HHS Public Access

Author manuscript

Neurosci Inform. Author manuscript; available in PMC 2024 March 01.

Published in final edited form as:

Neurosci Inform. 2024 March ; 4(1): . doi:10.1016/j.neuri.2023.100155.

A curious case of retrogenesis in language: Automated analysis of language patterns observed in dementia patients and young children

Changye Li[a,*], **Jacob Solinsky**[b], **Trevor Cohen**[c], **Serguei Pakhomov**[b]

[a]Institute of Health Informatics, University of Minnesota, Minneapolis, 55455, MN, USA

[b]College of Pharmacy, University of Minnesota, Minneapolis, 55455, MN, USA

[c]Division of Biomedical Informatics and Medical Education, University of Washington, Seattle, 98195, WA, USA

Abstract

Introduction: While linguistic retrogenesis has been extensively investigated in the neuroscientific and behavioral literature, there has been little work on retrogenesis using computerized approaches to language analysis.

Methods: We bridge this gap by introducing a method based on comparing output of a pre-trained neural language model (NLM) with an artificially degraded version of itself to examine the transcripts of speech produced by seniors with and without dementia and healthy children during spontaneous language tasks. We compare a range of linguistic characteristics including language model perplexity, syntactic complexity, lexical frequency and part-of-speech use across these groups.

Results: Our results indicate that healthy seniors and children older than 8 years share similar linguistic characteristics, as do dementia patients and children who are younger than 8 years.

Discussion: Our study aligns with the growing evidence that language deterioration in dementia mirrors language acquisition in development using computational linguistic methods based on

*Corresponding author. lixx3013@umn.edu (C. Li).

Author contributions

All authors attest that they meet the current International Committee of Medical Journal Editors (ICMJE) criteria for Authorship. **Changye Li**: conceptualization, investigation, methodology, software, writing - original draft, writing - review & editing; **Jacob Solinsky**: software, writing - original draft, writing - review & editing; **Trevor Cohen**: methodology, funding acquisition, writing original draft, writing - review & editing; **Serguei Pakhomov**: conceptualization, methodology, funding acquisition, writing - original draft, writing - review & editing.

Human and animal rights

The authors declare that the work described has not involved experimentation on humans or animals.

Informed consent and patient details

The authors declare that this report does not contain any personal information that could lead to the identification of the patient(s) and/or volunteers.

Declaration of competing interest

The authors declare that they have no known competing financial or personal relationships that could be viewed as influencing the work reported in this paper.

NLMs. This insight underscores the importance of further research to refine its application in guiding developmentally appropriate patient care, particularly in early stages.

Keywords

Retrogenesis; Alzheimer's disease; Computational linguistics; Natural language processing

1. Introduction

Retrogenesis refers to the observation that the loss of life skills caused by Alzheimer's Disease happens in the reverse order of their acquisition in childhood [1,2]. The evidence for retrogenesis collected up to this point can be categorized into behavioral and neuroscientific categories with the former obtained via longitudinal observational studies of child development and aging and the latter comprised of brain imaging studies. These studies can be further subcategorized by the type of brain phenomena measurable with modern imaging instrumentation and techniques including brain wave activity measured with electroencephalography (EEG), brain activation measured by hemodynamic response in functional magnetic resonance imaging (fMRI). The resulting data support several analytical techniques, including network analysis of functional connectivity and localization and differentiation of brain activity associated with development and deterioration of human functions and abilities.

The *behavioral* evidence for retrogenesis has been established in several studies showing inverse patterns of general functional and cognitive abilities development/deterioration in children and patients with Alzheimer's Disease [3,4] as well as language-specific patterns in lexico-semantic development/deterioration [5]. Pozueta et al. [4] demonstrated that patterns of concept acquisition in children were the mirror image of the loss in patients with semantic dementia. Reisberg [6] collated 16 successive functional stages and sub-stages of loss of capacity observed in Alzheimer's Disease patients into a Functional Assessment Staging (FAST) procedure extensively documented as having high concurrent and criterion validity with other independent cognitive measurements of Alzheimer's Disease in previous studies [7–9]. The sequence of functional loss captured by FAST appears to mirror the sequence of stages in normal human development (i.e., developmental ages) [10–12].

Much of the *neuroscientific* evidence comes from studies of human language (mostly language comprehension due to limitations of imaging technology and its sensitivity to motion artefacts present during language production). Several large brain regions including the left posterior superior temporal gyrus and sulcus (STG/STS) and left inferior frontal cortex (IFC) have been reported to be crucial for language comprehension (for a review, see Friederici [13]). Neuroimaging studies [14,15] further supported the argument that language comprehension in adults is lateralized specifically in the left hemisphere; however, Olulade et al. [16] found that the degree of language lateralization gradually shifts from both hemispheres to the left hemisphere during language development in children between age 5 and 11 years. Similarly, Meinzer et al. [17] observed that senior adults recruit additional right frontal areas to perform semantic tasks, indicating a more bilateral and reversed pattern compared to those at developmental ages (i.e., from infants to adolescents). Similar patterns

are observed in dementia patients as well, especially in the bilateral superior frontal gyrus (SFG) and inferior frontal gyrus (IFG) [18].

As children develop language (and other) abilities, their brains appear to undergo a process of specialization by which more localized brain regions become responsible for specific functions [19]. In Alzheimer's Disease, this process appears to be reversed - as the brain undergoes the process of neurodegeneration, the brain areas that once specialized in specific functions are no longer able to support these functions effectively and other brain areas are recruited to compensate for the deficits, which results in an apparent pattern of generalization (vs. specialization). Furthermore, this phenomenon of brain plasticity has been observed in the developing brains of children who sustained focal lesions in the left hemisphere and were also able to compensate by recruiting undamaged parts of the brain to support language function albeit with some residual deficits persisting into adulthood [20]. In addition to structural and functional imaging studies, more recent work has focused on investigations of functional brain connectivity (FC) which examine temporal correlations in spontaneous neural activation patterns (hemodynamic, electro- or magneto-encephalography) of activity within and between various brain regions (for a review, see Van Den Heuvel and Pol [21]).

