Analyzing Brain Age THROUGH Entity-Centric Analysis: Narratives in Picture Description Tasks

Building an advanced framework for linguistic analysis of cognitive decline

INTRODUCTION

Brain Aging

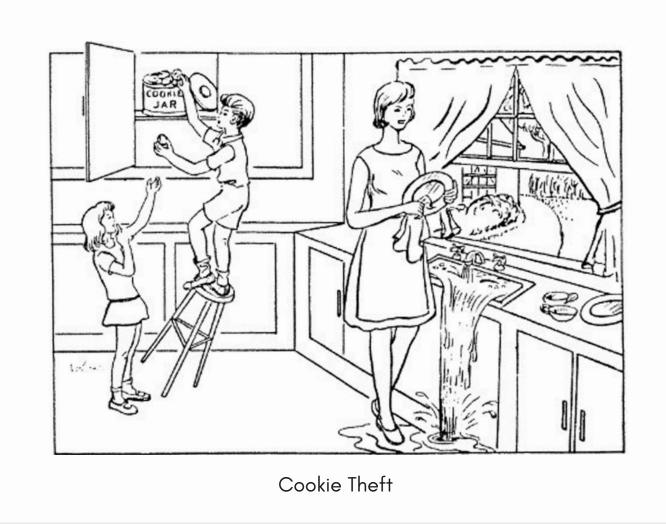
• Brain Aging is the process by which the brain change over time. Brain Age is distinct from biological age, is predicted using machine learning models based or MRI images.

2. Picture Description Tasks

• The subject is asked to "describe everything you see in this picture". Their response is transcribed for later analysis. For this project, the picture in question was "Cookie Theft", shown below.

3. Previous work

 Previous work analyzing Cookie Theft only used simple metrics like parts of speech tagging and keyword analysis, and Cookie Theft has never been used in the context of Brain Aging.



EXAMPLE PIPELINE

Entity List:	okay.
0 - woman/mom	um it 's in the <mark>kitchen</mark> .
1 - boy/brother	the water is on the uh I w water is running on the floor.
2 - girl/sister	looks like mom is doing the dishes.
	that looks like <mark>her daughter</mark> .
18 - window	and ano uh an another kid.
19 - curtains	it looks like <mark>her son</mark> .
20 - woman's dress	
21 - woman's shoes	and uh yeah Mom is wiping that dish uh plate.
22 - woman's hair	and the the boy is trying to get the cookie from the coo the cabinet.
42 - <mark>kitchen</mark>	the top came off.
 47 - cabinet handles	and he is on the chair, but the um the stool is about to t
48 - garden outside	and the girl is on the floor and um s uh trying to reach t cookie.
52 total Entities	maybe the uh h her brother is giv uh giving the cookie t
Ent	ity List + Annotated Transcript

Entity List + Annnotated Iranscript

METHODOLOGY

A three-dimensional approach to analyze discourse in the context of brain aging:

