Do disfluencies increase with age? Evidence from a sequential corpus study of disfluencies

Article *in* Psychology and Aging · March 2023

DOI: 10.1037/pag0000741

3 authors:

Eleonora Beier
University of California, Davis

Suphasiree Chantavarin
Chulalongkorn University

Fernanda Ferreira
University of California, Davis

Do disfluencies increase with age?
Evidence from a sequential corpus study of disfluencies

Eleonora J. Beier[1]
Suphasiree Chantavarin[1, 2]
Fernanda Ferreira[1]

[1]Department of Psychology, University of California Davis
[2]Faculty of Psychology, Chulalongkorn University

Author Note

All data and analysis scripts for this study are available on the Open Science Framework (https://osf.io/ew26p/)(Beier et al., 2023). The results of this study have been presented as a talk at the Psychonomic Society 61st Annual Meeting and at the 2019 California Meeting on Psycholinguistics. The present manuscript has not been submitted for publication elsewhere. We would like to give special thanks to all of the research assistants who coded for the presence of disfluencies and transcribed all interviews in our corpus; in alphabetical order by first name: Adrian Zhou, Anirudh Murugesan, Anya Grosskopf, Cassiopeia Pryor, Emily Strand, Francisco Correira, Isaias Ceballos III, Kevin Ly, Leslie Wang, Nicole Mehring, Sidney Faust, Wallace Lau. We would also like to thank Debra Long for her helpful suggestions, Emilio Ferrer, Mijke Rhemtulla and Philippe Rast for their comments on the statistical analyses, as well as the UC Davis language and cognitive science communities that provided us with useful feedback on this work. We acknowledge funding from the National Science Foundation grant number BCS16-50888 awarded to FF, National Science Foundation GRFP number 1650042 awarded to EB, and the Chulalongkorn University grant number CU_GIF_62_01_38_01 awarded to SC.

Abstract

Speech disfluencies such as repeated words and pauses provide information about the cognitive systems underlying speech production. Understanding whether older age leads to changes in speech fluency can therefore help characterize the robustness of these systems over the lifespan. Older adults have been assumed to be more disfluent, but current evidence is minimal and contradictory. Particularly noteworthy is the lack of longitudinal data that would help establish whether a given individual's disfluency rates change over time. This study examines changes in disfluency rates through a sequential design with a longitudinal component, involving the analysis of 325 recorded interviews conducted with 91 individuals at several points in their lives, spanning the ages of 20 to 94 years. We analyzed the speech of these individuals to assess the extent to which they became more disfluent in later interviews. We found that with older age, individuals spoke more slowly and repeated more words. However, older age was not associated with other types of disfluencies such as filled pauses (*uh*'s and *um*'s) and repairs. Overall, this study provides evidence that although age itself is not a strong predictor of disfluencies, age leads to changes in other speech characteristics among some individuals (i.e., speech rate and indicators of lexical and syntactic complexity), and those changes in turn predict the production of disfluencies over the lifespan. These findings help resolve previous inconsistencies in this literature and set the stage for future experimental work on the cognitive mechanisms underlying changes in speech production in healthy aging.

Public Significance Statement

This study measures changes in speech fluency over the lifespan sequentially using a novel corpus of 91 individuals interviewed at multiple times in their lives. We found that while older adults spoke more slowly and repeated more words, they were not more likely to produce filled pauses (*uh*'s and *um*'s) and repairs. Overall, we suggest that age is not as strong a predictor of disfluencies as are other individual differences in speech characteristics.

Keywords: *speech production; speech; disfluencies; cognitive aging; sequential corpus analysis*

Introduction

Although language functions are among the best preserved cognitive abilities in older adults, some evidence suggests that speaking becomes more difficult as we age. Older speakers have been theorized to become more disfluent – they may produce more *uh*’s, *um*’s, and word repetitions than they did when they were younger as they try to gain time to retrieve the word or structure they wish to articulate. These potential patterns are significant because disfluencies provide information concerning the robustness of the cognitive systems underlying speech production[1], such as those involved in memory retrieval and sequential decision-making (Clark & Wasow, 1998; Engelhardt et al., 2012; Engelhardt et al., 2010). Studying patterns of disfluency over the lifespan may reveal how these cognitive processes change in healthy aging, due, for example, to a significant decline in processing speed (Salthouse, 1996; Salthouse & Meinz, 1995). However, to date it is unclear whether disfluencies increase, decrease, or remain stable with age. This is because previous studies have investigated this question cross-sectionally and have used inconsistent classifications of disfluencies, leading to mixed results. In this sequential investigation of speech fluency—using a novel methodology that exploits the availability of online media—we overcome the limitations of previous cross-sectional designs by analyzing speech samples from people who have been interviewed multiple times in their lives. Using a new corpus of 325 interviews with 91 individuals, spanning the ages of 20 to 94 years, we address the question of whether individual speakers become more disfluent as they age, using an approach that allows for greater statistical power and fewer biases compared to cross-sectional designs.

Speech production as a window into cognitive aging

The language system—and its malfunctions—have long been used to study the many domain-general mechanisms supporting language comprehension and production (Clark & Wasow, 1998; Engelhardt et al., 2010; Engelhardt et al., 2012). These include working memory, long-term memory retrieval, perceptual processing, attention, executive functioning, planning and linearization of complex thought. In particular, speech disfluencies are a powerful tool for investigating the cognitive systems that support discourse production over the lifespan. For example, tip-of-the-tongue (TOT) states, which often lead to slower speech rate, reveal that speaking requires retrieving phonological, semantic, and grammatical representations at separate stages of processing. Numerous empirical studies have shown age-related deficits in retrieving lexical and phonological representations from memory. Older adults experience more tip-of-the-tongue (TOT) states than young adults do (Burke et al., 1991; Rastle & Burke, 1996; James & Burke, 2000), and picture naming studies have shown that older adults tend to be slower and less accurate than young adults at producing correct names for pictures (Nicholas et al., 1985; Rogalski et al., 2011). Clark and Wasow (1988) argue that disfluencies often reflect high-level decisions (e.g., retrieval of a concept from memory) that have downstream, lower-level consequences (e.g., a repeated word). In line with this view, word repetitions might be another strategy—in addition to slowing down speech or pausing—for buying time during production, permitting the speaker to accommodate early linguistic encoding of a concept that has been retrieved from memory (Clark & Wasow, 1998). A decline in inhibitory control systems over age (Engelhardt et al., 2010; Engelhardt et al., 2012) could also result in increased production of repairs and corrections in discourse. Thus, the trajectory of disfluencies over the lifespan is indicative of how these cognitive processes in the language system change as speakers age. In addition, because different kinds of disfluencies may arise from distinct cognitive and linguistic sources, studying changes in disfluencies over age allows investigators to isolate the specific aspects of cognition that are sensitive to aging effects.