Language acquisition is the developmental process of learning how to recognize, comprehend, produce and manage complex language. The neural basis of language acquisition and development has been extensively investigated with task-based fMRI studies [22,23] showing increased left lateralization in developmental ages. Specifically, children were found to have increased activity in the inferior/middle frontal, middle temporal, and angular gyri in the left hemisphere whereas decreased activity in the left posterior, superior frontal, and right anterior as their age increases. Such changes in the corresponding brain regions lead to improved language skills [24,22] as they affect the ability to retrieve and manipulate semantic information [25]. Other studies also indicate similar relationships between brain activity and language development. Kadis et al. [26] observed that the increased activity in the left inferior frontal gyrus and decreased activity in the right inferior frontal gyrus lead to better syntactic skills as childhood age increases. Similarly, Holland et al. [27] observed that children perform better on fluency tests as they age, and that this corresponds with increased brain activity in the left hemisphere and the degree of lateralization. These improved verbal fluency skills are further supported by work by Matute et al. [24] and Szaflarski et al. [22], showing that children score better (i.e., generate more words) as their age increases. Furthermore, Szaflarski et al. [22] found that language lateralization slowly decreases with age, which is associated with worse test scores.

The other side of retrogenesis – language deterioration – concerns a gradual reduction of language ability. As dementia progresses, it impairs multiple brain regions involved in the retrieval and manipulation of semantic information, resulting in a reversal of the process of language acquisition, with observable language deficits. Duarte et al. [28] found that dementia patients suffer from frontal, parietal and temporal lobe atrophies, which negatively impact performance on many language-related tasks [29]. Previous studies [30–32] indicate that dementia patients have lower scores for verbal fluency tasks compared to controls. Another widely-used diagnostic task for detecting language anomalies is the "Cookie Theft"

picture description (Fig. 3) task from the Boston Diagnostic Aphasia Examination [33]. Studies have identified subtle language differences in picture descriptions by dementia patients including overuse of pronouns [34], repetition and diminished syntactic complexity [35], and reliance on words with higher lexical frequency [34].

Benítez-Burraco and Ivanova [36] recently present evidence supporting the hypothesis of retrogenesis from healthy aging adults and children. However, there have been no previous investigations into retrogenesis utilizing computerized approaches for language analysis in dementia and developmental ages. Computerized approaches to conversational speech and language analysis present an appealing alternative to traditional assessments of isolated linguistic abilities (e.g., verbal fluency tests), which can be labor-intensive, time-consuming, and insensitive to complex linguistic patterns [37]. Modern neural language modeling approaches have been recently introduced for the characterization and quantification of linguistic anomalies in dementia and other neurodegenerative disorders [38–41]. In prior work, we have demonstrated that a pre-trained neural language model (NLM), GPT-2 [42] can be artificially impaired to simulate language deficits typical of patients with dementia resulting in a degraded GPT-2 model (GPT-D) [41]. In the current study, we hypothesize that computerized neural language modeling can be used to elucidate the phenomenon of retrogenesis in a quantifiable and reproducible fashion by providing access to several speech and language characteristics derived with computational linguistic methods.

Based on this prior work, we anticipated that GPT-2, which has been pre-trained on text produced by presumably cognitively healthy individuals, would be more “confused” by language produced by younger children and participants with dementia. In contrast we hypothesized that GPT-D, which has been deliberately degraded to induce linguistic anomalies that occur in dementia, would be more “confused” by transcripts produced by older children and healthy seniors, thereby showing that healthy seniors and older children share similar language characteristics on the one hand, and dementia patients and younger children on the other.

Our contributions are as follows: a) we applied a novel analytic technique leveraging NLMs for a comparative analysis of transcribed spontaneous speech from older participants with and without dementia, and children ranging in age from 5 to 11; b) our results indicate that healthy seniors and older children share similar linguistic characteristics on the one hand; as do dementia patients and younger children on the other; c) our study demonstrates that language deterioration observed in dementia patients mirrors aspects of language acquisition and can be estimated with computerized approach using NLMs.

2. Materials and methods

2.1. Dataset

The Pitt corpus from Dementia Bank[1] contains recordings and transcripts of responses to the “Cookie Theft” task (Fig. 3 in Appendix A) and cognitive test results. We focused on

[1] http://talkbank.org/DementiaBank/.

participants assessed with probable Alzheimer's Disease, with data from 169 patients and 99 healthy individuals.

Test of Narrative Language (TNL) [43][2] contains assessments of narrative language acquisition from school-aged children (5 to 11 years old) using story comprehension and retelling tasks. It includes data from 173 children with specific language impairment and 498 healthy children who participated in the storytelling task. We focused on the transcripts from the single picture description task for subsequent analysis, as it is most similar to the "Cookie Theft" task.

The detailed dataset characteristics of the Pitt corpus and TNL are provided in Table 1.

2.1.1. Text pre-processing—We performed basic pre-processing of both verbatim transcripts using TRESTLE (**T**oolkit for **R**eproducible **E**xecution of **S**peech **T**ext and **L**anguage **E**xperiments)[3] [44] to remove references to speech artifacts and speech that did not belong to the participant. Some participants in the Pitt corpus produced multiple picture descriptions on separate study visits. We performed participant-level analysis by averaging our calculations across all available transcripts for each participant.

2.2. Transformer-based model and paired-perplexity approach

The attention mechanism [45], as implemented within the transformer architecture, [46] has revolutionized language modeling. Auto-regressive transformer NLMs with attention are multilayer networks trained to predict words based on preceding context. The attention mechanism involves matrices that learn associations between predicted words and specific words in the preceding context. Effective transformer models require extensive training on large corpora, making pre-trained general-purpose models a useful starting point for specific tasks. Perplexity (PPL) is an intrinsic evaluation metric that measures how well a probability model predicts a sample [47] (see Equation (1)). Lower perplexity indicates better prediction on unseen data.

$$PPL(S) = P(S)^{-\frac{1}{n}} \tag{1}$$

where S represents a sequence of words $(w_1, ..., w_n)$ containing n words, and $P(S)$ denotes the conditional probability assigned to sequence S as shown in Equation (2):

$$\begin{aligned} P(S) &= P(w_1 \mid w_0) \cdots P(w_n \mid w_0 w_{1 \cdots w_{N-1}}) \\ &= \prod_{k=1}^{n} P(w_k \mid w_0 \cdots w_{k-1}) \end{aligned} \tag{2}$$

[2] https://childes.talkbank.org/access/Clinical-MOR/Gillam.html.
[3] https://github.com/LinguisticAnomalies/harmonized-toolkit.