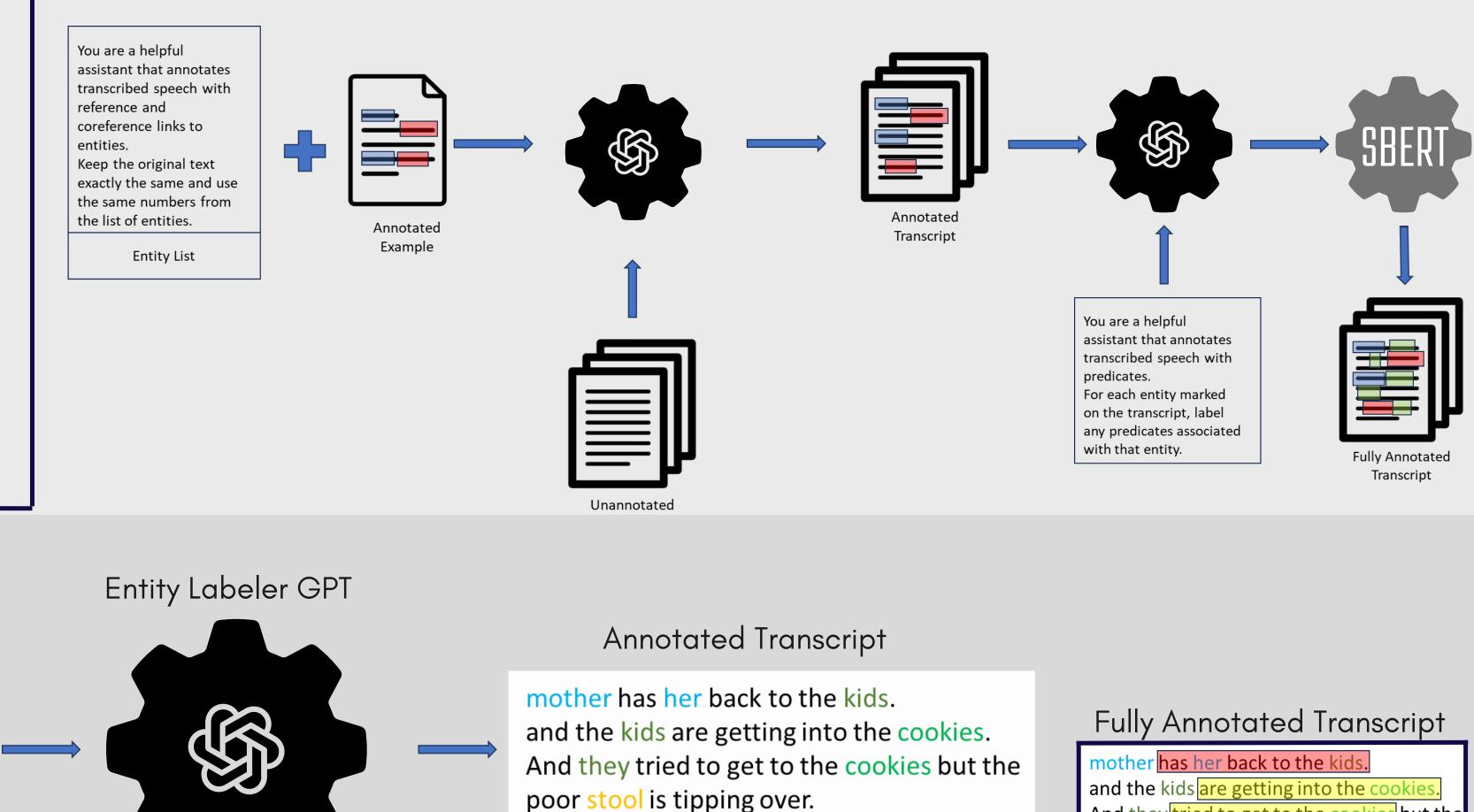
1. Lexical Richness

Analyzing word-level metrics including frequency, imageability, familiarity, and age of acquisition, helps us to measure the sophistication of language use among participants. 2. Syntactic Complexity

Using the Emory Language and Information Toolkit (ELIT), we quantify syntactic complexity by counting the number of descendents of each noun and verb in both dependency and constituency parse trees.

3. Semantic Content

We focus on the identification and linking of entities and predicates. Using GPT-4, we first identify mentions and links of entities in the picture "Cookie Theft", and then extract predicates associated with those entities. To link predicates, we cluster them with Sentence-BERT.



okie jar in

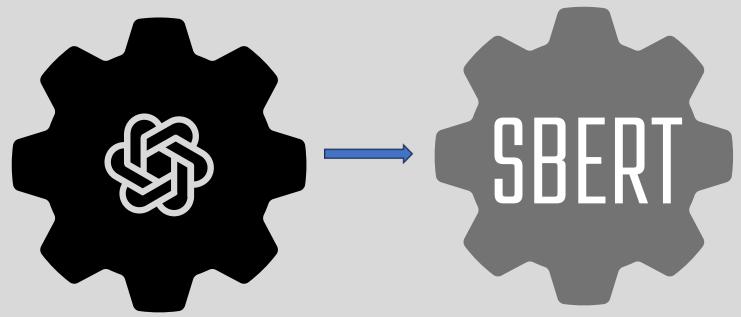
fall. the the

to <mark>her</mark>.

