



Lexical diversity and pausing in very old age: insights from semi-spontaneous narratives

Sonja Alantie, Tanja Makkonen & Kati Renvall

To cite this article: Sonja Alantie, Tanja Makkonen & Kati Renvall (17 Feb 2026): Lexical diversity and pausing in very old age: insights from semi-spontaneous narratives, *Aphasiology*, DOI: [10.1080/02687038.2026.2629419](https://doi.org/10.1080/02687038.2026.2629419)

To link to this article: <https://doi.org/10.1080/02687038.2026.2629419>



© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 17 Feb 2026.



Submit your article to this journal [↗](#)



Article views: 135



View related articles [↗](#)



View Crossmark data [↗](#)

Lexical diversity and pausing in very old age: insights from semi-spontaneous narratives

Sonja Alantie ^{a,b}, Tanja Makkonen ^{b,c,d} and Kati Renvall ^a

^aDepartment of Psychology and Speech–Language Pathology, University of Turku, Turku, Finland; ^bSpeech–Language Pathology, Tampere University Hospital, Tampere, Finland; ^cOpen University and Continuous Learning, Speech–Language Pathology, University of Turku, Turku, Finland; ^dFaculty of Medicine and Health Technology, Tampere University, Tampere, Finland

ABSTRACT

Background: Aging alters various aspects of language output. However, health care professionals lack knowledge about language skills in the vastly growing very old (VO) population. Narratives provide a lucrative method for examining language in ageing as they are sensitive to neurological changes.

Aims: We study the normative lexical diversity and pauses in healthy 80–100-year-old speakers' narratives and investigate whether they are predicted by physio-anatomical, socio-cognitive or linguistic background variables.

Methods: This cross-sectional study comprises fifty VO Finnish speakers, who produced a narrative from a wordless cartoon strip. The lexical diversity of the narratives was analysed by number of different words (NDW), type-token-ratio (TTR), words by grammatical class and noun-verb ratio. Pausal variables consisted of pause time ratio, pause frequency, mean pause duration and mean syllables between pauses. Background variables included age, dentition (natural teeth vs. dentures), educational level and linguistic skills by the Western Aphasia Battery Aphasia quotient (WAB-AQ), semantic and phonemic fluency and the Boston Naming Test (BNT). The association between lexical diversity, pausing and the background variables was studied by multiple regression analysis.

Results: Both lexical diversity and pausing showed notable individual variation. Higher NDW was linked to natural teeth among lexical diversity variables. NDW was also positively predicted by the highest educational level and WAB-AQ but negatively associated with BNT scores. Within the pause variables, pause time ratio was predicted by age as the oldest speakers produced the lowest proportions of pauses.

Conclusions: This study offers normative data on multiple variables of the VO speakers' language use. It also provides insight into the interconnections between ageing, language and personal factors. As natural teeth and high education were associated with diverse vocabulary, we recommend that clinicians assess the participants' dental status and educational history. The contrasting effects of the

ARTICLE HISTORY

Received 26 April 2025

Accepted 4 February 2026

KEYWORDS

Lexical diversity; pause; language; narrative; ageing

CONTACT Sonja Alantie  sonja.e.alantie@utu.fi; sonja.alantie@pirha.fi  Speech–Language Pathology, Tampere University Hospital, P.O. Box 2000, Neurology Ward, Tampere 33521, Finland

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

WAB-AQ and BNT on lexical diversity underscore the need to acknowledge that lexical abilities may appear differently in formal tests and in narration. The finding that the oldest of the speakers used less time pausing challenges the assumption about pausing consistently increasing with age. This study adds to the missing knowledge on ageing and language. It provides health care professionals more expertise in identifying typical age-related variety. Nevertheless, to enable narrative evaluation in routine clinical practice more international research on typical old and adult speakers is required as well as developments in automated analysis of spoken language.

Introduction

Population worldwide is vastly ageing (Statistics Finland, 2024; United Nations, 2025), which makes an increasingly large portion of people prone to acute and degenerative neurological diseases (Niccoli & Partridge, 2012). In this study, we adopt the terminology used by Forman et al. (1992) referring to individuals 80 years and older as very old (henceforth VO) and use the term “old” as a broad classification encompassing individuals from 60 years onwards. Within the next 30 years, even the number of VO individuals is expected to more than triple and outnumber infants. This will eventually also affect the age distribution of speech-language pathologists’ clientele. Ageing is shown to affect cognitive functions, sentence comprehension and various aspects of language production including speech tempo, fluency, word retrieval and sentence formation (Baciu et al., 2021; Cahana-Amityay & Albert, 2014). Nevertheless, there is little language and culture specific knowledge on how the VO age affects language output, which renders differentiation between typical and neuropathological linguistic changes challenging. It also places excessive responsibility on intuitive decision-making in speech therapy assessments.

Factors behind language variation in ageing

Old individuals are not a heterogeneous group defined by solely by their age. There is a complex interplay of demographic factors influencing language skills in aging populations (Alantie et al., 2022; Baciu et al., 2021). Physio-anatomical, social, cultural, environmental and cognitive factors contribute to successful skill maintenance (Baciu et al., 2021; Rowe & Kahn, 1987; Uotinen et al., 2003). Advantageous genes, good overall health, high educational level, active lifestyle and functional cognitive reserve are generally connected to intact language performance in ageing (Baciu et al., 2021). When it comes to age related physio-anatomical changes, the loss of natural teeth and hearing are common. A growing body of research connects good dental health (Alantie et al., 2022; Li et al., 2023) and functional hearing (Huang & Lin, 2024) with the resilience to cognitive-linguistic decline.

While the link between dental health and cognition is not a new concept, it is still poorly recognised in the discipline of speech-language pathology. In short, tooth loss and dental diseases are associated with increased risk of executive-cognitive decline, likely due to their interrelation with malnutrition and generalised and neural inflammation (Dos Santos Nogueira et al., 2023; Li et al., 2023). As age-related linguistic changes result from

cognitive alterations, there is also a link between dental health and language skills. For example, in a previous study on very old Finnish speakers, poor dentition was found to predict a decline in formal language test performance (Alantie et al., 2022). This connection was seen in more demanding language tasks that require executive functions. Furthermore, it is noteworthy that the relationship between poor dentition and weaker language skills could not be explained by motor-speech deficiencies in the study.

Narrative tasks as a tool for language assessment

One of the most common ways to assess the spectrum of spoken language domains is the use of narrative tasks. With narrative tasks we refer to spoken or written accounts of connected events, which are produced with minimal external prompts (spontaneous narrative) or prompted for example by memory (narrative retell) or a picture sequence (semi-spontaneous narrative). Narrative tasks help to induce connected speech (continuous stretched of utterances) yielding rich samples of naturalistic language production (Angelopoulou, Kasselimis, Varkanitsa, et al., 2024). Most narratives share a similar structure called a story grammar including at least orientation (introduction of the setting, characters, and context), complication (the central problem, event, or conflict that disrupts the normal situation), resolution (how the complication is solved) and coda (a closing statement that may provide a moral) (Kemper et al., 1990).

With its multifaceted macrostructural (story grammar and content) and microstructural (productivity, lexicon, grammar, cohesive devices and fluency) elements, narrative speech is considered one of the most complex discourses of language (Kemper et al., 1990). Narrative skills develop, refine and remain until very late in life (Ulatowska et al., 1998). Narratives are always influenced by their cultural and personal context (Bruner, 2010). When people narrate, they inevitably shape their experiences into verbal events based on their perspective, language structure, and their own knowledge of language. Therefore, narratives connect to the speaker's social history as well as current language skills. Narrative abilities can even serve as the earliest indicators of neurological disorders (Drummond et al., 2015; Harris et al., 2008; Roark et al., 2011). Narrative tasks are, thus, sensitive to a person's age, neurological health and socio-cultural background making them ideal for studying the language output of VO speakers.

In both clinical language assessments and research, narrative tasks and other forms of connected speech, such as picture description, conversation, or monologue, can effectively complement comprehensive test batteries (Roark et al., 2011). However, not all languages have established practices for studying adult language in clinical settings, automated analysis methods for narrative speech or norms for interpreting the performance (e.g., Aldarmaki et al., 2022; Norvik et al., 2022; Pernambuco et al., 2017). This is also the case in Finnish language of which this study concerns. Nevertheless, speech-language pathologists in Finland and elsewhere often conduct narrative tasks to patients of different ages to gain at least restricted amount of qualitative information. Connected speech in Finland is commonly evaluated by the classical picture description tasks "Picnic scene" in the Finnish version of the Western Aphasia Battery (WAB) (Kertesz et al., 2005) and "Cookie theft" in the Finnish version of the Boston Diagnostic Aphasia Examination (Laine, Niemi, et al., 1997). Finnish speech-language pathologists also often utilize a wordless cartoon strip, the so-called Scarecrow cartoon ("Fugleskraemsel går amok"

by Henning Dahl Mikkelsen, n.d., see Korpijaakko-Huuhka, 2003) to elicit semi-spontaneous narratives, although it has not been validated on healthy adult Finnish speakers and there are no norms for its interpretation. The scarecrow cartoon was first introduced as an elicitation method in a thesis concerning communication skills in people with aphasia (Korpijaakko-Huuhka, 1995, 2003) and it has since been employed in the Nordic countries to study motor-speech and language skills in people with and without neurological conditions (see e.g., Høeg et al., 2023; Makkonen et al., 2018; Penttilä et al., 2018).

In this study, we investigate how VO speakers produce two microstructural elements of narrative speech. We investigate lexical diversity and pausing in the semi-spontaneous narratives of 80–100-year-old neurologically healthy speakers. The first objective is to offer missing normative data on the typical language use of VO speakers by using simple variables serving as proxies for the more complex components of language. The variables (number of different words [NDW], type-token-ratio [TTR], words by grammatical class, noun-verb ratio, pause time ratio, pause frequency, mean pause duration and mean syllables between pauses) have been chosen based on their sensitivity to neural change as well as their technical feasibility in clinical work. The other objective is to investigate which physio-anatomical (here age and dentition), socio-cognitive (here educational level) and linguistic factors (language test performance) might predict lexical diversity and pausing conduct. This study is a part of project that examines motor-speech and language skills and their contributing background factors in very old Finnish speakers. We will next provide a general overview of what is already known about lexical diversity and pausing in the connected speech of aged speakers.