The *paired perplexity* paradigm [39] employs two models to construct perplexity pairs for transcripts generated by healthy controls and dementia participants, respectively. Specifically, $LM_{dementia}$ is trained on all transcripts from the dementia group, while $LM_{control}$ represents speech from healthy controls using their respective transcripts. Both models evaluate held-out transcripts, resulting in perplexity values $PPL_{control}$ and $PPL_{dementia}$. Transcripts are classified through binary decisions based on the threshold $PPL_{diff} = PPL_{control} - PPL_{dementia}$ when both models have the same equal error rate (EER). Intuitively, models trained on language from cognitively healthy participants should be surprised by language from dementia patients, and vice versa. The difference between paired perplexity values from "cognitively healthy" and "dementia" language models yields state-of-the-art results [39,40] in identifying dementia using a single interpretable NLM-derived feature.

2.3. GPT-2 and GPT-D model

GPT-2 is an auto-regressive Transformer model with 12 layers and 12 attention heads per layer. Each attention head consists of key, query and value matrices to estimate the relationships between different elements in a sequence. These matrices transform representations of units of an input sequence as they propagate from layer to layer, to expose different aspects of the information that they contain. The query and key matrices are used to determine the importance of other tokens to a token of interest. The value matrix exposes the linguistic information to be attended to. Representations are composed from token-specific outputs of the value matrix based on the attention scores produced by the querykey interactions.

GPT-2 is pre-trained on language presumably produced by cognitively healthy individuals. However, the lack of a corresponding large dataset from cognitively impaired participants poses challenges when applying the paired perplexity paradigm. To address this, a previous study [41] introduced GPT-D, in which masks portions of parameters in the value matrices. Pairing GPT-2 and GPT-D produced robust performance in identifying transcripts produced by dementia patients, both in responses to cognitive tasks and casual conversation. GPT-D also was shown to mimic language behaviors commonly found in dementia patients.

2.4. Syntactic complexity & mean log lexical frequency

Syntactic complexity is a crucial measurement for various language-related tasks, including identifying dementia and Mild Cognitive Impairment from writing [48] and speech [49], and evaluating children's language acquisition [50]. Three main methods are used to calculate syntactic complexity: a) the Yngve [51] method; b) the Frazier [52] method; and c) the syntactic dependency length (SDL) between lexical items in the sentence [53–55].

The Yngve method determines the Yngve depth (YD) by analyzing the tree representation of a syntactically parsed sentence. It scans the tree from top-down, right-to-left, assigning a score of 0 to the rightmost branch and incrementing by 1 for each level towards the left. Each word, representing a leaf node, receives a score that accumulates from the root node to the word. In contrast, the Frazier score emphasizes syntactical tree depth. It assigns a score of 1 to words connected directly to the root and 1.5 to leaf nodes whose branches are the

lowest node not situated as the leftmost child of its parent, continuing until no such node is found. This process yields the Frazier depth (FD). Unlike Yngve and Frazier, which focus on tree breadth or depth, the SDL method calculates the non-hierarchical length of grammatical dependencies between lexical items in the sentence.

We estimated the following scores for each transcript from Pitt corpus and TNL using the Computerized Linguistic Analysis System (CLAS) [48], a fully automatic system developed using the AllenNLP library [56]. Besides the syntactic complexity metrics, we also computed the averaged POS tags count and number of clauses for each transcript.

2.5. Evaluation

We adapted the paired perplexity paradigm, to evaluate transcripts from both the Pitt corpus and TNL datasets using GPT-2 and GPT-D, respectively. The generated PPLs from each model were normalized using Equation (3), where i represents the i-th transcript out of m total transcripts from the corresponding dataset.

$$\left|PPL\right| = PPL_i - \frac{\sum_{i=1}^{i=m} PPL_i}{m} \tag{3}$$

We computed the normalized syntactic complexity, including Yngve, Frazier, and SDL, for each transcript, denoted as $\bar{C}_t$, using Equation (4). Here, $score_j$ represents the score for each word in a sentence containing a total of w words, and we calculated the sum of scores for each word on the sentence-level, c_i, and then obtained the average complexity by summing the scores on the sentence level and dividing by the total number of sentences n.

$$\begin{aligned}\bar{C}_t &= \frac{\sum_{i=1}^{n} c_i}{n} \\ &= \frac{\sum_{i=1}^{n} 1\left(\sum_{j=1}^{w} score_j\right)}{n}\end{aligned} \tag{4}$$

We defined the age segmentation of “younger” and “older” children is defined as the intersection of normalized PPLs between GPT-2 and GPT-D. Intuitively, GPT-D should have higher perplexity on transcripts generated by “older” children and healthy seniors with higher syntactic complexity, resulting in higher PPLs. Conversely, GPT-D should be less perplexed by transcripts generated by “younger” children and dementia patients, leading to lower PPLs. Conversely, GPT-2 is expected be perplexed by transcripts from “younger” children and dementia patients, producing higher PPLs, while it is expected to yield lower PPLs for transcripts from “older” children and healthy elders.

3. Results

As illustrated in Fig. 1, with children younger than 8 years old categorized as “younger” and those older than 8 years old as “older.” As shown in Fig. 2, dementia patients exhibit lower syntactic complexity than healthy elders across various measures, including the number of clauses, FD, YD, and SDL. A similar trend is observed for children, with older children displaying more complex linguistic behaviors. Additionally, we found that dementia patients and younger children demonstrate similar levels of syntactic complexity, while healthy elders remain at a similar level to older children. We tested these findings for significance using a two-sided Welch’s t-test (alpha = 0.01) of the average syntactic complexity measures of all conceivable subset combinations of younger/older children from TNL and dementia/healthy seniors from the Pitt corpus.