Competing theories attribute the age-related changes in speech production to different causes. Production impairments may be due to a deterioration of domain-general cognitive functions, such as reduced processing speed (Salthouse, 1996; Salthouse & Meinz, 1995) and working memory capacity (Rayner et al., 2006; Stine-Morrow et al., 2006). Other theories attribute the changes in production to age-related declines specifically in language functions. Word retrieval difficulty, for example, may be a consequence of age-related deficits in the ability to inhibit competing words (e.g., Healey et al., 2013). Alternatively, the Transmission Deficit Hypothesis proposes that word retrieval difficulty may occur because the representations for the form and sound of words that have not been frequently or recently retrieved become weakened in the aging brain, reducing the efficacy and speed of transmission and priming from one representation to another (Burke et al., 2000; Mackay & Burke, 1990; Burke & Shafto, 2008). A third competing account is that word retrieval difficulty may not reflect cognitive decline per se, but rather the cost of having to sort through more information stored in memory, an expected outcome of learning across the lifespan (Ramscar et al., 2014; but see Shafto et al., 2017). According to this view, older adults' performance on psychometric tests may exemplify the tradeoff between amount of information accumulated and ease of retrieval, rather than a deterioration of neurocognitive functions.

To distinguish among these accounts, it is necessary to first determine the extent to which different aspects of speech production and fluency vary—or remain stable—over the lifespan. Answering this question would provide insights into the relevance of domain-general cognitive functions to the language system, and conversely, observing age-related changes in language processing is an additional tool for the study of cognitive aging.

Disfluencies over the lifespan

Disfluencies have been identified at different levels of production, such as prosodic (e.g., improper stress), lexical (e.g., repetitions), and syntactic (e.g., phrase revisions). Other types of disfluencies include filled pauses (e.g., *um, uh*), lexical fillers (e.g., *well, you know*) and silent pauses. In addition, fluency can be assessed by measuring a person's speech rate, or the speed with which they talk. In this study, we focus on filled pauses (*um, uh*), repeats (e.g., *went to the the store*), and repairs (e.g., *I think- I believe that…*), as well as speech rate. These measures were selected because previous studies have found significant age-related changes in these fluency measures, indicating that older adults speak more slowly (Horton et al., 2010; Castro & James, 2014; Gordon et al., 2019) and produce more filled pauses, repeats, and repairs than younger adults do (e.g., Bortfeld et al., 2001; Metz & James, 2019; Dennis & Hess, 2016). Silent pauses were not examined because our speech samples often contained overlapping talkers and background music, making the identification of silence almost impossible.

As speech disfluencies are thought to arise from disruptions in the cognitive mechanisms underlying production, it is reasonable to assume that decline in these cognitive processes with age would also lead older speakers to produce more disfluencies. In fact, evidence supports this hypothesis (e.g., Mortensen et al., 2006; Burke & Shafto, 2008). Older adults have been found to produce more filled pauses (consistently defined across previous studies as the count of *uh*'s and *um*'s), with a marked increase occurring around age 60 as compared to middle-aged and young adults, observed in picture description tasks (James et al., 2018; Spieler & Griffin, 2007) as well as in conversational speech and narratives (Bortfeld et al., 2001; Manning & Monte, 1981; Horton et al., 2010; Metz & James, 2019). Compared to middle-aged and young adults, older adults may also be more likely to repeat sounds, words, or short phrases than middle-aged and young adults do when describing pictures (Le Dorze & Bedard, 1998; Dennis & Hess, 2016; Castro & James, 2014) and during conversational speech (Bortfeld et al., 2001). Older adults also show a greater tendency to repair their utterances in picture description tasks (Shewan & Henderson, 1988; Spieler & Griffin, 2007; Dennis & Hess, 2016; Castro & James, 2014) and in conversational speech (Schmitter-Edgecombe et al., 2000; Bortfeld et al., 2001). Lastly, older adults' speech becomes slower with age during conversation (Duchin & Mysak, 1987; Horton et al., 2010) and in picture description tasks (Castro & James, 2014). Overall, age-related increases in disfluency rates may partly reflect adaptive strategies to cope with reduced speed and accuracy in retrieving words from the lexicon (Warren et al., 2018; but see Kave & Goral, 2017). Older adults may, for instance, slow down speech or repeat words to allocate more time for word retrieval. Beyond word-level processing, increased disfluency rates may also reflect deficits in planning the content and syntactic structure of utterances (Bortfeld et al., 2001).

In contrast to these findings, however, other studies have reported that disfluencies do not significantly increase with age, particularly for repeats and repairs. Some studies have reported an increase in repairs but not repeats (Schmitter-Edgecombe et al., 2000), while other studies did not find an age-related increase in repeats and repairs in picture description tasks (Cooper, 1990; Duchin & Mysak, 1987) and in open-ended speech (Duchin & Mysak, 1987; Metz & James, 2019). Overall, converging evidence indicates that the rate of filled pauses increases with age, but evidence is mixed regarding the rate of repeats and repairs across age.