Lexical diversity in typical and pathological ageing

Lexical diversity refers to the range of distinct words used in a language output sample (Kintz et al., 2016). It entails both the knowledge of vocabulary (word knowledge and vocabulary size) and the capacity to demonstrate it (lexical access and retrieval) (Chapelle, 1994; Fergadiotis & Wright, 2011). It integrates various aspects of cognitive ability and language processing strategies, as individual words are selected and retrieved to construct sentences and larger discourse structures (Chapelle, 1994; Fergadiotis & Wright, 2011). From an ageing perspective, the manifestation of lexical diversity is a relevant topic. According to the international body of research, word knowledge improves even until the seventh decade of life remaining stable or declining only slightly during very old age (Kavé, 2024; Kavé et al., 2010; Lövdén et al., 2004; Salthouse, 2019). The robustness of word knowledge is in stark contrast with active vocabulary output, as the most common cognitive change in ageing is the decline in word retrieval (Kavé & Yafé, 2014; Kavé et al., 2010). Therefore, lexical diversity is the outcome of two realms of language that respond differently to typical ageing.

In addition to a speaker's age, lexical diversity is influenced by the speaker's neurological health and discourse type (Fergadiotis & Wright, 2011; Fergadiotis, Wright, & Capilouto, 2011; Stark, 2019; Zozuk, 2023). For example, in picture description tasks, increasing age in healthy speakers has been connected to lower lexical diversity and tendency to exhibit familiar, high frequency words (Cho et al., 2021; Fergadiotis, Wright, & Capilouto, 2011). Contrastively, in spontaneous monologues and conversational speech,

older speakers show higher lexical diversity compared to younger speakers (Horton et al., 2010; Kemper & Sumner, 2001; Moscoso Del Prado Martín, 2017). Unfortunately, we were unable to find studies using picture sequences or cartoons to investigate the overall lexical diversity in typical younger and older adult speakers. In adults and old individuals with neurological conditions such as aphasia, mild cognitive impairment, Alzheimer's disease, or frontotemporal degeneration, overall lexical diversity is typically lower than that of healthy speakers in connected speech, including spontaneous narratives, picture sequence narratives and picture descriptions (Coppieters et al., 2024; Fergadiotis & Wright, 2011; Williams et al., 2023; Zozuk, 2023).

Lexical diversity by grammatical word class

Lexical diversity can also be examined based on the distribution of produced words across grammatical classes. Different grammatical distributional patterns can occur by age and neurological condition (Armstrong, 2000; Cho et al., 2021; Coppieters et al., 2024; Høeg et al., 2023; Kavé et al., 2009; Williams et al., 2023). Regarding typical ageing, Cho et al. (2021) examined picture descriptions produced by 52–89-year-old ($n = 37$) and 18–22-year-old ($n = 76$) speakers. The older speakers produced significantly more verbs, pronouns and fillers and fewer nouns, conjunctions, prepositions and determiners than the young participants. The nouns and verbs of the older speakers were also more familiar in quality. Høeg et al. (2023) also reported a statistically significant difference in noun usage between two healthy old age groups. In their study on linguistic variables in picture description and cartoon narration, they found that neurologically healthy speakers aged 76 years and above produced a lower frequency of nouns as well as fewer noun types compared to 59–75-year-old speakers.

When it comes to pathological ageing, for example speakers with different types of aphasia (Armstrong, 2000), Alzheimer's disease (Williams et al., 2023) and frontotemporal degeneration (Coppieters et al., 2024) demonstrate each their own conducts of word class usage. For example, compared to healthy peers, speakers with non-fluent agrammatic aphasia are found to use more nouns than verbs while speakers with fluent aphasia use more verbs than nouns, but the frequency and diversity of the verbs may still be reduced compared to healthy peers (Armstrong, 2000). People with Alzheimer's disease utilise fewer and less diverse nouns and rely more on pronouns (Williams et al., 2023). Additionally, with increasing age, people with Alzheimer's disease produce nouns of lower frequency whereas healthy speakers have been shown to use nouns of higher frequency. In many languages, individuals with primary progressive aphasia have been found to differ from healthy peers in their noun-verb ratio, that is the proportion of nouns relative to verbs they produce in connected speech (Coppieters et al., 2024). These findings on typical and atypical ageing highlight the clinical potential in accounting for differences in word class distribution alongside with overall lexical diversity.

Although lexical diversity has the potential to indicate age related change or impairment, the clinical application of the knowledge is difficult. There is often little to no research-based information available that is cultural-linguistically appropriate, and adapting research findings from other languages requires careful consideration and strong professional expertise. Currently, only few studies in the international literature include substantial samples of very old speakers. Often the number of VO speakers within a larger

group is also unspecified. Additionally, the utilised discourse types and the measures of lexical diversity are yet sporadic.

Pausing

There is increasing evidence that temporal aspects of speech, especially the frequency and duration of silent pauses can serve as means for investigating the cognitive processes involved in speech production (Angelopoulou, Kasselimis, Varkanitsa, et al., 2024). Such processes include articulation, word access and retrieval, self-monitoring, functions of episodic or semantic memory and the planning and organization of speech output (Angelopoulou, Kasselimis, Varkanitsa, et al., 2024; Bóna, 2014). Like lexical diversity, pausing behaviour is responsive to age, neurological health and discourse genre (Angelopoulou, Kasselimis, Goutsos, et al., 2024; Angelopoulou, Kasselimis, Varkanitsa, et al., 2024; Bóna, 2014; Keszler & Bóna, 2019; Lofgren & Hinzen, 2022; Mack et al., 2015; Martins & Andrade, 2011).

Increasing age is generally suggested to result in more frequent (Bóna, 2014; Martins & Andrade, 2011) and longer pausing (Cho et al., 2021; Keszler & Bóna, 2019) in connected speech. Pausing frequency in VO speakers has been reported for example in the cross-sectional studies by Bóna (2014) and Martins and Andrade (2011), in which 66–99-year-old and over 80-year-old speakers produced pauses more frequently than their junior speakers. However, the age-related changes in pause frequency include high individual variability as shown in the longitudinal study by Keszler and Bóna (2019). They examined speakers ($N=7$) three times from middle age to old or VO age discovering that the increase in pause frequency was not linear over time in all speakers. More coherent changes were measured in the mean duration of pauses over time, although even that did not always develop linearly. An important finding was that the pausal variables did not increase for all speakers. In fact, some VO speakers retained speech like that of young adults.

Research aims and questions

In this study we produce normative data on VO individuals by providing multiple metrics of lexical diversity and pauses in narrative speech. We also examine demographic and linguistic background variables as potential predictors of lexical diversity and pausing, as various factors other than the age have been shown to influence output performance in the elderly (Alantie et al., 2022). This is one of the first studies in the field of speech-language pathology investigating directly the connection between dentition and language skills. The results of this study will contribute to forming age-accurate and up-to-date performance profiles of VO speakers. The knowledge will help to distinguish features of typical ageing. Studying the tendencies of VO speakers will also bring insight into the realm of life-long trajectories of language development and change.

This study addresses the following questions

- (1) What is the lexical diversity (number of different words [NDW], type-token-ratio [TTR], words by grammatical class and noun-verb ratio) in semi-spontaneous narratives produced by very old speakers?
- (2) What is the proportion and frequency of pauses (pause time ratio, pause frequency, mean pause duration and mean syllables between pauses) in very old speakers' semi-spontaneous narratives?
- (3) Do the very old speakers' physio-anatomical (age and dentition), socio-cognitive (educational level) and linguistic (language test performance) background factors predict lexical diversity or pausing behavior in the narratives?

Methods

Participants

The inclusion criteria for this cross-sectional study were (a) age of at least 80 years (b) native speaker of Finnish, (c) no diagnosed or suspected dementing illness, (d) no diagnosed speech or language disabilities, and (e) a Mini-Mental State Examination test (MMSE) score of 24–30 (Folstein et al., 1975; Hänninen et al., 2010). Participants were recruited through open advertisements and visits to senior groups and homes and by spreading the word in the communities. The study comprises altogether fifty community dwelling participants (age range 80–100 years, females $n = 41$, males $n = 9$). The participants required non or moderate domestic help. Most of them lived alone ($n = 35$) and the rest with a significant other ($n = 15$). All of them reported interacting with other people outside their home daily ($n = 33$) or at least weekly ($n = 17$). Ethical approval for this study was obtained from the Regional Ethics Committee of the Tampere Human Sciences. All the participants provided their written informed consent of participation.

Demographic factors

In this study we used three demographic predictor variables, namely age, educational level and dental status to represent physio-anatomical and socio-cognitive status (Table 1). The information was obtained by the participants' self-reports. Education here served as a proxy for a combination of socio-cognitive factors encompassing cognitive skills, intelligence, environmental circumstances and personal characteristics (see Alantie et al., 2022). Education was distributed into three levels. Over half of the participants ($n = 28$) wore complete or partial dentures,

Table 1. Demographic variables of the very old speakers ($N = 50$).

Age (years)	MMSE	Educational level	Dentition
Range 80–100	Range 24–30	Primary $n = 18$	Natural $n = 22$
$M = 84.8$	$M = 27.6$	Secondary $n = 20$	Dentures $n = 28$
$SD = 4.1$	$SD = 1.8$	Tertiary $n = 12$	

Note. M = mean value; SD = standard deviation; MMSE = Mini-Mental State Examination; primary = primary education or lower secondary education; secondary = upper-secondary education or post-secondary non-tertiary education; tertiary = Bachelor's or equivalent level or Master's or equivalent level education.

while the rest had adequate natural dentition. Additionally, poor hearing was reported by 24 out of 50 participants, most of whom used a hearing aid ($n = 14$). Hearing was, however, left out of the predictor variables, because it was not statistically significant in the preliminary analyses. Additionally, since hearing was screened only through self-evaluation, it was not considered a reliable measure of auditory competence.