As shown in Table 2, there were no significant differences in syntactic complexity (FD, YD, and SDL) between transcripts from dementia patients and younger children. Similarly, no differences were found in syntactic complexity between healthy seniors and older children. Although we observed significant differences in the average counts of adjectives and pronouns between dementia patients and younger children, as well as the count of averaged adjectives between healthy seniors and older children (as detailed in Table 3), no similar patterns emerged for the remaining parts-ofspeech (POS) tags. Significant differences were also observed on the within-group comparison. Specifically, Table 2 shows significant differences in the number of clauses, FD, YD, and SDL between the transcripts produced by dementia patients and healthy seniors, as well as between younger and older children. Similarly, we found that the majority of POS tags for the within-group comparison remained significantly different for dementia patients and healthy seniors, as well as younger children and older children.

4. Discussion

Our key findings are summarized as follows: a) automated methods, including state-of-the-art NLP models, reflect retrogenesis in language samples produced by children and elders, particularly dementia patients; b) with GPT-D, children under the age of 8 exhibit linguistic behaviors similar to dementia patients, and this pattern also applies to children above the age of 8 and healthy seniors.

The current work applies GPT-D to the language behaviors of “younger” children. A previous study [40] showed that control and dementia LMs react differently to narratives with higher and lower lexical frequencies, respectively. The current work confirms this finding, and shows that it applies in the context of language from children also. Our analysis of syntax is also consistent with prior research, indicating that dementia patients exhibit reduced syntactic complexity [57,35], while children show increased syntactic complexity as they age [58,59].

With the paired perplexity paradigm, GPT-D reveals a comparable pattern in linguistic behaviors used by dementia patients and younger children, while GPT-2 suggests a similar pattern in linguistic behaviors used by healthy elders and older children. As such, our

findings provide automatically estimated evidence of retrogenesis, indicating a possible reversed mirror relationship between language deterioration observed in Alzheimer's Disease and language acquisition in developmental ages.

Fig. 2 and Table 2 provide additional support for the language retrogenesis effect between dementia patients and younger children, indicating a comparable level of syntactic complexity. Likewise, similar patterns are noted between healthy elders and older children. Our results (Table 3) indicate that both younger children and dementia patients exhibit comparable overuse of pronouns previously observed in dementia [34].

The paired perplexity paradigm provides an automated and objective measurement of changes in linguistic behavior observed in retrogenesis using NLMs. It enhances our understanding of the intricate relationship between language deterioration and cognitive decline, offering a comprehensive view of the underlying neural processes during language development and degeneration. While further research on longitudinal and diverse corpora is needed, exploring language deterioration and language acquisition with NLMs may inform the development of targeted interventions for language-related conditions in both dementia patients and children.

4.1. Limitations

The presented work has several limitations. Firstly, all datasets are in American English, which may introduce language bias influencing the language patterns between children and dementia patients. Additionally, the content differences in picture description tasks used in Pitt corpus and TNL could potentially impact the observed language patterns. The age segmentation of "younger" and "older"applies only to the TNL dataset, and we did not interpret that cognition and language ability necessarily change monotonically with age. While GPT-2 and GPT-D reflect the linguistic behaviors of healthy or dementia individuals, and older or younger children, respectively, they should not be considered as accurate, comprehensive neural representations of human cognition and language.

5. Conclusion

Our study demonstrates the potential of using state-of-the-art NLP models and the concept of retrogenesis to monitor language acquisition and identify linguistic changes associated with dementia patients and younger children. Our study provides additional evidence for the neuroimaging studies for better understanding the cognitive and linguistic similarities between younger children and dementia patients. Our work contributes to the growing understanding of language acquisition and deterioration in dementia, offering promising prospects for early detection and tailored care interventions. Future research is needed to validate these findings through longitudinal studies and diverse language samples to strengthen the reliability and generalizability of our results.

Acknowledgement

This research was supported by a grant from the National Institute on Aging (AG069792).

Appendix A

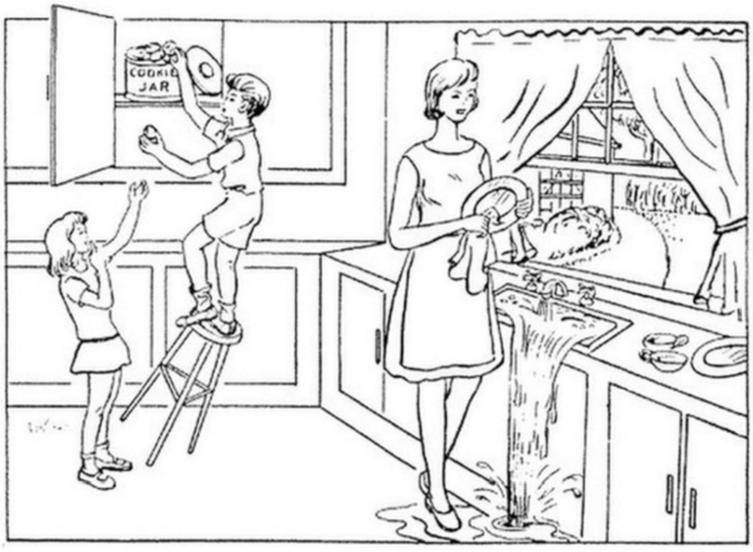

Fig. 3.
Cookie Theft picture stimulus. Participants are asked to describe everything they see going on in the picture.

References

[1]. Reisberg B, Kenowsky S, Franssen EH, Auer SR, Souren LE, Towards a science of Alzheimer's disease management: a model based upon current knowledge of retrogenesis, Int. Psychogeriatr 11 (1999) 7–23, 10.1017/s1041610299005554. [PubMed: 10189596]

[2]. Reisberg B, Franssen EH, Hasan SM, Monteiro I, Boksay I, Souren LE, Kenowsky S, Auer SR, Elahi S, Kluger A, Retrogenesis: clinical, physiologic, and pathologic mechanisms in brain aging, Alzheimer's and other dementing processes, Eur. Arch. Psychiatry Clin. Neurosci 249 (1999) S28–S36, 10.1007/pl00014170.