These discrepancies in the literature concerning age-related changes in disfluencies may be due to a number of factors. One is the inconsistency in the topics and settings of discussions, which have been shown to influence speech fluency (Castro & James, 2014; James et al., 2018; Kemper et al., 2003; Metz & James, 2019). Another potential contributing factor is that different studies have focused on different subsets of disfluencies, and they also tend to classify and define disfluencies in different ways. For instance, the operational definition of "repeats" varies across studies, which may (or may not) include or distinguish between the repetition of a single word (*the the;* Cooper, 1990), part-words (*I went t-t-to the store;* Metz & James, 2019; Duchin & Mysak, 1987), or phrases longer than one word (*the store, the store that sells candles*; Dennis & Hess, 2016). Given the discrepancies in how disfluencies are categorized and measured across different studies, it is difficult to draw conclusions about age-related changes in disfluency rates. In fact, it may be counterproductive to analyze disfluencies as a monolithic category because different types of disfluencies may arise from changes in different underlying cognitive mechanisms (Bortfeld et al., 2001; Castro & James, 2014). For example, speech fillers at different sentential locations might indicate problems in different stages of speech production, such as conceptual, lexical, phonological, or articulation difficulties (Fraundorf & Watson, 2014; Hartsuiker & Notebaert, 2010; Metz & James, 2019). Similarly, filled and unfilled pauses may trade off, as speakers can choose to gain time while planning the next utterance by either staying silent or by producing a filler sound like *uh*.

A third reason for the mixed results may be that previous studies employed cross-sectional designs, which only permit comparison of disfluency rates between different groups of individuals. Because the people at each age are different in cross-sectional comparisons, the observed differences in speech fluency may result from inter-individual characteristics other than age, such as preferred strategies for coping with difficulties during speech production. Moreover, in cross-sectional studies it is difficult to disentangle the effects of aging from generational differences between the birth cohorts in the sample. Compared to earlier-born cohorts at the same age, later-born cohorts perform better at baseline and also show less cognitive decline across time on measures of language, executive functioning, attention, processing speed, and verbal memory (Dodge et al., 2014, 2017). It is therefore possible that the apparent decrease in speech fluency in cross-sectional studies indicates a between-generation, positive cohort effect favoring later-born cohorts, rather than a within-person effect of aging per se.

To circumvent these issues relating to cross-sectional comparisons, a study with a longitudinal component is advantageous, as it enables a comparison of disfluency rates among the same individuals as they age. In general, longitudinal studies are scant compared to cross-sectional studies because they require enormous amounts of time and resources to conduct, and the sample size and data may decrease as participants drop out of the study over time. A few longitudinal studies have investigated the relationship between age and speech production, including phonetic characteristics (e.g., f0 and formant frequencies; Reubold et al., 2010; Gahl et al., 2014; Gahl & Baayen, 2019), accent (Harrington et al., 2000), and grammatical complexity and propositional content (Kemper et al., 2001). However, these studies primarily examined age ranges limited to middle adulthood only (Gahl et al., 2014; Gahl & Baayen, 2019), older adults only (Kemper et al., 2001; Berisha et al., 2015), or were limited to a very small sample size (e.g., between one and five speakers, Harrington et al. 2000; Reubold et al., 2010). Other longitudinal studies have used the picture naming task to assess lexical retrieval ability over the lifespan (Connor et al., 2004; Zec et al., 2005; Kave et al., 2010), but there has not been any longitudinal investigation of speech disfluencies in particular.

The current study

The goal of the present study is to examine age-related changes in speech fluency through a sequential design with a longitudinal component, rather than through a purely cross-sectional design, and by determining the independent effect of age on different types of disfluencies. Our study takes advantage of the ubiquity of publicly available interviews in online media to permit a sequential investigation of disfluencies in conversational speech. We generated a corpus of spoken language across the lifespan (20 to 94 years) by focusing on individuals who have been interviewed at several points in their lives and whose interviews are publicly available. This approach presents a unique opportunity to examine how fluency changes as individuals age by introducing a longitudinal component which alleviates the limitations of cross-sectional designs (Lu et al., 2013; Xu et al., 2014; Deary et al., 2009; Whitbourne, 2019).

We used this novel corpus to assess how disfluencies and speech rate vary for each individual as they age. Beyond these main effects, we also performed exploratory analyses to determine whether specific aspects of language function (e.g., content word frequency) vary with age and influence each other, to identify the potential mechanisms responsible for the observed changes in production. While the design of this corpus study allows us to observe age-related changes in speech production sequentially and alleviates the limitations of cross-sectional designs, it also introduces an important limitation as it precludes the possibility of directly measuring performance associated with the cognitive domains that may underlie fluency. Our exploratory analyses address this limitation by measuring aspects of the interviewees' speech that may also reflect changes in the cognitive mechanisms underlying speech production (e.g., retrieval of information, and utterance planning), such as sentence length and lexical diversity. We additionally measured content word frequencies as a proxy for vocabulary size and language experience, another potential contributing factor to speech fluency (Ramscar et al., 2014). We selected these particular measures of speech characteristics due to their relevance in a previous cross-sectional study of speech production. Lexical diversity and sentence length were found to correlate both with age and with the rate of filled pauses (Horton et al., 2010). While word frequency was not related to filled pauses, it is possible that this variable may predict the rate of other types of disfluencies (e.g. repeats and repairs) that were not measured in Horton et al. (2010). Thus, we assessed whether these aspects of production predict each type of disfluency, and whether they themselves vary as a function of age and speech rate. This allows us to test the idea that different disfluency types may emerge through dissociable underlying cognitive mechanisms, such as a decrease in processing speed (Salthouse, 1996; Salthouse & Meinz, 1995), difficulty with word retrieval (Burke et al., 1991; Nicholas et al., 1985), or increased vocabulary size (Ramscar et al., 2014). Lastly, we performed an exploratory analysis on a subset of our sample comprising only older adults who were at least 60 years old at the time of their last interview, in order to draw a more direct comparison between our findings and those of cross-sectional studies of older adults.