Language performance in clinical tests

We used four common linguistic tests as predictor variables (for the test scores see Table 2). We conducted the Finnish version of the Western Aphasia Battery (WAB) (Kertesz et al., 2005), from which we utilized the Aphasia Quotient (AQ, max 100) as the summary value of general competence in oral expression and auditory comprehension. To assess the flexibility of time-pressured lexical access and cognitive executive functions, we conducted a semantic animal category task and a phonemic word-fluency task with initial letter/k/ (see Alantie et al., 2022). We incorporated the semantic and phonemic tasks based on previous research findings, which suggest that both semantic and phonological language mechanisms and their interaction can account for age-related changes in word production (Baciu et al., 2021). For vocabulary knowledge, access and retrieval in pictorial confrontation naming we used the Finnish version of the Boston Naming Test (BNT) (Laine, Koivuselkä-Sallinen, et al., 1997).

Lexical diversity

To assess lexical diversity, we conducted a narration task with the wordless Scarecrow cartoon strip (see Korpijaakko-Huuhka, 1995, 2003). In the cartoon the main character is sowing seeds in the ground and dreams of an abundant yield when suddenly birds come to eat the seeds. The annoyed main character then fetches building material from a shed and erects a scarecrow on the field. The birds are, however, unbothered by the scarecrow so the main character solves the problem by pulling the scarecrow up from the ground and whisks the birds away with it. This cartoon strip enables the production of all the classical elements of story grammar – orientation, complication, resolution and coda (Korpijaakko-Huuhka, 1995). The participants were asked to tell the story depicted in the nine-part picture sequence in their own words in such precision that a listener who was not looking at the pictures could follow the story. The participants were given any preparation time they needed. We chose to use a semi-spontaneous instead of a fully spontaneous narrative as it poses equal cognitive-linguistic requirements and

Table 2. Comprehensive language performance, verbal fluency and confrontation naming in the very old speakers ($N = 50$).

Variable	WAB-AQ	Semantic fluency	Phonemic fluency	BNT
Range	84.9–99.7	8–38	2–26	27–59
<i>M</i> / <i>SD</i>	95.7/3.4	21.5/2.8	13.8/5.6	48.6/6.8
Mode	99	24	10	51

Note. *M* = mean value, *SD* = standard deviation, WAB-AQ = Western Aphasia Battery Aphasia Quotient, BNT = Boston Naming Test.

opportunities for all the speakers. The same cartoon strip is in active clinical use by Finnish clinical speech-language pathologists, and it has been used at least in Finnish and Norwegian studies to examine motor-speech and language skills in populations with and without neurological conditions (see e.g., Høeg et al., 2023; Makkonen et al., 2018; Penttilä et al., 2018).

Finnish is typologically an agglutinative language and rich in suffixal affixation. In addition, Finnish has many dialects, and the spoken dialectal varieties differ greatly from the standard written Finnish. The dialects used by the participants of this study originate mainly from two regions Pirkanmaa and Satakunta areas. Finnish lacks computer programs for conducting comprehensive automated linguistic analysis of spoken language samples. Therefore, we analysed lexical diversity by combining manual work with some applied automated language tools as described below.

We audio-recorded the narratives with a Zoom-recorder and a head-set microphone and annotated them orthographically with PRAAT (Boersma & Weenink, 2023). From the annotation we discarded the parts which were not related to the story macrostructure that is the parts that did not contribute to the story grammar (orientation, complication, resolution or coda). We made this choice to focus specifically on the vocabulary elicited by the topics of this comic strip and its visual prompts and to avoid attributing greater lexical diversity to speakers who digressed, compared to those who stayed on the topic or the instruction of the task which was to tell the story presented by the pictures. Based on our clinical experience this facilitates targeting lexical analysis also in clinical practice.

All removed occurrences were jointly reviewed by two researchers. In cases of ambiguity, discrepancies were discussed and resolved through consensus. The removed cases included metadiscourse such as speech directed to the researcher (e.g., *täytyyks tästä vielä muuta sanoo* “do I have to say something else about this?”, *kuule, tää menee yli mun hilseen* “listen, this goes over my head”), speech concerning their own performance (e.g., *ai eiks aika oikein tullu* “oh, didn’t it come quite correct”) and verbose speech (e.g., *minäkin vein vanhan takkini tonne seisomaan vattupuskan telineen päälle että aattelin että se pelkää että mä seison siellä* “I also took my jacket there to hang on the rack that supports the raspberry bushes and thought that it would fear it was me standing there”). These types of utterances were excluded from the narrations of 38 (out of 50) speakers, and the mean proportion of excluded utterances per speaker was 15.5%.

We also omitted incomplete words like false starts, and immediate repetitions of words so that the repeated word was counted as one item (e.g., *ja ja ja ja* “and and and and”, *varmaan varmaan* “probably probably”). We analysed words as single items although some verb tenses are formed with two parts such as the present perfect (e.g., *on mennyt* “have gone”). Therefore, it should be noted that rich use of these tenses would increase the total number of verbs but not the number of different verb lemmas. Most of the stories were, however, told in the present tense with none or some occurrences of the imperfect or present perfect tense.

To count distinct words, we utilised the corpus tool AntConc (Anthony, 2024) to extract all different word forms and their frequencies in a list form. Every word item on the list was manually marked for their basic lemma form. All words were also labelled for their grammatical class either as noun, verb, adjective, adverb, pronoun, conjunction, interjection or numeral. To ensure the correctness of the lemma form and grammatical class, we used free Finnish online grammatical parsers Sparv 2 (Språkbanken Text, n.d.), Fintag

(Kielipankki, n.d.) and Fintwol (Lingsoft, n.d.), which sometimes offered different results. In ambiguous cases, we relied on the word's context within the speaker's narrative and selected the label based on our sense of language. We also had to verify the meanings of some more unfamiliar dialectal lexical variations using diverse online sources. After this procedure of lemmatising and categorizing all the words in the narratives we had a database of all the different words by grammatical class and the number of their occurrences by speaker and whole group.

Pauses

To study pausing behaviour we detected all pauses in the narratives in PRAAT (Boersma & Weenink, 2023) by delimiting the boundaries of pauses in a pause tier under the spectrogram to mark speaking and pausing time in milliseconds (ms). Although content that was not related to the story was extracted from the analysis of lexical diversity, pauses were analysed from the complete narration task without discarding pauses at the boundaries or within the excluded lexical segments. To automatically extract information about pause duration, frequency and distribution we used a corpus tool designed and coded for the purpose of this study. Pauses in this study comprise silence, hesitation sounds (/‘‘/er’/mm’/um’) and non-verbal sounds like coughing, laughter, or audible inhalation – that is to say other than articulated speech that carries semantic content. Filler words (niinku “like”, tuota noin “so”) or interjection (juu “okay”) were not included in pauses. Pauses of at least 200 milliseconds were included in the analysis (Niimi & Nishio, 2001). Adopting a predetermined time threshold for pauses has been widely applied in the field (Angelopoulou, Kasselimis, Goutsos, et al., 2024). Although this approach has encountered some justifiable concerns, as a universal cultural-linguistic and context-free threshold does not exist, we chose to use a threshold. We base the decision to use a threshold on the indication that pauses exceeding 200 ms may serve cognitive-communicative functions rather than merely reflecting physical articulation movements (see Angelopoulou, Kasselimis, Goutsos, et al., 2024).

Lexical and pausal response variables

We extracted metrics on lexical diversity and pausing behaviour to test whether they were predicted by the speakers' background variables. Although numerous measures of lexical diversity and pausing exist, we chose variables that are sufficiently practical to be calculated even in clinical settings. The response variables we selected have been found sensitive to linguistic-cognitive changes in previous studies as explained below. For the variables and their definition see Table 3.

The number of different words (NDW) and type-token ratio (TTR) both measure lexical diversity and tend to vary according to sample length (Fergadiotis & Wright, 2011). When a narrative sample lengthens NDW is likely to increase as the probability to introduce new words increases. With longer samples, TTR, on the contrary, is likely to decrease as the probability to repeat words increases. Both NDW and TTR have been criticised by arguing that the comparison between different speakers is confounded by the sample length (Fergadiotis & Wright, 2011; Jarvis, 2013). However, we find an interest in the sample length itself given the sample is generated by the same incentive across all participants, in

Table 3. Lexical diversity and pause metrics with definition.

Variable	Definition
Number of different words (NDW)	Number of words with distinct lemmas or types
Type-token ratio (TTR)	Number of different word lemmas (types) divided by the number of all words (tokens)
Noun-verb ratio	The number of all nouns divided by the number of all verbs
Pause time ratio	Total pausal duration in seconds divided by the total narration time
Pause frequency	Total number of pauses divided by the total number of words
Mean pause duration	Total pausal duration in seconds divided by the number of pauses
Mean syllables between pauses	The mean number of syllables occurring between two successive pauses

our case the same cartoon strip. Additionally, we find that truncating the storyline for even size samples for each participant would disregard the narration's integrity (Fergadiotis & Wright, 2011). Some of the benefits of NDW and TTR are that they are useful for short spoken samples as they are straightforward to compute (no need for parameter choices like window size) and give immediate sense of how varied the vocabulary is without averaging. NDW and TTR are both very traditional variables enabling comparison to studies conducted over time. They have both proved useful in detecting change in language abilities and in distinguishing speaker groups also across different languages (e.g., Coppieters et al., 2024; Fergadiotis & Wright, 2011; Fergadiotis, Wright, & Capilouto, 2011; Harris Wright et al., 2003; Jarvis, 2013; Zozuk, 2023).

Noun-verb ratio (NVR) was included in our study merely to depict the proportional semantic relationship of nouns and verbs as this proportion may vary because of neurological changes (Coppieters et al., 2024). Therefore, VO speakers may exhibit characteristic noun-verb ratios. In international studies, noun-verb ratio has also been used as a proxy for syntactic complexity (e.g., Stark, 2019), but there are no Finnish studies confirming its usefulness in this sense.