[3]. Rubial-Alvarez S, de Sola S, Machado M-C, Sintas E, Boehm P, SánchezBenavides G, Langohr K, Muniz R, Peña-Casanova J, The comparison of cognitive and functional performance in children and Alzheimer's disease supports the retrogenesis model, J. Alzheimer's Dis 33 (2013) 191–203, 10.3233/JAD-2012-121123. [PubMed: 22948284]

[4]. Pozueta A, Lage C, García-Martínez M, Kazimierczak M, Bravo M, López-García S, Fernández-Rodríguez A, Riancho J, González-Suárez A, Vázquez-Higuera JL, Ángel Cano-Abascal F. Martínez-Dubarbie, de ArcochaTorres M, Jiménez-Bonilla J, Banzo I, Rodríguez-Rodríguez E, Sánchez-Juan P, A snake that bites its own tail. Acquisition and loss of concepts in children and semantic dementia patients through the analysis of drawings, Cortex 128 (2020) 162–173, 10.1016/j.cortex.2020.03.007, https://www.sciencedirect.com/science/article/pii/S0010945220301088. [PubMed: 32361266]

[5]. Simoes Loureiro I, Lefebvre L, Retrogenesis of semantic knowledge: Comparative approach of acquisition and deterioration of concepts in semantic memory, Neuropsychology 30 (2016) 853, 10.1037/neu0000272. [PubMed: 26913484]

[6]. Reisberg B, Functional assessment staging (FAST), Psychopharmacol. Bull 24 (1988) 653–659. [PubMed: 3249767]

[7]. Franssen EH, Reisberg B, Neurologic markers of the progression of Alzheimer's disease, Int. Psychogeriatr 9 (1997) 297–306, 10.1017/s1041610297005036. [PubMed: 9447450]

[8]. Bobinski M, Wegiel J, Wisniewski H, Tarnawski M, Mlodzik B, Reisberg B, De Leon M, Miller D, Atrophy of hippocampal formation subdivisions correlates with stage and duration of Alzheimer disease, Dement. Geriatr. Cogn. Disord 6 (1995) 205–210, 10.1159/000106948.

[9]. Bobinski M, Wegiel J, Tarnawski M, Bobinski M, Reisberg B, De Leon MJ, Miller DC, Wisniewski HM, Relationships between regional neuronal loss and neurofibrillary changes in the hippocampal formation and duration and severity of Alzheimer disease, J. Neuropathol. Exp. Neurol 56 (1997) 414–420, 10.1097/00005072-199704000-00010. [PubMed: 9100672]

[10]. Reisberg B, Dementia: a systematic approach to identifying reversible causes, Geriatrics (Basel) 41 (1986) 30–46.

[11]. Reisberg B, Functional degenerative stages in dementia of the Alzheimer's type appear to reverse normal human development, Biol. Psychiatry 7 (1986) 1319–1321, 10.1177/153331750201700411.

[12]. Reisberg B, Pattschull-Furlan A, Franssen E, Sclan SG, Kluger A, Dingcong L, Ferris SH, Dementia of the Alzheimer type recapitulates ontogeny inversely on specific ordinal and temporal parameters, in: Neurodevelopment, Aging and Cognition, Springer, 1992, pp. 345–369.

[13]. Friederici AD, The neural basis of language development and its impairment, Neuron 52 (2006) 941–952, 10.1016/j.neuron.2006.12.002. [PubMed: 17178399]

[14]. Knecht S, Deppe M, Dräger B, Bobe L, Lohmann H, Ringelstein E-B, Henningsen H, Language lateralization in healthy right-handers, Brain 123 (2000) 74–81, 10.1093/brain/123.1.74. [PubMed: 10611122]

[15]. Parker GJ, Luzzi S, Alexander DC, Wheeler-Kingshott CA, Ciccarelli O, Ralph MAL, Lateralization of ventral and dorsal auditory-language pathways in the human brain, NeuroImage 24 (2005) 656–666, 10.1016/j.neuroimage.2004.08.047. [PubMed: 15652301]

[16]. Olulade OA, Seydell-Greenwald A, Chambers CE, Turkeltaub PE, Dromerick AW, Berl MM, Gaillard WD, Newport EL, The neural basis of language development: Changes in lateralization over age, Proc. Natl. Acad. Sci 117 (2020) 23477–23483, 10.1073/pnas.1905590117. [PubMed: 32900940]

[17]. Meinzer M, Flaisch T, Wilser L, Eulitz C, Rockstroh B, Conway T, Gonzalez-Rothi L, Crosson B, Neural signatures of semantic and phonemic fluency in young and old adults, J. Cogn. Neurosci 21 (2009) 2007–2018, 10.1162/jocn.2009.21219. [PubMed: 19296728]

[18]. Paek EJ, Murray LL, Newman SD, Neural correlates of verb fluency performance in cognitively healthy older adults and individuals with dementia: a pilot fMRI study, Front. Aging Neurosci 12 (2020) 73, 10.3389/fnagi.2020.00073. [PubMed: 32265685]

[19]. Brown TT, Lugar HM, Coalson RS, Miezin FM, Petersen SE, Schlaggar BL, Developmental changes in human cerebral functional organization for word generation, Cereb. Cortex 15 (2005) 275–290, 10.1093/cercor/bhh129. [PubMed: 15297366]

[20]. MacWhinney B, Feldman H, Sacco K, Valdes-Perez R, Online measures of basic language skills in children with early focal brain lesions, Brain Lang. 71 (2000) 400–431, 10.1006/brln.1999.2273. [PubMed: 10716870]

[21]. Van Den Heuvel MP, Pol HEH, Exploring the brain network: a review on resting-state fMRI functional connectivity, Eur. Neuropsychopharmacol 20 (2010) 519–534, 10.1016/j.euroneuro.2010.03.008. [PubMed: 20471808]