Overall, this sequential corpus study assessed whether different types of speech disfluencies vary with age and as a function of other changes in speech characteristics (i.e., speech rate, utterance length, lexical diversity, and word frequency). As we further argue in the Discussion, drawing a clearer picture of changes in fluency over the lifespan is the first step towards a better understanding of the age-related cognitive processes that underlie the language system and what speech production reveals about cognitive aging.

Method

Transparency and openness

All data and analytic code are available on the Open Science Framework (OSF) (see Author Note; see Results section for details about software packages used). The following sections describe the process through which we arrived at our sample size, including data exclusions, as well as all measures employed. There are no materials or manipulations to report due to the nature of this corpus study. The study design, hypotheses, and analytic plan were not preregistered.

Interview selection

Data were obtained from interviews of 91 individuals, accessed through publicly available websites such as YouTube or online talk shows. Interviewees were drawn from a variety of occupations, categorized as: actors, athletes, authors, business leaders, composers and musicians (instrumental), directors, fashion designers, politicians, singers, and TV anchors. This improved the generalizability of our findings, as famous individuals from a single occupation (e.g., only politicians) may display peculiarities in their speech due to training or experience. The age range in the sample was 20 to 94 years (see Table 1 for the sample demographics). The University of California Davis Institutional Review Board declared this research exempt from IRB review.

The selection of potential interviewees and interviews was performed by coders who did not know the study's hypotheses, to avoid selection bias. The coders were instructed to select interviewees for whom they could find a minimum of three publicly available interviews that were at least seven years apart. All selected interviewees were (as far as we are aware and could ascertain through media sources) native English speakers (regardless of nationality or accent), and all interviews were conducted in English.

Additionally, selected interviews had to contain at least two minutes of the interviewee's speech. Due to the nature of the interviews, the interviewee's speech was often intermixed with speech from other individuals in the form of interruptions and conversational overlap (e.g., the interviewer or other interviewees). Therefore, the coders marked the onset and offset of speech segments where the interviewee spoke alone for a minimum of five seconds. Speech segments ended when the interviewee was interrupted or was silent for more than five seconds. At this stage, very brief interruptions on the part of the interviewer (e.g., back-channel responses such as “yeah” in response to the interviewee’s speech) were included as part of the interviewee’s speech segments, but these segments were not analyzed as part of the interviewees’ speech. The duration of the interviewee’s speech for each interview was calculated by summing the length of the interviewee’s speech segments. Longer interviews were cropped so that the total length of the interviewee’s speech was approximately two minutes. Interviews with less than two minutes of interviewee speech were excluded from the analyses.

Prior to this process, a total of 102 interviewees was selected for a total of 375 interviews. Twenty-eight interviews (11 interviewees) were excluded from the analyses for one of the following reasons: The interview contained less than two minutes of the interviewee’s speech; interview year and interviewee’s age were not specified; the interview took place less than seven years from other selected interviews; or the publicly available interview was no longer available. Thus, our analyses included 91 interviewees and a total of 325 interviews (see Table 2 for a summary of the corpus characteristics, and Table 3 for the full list of selected interviews).

Table 1

Demographics of the interviewees in the speech corpus

Variable	N=91	%
Gender		
Female	28	30.77
Male	63	69.23
Occupation		
Actors/Actresses	19	20.88
Athletes	7	7.70
Authors	7	7.70
Business leaders	7	7.70
Composers/Musicians	7	7.70
Directors	9	9.89
Fashion designers	2	2.20
Politicians	11	12.09
Singers	19	20.88
TV Anchors	3	3.30

Table 2

Summarized characteristics of the speech corpus

Variable	*M*	*SD*	Range
Number of Interviews Per Person	3.57	0.85	3-6
Age at Interview (Years)	53.13	15.78	20-94
Total Time Span Per Person (Years)	28.04	9.84	13-51
Time Span Between Interviews (Years)*	10.91	5.19	5-36
Age At First Interview (Years)	38.17	11.17	20-67

**Note*. The lower range is 5 because a small subset (3%) of the interviews had a time span of 5-6 years between interviews. All remaining interviews were at least 7 years apart.

Table 3

List of interviewees (N = 91) and interviews (N = 325) in the analyses

Interviewees by Occupation	Number of Interviews	Interview Age	Interview Year
Actors/Actresses			
Betty White	4	43, 64, 78, 94	1965, 1986, 2000, 2016
Christie Brinkley	3	31, 40, 61	1985, 1994, 2015
Diane Lane	3	35, 45, 52	2000, 2010, 2017
Harrison Ford	5	35, 43, 51, 60, 74	1977, 1985, 1993, 2002, 2016
Hugh Grant	3	35, 42, 55	1995, 2002, 2016
Ian McKellen	3	45, 68, 77	1984, 2007, 2016
Jennifer Tilly	4	35, 40, 47, 55	1993, 1998, 2005, 2013
Kelsey Grammer	3	37, 53, 61	1992, 2008, 2016
Lily Tomlin	5	36, 45, 61, 70, 77	1975, 1984, 2000, 2009, 2016
Marisa Tomei	3	29, 45, 52	1993, 2009, 2016
Meryl Streep	4	34, 46, 54, 68	1983, 1995, 2003, 2017
Robert DeNiro	4	38, 48, 67, 74	1981, 1991, 2010, 2017
Russell Crowe	4	30, 37, 44, 51	1994, 2001, 2008, 2015
Sarah Jessica Parker	3	28, 44, 51	1993, 2009, 2016
Sela Ward	3	47, 54, 60	2003, 2010, 2016
Shirley MacLaine	3	43, 56, 66	1977, 1990, 2000
Tom Hanks	3	32, 39, 57	1988, 1995, 2013
Tom Selleck	3	58, 65, 72	2003, 2010, 2017
Warren Beatty	3	61, 71, 79	1998, 2008, 2016
Athletes			
Billie Jean King	3	30, 66, 73	1973, 2010, 2017
Caitlyn Jenner	3	28, 35, 66	1978, 1984, 2015
Carl Lewis	3	35, 49, 56	1996, 2010, 2017
Larry Bird	3	29, 36, 60	1985, 1992, 2017
Magic Johnson	4	21, 36, 41, 57	1980, 1992, 2000, 2017
Michael Jordan	3	21, 32, 50	1984, 1995, 2013
Pete Sampras	3	31, 38, 46	2002, 2009, 2017
Authors			
Hunter S. Thompson	3	40, 50, 60	1977, 1987, 1997
J. K. Rowling	3	34, 42, 50	1999, 2007, 2015
Kurt Vonnegut	4	56, 65, 74, 83	1978, 1987, 1996, 2005
Margaret Atwood	4	55, 62, 70, 77	1994, 2001, 2009, 2017