The distribution and proportion of pauses in connected speech reflect underlying complex language functions like word access and retrieval, self-monitoring, memory, speech planning and organization (Angelopoulou, Kasselimis, Varkanitsa, et al., 2024; Bóna, 2014; Mack et al., 2015). To view pausal behaviour comprehensively with measurable metrics, we selected three variables used in previous studies: pause frequency (Bóna, 2014; Martins & Andrade, 2011), pause ratio (Keszler & Bóna, 2019; Mack et al., 2015) and mean pause duration (Keszler & Bóna, 2019). We also introduced an experimental variable, the mean number of syllables between pauses, which captures the continuous flow of speech output between cognitive disruptions but can be easily measured without observing semantic content or syntax (in contrast to for example analysing the mean length of utterance).

Procedure

Participants were assessed by the first author either in their home or at the university according to their preference. The administration of the study protocol was constant across participants, and it was executed within a single session taking altogether 1–2 hours. The protocol entailed (1) an interview of the educational, social and health status (including dentition), (2) the BNT (Laine, Koivuselkä-Sallinen, et al., 1997), (3) a semi-spontaneous narrative elicited by the Scarecrow cartoon strip, (4) oral diadochokinetic

rate tasks, (5) WAB (including semantic fluency tasks) and added phonemic fluency task (with initial k-letter) (6) MMSE and (7) spontaneous personal narrative. Tasks (4) and (7) are not included in this study.

Statistical analyses

To investigate the relationship between background predictors and lexical diversity and pausing, generalized regression models were applied based on the nature of each outcome variable. The predictor variables included age, dentition (dentures vs. natural teeth), educational level (primary, secondary, tertiary), WAB-AQ, semantic fluency, phonemic fluency, and BNT. Lexical diversity measures included number of different words (NDW), type-token ratio (TTR), and noun-verb ratio. Pausal measures included pause time ratio, pause frequency, mean pause duration, and mean syllables between pauses. Variable means, standard deviations, and correlations with confidence intervals are visible in Appendix (Table A1)

For NDW, a negative binomial regression model was used due to overdispersion in count data. TTR was analyzed using beta regression with a logit link function since it is a proportion bound between 0 and 1. Noun-verb ratio, a skewed continuous variable, was examined using gamma regression with a log link function.

For pausal variables, pause time ratio and pause frequency were modelled using beta regression, while mean pause duration was analysed using Poisson regression, and mean syllables between pauses were modeled using gamma regression with a log link function. Effect sizes were expressed as exponentiated coefficients: Incidence Rate Ratios (*IRR*) for Poisson and negative binomial models, and Odds Ratios (*OR*) or *Exp(B)* for beta and gamma regressions, depending on the model link function. Because continuous predictors were standardized, these effect sizes represent the multiplicative change in the outcome associated with a one standard deviation (*SD*) increase in the predictor. All effects are reported with 95% confidence intervals and p-values.

Results

Lexical diversity in the narratives

All the participants were able to generate a narrative by a wordless picture sequence. The overall lexical diversity of the narratives is expressed using different metrics in Table 4. The speakers demonstrated substantial within-group variability in their total word production, as well as in NDW, TTR, and NVR. According to the group's mean TTR, VO speaker utilised 70% of the words once while 30% of the words were recurrences.

Table 4. All words, number of different words (NDW) and type-token ratio (TTR) in the very old speakers' narratives.

Variable	Words (N)	NDW (n)	TTR (%)
Range	31–274	22–157	54.4–86.7
Mean/ <i>SD</i>	94/41.8	64.3/24.1	70/7.4
Median/Mode	86.5/105	60/51	69.9/66.6

Note. *SD* = standard deviation.

Table 5. Grammatical word class distribution and noun-verb ratio (NVR) in the very old speakers' narratives.

Words (N)	Noun	Verb	NVR (%)	Adjective	Adverb	Pronoun	Conjunction	Interjection	Numeral
Range	4–39	7–64	26.3–155.6	0–12	1–72	0–59	4–34	0–10	0–3
Mean/ <i>SD</i>	15.7/ 7.7	22.5/ 10.2	74.6/28	3.2/2.8	17/10.2	21.5/ 11.2	13.1/6	1.4/1.9	0.2/0.5
Median/ Mode	13.5/10	19/19	69.4/100	3/1	15.5/15	20/15	12/10	1/1	0/0
Lemmas (n)									
Range	4–32	4–44		0–12	1–35	0–30	2–7	0–5	0–3
Mean/ <i>SD</i>	13.8/ 6.5	17.9/7		3/2.6	11.4/ 5.7	12.7/5.5	4.2/1.7	1.1/1.2	0.2/0.5
Median/ Mode	12/9	17/12		2.5/1	10/8	12/7	4/3	1/1	0/0

Note. *SD* = standard deviation.

The distribution of all word tokens by grammatical class shows the predominance of verbs (23%) and pronouns (23%), followed by adverbs (18%), nouns (17%) and conjunction (14%). Adjectives (3%) and interjections were used infrequently (2%) and the use of numerals was rare (0%). Grammatical class distribution is visible in Table 5. Some speakers did not produce adjectives, pronouns, interjections or numerals at all, while a different degree of nouns, verbs, adverbs and conjunctions were always present in the narratives. On average, VO speakers produced more verbs than nouns (noun-verb ratio $M = 74.6$), but there were also speakers who displayed noun prominence or even production of nouns and verbs.

Pausing patterns in the narratives

The VO speakers' pausing behaviour is characterised by vast within-group variation in both the temporal and distributional aspects of pausing including the total number of pauses, pause time relative to narration time (pause time ratio), the number of pauses relative to number of words (pause frequency), the average length of pauses (mean pause duration) and how often pauses occurred in the speech flow (mean syllables between pauses) (Table 6). On average, more than a third of the narration time was spent pausing. The general mean duration of a pause was 1 second and pauses occurred approximately every 10 syllables. The number and frequency of pauses varied according to narration length, but none of the speakers told the story without producing 200 ms pauses or longer.

Table 6. Pause distribution and time in the very old speakers' narratives.

Variable	Pauses (N)	Pause time ratio (%)	Pause frequency (%)	Mean pause duration (s)	Mean syllables between pauses (n)
Range	8–74	15.9–57.3	11.4–41.9	0.4–2.9	3.9–17
Mean/ <i>SD</i>	22.4/14	33.6/10.7	20.3/6	1.1/0.5	9.6/2.8
Median/Mode	18/10	29.5/15.9	18.7/28.6	1/0.4	9.4/3.9

Note. *SD* = standard deviation.

Background variable predictability in relation to lexical diversity and pausing

Multiple regression models were performed to examine the contribution of age, dentition, educational level and language skills (WAB-AQ, semantic and phonemic fluency, BNT) on lexical diversity and pausing behaviour in the semi-spontaneous narratives of the VO speakers. We found statistically significant predictors for NDW and pause time ratio, but none of the background variables were associated with TTR, noun-verb ratio, pause frequency, mean pause duration and mean syllables between pauses. For the regression models, see [Tables 7 and 8](#).

Four predictors showed significant associations with NDW. Individuals with dentures produced significantly fewer different words compared to those with natural teeth (IRR = 0.80, 95% CI: 0.73–0.88, $p < .001$). Higher WAB-AQ scores were associated with increased lexical diversity (IRR = 1.20, 95% CI: 1.13–1.28, $p < .001$), while higher BNT scores were associated with reduced lexical diversity (IRR = 0.82, 95% CI: 0.76–0.88, $p < .001$). In addition, individuals with the highest (tertiary) education produced significantly more different words than those with the lowest (primary) education (IRR = 1.21, 95% CI: 1.09–1.34, $p < .001$). The regression analysis of pause ratio identified age as the only significant predictor (OR = 0.85, 95% CI: 0.74–0.98, $p = .021$). As age increases, the pause ratio decreases suggesting that older individuals tend to have a lower proportion of pause time in their narratives.

Discussion

Normative lexical diversity and pausing in the VO speakers' narratives

In this study we investigated lexical diversity and pausing in semi-spontaneous narratives produced by very old (VO) speakers. We also studied whether different physio-anatomical, socio-cognitive and linguistic background factors, namely age, dentition, educational level and language test performance predicted lexical production and pausing behaviour.

All 80–100-year-old participants were able to produce a story of some sort by looking at a wordless picture sequence. All metrics of lexical diversity (number of different words [NDW], type-token-ratio [TTR], words by grammatical class and noun-verb ratio) as well as pausing (pause time ratio, pause frequency, mean pause duration and mean syllables between pauses) conveyed a prominent inter-individual variability. This finding of notable heterogeneity conforms to previous studies of age related cognitive-linguistic alterations (e.g., Baciu et al., 2021; Keszler & Bóna, 2019).

Comparing older and younger populations is the common method for examining age-related changes, and within this framework, a universal decline in cognitive-linguistic functions is typically observed (Baciu et al., 2021). Unfortunately, research on language production for example in neurologically healthy Finnish adult and elderly populations is inadequate and consists almost exclusively of narrow control samples from pathological studies. The age range of the normative control samples is also often very wide, so the sampling does not target any particular age group or depict any particular stage of the adult language development. Considering the subject-matter of this study, we can only compare the mean duration of pauses of the VO speakers to that of a predominantly younger Finnish group aged 18–89 years ($N = 70$) (Penttilä et al., 2018). In the study of Penttilä et al. (2018) healthy adult speakers had a mean pause duration of 1.21 seconds

Table 7. Regression results for predicting lexical diversity.