[22]. Szaflarski JP, Holland SK, Schmithorst VJ, Byars AW, fMRI study of language lateralization in children and adults, Hum. Brain Mapp 27 (2006) 202–212, 10.1002/hbm.20177. [PubMed: 16035047]

[23]. Szaflarski JP, Schmithorst VJ, Altaye M, Byars AW, Ret J, Plante E, Holland SK, A longitudinal functional magnetic resonance imaging study of language development in children 5 to 11 years old, Ann. Neurol 59 (2006) 796–807, 10.1002/ana.20817. [PubMed: 16498622]

[24]. Matute E, Rosselli M, Ardila A, Morales G, Verbal and nonverbal fluency in Spanish-speaking children, Dev. Neuropsychol 26 (2004) 647–660, 10.1207/s15326942dn2602_7. [PubMed: 15456689]

[25]. Mummery CJ, Patterson K, Wise RJ, Vandenbergh R, Price C, Hodges J, Disrupted temporal lobe connections in semantic dementia, Brain 122 (1999) 61–73, 10.1093/brain/122.1.61. [PubMed: 10050895]

[26]. Kadis DS, Pang EW, Mills T, Taylor MJ, McAndrews MP, Smith ML, Characterizing the normal developmental trajectory of expressive language lateralization using magnetoencephalography, J. Int. Neuropsychol. Soc 17 (2011) 896–904, 10.1017/S1355617711000932. [PubMed: 21813032]

[27]. Holland SK, Plante E, Byars AW, Strawsburg RH, Schmithorst VJ, Ball WS Jr., Normal fMRI brain activation patterns in children performing a verb generation task, NeuroImage 14 (2001) 837–843, 10.1006/nimg.2001.0875. [PubMed: 11554802]

[28]. Duarte A, Hayasaka S, Du A, Schuff N, Jahng G-H, Kramer J, Miller B, Weiner M, Volumetric correlates of memory and executive function in normal elderly, mild cognitive impairment and Alzheimer's disease, Neurosci. Lett 406 (2006) 60–65, 10.1016/j.neulet.2006.07.029. [PubMed: 16904823]

[29]. Stuss DT, Alexander MP, Floden D, Binns MA, Levine B, McIntosh AR, Rajah N, Hevenor SJ, Fractionation and localization of distinct frontal lobe processes: Evidence from focal lesions in humans, in: Principles of Frontal Lobe Function, Oxford University Press, 2002, pp. 392–407.

[30]. Davis C, Heidler-Gary J, Gottesman R, Crinion J, Newhart M, Moghekar A, Soloman D, Rigamonti D, Cloutman L, Hillis A, Action versus animal naming fluency in subcortical dementia, frontal dementias, and Alzheimer's disease, Neurocase 16 (2010) 259–266, 10.1080/13554790903456183. [PubMed: 20104387]

[31]. Gomez RG, White DA, Using verbal fluency to detect very mild dementia of the Alzheimer type, Arch. Clin. Neuropsychol 21 (2006) 771–775, 10.1016/j.acn.2006.06.012. [PubMed: 17011743]

[32]. Alegret M, Peretó M, Pérez A, Valero S, Espinosa A, Ortega G, Hernández I, Mauleón A, Rosende-Roca M, Vargas L, et al. , The role of verb fluency in the detection of early cognitive impairment in Alzheimer's disease, J. Alzheimer's Dis 62 (2018) 611–619. [PubMed: 29480180]

[33]. Goodglass H, Kaplan E, Weintraub S, BDAE: The Boston Diagnostic Aphasia Examination, Lippincott Williams & Wilkins, Philadelphia, PA, 2001.

[34]. Almor A, Kempler D, MacDonald MC, Andersen ES, Tyler LK, Why do Alzheimer patients have difficulty with pronouns? Working memory, semantics, and reference in comprehension and production in Alzheimer's disease, Brain Lang. 67 (1999) 202–227, 10.3233/JAD-170826. [PubMed: 10210631]

[35]. Hier DB, Hagenlocker K, Shindler AG, Language disintegration in dementia: Effects of etiology and severity, Brain Lang. 25 (1985) 117–133, 10.1016/0093-934X(85)90124-5, https://www.sciencedirect.com/science/article/pii/0093934X85901245. [PubMed: 2411334]

[36]. Benítez-Burraco A, Ivanova O, Revisiting the hypothesis of language retrogenesis from an evolutionary perspective, Neuropsychology (2023).

[37]. Giles E, Patterson K, Hodges JR, Performance on the Boston cookie theft picture description task in patients with early dementia of the Alzheimer's type: missing information, Aphasiology 10 (1996) 395–408, 10.1080/02687039608248419.

[38]. Klumpp P, Fritsch J, Noeth E, ANN-based Alzheimer's disease classification from bag of words, in: Speech Communication, 13th ITG-Symposium, 2018, pp. 1–4.

[39]. Fritsch J, Wankerl S, Nöth E, Automatic diagnosis of Alzheimer's disease using neural network language models, in: ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 5841–5845.

[40]. Cohen T, Pakhomov S, A tale of two perplexities: Sensitivity of neural language models to lexical retrieval deficits in dementia of the Alzheimer's type, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, Association for Computational Linguistics, 2020, pp. 1946–1957, https://aclanthology.org/2020.acl-main.176.

[41]. Li C, Knopman D, Xu W, Cohen T, Pakhomov S, GPT-D: Inducing dementiarelated linguistic anomalies by deliberate degradation of artificial neural language models, in: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Dublin, Ireland, 2022, pp. 1866–1877, https://aclanthology.org/2022.acl-long.131.

[42]. Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I, Language models are unsupervised multitask learners, OpenAI Blog 1 (2019) 9.

[43]. Gillam RB, Pearson NA, TNL: Test of Narrative Language, PRO-ED, Austin, TX, 2004.

[44]. Li C, Xu W, Cohen T, Michalowski M, Pakhomov S, Trestle: Toolkit for reproducible execution of speech, text and language experiments, AMIA Summits Transl. Sci. Proc 2023 (2023) 360. [PubMed: 37350929]

[45]. Bahdanau D, Cho K, Bengio Y, Neural machine translation by jointly learning to align and translate, arXiv preprint, arXiv:14090473, 2014.