Ray Bradbury	3	48, 65, 81	1968, 1985, 2001
Stephen King	5	39, 46, 54, 62, 60	1986, 1993, 2001, 2009, 2015
Toni Morrison	3	62, 72, 81	1993, 2003, 2012
Business Leaders			
Bill Gates	3	47, 55, 61	2002, 2010, 2017
Donald Trump	6	33, 41, 48, 57, 64, 71	1979, 1987, 1994, 2003, 2010, 2017
John Chambers	3	52, 59, 67	2001, 2008, 2016
Larry Ellison	3	51, 62, 69	1995, 2006, 2013
Michael Bloomberg	3	56, 68, 75	1998, 2010, 2017
Rupert Murdoch	4	37, 58, 79, 86	1968, 1989, 2010, 2017
Steve Jobs	3	25, 35, 44	1980, 1990, 1999
Composers and Musicians			
Bob Dylan	3	24, 39, 46	1965, 1980, 1987
Charlie Watts	4	25, 35, 50, 71	1966, 1976, 1991, 2012
Glenn Gould	3	27, 34, 49	1959, 34 1966, 49 1981
John Williams	3	57, 67, 84	1989, 1999, 2016
Kirk Hammett	4	22, 29, 42, 54	1984, 1991, 2004, 2016
Leonard Bernstein	3	30, 48, 65	1948, 1966, 1983
Yo-Yo Ma	3	46, 53, 61	2001, 2008, 2017
Directors			
David Lynch	4	45, 54, 64, 71	1991, 2000, 2010, 2017
Francis Ford Coppola	5	40, 49, 58, 68, 76	1979, 1988, 1997, 2007, 2015
George Lucas	4	27, 47, 55, 72	1971, 1991, 1999, 2016
James Cameron	4	35, 43, 56, 62	1989, 1997, 2010, 2017
Martin Scorsese	5	28, 40, 54, 61, 74	1970, 1982, 1996, 2003, 2017
Quentin Tarantino	3	34, 47, 54	1997, 2010, 2017
Spike Lee	4	36, 43, 52, 59	1993, 2000, 2009, 2016
Steven Spielberg	6	29, 36, 43, 51, 62, 70	1975, 1982, 1989, 1997, 2008, 2016
Tim Burton	3	41, 51, 58	1999, 2009, 2016
Fashion Designers			
Karl Lagerfeld	3	66, 76, 83	2000, 2010, 2017
Vivienne Westwood	3	50, 69, 76	1991, 2010, 2017
Politicians			
Al Franken	3	32, 59, 66	1983, 2010, 2017
Bernie Sanders	4	47, 54, 61, 75	1988, 1995, 2002, 2016
Chris Christie	3	38, 48, 54	2000, 2010, 2017
Condoleezza Rice	3	48, 56, 62	2002, 2010, 2017
George W. Bush	3	54, 61, 69	2000, 2007, 2015
Hillary Clinton	4	32, 49, 56, 69	1979, 1996, 2003, 2017
Jerry Brown	3	37, 72, 79	1975, 2010, 2017
Jimmy Carter	4	56, 70, 83, 91	1980, 1994, 2007, 2015
John McCain	3	67, 74, 80	2003, 2010, 2017
Ron Paul	5	53, 60, 67, 75, 81	1988, 1995, 2002, 2010, 2017
Ronald Reagan	4	56, 64, 75, 80	1967, 1975, 1986, 1991
Singers			
Annie Lennox	3	30, 49, 60	1984, 2003, 2014
Barbra Streisand	5	27, 37, 44, 54, 72	1969, 1979, 1986, 1996, 2014
Bono	3	23, 46, 56	1983, 2006, 2016

Bruce Springsteen	4	29, 36, 58, 67	1978, 1985, 2007, 2017
Chuck Berry	3	61, 83, 90	1987, 2010, 2017
Dion DiMucci	3	50, 68, 77	1989, 2007, 2016
Dolly Parton	4	31, 53, 61, 70	1977, 1999, 2007, 2016
Eric Burdon	3	41, 61, 75	1982, 2002, 2016
Grace Slick	3	45, 59, 76	1984, 1998, 2015
Kate Bush	3	20, 27, 35	1978, 1985, 1993
Mick Jagger	6	22, 32, 39, 47, 59, 67	1965, 1975, 1982, 1990, 2002, 2010
Nina Simone	3	36, 51, 66	1969, 1984, 1999
Paul Simon	6	29, 36, 45, 52, 65, 75	1970, 1977, 1986, 1993, 2006, 2017
Robbie Robertson	3	45, 62, 74	1988, 2005, 2017
Sheryl Crow	3	35, 43, 55	1997, 2005, 2017
Steven Tyler	3	39, 49, 67	1987, 1997, 2015
Stevie Nicks	5	29, 37, 53, 60, 68	1977, 1985, 2001, 2008, 2016
Sting	5	33, 40, 49, 56, 65	1984, 1991, 2000, 2007, 2017
Ted Nugent	4	34, 52, 62, 68	1982, 2000, 2010, 2017
TV Anchors			
Ellen DeGeneres	3	40, 48, 59	1998, 2006, 2017
Jerry Springer	3	54, 66, 72	1998, 2010, 2016
Oprah Winfrey	3	37, 53, 63	1991, 2007, 2017

Disfluency coding

Audio files for each interview were downloaded and disfluencies were coded using Praat software (Boersma & Weenink, 2021) by coders who were unaware of the purposes of the study. The coders listened to the speech segments (averaging to about 2 minutes of interviewee's speech for each interview) and were instructed to code the following types of disfluencies. Each disfluency type was coded by two independent coders. Only disfluencies that were marked by both coders were included in the analyses.