Predictor	Lexical diversity variables											
	NDW				TTR				NVR			
	Est.	IRR	95% CI	p	Est.	OR	95% CI	p	Est.	Exp(B)	95% CI	p
Age	-0.02	0.98	0.94-1.03	0.497	-0.02	0.98	0.87-1.10	0.730	0.05	1.05	0.92-1.21	0.453
Dentition	-0.22	0.80	0.73-0.87	<0.001*	0.19	1.20	0.96-1.51	0.103	0.03	1.03	0.79-1.34	0.825
Education												
<i>Secondary vs. primary</i>	-0.07	0.93	0.85-1.03	0.146	-0.05	0.95	0.74-1.22	0.682	-0.27	0.76	0.57-1.01	0.071
<i>Tertiary vs. primary</i>	0.19	1.21	1.09-1.34	<0.001*	0.05	1.06	0.80-1.40	0.708	-0.02	0.98	0.70-1.36	0.917
WAB-AQ	0.19	1.20	1.13-1.28	<0.001*	-0.02	0.98	0.84-1.15	0.806	-0.07	0.93	0.77-1.13	0.447
BNT	-0.20	0.82	0.77-0.87	<0.001*	0.03	1.03	0.88-1.20	0.732	0.13	1.14	0.94-1.38	0.168
Semantic fluency	-0.01	0.99	0.93-1.04	0.659	0.06	1.06	0.91-1.23	0.459	0.11	1.11	0.93-1.34	0.232
Phonemic fluency	-0.06	0.94	0.90-0.98	0.008	0.05	1.05	0.93-1.18	0.457	0.01	1.01	0.88-1.15	0.940

Note. NDW = number of different words; TTR = Type-token ratio; NVR = Type-token ratio; WAB-AQ = Western Aphasia Battery Aphasia Quotient; BNT = Boston Naming Test; Est. = Standardized estimate; IRR = incidence rate ratio (negative binomial); OR = odds ratio (Beta & Poisson regression); Exp(B) = exponentiated coefficient (gamma regression); CI = confidence interval. Dentition is coded as dentures vs. natural teeth. Continuous predictors are standardized (z-scores), so exponentiated effect sizes represent the multiplicative change in each outcome associated with a one standard deviation increase in the predictor. * $p < .05$.

Table 8. Regression results for predicting pausing behaviour.

Predictor	Pause time ratio			Pause frequency			Mean pause duration			Mean syllables between pauses			
	Est.	OR	95% CI	Est.	OR	95% CI	Est.	Exp(B)	95% CI	Est.	IRR	95% CI	p
Age	-0.16	0.85	0.74-0.98	-0.02	0.98	0.88-1.11	0.709	0.90	0.79-1.04	0.136	1.05	0.94-1.17	0.396
Dentition	-0.14	0.87	0.68-1.11	0.04	1.04	0.83-1.31	0.700	0.87	0.66-1.16	0.305	0.99	0.79-1.23	0.93
Education													
<i>Secondary vs. primary</i>	0.00	1.00	0.76-1.32	-0.06	0.94	0.74-1.19	0.591	0.99	0.74-1.33	0.951	1.07	0.83-1.39	0.622
<i>Tertiary vs. primary</i>	-0.07	0.93	0.67-1.30	0.07	1.08	0.83-1.41	0.579	0.93	0.67-1.30	0.667	0.96	0.72-1.27	0.785
WAB-AQ	-0.13	0.87	0.73-1.03	-0.10	0.91	0.79-1.06	0.195	0.93	0.77-1.13	0.419	1.06	0.89-1.28	0.441
BNT	-0.09	0.91	0.76-1.09	0.00	1.00	0.85-1.19	0.960	0.91	0.76-1.09	0.348	1.06	0.89-1.26	0.481
Semantic fluency	-0.14	0.87	0.74-1.02	-0.05	0.95	0.82-1.11	0.479	0.92	0.78-1.10	0.359	1.07	0.91-1.25	0.401
Phonemic fluency	-0.04	0.96	0.83-1.11	-0.01	0.99	0.88-1.12	0.822	0.96	0.83-1.11	0.605	1.01	0.89-1.15	0.833

Note. Est. = Standardized estimate; IRR = incidence rate ratio (negative binomial); OR = odds ratio (Beta & Poisson regression); Exp(B) = exponentiated coefficient (gamma regression); CI = confidence interval; WAB-AQ = Western Aphasia Battery Aphasia Quotient; BNT = Boston Naming Test; Dentition is coded as dentures vs. natural teeth. Continuous predictors are standardized (z-scores), so exponentiated effect sizes represent the multiplicative change in each outcome associated with a one standard deviation increase in the predictor. * $p < .05$.

(range 0.1–5.8, $SD = 0.98$) in the same cartoon task as we used. As the VO speakers in the present study produced a mean pause duration of 1.1 seconds (range 0.4–2.9, $SD = 0.5$) we cannot assert that an age-related decline exists.

By looking at the grammatical class distribution of the narrations, we can make some coarse interpretation of the narrative trends by VO speakers. Verbs (23%) were one of the two most current grammatical classes indicating that VO speakers did not just name visual objects but reported of concrete and/or abstract actions to display events. Nouns (17%) were outnumbered by pronouns (23%), which can be explained by referencing to an introduced topic with a pronoun. Nevertheless, without analysing the context we cannot tell if the referencing was always appropriate or whether pronouns have been used to replace nouns for example because of word finding difficulties. The low occurrence of adjectives (3%) might suggest that these narratives were not very descriptive. However, the rich use of adverbs (18%) provided more detail and context to the sentences, making them more specific and perhaps descriptive in that sense. The relatively frequent appearance of conjunctions (14%) indicates an emphasis on sentence connectivity and syntactical coherence like temporal or causal relationships, which are integral to narrative flow. The relative proportion of the word classes can be induced by the nature of the elicitation task itself, but without further research we are unable to state whether the distribution would be significantly different in younger speakers. It should be noted that we are unable to say much about the quality of the stories based just on the microstructures we studied. In the future studies, a comprehensive examination of the narratives should include an analysis of the macrostructure, the semantic context of the words and pauses and the analysis of possible errors occurring at the macro- and microlevels. There is evidence that many aspects of narrative abilities decline with age regardless of cultural-linguistic background and many of the chances can be tracked to altered executive functions (e.g., Juncos-Rabadan, 1996; Kemper et al., 1990).

The significance of background factors

The regression analysis gave interesting results about the association between some of the predictor and response variables. Number of different words (NDW) was predicted by dental status, educational level and language test performance. Speakers with dentures produced lower NDW in their narratives compared to speakers with natural teeth. If we are to interpret a high NDW as an indication of preserved word-knowledge and successful word-retrieval, the connection between having natural teeth and high NDW is understandable. The result is likely explained by the fact that denture users have poorer oral health which is associated with compromised central nervous system health (Marcenes et al., 2013) and, in turn, poorer cognitive capacity (Okamoto et al., 2010) and further poorer linguistic functions (Cahana-Amitay & Albert, 2014).

Tertiary or the highest level of education was associated with higher NDW compared to primary or the lowest educational level. High educational attainment has generally been considered as an efficient buffer against language decline (Baciu et al., 2021). Although, high NDW does not straightforward imply the quality of the narration per se, it nevertheless demonstrates the speaker's access to a varied vocabulary in the task. In a previous study by Alantie et al. (2022), higher than primary education was also positively associated

with better WAB performance and semantic fluency in certain VO speakers that were under the age of 85–86 years or who had adequate natural teeth.

Additionally, higher WAB-AQ scores were statistically significantly associated with higher NDW in the current study. Better comprehensive language skills were, thus, reflected in richer vocabulary usage in semi-spontaneous narratives. As an unforeseen finding, the speakers with higher BNT scores, however, displayed lower NDW in their narratives. It is possible that speakers who were capable of precise picture naming told the story more concisely and accurately, using strategically consistent word choices. This perhaps reduced the need for synonym variation, reformulations and circumlocutions and, thus, resulted in lower occurrence of different words in the narration. The naming task (the BNT) and the narrative were conducted consecutively. Nevertheless, a semantic priming effect that would enable participants to access words in the narrative task more efficiently after the confrontation naming task is, however, unlikely. The two tasks do not involve same items or concepts, which makes the circumstances improbable for a priming-effect. Another explanation might be, that speakers who performed well in confrontation naming also produced more naming-focused or declarative speech in cartoon narration. It may have been rather a question of strategic transfer or even executive rigidity, if the participants were first simply naming object and then, in fact, carried out the same shortness or conciseness of expression to the narrative task.

The typicality of our finding is difficult to interpret. On average, the VO speakers had lower scores in the BNT than younger adult speakers in a prior data (VO speakers: BNT $M = 48.6$, cf. Manninen et al., 2015, $N = 62$, age range 18–65 years: $M = 5.4$ years). We were unable to find studies examining the relationship between the BNT or other picture naming tasks and NDW in connected speech in healthy adult and elderly populations for comparison. In some studies of speakers with aphasia, picture naming skills have been found to have a positive association with lexical diversity measures (Boucher et al., 2022). This association is opposite to our findings indicating perhaps a difference between damaged and healthy aged language. Overall, the relationship between naming and connected speech is yet unclear as shown by studies on aphasia (e.g., Biran et al., 2024; Mason & Nickels, 2022). As Mason and Nickels (2022) have suggested, whether a lexical item can be used in connected speech might not only be related to the availability of that item in a picture naming setting, as in the BNT, but also to the availability of the other lexical items required for the sentences in connected speech tasks. In summary, there may be several different linguistic-cognitive reasons behind the discrepancy between confrontation naming and vocabulary use in narrative speech.

Another significant relationship was discovered between pause time ratio and age within the range of 80–100 years. As age increases, the pause ratio decreases. Thus, the older the individuals are the less time pausing they spend during narration. This challenges the assumption that aging universally results in a higher proportion of pause time during connected speech (Bóna, 2014; Cho et al., 2021; Keszler & Bóna, 2019; Martins & Andrade, 2011). Our result may correspond better to (Keszler & Bóna's, 2019) finding that pausing variables do not always develop linearly with age. There are at least two possible explanations for this surprising trend in our study. First, as inhibitory executive functions often weaken by ageing (Cahana-Amitay & Albert, 2014) the oldest of the VO speakers may not exploit pausing for planning, processing, monitoring or correcting their language output as much as their junior VO speakers. Silent pauses may be replaced for example by

circumlocutions or repetitions. Second, since the debilitating effects of aging can be mitigated by a range of life circumstances (Baciu et al., 2021), the oldest VO speakers in this study may be those who have successfully delayed age-related decline. Their pause time ratio may assimilate that of much younger speakers. Both these tentative explanations would align with the notion that typical aging can be either usual or successful in terms of physical and cognitive function (Rowe & Kahn, 1987; Uotinen et al., 2003). Thus, it is possible that the oldest speakers in this study, were individuals who had effectively preserved their cognitive reserve and linguistic abilities. However, this possibility remains highly speculative because the idea cannot be substantiated without larger longitudinal research data.

