorbidity with dysarthria	-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

UTTERANCES

Aphasiabank Transcription: Cindessa uh was a poor boy.

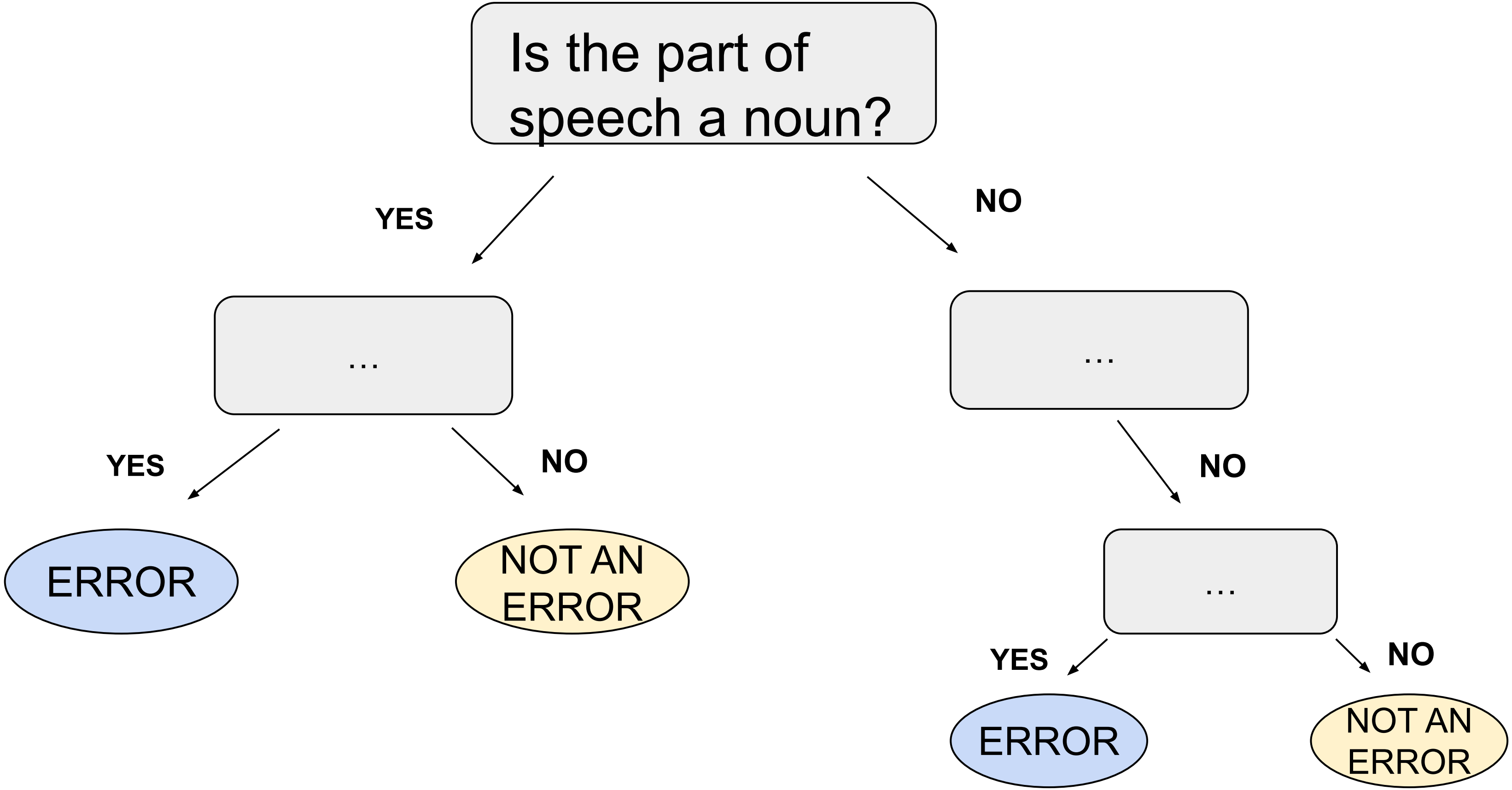
Target sentence: Cinderella was a poor girl.

features from aphasiabank	
error tag	error
intended word	girl
part of speech	noun
location	5

features from psycholinguistic database	
imageability	100
frequency	6.53
age of acquisition	320
length of word (grapheme number)	4

DECISION TREES

Target sentence: Cinderella was a poor girl.