Filled pauses were defined as *uh*'s and *um*'s produced by the interviewees. Out of a total of 3,698, 2,880 (78%) filled pauses were detected by both coders.

Repeats were defined as single words or short phrases (not more than three words) uttered more than once consecutively by the speaker (e.g., *I think that- I think that ...)*. If the words were repeated more than once, the onset of each repetition was marked as an independent repeat (e.g., *I- I- I think* would constitute two repeats). Partial-word repetitions of syllables were also included as repeats (e.g., *my mo-my mother*... would be categorized as one repeat). Repeated words with intervening filler words (e.g., *I* *think- uh, I think...*) were also categorized as repeats. Repeats always contained the same exact words, and not a reformulation of the same idea using different words (e.g., *I think- uh, believe*). Out of a total of 389, 272 (70%) repeats were detected by both coders.

Repairs were defined as single words or short phrases (not more than three words) that immediately replaced a previously uttered word or short phrase (e.g., *I think- I believe that...*), fitting within the context of the sentence. Repairs were coded only if they could still fit with the sentence if swapped with the reparandum (the words to be repaired). Repairs with intervening filler words (e.g., *the cat- uh I mean, the dog*) were also included as repairs, though the onset was set to the start of the repair and not of the filler words. Additionally, if the entire sentence was abandoned mid-sentence and a new sentence started, the onset of the new sentence was marked as an abandonment. For the purposes of the current analyses, we counted abandonments as repairs. Out of a total of 180, 90 (50%) repairs were detected by both coders. The agreement between coders was lower than expected given that both coders had been given the same instructions. We discuss this limitation and its potential implications in the Discussion.

Speech rate

The speech rate for each interview was obtained by dividing the number of words uttered by the interviewee by the length of time the interviewee spoke. First, coders unaware of the study's hypotheses transcribed the speech segments for each interview. Short intervening comments or questions from the interviewer were also transcribed but were marked as belonging to someone other than the interviewee and were removed when calculating speech rate. We additionally removed unintelligible words, non-words (e.g., laughs) and punctuation. Based on previous studies (e.g., Bortfeld et al., 2001), disfluencies were included as part of the word count.

To achieve a more accurate estimate of the duration of the interviewee's speech for each interview, coders used Praat (Boersma & Weenink, 2021) to mark the onset and offset of each intervening utterance by the interviewer or others. The duration of these intervening utterances was subtracted from the total duration of the speech segments for each interview, resulting in the amount of time containing only speech by the interviewee. Speech rate was calculated by dividing the word count by the duration (in seconds) of the interviewee’s speech for each interview, leading to a measure of words per second.

While our main measure of speech rate was words per second (consistent with past research, e.g., Horton et al., 2010), we also performed all analyses using syllables per second, which has also been used as a measure of speech rate in previous studies (e.g., Duchin & Mysak, 1987; Shewan & Henderson, 1988). The number of syllables in each word produced by the interviewee was determined using the *nsyllable* function in the R package *quanteda* (version 2.0.1, Benoit et al., 2018). The total number of syllables produced by the interviewee in each interview was then divided by the duration of the interviewee’s speech, as described above, to obtain a measure of syllables per second. Using syllables per second as a measure of speech rate led to the same results for the majority of our models as using words per second, with two exceptions, which we report in the Results. All analyses on syllable rate can be obtained through the OSF (see Author Note).

Exploratory analyses

We performed exploratory analyses on a number of variables of interest to assess whether aspects of speech other than disfluencies and speech rate varied as a function of age, and to determine how these factors interacted. Interview transcriptions were edited so that they contained only speech from the interviewee. All transcribed non-words (including *uh*’s and *um*’s) were removed, though punctuation was maintained. The resulting transcriptions were analyzed using Coh-Metrix 3.0 (Graesser et al., 2004; McNamara et al., 2014), which has been utilized extensively as an automated tool to acquire text and discourse characteristics (e.g., Kemper et al., 2010; Rabaglia & Salthouse, 2011). In particular, we extracted information about content word frequencies (acquired from the CELEX corpus), sentence length, and lexical diversity (calculated as the type/token ratio). The measure of average sentence length is an approximation because the sentence boundaries in subjects’ speech were necessarily determined by the coders who transcribed the interviews.

Additionally, to allow a more direct comparison to previous cross-sectional findings on disfluency changes in older adults, an exploratory analysis was conducted that excluded interviewees whose final interview age was under 60. This process removed 24 interviewees (26% of the original sample of 91 interviewees), leaving 248 observations in this exploratory analysis.

Statistical analyses

Our primary dependent variable was the ratio of disfluencies (filled pauses, repeats, and repairs, respectively) over the total number of words for each interview. We conducted separate analyses for each disfluency type, as well as an aggregate analysis that combined all disfluency types. To analyze the disfluency ratio data (bounded between 0 and 1), we ran fractional logistic mixed effects models using the *lme4* package (version 1.1-25, Bates et al., 2015) in R (version 4.0.3, R Core Team, 2020) and included by-subject (interviewee) variability as a random effect. All other analyses (on non-proportion data) were conducted using linear mixed effects models. The *DHARMa* package (version 0.4.6, Hartig, 2022) was used to check the distributional assumptions for the models. Results from these simulated residuals confirmed that there were no significant issues of over- or under-dispersions from the expected distribution, outliers, nor quantile deviations, indicating that the models adequately fit the data.