Another influencing factor for the decreased pause time ratio in the oldest of the VO speakers might lie in the language-specific adaptation and compensation processes. According to Gerstenberg et al. (2018) the development in pausal behavior in German speakers differed from that of French speakers. The articulation rate of the German speakers decreased within a ten-year follow-up, whilst the French speakers showed a steady increase. This indicated a difference in the occurrence of pauses and thus a different cognitive processing of speech production. Unfortunately, we lack similar longitudinal studies of pauses in spoken Finnish to compare with. We can but speculate whether the decrease in pause time ratio, as well as the other findings of this study, are characteristic traits of age-related Finnish language use. It would be interesting to study the elderly Swedish (the other official language in Finland) speaking population in Finland to discover whether they present similar pausing and lexical patterns. Overall, to better understand the interplay between language and aging, it would be essential to gather comprehensive data also in other minority and underrepresented languages.

Given the relatively small sample size of our study, the statistical result discussed above may reflect the personal narrative styles as well as age-typical characteristics. Both cross-sectional and longitudinal studies on VO speakers with larger samples are required in the future. It would also be important to include qualitative content analysis of the narrations to complement these results in the future.

Limitations of this study

While this study offers new information about vocabulary use and pauses in the very old speakers, the following limitations should be acknowledged. The participants' hearing was examined by mere self-evaluation. In future studies, hearing thresholds should be clinically measured for more reliable assessment. In this study, a comic strip task was used as an elicitation method, although it has not been validated for standardized testing purposes. Content unrelated to the comic strip narrative was removed from the storytelling data in the analyses of lexical diversity. Nevertheless, we acknowledge the importance of metadiscourse in spoken language research. Elements such as digressions, expressions of uncertainty, requests for confirmation, or self-directed speech may offer valuable insights into processes associated with aging related cognitive-linguistic change. Therefore, investigating metadiscoursal features as an independent research focus would be essential in future studies.

The type-token ratio (TTR) has limitations – primarily its sensitivity to text length. As text grows, repeated words accumulate, causing TTR to sink even if the vocabulary is rich

and flexible. This is why it is not ideal for comparing samples of different sizes. A more modern and stable way to study lexical diversity would have been for example the moving average TTR (MATTR), which reduces sample length bias by calculating TTR over a fixed-size sliding window across the text and then averaging the results (Jarvis, 2013). However, even MATTR is not completely independent of sample size (Jarvis, 2013). Moreover, it is impractical to calculate by hand in clinical practice or even in research as many minority languages lack accessible computerized analyzing methods for this, and using an open artificial intelligence software for patient data is prohibited for ethical reasons.

The same variables and analytical approaches have also not been examined using this elicitation tool in previous research. Therefore, direct comparability with earlier data is not possible. This, nevertheless, reflects everyday practice in clinical speech-language therapy and research in minority and underrepresented languages (e.g., Norvik et al., 2022; Pommée et al., 2022). We have not performed inter-rater analysis for the lexical or pause variables. However, the analysis has been supported by consensus process among the authors and with the help of online grammatical parsers.

The sample size of this study is relatively modest and the gender-distribution is imbalanced with an overrepresentation of women (females $n = 41$, males $n = 9$). The sample, however, comprises all the volunteer participants that fulfilled the inclusion criteria within the recruitment period. While women outnumber men in older age groups in most societies, the female-to-male ratio still differs from normative populations. Thus, the generalizability of this data for real-life populations and for especially male speakers is limited. The limited sample size may also reduce statistical power, elevating the risk of Type II errors, and potentially contributing to the lack of significant associations observed for several predictors. Furthermore, this study does not offer comparison between neurologically healthy speakers and clinical groups. In the absence of more comprehensive language- and culture-specific samples, this data can nevertheless be used as reference in clinical work and in future studies involving growing populations of neuropathological and bi/multilingual speakers.

Conclusion

This study offers missing normative data on the different metrics of lexical diversity, pausing and their predictors in semi-spontaneous narratives by very old (VO) Finnish speakers. Both lexical diversity and pausing behaviour were characterised by marked individual variability. Thus, we can assume that the sample, although modest in size, captured a fair spectrum of typical ageing.

The novelty of this study is the association found between intact dentition and lexical diversity, which indicates the usability of dental status as a proxy for neural or systemic health and cognition. Lexical diversity was also positively predicted by high educational level and good comprehensive language skills while it was negatively predicted by good confrontation naming. This discrepancy between the two linguistic predictor variables suggests that lexical knowledge, word choices or word retrieval may manifest differently depending on task type. Within the range of 80–100 years, the oldest participants in this study had lower proportion of pauses than the younger ones. This contradicts the

expectations based on prior international studies. The result may suggest pausing patterns do not appear linearly by age in cross-sectional views.

We hope that clinical speech-language pathologists and researchers alike will be able to utilize the data presented in this study to benchmark the performance of their own patients and participants, thereby reducing reliance on intuition and experiential knowledge alone. This will increase the expertise in identifying typical linguistic changes. The trends we observed might also help specialists in those societies, who still lack native norm bases of their own. More longitudinal and cross-sectional research is required on typical adult and old populations. However, the full clinical potential of the results can only be realized once automatic speech recognition programs and automated speech analysis tools become available to clinicians in multiple languages, helping to reduce manual work and accelerate and standardize data analysis.

Acknowledgements

We want to thank Professor Jukka Tyrkkö for his expert assistance in coding the corpus tool for this study and statistician Ali Moazami Goodarzi for his expert contribution. We would like to thank cordially all the participants for their invaluable time and effort.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project was supported by research grants funded by the Finnish Cultural Foundation Pirkanmaa Regional Fund and Emil Aaltonen Foundation.

ORCID

Sonja Alantie  <http://orcid.org/0000-0002-2090-2023>

Tanja Makkonen  <http://orcid.org/0000-0001-6358-5876>

Kati Renvall  <http://orcid.org/0000-0002-3329-4602>

Data availability statement

The dataset utilized in this article is not available for public access. For data sharing request please contact the first author.

References

- Alantie, S., Tyrkkö, J., Makkonen, T., & Renvall, K. (2022). Is old age just a number in language skills? Language performance and its relation to age, education, gender, cognitive screening, and dentition in very old Finnish speakers. *Journal of Speech, Language, and Hearing Research*, 65(1), 274–291. https://doi.org/10.1044/2021_JSLHR-21-00178
- Aldarmaki, H., Ullah, A., Ram, S., & Zaki, N. (2022). Unsupervised automatic speech recognition: A review. *Speech Communication*, 139, 76–91. <https://doi.org/10.1016/j.specom.2022.02.005>

- Angelopoulou, G., Kasselimis, D., Goutsos, D., & Potagas, C. (2024). A methodological approach to quantifying silent pauses, speech rate, and articulation rate across distinct narrative tasks: Introducing the connected speech analysis protocol (CSAP). *Brain Sciences*, 14(5), 466. <https://doi.org/10.3390/brainsci14050466>
- Angelopoulou, G., Kasselimis, D., Varkanitsa, M., Tsolakopoulos, D., Papageorgiou, G., Velonakis, G., Meier, E., Karavassilis, E., Pantoleon, V., Laskaris, N., Kelekis, N., Tountopoulou, A., Vassilopoulou, S., Goutsos, D., Kiran, S., Weiller, C., Rijntjes, M., & Potagas, C. (2024). Investigating silent pauses in connected speech: Integrating linguistic, neuropsychological, and neuroanatomical perspectives across narrative tasks in post-stroke aphasia. *Frontiers in Neurology*, 15, 1347514. <https://doi.org/10.3389/fneur.2024.1347514>
- Anthony, L. (2024). AntConc (Version 4.3.1) [Computer Software]. Waseda University. <https://www.laurenceanthony.net/software/AntConc>.
- Armstrong, E. (2000). Aphasic discourse analysis: The story so far. *Aphasiology*, 14(9), 875–892. <https://doi.org/10.1080/02687030050127685>
- Baciu, M., Banjac, S., Roger, E., Haldin, C., Perrone-Bertolotti, M., Loevenbruck, H., & Démonet, J.-F. (2021). Strategies and cognitive reserve to preserve lexical production in aging. *GeroScience*, 43(4), 1725–1765. <https://doi.org/10.1007/s11357-021-00367-5>
- Biran, M., Ben-Or, G., & Yihye-Shmuel, H. (2024). Word retrieval in aphasia: From naming tests to connected speech and the impact on well-being. *Aphasiology*, 38(4), 738–757. <https://doi.org/10.1080/02687038.2023.2228017>
- Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer* (Version 6.3.13) [Computer software]. <http://www.praat.org/>
- Bóna, J. (2014). Temporal characteristics of speech: The effect of age and speech style. *The Journal of the Acoustical Society of America*, 136(2), EL116–EL121. <https://doi.org/10.1121/1.4885482>
- Boucher, J., Marcotte, K., Brisebois, A., Courson, M., Houzé, B., Desautels, A., Léonard, C., Rochon, E., & Brambati, S. M. (2022). Word-finding in confrontation naming and picture descriptions produced by individuals with early post-stroke aphasia. *The Clinical Neuropsychologist*, 36(6), 1422–1437. <https://doi.org/10.1080/13854046.2020.1817563>
- Bruner, J. (2010). Narrative, culture, and mind. In D. Schiffrin, A. DE. Fina, & A. Nylund (Eds.), *Telling stories: Language, narrative, and social life* (pp. 45–49). Georgetown University Press.
- Cahana-Amitay, D., & Albert, M. L. (2014). Brain and language: Evidence for neural multifunctionality. *Behavioural Neurology*, 2014(1), 1–16. <https://doi.org/10.1155/2014/260381>
- Chappelle, C. A. (1994). Are C-tests valid measures for L2 vocabulary research? *Second Language Research*, 10(2), 157–187. <https://doi.org/10.1177/026765839401000203>
- Cho, S., Nevler, N., Shellikeri, S., Parjane, N., Irwin, D. J., Ryant, N., Ash, S., Cieri, C., Liberman, M., & Grossman, M. (2021). Lexical and acoustic characteristics of young and older healthy adults. *Journal of Speech, Language, and Hearing Research*, 64(2), 302–314. https://doi.org/10.1044/2020_JSLHR-19-00384
- Coppieters, R., Bouzigues, A., Jiskoot, L., Montembeault, M., Tee, B. L., Rohrer, J. D., Bruffaerts, R., Russell, L. L., Foster, P. H., Ferry-Bolder, E., Convery, R., Bocchetta, M., Cash, D., Goldsmith, S., Samra, K., Thomas, D. L., van Swieten, J. C., Seelaar, H., Poos, J., . . . Remes, A. (2024). A systematic review of the quantitative markers of speech and language of the frontotemporal degeneration spectrum and their potential for cross-linguistic implementation. *Neuroscience & Biobehavioral Reviews*, 167, 105909. <https://doi.org/10.1016/j.neubiorev.2024.105909>
- Cumming, G. (2014). The new statistics: Why and how. *Psychological science*, 25(1), 7–29. <https://doi.org/10.1177/0956797613504966>
- Dos Santos Nogueira, D. M., De Matos Vieira, A. I. C., & De Azevedo Reis, E. (2023). Oral-motor function in older people: Exploring new dimensions of frailty. *Journal of Oral Rehabilitation*, 50(9), 802–811. <https://doi.org/10.1111/joor.13492>
- Drummond, C., Coutinho, G., Fonseca, R. P., Assunção, N., Teldeschi, A., de Oliveira-Souza, R., Moll, J., Tovar-Moll, F., & Mattos, P. (2015). Deficits in narrative discourse elicited by visual stimuli are already present in patients with mild cognitive impairment. *Frontiers in Aging Neuroscience*, 7, 96. <https://doi.org/10.3389/fnagi.2015.00096>