HISTORY OF MODELS FOR CURRENT PROJECT

01

Preliminary detection model at word level

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

02

Final detection model at word level

Included previously excluded features

03

Preliminary detection model at sentence level

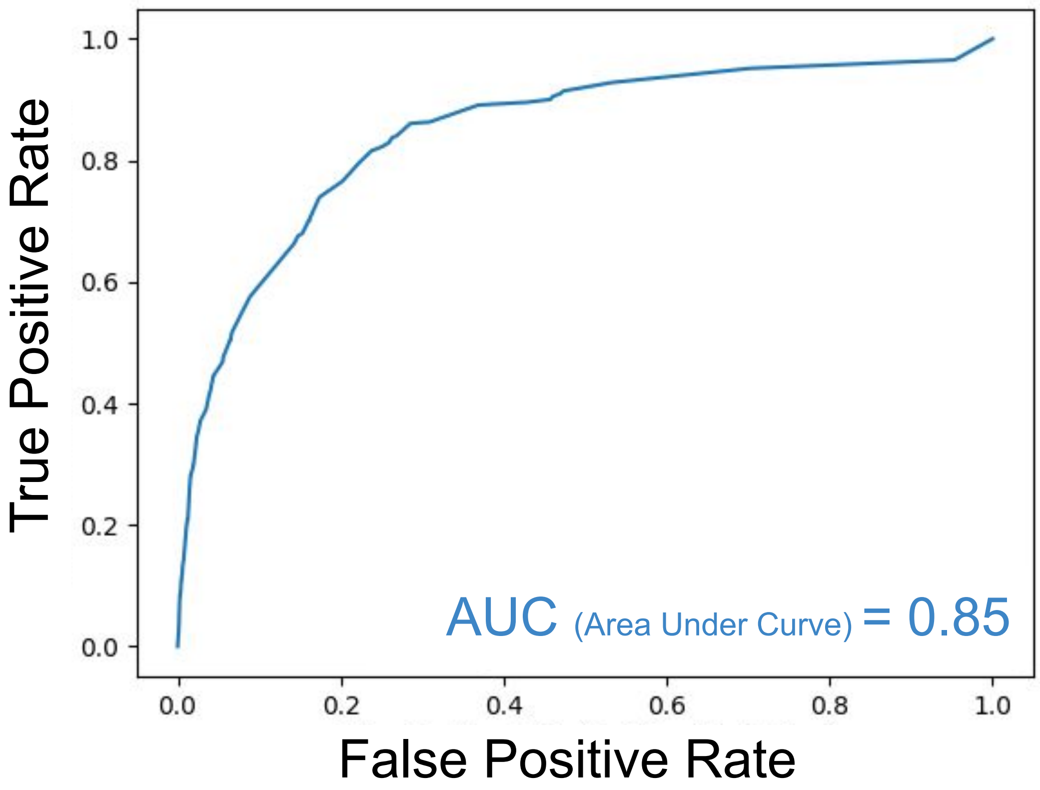
Included aggregate features of utterance and individual word features

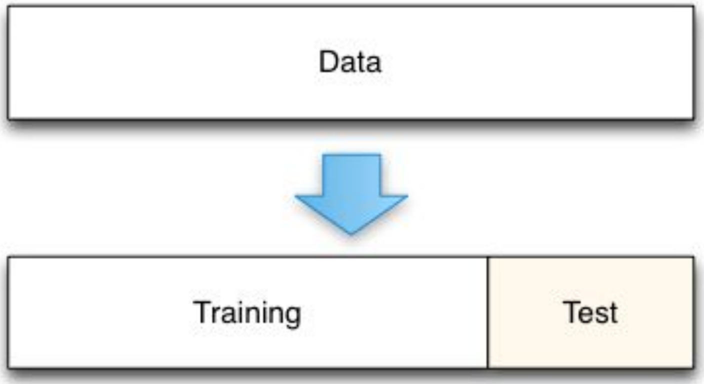
02

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





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

03 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

04

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

05

Classification model at the sentence level

Could include non core word features such as gestures, pauses, filler words, etc.

06

Prediction model for types of error

Given a speech sample, this model could predict the most likely errors a PWA may produce

07

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

Bird, H., Howard, D., & Franklin, S. (2003). Verbs and nouns: The importance of being imageable. *Journal of Neurolinguistics*, *16*(2–3), 113–149. https://doi.org/10.1016/S0911-6044(02)00016-7

Butterworth, B., Howard, D., & McLoughlin, P.J. (1984). The semantic deficit in aphasia: The relationship between semantic errors in auditory comprehension and picture naming. *Neuropsychologia, 22*, 409-426.

Dalton, S. G. H., Shultz, C., Henry, M. L., Hillis, A. E., & Richardson, J. D. (2018). Describing Phonological Paraphasias in Three Variants of Primary Progressive Aphasia. *American Journal of Speech-Language Pathology*, *27*(1 Suppl), 336–349. https://doi.org/10.1044/2017_AJSLP-16-0210

Le, Duc & Licata, Keli & Mower Provost, Emily. (2018). Automatic Quantitative Analysis of Spontaneous Aphasic Speech. *Speech Communication.* 100. 10.1016/j.specom.2018.04.001.

MacWhinney, B., Fromm, D., Forbes, M. & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology, 25,*1286-1307.