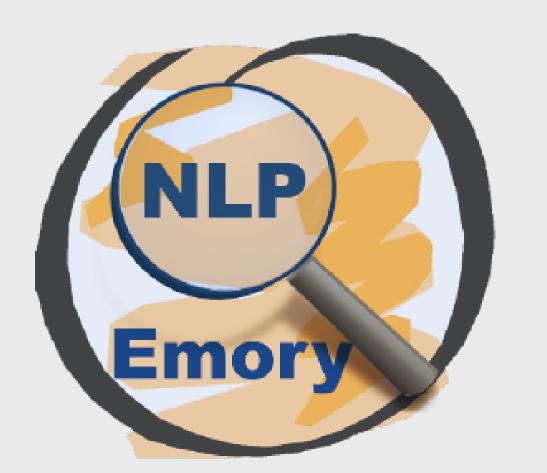
mother has her back to the kids. and the kids are getting into the cookies. And they tried to get to the cookies but the poor stool is tipping over.

Unannotated Transcript

Predicate Labeler GPT





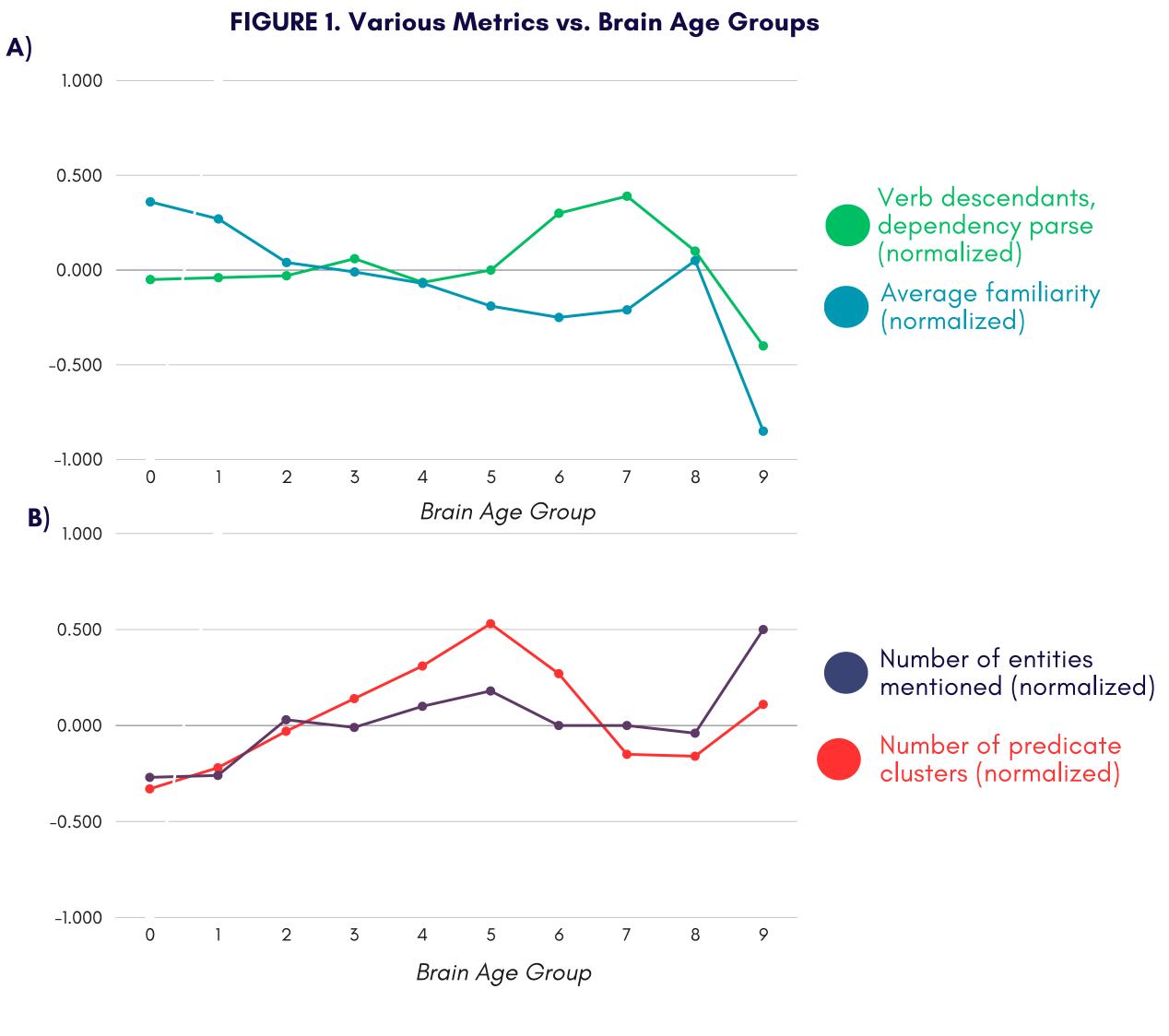


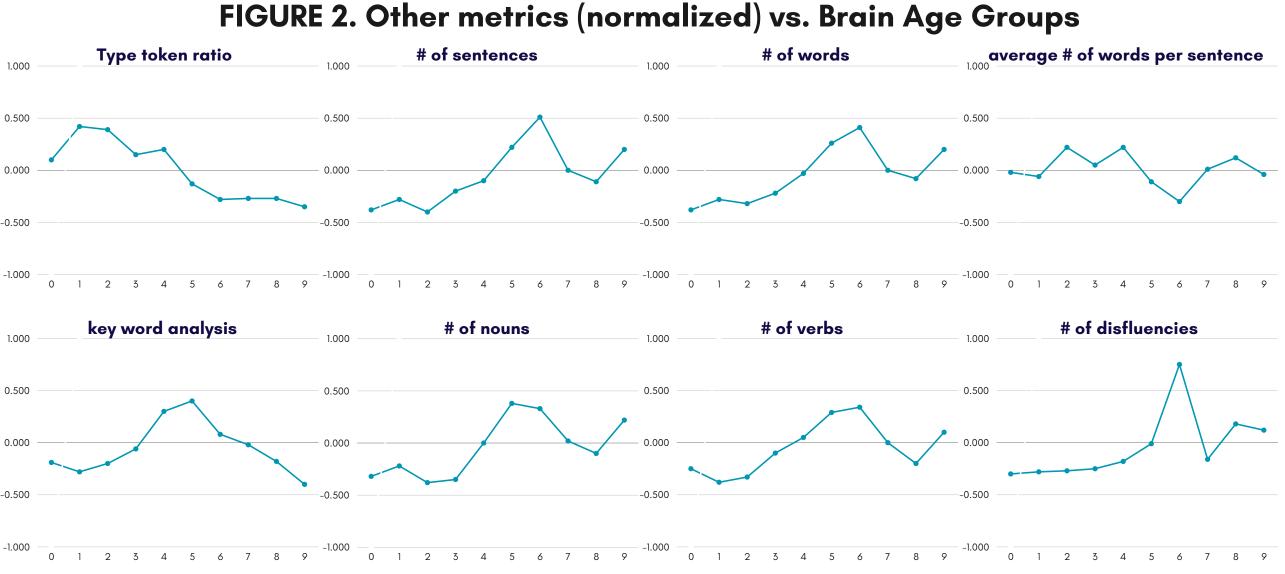


And they<mark>tried to get to the cookies</mark> but the ol is tipping over. poor sto Predicate Clusterer

RESULTS

Participants were sorted by brain age into 10 groups, to optimize for smoother graphs. Word level metrics indicated that participants in older age groups tended to use words with lower familiarity scores (A). Contrastingly, the number of verb descendents from each noun and verb increased slightly, then plummeted (A), suggesting a reduction in syntactic complexity. Older age groups tended to use a broader range of entities, but fewer predicates (B). This could indicate they discuss more entities without going into the same level of detail. Other results are displayed in Figure 2.





ACKNOWLEDGEMENTS

- this research possible.

• We would like to thank the Aging Brain Cohort (ABC) team for providing access to the dataset that made

• E.G. was supported by the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health (NIH) under Award Numbers UL1TR002378 and KL2TR002381.