Results

Table 4 presents the inter-correlations between all variables of interest (primary and exploratory). While the table provides a preliminary look at the relationship between variables, we base our conclusions on the results of mixed effects models that individually assessed each relationship of interest while accounting for other fixed and random effects in each model, as described in the following sections.

Table 4
Inter-correlations between all variables of interest.

	Age	Speech Rate	Filled Pauses Ratio	Repeats Ratio	Repairs Ratio	Word Frequency	Sentence Length	Lexical Diversity
Age	1							
Speech Rate	-0.2	1						
Filled Pauses Ratio	0.05	-0.28	1					
Repeats Ratio	0.01	-0.09	0.26	1				
Repairs Ratio	0.01	0.01	0.06	0.2	1			
Word Frequency	-0.06	0.21	-0.03	0.08	0.22	1		
Sentence Length	0.02	0.14	0.13	0.06	-0.01	-0.05	1	
Lexical Diversity	0.2	-0.44	0.13	-0.02	-0.23	-0.49	-0.07	1

Speech rate and age

As prior work has shown that older age is associated with slower speech (e.g., Horton et al., 2010; Gordon et al., 2019), we ran a mixed effects model on speech rate, with interviewee age as a fixed effect and a random slope, and interviewee as a random intercept. Consistent with prior research, speech rate decreased over age, whether measured as words per second ($\beta = -0.005$, $SE = 0.002$, $p = 0.005$; see Figure 1) or syllables per second ($\beta = -0.007$, $SE = 0.002$, $p = 0.005$).

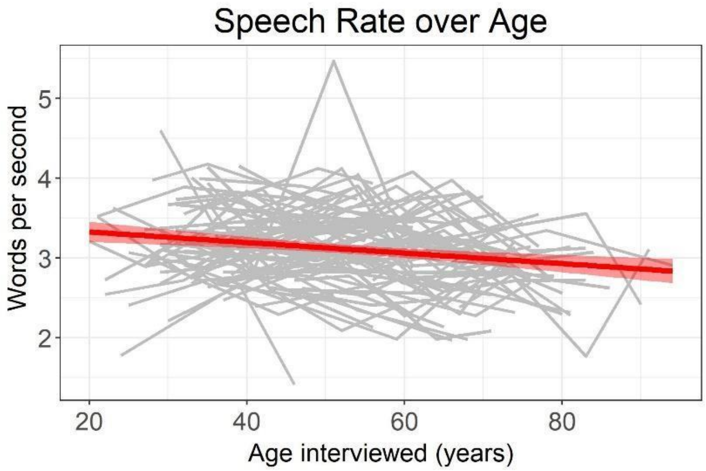

Figure 1. Effect of age on speech rate. Gray lines represent data for each individual. Red line represents best fit. Shaded red area represents Standard Error.

Disfluencies and age

A logistic mixed effects model with fixed effects of interviewee age, speech rate and their interaction, random slopes for age and speech rate, and random intercepts for interviewee was run separately on filled pauses, repeats, repairs, and the sum of all disfluency types. Because speech rate co-varied with age in our data and has been found to be related to disfluencies (Duchin & Mysak, 1987; Horton et al., 2010), it was important to control for the effect of this variable by including it as a predictor. We found that increased age resulted in significantly more repeats (see Figure 2 and Table 5 for model structure and output), while age did not significantly predict the disfluency ratio for filled pauses, repairs or all disfluencies combined. Slower speech was associated with more filled pauses but not with repeats, repairs or all disfluencies combined. When syllable rate was used rather than word rate, the models' results were the same except that the effect of age on repeats was no longer significant (p = 0.055). Additionally, an exploratory analysis (visit OSF for more details; see Author Note) revealed that when speech rate was not controlled for, the effect of age on repeats was no longer significant (p = 0.069), nor was there an effect of age on the other disfluency types. Since the effect of age on repeats was small ($\beta = 0.005$), these exploratory results suggest that whether an effect of age is found might depend on whether other relevant factors, such as speech rate, are controlled for, as well as how these factors are measured (word rate or syllable rate). This may help explain inconsistencies in previous studies, as speech rate has not always been considered.

Table 5

Model output for main models on disfluencies as a function of Age and Speech Rate

	Filled Pauses			**Repeats**			**Repairs**			**All Disfluencies**		
Predictors	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Age	-0.002	0.002	0.240	0.005	0.002	**0.029**	0.002	0.003	0.509	0.001	0.002	0.521
Speech Rate	-0.299	0.071	**<0.001**	0.071	0.096	0.456	0.059	0.102	0.558	-0.103	0.062	0.098
Age* Speech Rate	-0.006	0.004	0.108	0.001	0.005	0.768	0.004	0.006	0.444	-0.002	0.003	0.424

Note. P-values in bold are significant at the 0.05 level. In cases of non-convergence, the random effects structure was simplified by iteratively dropping the random effects associated with the smaller variance until the model converged (Barr et al., 2013). This did not lead to any changes in the significance of the fixed effects. Final model structures:

(1) Filled pause ratio ~ Age*Speech rate + (1 | Interviewee) + (0 + Age | Interviewee) + (0 + Speech rate | Interviewee)
(2) Repeats ratio ~ Age*Speech rate + (1 | Interviewee) + (0 + Age | Interviewee) + (0 + Speech rate | Interviewee)
(3) Repairs ratio~ Age*Speech rate + (1 | Interviewee) + (0 + Age | Interviewee) + (0 + Speech rate | Interviewee)
(4) All disfluencies ratio ~ Age*Speech rate + (1 | Interviewee) + (0 + Age | Interviewee) + (0 + Speech rate | Interviewee)