- Fergadiotis, G., & Wright, H. H. (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology*, 25(11), 1414–1430. <https://doi.org/10.1080/02687038.2011.603898>
- Fergadiotis, G., Wright, H. H., & Capilouto, G. J. (2011). Productive vocabulary across discourse types. *Aphasiology*, 25(10), 1261–1278. <https://doi.org/10.1080/02687038.2011.606974>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
- Forman, D., Berman, A., McCabe, C., Baim, D., & Wei, J. (1992). Ptca in the elderly: The “young-old” versus the “old-old”. *Journal of the American Geriatrics Society*, 40(1), 19–22. <https://doi.org/10.1111/j.1532-5415.1992.tb01823.x>
- Gerstenberg, A., Fuchs, S., Kairet, J. M., Schröder, J., & Frankenberg, C. (2018). A cross-linguistic, longitudinal case study of pauses and interpausal units in spontaneous speech corpora of older speakers of German and French. *Speech Prosody*, 2018, 211–215. <https://doi.org/10.21437/SpeechProsody.2018-43>
- Hänninen, T., Pulliainen, V., Sotaniemi, M., Hokkanen, L., Salo, J., Hietanen, M., Pirttilä, T., Pöyhönen, M., Juva, K., Remes, A., & Erkinjuntti, T. (2010). Early detection of cognitive changes in memory diseases: New cut-off scores for the Finnish version of CERAD neuropsychological battery. *Duodecim; Lääketieteellinen Aikakauskirja*, 126(17), 2013–2021.
- Harris, J. L., Kiran, S., Marquardt, T. P., & Fleming, V. B. (2008). Communication wellness Check-up©: Age-related changes in communicative abilities. *Aphasiology*, 22(7–8), 813–825. <https://doi.org/10.1080/02687030701818034>
- Harris Wright, H., Silverman, S., & Newhoff, M. (2003). Measures of lexical diversity in aphasia. *Aphasiology*, 17(5), 443–452. <https://doi.org/10.1080/02687030344000166>
- Høeg, N., Norvik, M. I., Jensen, B. U., Haaland-Johanssen, L. E. W., & Lind, M. (2023). Linguistic variables in picture and cartoon descriptions by neurologically healthy Norwegian adults. In *Nordic Aphasia Conference 2023*. <https://doi.org/10.13140/RG.2.2.15741.95202>
- Horton, W. S., Spieler, D. H., & Shriberg, E. (2010). A corpus analysis of patterns of age-related change in conversational speech. *Psychology and Aging*, 25(3), 708–713. <https://doi.org/10.1037/a0019424>
- Huang, A. R., & Lin, F. R. (2024). Hearing loss and dementia in older adults: A narrative review. *Journal of the Chinese Medical Association*, 87(3), 252–258. <https://doi.org/10.1097/JCMA.0000000000001042>
- Jarvis, S. (2013). Capturing the diversity in lexical diversity. *Language Learning*, 63(s1), 87–106. <https://doi.org/10.1111/j.1467-9922.2012.00739.x>
- Juncos-Rabadan, O. (1996). Narrative speech in the elderly: Effects of age and education on telling stories. *International Journal of Behavioral Development*, 19(3), 669–685. <https://doi.org/10.1177/016502549601900313>
- Kavé, G. (2024). Vocabulary changes in adulthood: Main findings and methodological considerations. *International Journal of Language & Communication Disorders*, 59(1), 58–67. <https://doi.org/10.1111/1460-6984.12820>
- Kavé, G., Knafo, A., & Gilboa, A. (2010). The rise and fall of word retrieval across the lifespan. *Psychology and Aging*, 25(3), 719–724. <https://doi.org/10.1037/a0018927>
- Kavé, G., Samuel-Enoch, K., & Adiv, S. (2009). The association between age and the frequency of nouns selected for production. *Psychology and Aging*, 24(1), 17–27. <https://doi.org/10.1037/a0014579>
- Kavé, G., & Yafé, R. (2014). Performance of younger and older adults on tests of word knowledge and word retrieval: Independence or interdependence of skills? *American Journal of Speech-Language Pathology*, 23(1), 36–45. [https://doi.org/10.1044/1058-0360\(2013\)12-0136](https://doi.org/10.1044/1058-0360(2013)12-0136)
- Kemper, S., Rash, S., Kynette, D., & Norman, S. (1990). Telling stories: The structure of adults’ narratives. *European Journal of Cognitive Psychology*, 2(3), 205–228. <https://doi.org/10.1080/09541449008406205>
- Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging*, 16(2), 312–322. <https://doi.org/10.1037/0882-7974.16.2.312>

- Kertesz, A., Pietilä, M.-L., Lehtihalmes, M., Klippi, A., & Lempinen, M. (Eds.). (2005). *Western Aphasia Battery: Käsikirja [Finnish manual]*. Hogrefe Psykologien Kustannus.
- Keszler, B., & Bóna, J. (2019). Pausing and disfluencies in elderly speech - longitudinal case studies. *Proceedings of DiSS 2019*, 67–70. <https://doi.org/10.21862/diss-09-018-kesz-bona>
- Kielipankki. (n.d.). *Fintag: Finnish tagtools*. The Language Bank of Finland. <https://www.kielipankki.fi/tools/fintag/>
- Kintz, S., Fergadiotis, G., & Wright, H. H. (2016). Aging effects on discourse production. In H. H. Wright (Ed.), *Cognition, language and aging* (ss. 81–106). John Benjamins Publishing Company. <https://doi.org/10.1075/z.200.04kin>
- Korpijaakko-Huuhka, A.-M. (1995). *Kertomuksen koherenssi [The coherence of a story]* [Licentiate thesis, University of Helsinki, Department of Phonetics].
- Korpijaakko-Huuhka, A.-M. (2003). *Kyllä se lintupelotintaulujuttu siinä nyt on käsittelyssä: Afaattisten puhujien kielellisiä valintoja sarjakuvatehtävässä [Aphasic speakers' linguistic choices in a cartoon-story task]* [Doctoral dissertation, University of Helsinki]. Helda. <https://helda.helsinki.fi/items/43179710-bca2-4177-a471-fb1ebcf407c6>
- Laine, M., Koivuselkä-Sallinen, P., Hänninen, R., & Niemi, J. (1997). *Bostonin nimentätesti: Suomenkielinen versio [Boston Naming Test: Finnish version]*. Psykologien kustannus.
- Laine, M., Niemi, J., Koivuselkä-Sallinen, P., & Tuomainen, J. (1997). *Afasian ja liitännäihäiriöiden arviointi [The assessment of aphasia and related disorders]*. Psykologien kustannus.
- Li, L., Zhang, Q., Yang, D., Yang, S., Zhao, Y., Jiang, M., Wang, X., Zhao, L., Liu, Q., Lu, Z., Zhou, X., Gan, Y., & Wu, C. (2023). Tooth loss and the risk of cognitive decline and dementia: A meta-analysis of cohort studies. *Frontiers in Neurology*, 14, 1103052. <https://doi.org/10.3389/fneur.2023.1103052>
- Lingsoft. (n.d.). Fintwol morphological analyzer. <http://www2.lingsoft.fi/cgi-bin/fintwol>
- Lofgren, M., & Hinzen, W. (2022). Breaking the flow of thought: Increase of empty pauses in the connected speech of people with mild and moderate Alzheimer's disease. *Journal of Communication Disorders*, 97, 106214. <https://doi.org/10.1016/j.jcomdis.2022.106214>
- Lövdén, M., Ghisletta, P., & Lindenberger, U. (2004). Cognition in the Berlin Aging Study (BASE): The first 10 years. *Aging, Neuropsychology, and Cognition*, 11(2–3), 104–133. <https://doi.org/10.1080/13825580490510982>
- Mack, J. E., Chandler, S. D., Meltzer-Asscher, A., Rogalski, E., Weintraub, S., Mesulam, M.-M., & Thompson, C. K. (2015). What do pauses in narrative production reveal about the nature of word retrieval deficits in PPA? *Neuropsychologia*, 77, 211–222. <https://doi.org/10.1016/j.neuropsychologia.2015.08.019>
- Makkonen, T., Ruottinen, H., Puhto, R., Helminen, M., & Palmio, J. (2018). Speech deterioration in amyotrophic lateral sclerosis (ALS) after manifestation of bulbar symptoms. *International Journal of Language & Communication Disorders*, 53(2), 385–392. <https://doi.org/10.1111/1460-6984.12357>
- Manninen, R., Pietilä, M., Setälä, P., & Laitinen, V. (2015). *Kat-testi: Kielelliset arviointitehtävät: lievien häiriöiden määrittämiseksi aikuisilla [KAT Test: Linguistic assessment tasks to determine mild disorders in adults]*. Puheterapeuttien Kustannus Oy.
- Marcenes, W., Kassebaum, N. J., Bernabé, E., Flaxman, A., Naghavi, M., Lopez, A., & Murray, C. J. L. (2013). Global burden of oral conditions in 1990–2010. A systematic analysis. *Journal of Dental Research*, 92(7), 592–597. <https://doi.org/10.1177/0022034513490168>
- Martins, V. D. O., & Andrade, C. R. F. D. (2011). Estudo das pausas em idosos. *Revista Da Sociedade Brasileira De Fonoaudiologia*, 16(3), 344–349. <https://doi.org/10.1590/S1516-80342011000300017>
- Mason, C., & Nickels, L. (2022). Are single-word picture naming assessments a valid measure of word retrieval in connected speech? *International Journal of Speech-Language Pathology*, 24(1), 97–109. <https://doi.org/10.1080/17549507.2021.1966098>
- Moscoso Del Prado Martín, F. (2017). Vocabulary, grammar, sex, and aging. *Cognitive Science*, 41(4), 950–975. <https://doi.org/10.1111/cogs.12367>
- Niccoli, T., & Partridge, L. (2012). Ageing as a risk factor for disease. *Current Biology*, 22(17), R741–R752. <https://doi.org/10.1016/j.cub.2012.07.024>