[46]. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser LU, Polosukhin I, Attention is all you need, in: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R. (Eds.), Advances in Neural Information Processing Systems, vol. 30, Curran Associates, Inc., 2017, https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

[47]. Jelinek F, Mercer R, Bahl L, Baker J, Perplexity—a measure of the difficulty of speech recognition tasks, J. Acoust. Soc. Am 62 (1977).

[48]. Pakhomov S, Chacon D, Wicklund M, Gundel J, Computerized assessment of syntactic complexity in Alzheimer's disease: A case study of Iris Murdoch's writing, Behav. Res. Methods 43 (2011) 136–144, 10.3758/s13428-010-0037-9. [PubMed: 21287110]

[49]. Roark B, Mitchell M, Hollingshead K, Syntactic complexity measures for detecting mild cognitive impairment, in: Biological, Translational, and Clinical Language Processing, Association for Computational Linguistics, Prague, Czech Republic, 2007, pp. 1–8, https://aclanthology.org/W07-1001.

[50]. Lu X, Automatic measurement of syntactic complexity in child language acquisition, Int. J. Corpus Linguist 14 (2009) 3–28.

[51]. Yngve VH, A model and an hypothesis for language structure, Proc. Am. Philos. Soc 104 (1960) 444–466, http://www.jstor.org/stable/985230.

[52]. Frazier L, Syntactic complexity, in: Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives, 1985, pp. 129–189.

[53]. Gibson E, Linguistic complexity: Locality of syntactic dependencies, Cognition 68 (1998) 1–76, 10.1016/S0010-0277(98)00034-1. [PubMed: 9775516]

[54]. Lin D, On the structural complexity of natural language sentences, in: COLING 1996: The 16th International Conference on Computational Linguistics, Volume 2, 1996.

[55]. de Marneffe M-C, Manning CD, The Stanford typed dependencies representation, in: Coling 2008: Proceedings of the Workshop on Cross-Framework and Cross-Domain Parser Evaluation, Coling 2008 Organizing Committee, Manchester, UK, 2008, pp. 1–8, https://aclanthology.org/W08-1301.

[56]. Gardner M, Grus J, Neumann M, Tafjord O, Dasigi P, Liu NF, Peters M, Schmitz M, Zettlemoyer LS, AllenNLP: A deep semantic natural language processing platform, arXiv:1803.07640,2017.

[57]. Lyons K, Kemper S, Labarge E, Ferraro FR, Balota D, Storandt M, Oral language and Alzheimer's disease: A reduction in syntactic complexity, Aging Neuropsychol. Cogn 1 (1994) 271–281, 10.1080/13825589408256581.

[58]. Klecan-Aker JS, Hedrick DL, A study of the syntactic language skills of normal school-age children, Lang. Speech Hear. Serv. Schools 16 (1985) 187–198, 10.1044/0161-1461.1603.187.

[59]. Delage H, Frauenfelder UH, Syntax and working memory in typically-developing children: Focus on syntactic complexity, Lang. Interact. Acquis 10 (2019) 141–176.

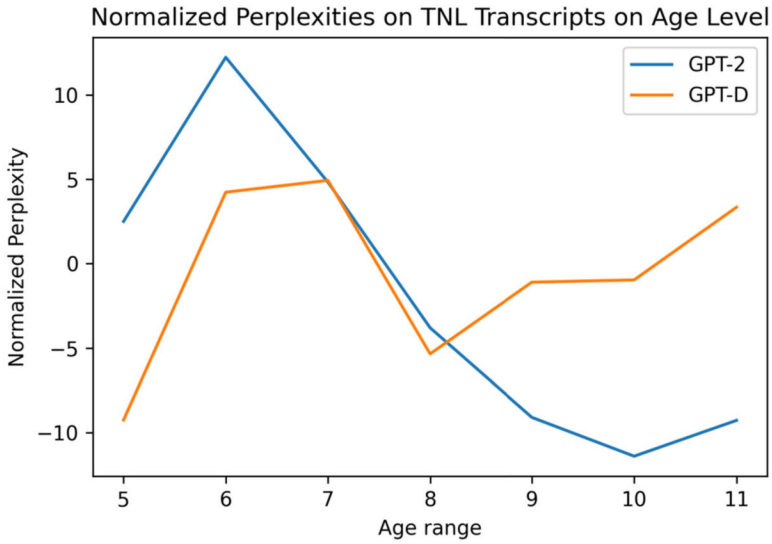

Fig. 1.
The allocation of normalized PPLs from TNL transcripts across ages.

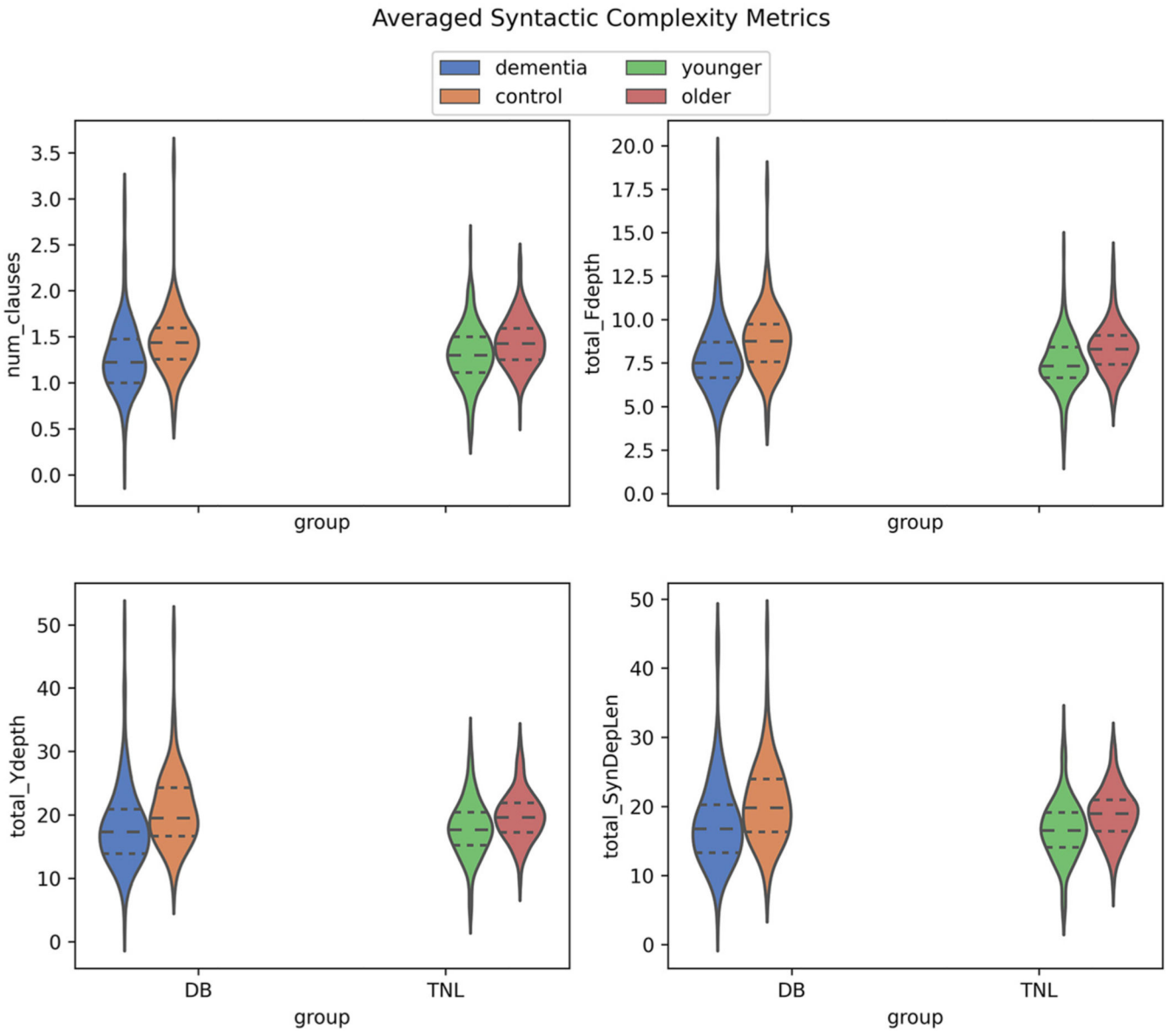

Fig. 2.

Within- and between-group means of syntactic complexity metrics, including number of clauses (top left), FD (top right), YD (bottom left), and SDL (bottom right) comparison between dementia patients and healthy elder in Pitt, younger and older children in TNL.

Table 1

Participant characteristics in Pitt and TNL datasets.

Dataset	Mean age (SD)	Number of transcripts	Transcript mean length (SD)
Pitt Control	64.22 (7.94)	242	126.51 (62.98)
Pitt dementia	71.55 (8.60)	235	116.40 (63.81)
TNL 5 y.o.	–	49	102.11 (59.47)
TNL 6 y.o.	–	100	102.18 (64.63)
TNL 7 y.o.	–	100	133.93 (94.69)
TNL 8 y.o.	–	101	173.72 (138.16)
TNL 9 y.o.	–	45	177.05 (93.69)
TNL 10 y.o.	–	50	224.88 (205.68)
TNL 11 y.o.	–	54	187.52 (75.57)

Table 2

The absolute mean differences in FD, YD and SDL measures and two-sided Welch's t-test p-values obtained from TNL and Pitt transcripts, where **bold** indicates differences with p-values above the alpha threshold of 0.01.

Subset	Number of clauses	FD	YD	SDL
	Absolute difference (p-value)			
Dementia vs. younger	0.03 (**0.34**)	0.31 (**0.09**)	0.07 (**0.90**)	0.82 (**0.14**)
Control vs. younger	0.14 (0)	1.31 (0)	2.84 (0)	3.57 (0)
Dementia vs. older	0.17 (0)	0.58 (0)	1.84 (0)	1.30 (0.01)
Control vs. older	(**0.97**)	0.42 (0.04)	0.93 (**0.14**)	1.45 (**0.02**)
Dementia vs. control	0.17 (0)	(0)	2.77 (0)	2.75 (0)
Younger vs. older	0.13 (0)	0.89 (0)	1.91 (0)	2.12 (0)

Table 3

The absolute mean differences in POS tag counts and two-sided Welch's t-test p-values obtained from TNL and Pitt transcripts, where **bold** indicates differences with p-values above the alpha threshold of 0.01. NOUN, ADJ, ADV, VERB, DET, IN, PRON, and PROPN represent noun, adjective, adverb, verb, determiner, preposition, pronoun, and proper noun, respectively.

Subset	NOUN	ADJ	ADV	VERB	DET	CONJ	IN	PRON	PROPN
	Mean difference (p-value)								
Dementia vs. younger	0.16 (0)	0.04 (**0.02**)	0.13 (0)	0.23 (0)	0.21 (0)	0.32 (0)	0.16 (0)	0.02 (0.01)	0.03 (**0.46**)
Control vs. younger	0.16 (0)	0.03 (**0.12**)	0.22 (0)	0.48 (0)	0.42 (0)	0.34 (0)	0.37 (0)	0.01 (**0.11**)	0.12 (0)
Dementia vs. older	0.1 (**0.02**)	0.07 (0)	0.3 (0)	(**0.91**)	0.33 (0)	0.31 (0)	0.09 (0)	0.04 (0)	0.22 (0)
Control vs. older	0.51 (0)	(**0.78**)	0.39 (0)	0.25 (0)	0.53 (0)	0.33 (0)	0.29 (0)	0.03 (0)	0.31 (0)
Dementia vs. control	0.41 (0)	0.07 (0)	0.09 (0)	0.25 (0)	0.21 (0)	0.01 (**0.66**)	0.2 (0)	0.01 (0)	0.09 (0.01)
Younger vs. older	0.05 (**0.13**)	0.03 (**0.11**)	0.17 (0)	0.23 (0)	0.11 (0)	0.01 (**0.68**)	0.07 (0)	0.02 (0)	0.20 (0)