- Niimi, M., & Nishio, S. (2001). Speaking rate and its components in dysarthric speakers. *Clinical Linguistics & Phonetics*, 15(4), 309–317. <https://doi.org/10.1080/02699200010024456>
- Norvik, M. I., Lind, M., & Jensen, B. U. (2022). Working with multilingual aphasia: Attitudes and practices among speech and language pathologists in Norway. *International Multilingual Research Journal*, 16(4), 273–290. <https://doi.org/10.1080/19313152.2021.2015935>
- Okamoto, N., Morikawa, M., Okamoto, K., Habu, N., Iwamoto, J., Tomioka, K., Saeki, K., Yanagi, M., Amano, N., & Kurumatani, N. (2010). Relationship of tooth loss to mild memory impairment and cognitive impairment: Findings from the Fujiwara-kyo study. *Behavioral and Brain Functions*, 6(1), 77. <https://doi.org/10.1186/1744-9081-6-77>
- Penttilä, N., Korpijaakko-Huuhka, A.-M., & Kent, R. D. (2018). Tavallista sujumattomuutta: aikuisten puheen sujuvuuden kvantitatiivinen analyysi [Typical fluency in Finnish adults]. *Puhe Ja Kieli*, 38(3), 153–173. <https://doi.org/10.23997/pk.77384>
- Pernambuco, L., Espelt, A., Magalhães, H. V., & de Lima, K. C. (2017). Recommendations for elaboration, transcultural adaptation and validation process of tests in speech, hearing and language pathology. *Codas (São Paulo)*, 29(3), Article e20160217. <https://doi.org/10.1590/2317-1782/20172016217>
- Pommée, T., Balaguer, M., Mauclair, J., Piquier, J., & Woisard, V. (2022). Assessment of adult speech disorders: Current situation and needs in French-speaking clinical practice. *Logopedics, Phoniatrics, Vocology*, 47(2), 1–15. <https://doi.org/10.1080/14015439.2020.1870245>
- Roark, B., Mitchell, M., Hosom, J.-P., Hollingshead, K., & Kaye, J. (2011). Spoken language derived measures for detecting mild cognitive impairment. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(7), 2081–2090. <https://doi.org/10.1109/TASL.2011.2112351>
- Rowe, J. W., & Kahn, R. L. (1987). Human aging: Usual and successful. *Science*, 237(4811), 143–149. <https://doi.org/10.1126/science.3299702>
- Salthouse, T. A. (2019). Trajectories of normal cognitive aging. *Psychology and Aging*, 34(1), 17–24. <https://doi.org/10.1037/pag0000288>
- Språkbanken Text. (n.d.). *Sparv: Språkbanken Text's text analysis tool*. University of Gothenburg. <https://spraakbanken.gu.se/en/tools/sparv>
- Stark, B. C. (2019). A comparison of three discourse elicitation methods in aphasia and age-matched adults: Implications for language assessment and outcome. *American Journal of Speech-Language Pathology*, 28(3), 1067–1083. https://doi.org/10.1044/2019_AJSLP-18-0265
- Statistics Finland. (2024). *Population structure*. https://stat.fi/tup/suoluk/suoluk_vaesto.html#vaeston-ikarakenne
- Ulatowska, H. K., Chapman, S. B., Highley, A. P., & Prince, J. (1998). Discourse in healthy old-elderly adults: A longitudinal study. *Aphasiology*, 12(7–8), 619–633. <https://doi.org/10.1080/02687039808249562>
- United Nations. (2025). Ageing. <https://www.un.org/en/global-issues/ageing>
- Uotinen, V., Suutama, T., & Ruoppila, I. (2003). Age identification in the framework of successful aging. A study of older Finnish people. *The International Journal of Aging and Human Development*, 56(3), 173–195. <https://doi.org/10.2190/6939-6W88-P2XX-GUQW>
- Williams, E., Theys, C., & McAuliffe, M. (2023). Lexical-semantic properties of verbs and nouns used in conversation by people with Alzheimer's disease. *PLOS ONE*, 18(8), e0288556. <https://doi.org/10.1371/journal.pone.0288556>
- Zozuk, N. Č. (2023). Lexical diversity and language impairment. *Journal of Linguistics/Jazykovedný Casopis*, 74(1), 301–309. <https://doi.org/10.2478/jazcas-2023-0047>

Appendix

Table A1. Variable means, standard deviations, and correlations with confidence intervals

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	84.84	4.06													
2. Education	1.88	0.77	-0.1												
3. Dentition	1.54	0.50	0.22	-0.30*											
			[-0.06, .47]	[-.54, -.03]											
4. WAB-AQ	95.68	3.41	-0.45**	.29*	-0.30*										
			[-.65, -.19]	[.01, .53]	[-.54, -.03]										
5. SemF	21.16	6.80	-0.39**	.42**	-0.36*	.67**									
			[-.60, -.12]	[.15, .62]	[-.58, -.09]	[.48, .80]									
6. PhonF	13.92	5.62	-0.07	.37**	-0.22	.46**	.58**								
			[-.34, .21]	[.11, .59]	[-.47, .07]	[.21, .66]	[.35, .74]								
7. BNT	48.56	6.84	-0.45**	.39**	-0.49**	.70**	.52**	.37**							
			[-.65, -.20]	[.12, .60]	[-.67, -.24]	[.52, .82]	[.28, .70]	[.10, .59]							
8. NDW	64.30	24.07	-0.12	0.14	-0.23	0.17	0.12	-0.06	-0.05						
			[-.39, .16]	[-.15, .40]	[-.48, .05]	[-.12, .43]	[-.17, .38]	[-.33, .23]	[-.32, .23]						
9. TTR	0.70	0.07	-0.12	0.11	0.1	0.14	0.18	0.2	0.11	-0.25					
			[-.38, .17]	[-.17, .38]	[-.19, .37]	[-.15, .40]	[-.10, .44]	[-.09, .45]	[-.17, .38]	[-.49, .03]					
10. NVR	0.75	0.28	-0.06	0.12	-0.19	0.11	0.19	0.14	0.19	-0.07	0.12				
			[-.33, .22]	[-.17, .38]	[-.44, .10]	[-.18, .37]	[-.10, .44]	[-.15, .40]	[-.09, .44]	[-.34, .21]	[-.16, .39]				
11. PTR	0.34	0.11	0.02	-0.30*	0.1	-0.49**	-0.48**	-0.41**	-0.37**	-0.14	-0.07	-0.17			
			[-.26, .29]	[-.53, -.02]	[-.19, .37]	[-.68, -.25]	[-.67, -.23]	[-.62, -.15]	[-.59, -.11]	[-.40, .15]	[-.34, .21]	[-.43, .11]			
12. PF	0.20	0.06	0.13	-0.1	0.15	-0.38**	-0.33*	-0.24	-0.27	0.13	0	0.18	0.55**		

(Continued)

Table A1. (Continued).

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
13. MPD	1.13	0.49	-0.01 [-.29, .27]	-0.22 [-.47, .06]	0.04 [-.24, .32]	-0.33* [-.56, -.06]	-0.33* [-.56, -.06]	-0.29* [-.53, -.01]	-0.28 [-.51, .00]	-0.26 [-.50, .02]	0.01 [-.27, .29]	-0.24 [-.48, .04]	.79** [.66, .88]	0.06 [-.22, .33]	
14. MSBP	9.12	2.90	-0.06 [-.33, .22]	0.19 [-.09, .44]	-0.17 [-.43, .11]	.38** [.11, .59]	.37** [.10, .59]	.30* [.02, .53]	.35* [.08, .57]	-0.23 [-.48, .05]	0.08 [-.20, .35]	-0.05 [-.33, .23]	-0.61** [-.76, -.41]	-0.91** [-.95, -.85]	-0.14 [-.40, .14]

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). WAB-AQ = Western Aphasia Battery – Aphasia Quotient; SemF = semantic fluency; PhonF = phonemic fluency; BNT = Boston Naming Test; NDW = number of different words; TTR = type-token ratio; NVR = noun-verb ratio; PTR = pause time ratio, PF = pause frequency, MPD = mean pause duration, MSBP = mean syllables between pauses.
*indicates $p < .05$. **indicates $p < .01$.