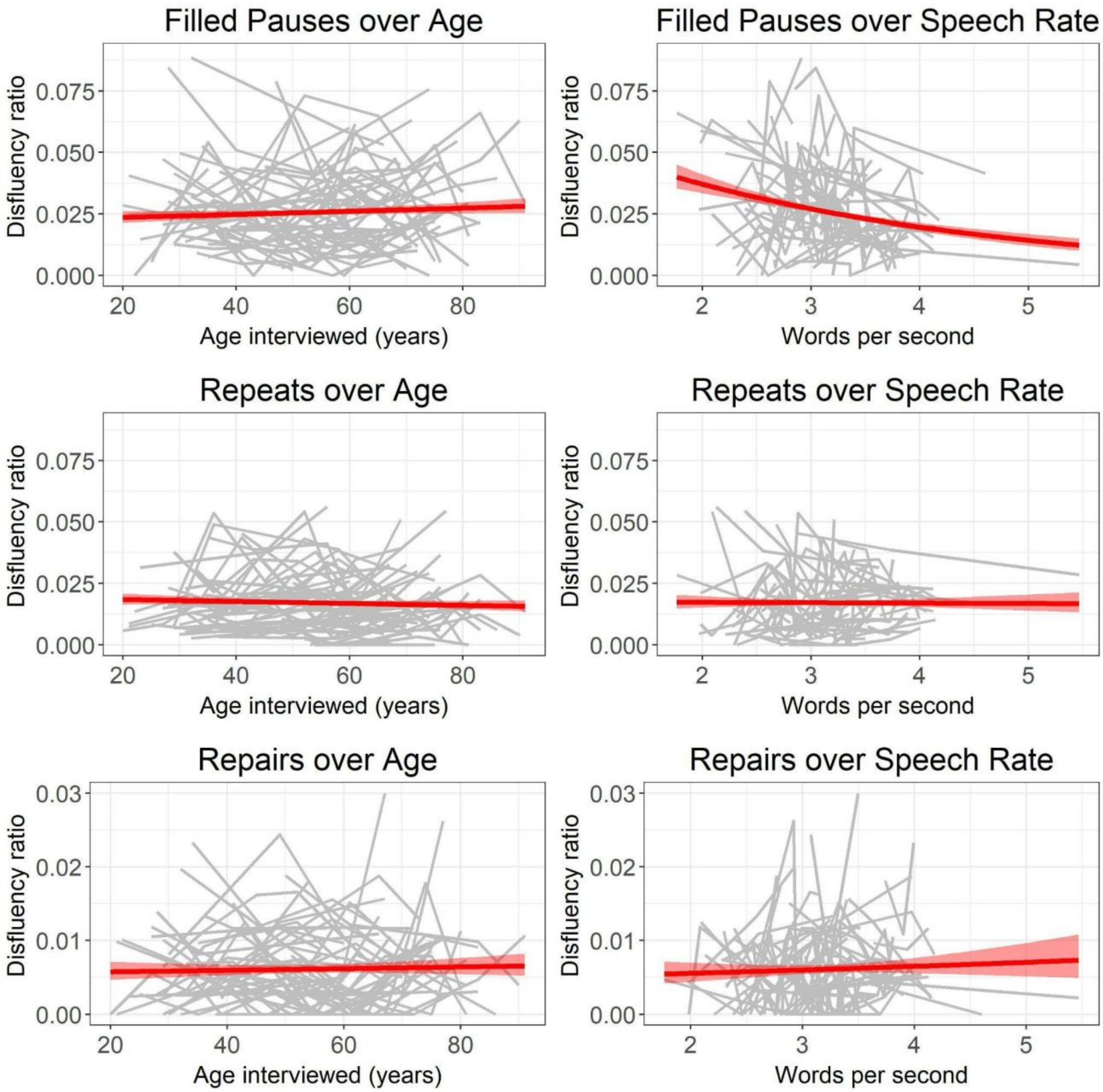

Figure 2. Effects of age and speech rate on each disfluency type. Y-axes represent the disfluency ratio separately calculated for each disfluency type (filled pauses, repeats, and repairs). Gray lines represent data for each individual. Red lines represent best fit. Shaded red areas represent Standard Error.

Exploratory analyses

We performed exploratory analyses on aspects of speech production that may reflect age-related changes in the underlying mechanisms of speech production, including content word frequencies, sentence length, and lexical diversity (more details described under Method). The *performance* package (version 0.10.2, Lüdecke et al., 2021) was used to verify that there were no issues of multicollinearity between the predictors in each model (all VIF values were below 3.60, and all tolerance values were above 0.28). Lastly, we performed an exploratory analysis on a subset of our sample comprising only older adults who were at least 60 years old at the time of their last interview.

Disfluencies and other aspects of production

Average content word frequency, sentence length, and lexical diversity were added as fixed effects in logistic mixed effects models also containing fixed effects of age, speech rate and their interaction, random slopes for age and speech rate, and random intercepts for interviewee. The model was run separately on each disfluency type and on all disfluencies combined. We found that longer sentence length was significantly associated with more filled pauses (see Figure 3 and Table 6 for model structure and output), higher content word frequency was associated with more repeats, and higher lexical diversity was associated with fewer repairs. The model on all disfluencies combined revealed that higher disfluency rates were associated with slower speech rate, higher content word frequency, and longer sentence length. However, these significant effects on all disfluencies combined are likely due to the independent effects of these factors on each type of disfluency, as revealed by our more specific analyses. These exploratory analyses suggest that changes in different types of disfluencies emerge from dissociable cognitive mechanisms, which are also reflected in other changes in speech production. This reiterates the importance of differentiating between disfluency types, and helps explain why previous cross-sectional studies on age-related changes in disfluencies have yielded inconsistent results.

Table 6

Model output for exploratory models on disfluencies as a function of other aspects of production

	Filled Pauses			Repeats			Repairs			All Disfluencies		
Predictors	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Age	-0.002	0.002	0.190	0.004	0.002	**0.049**	0.003	0.003	0.368	0.001	0.002	0.689
Speech Rate	-0.343	0.070	**<0.001**	0.002	0.093	0.985	-0.133	0.104	0.205	-0.179	0.058	**0.002**
Word Frequency	1.283	1.382	0.353	4.811	1.661	**0.004**	4.119	2.522	0.102	2.590	1.068	**0.015**
Sentence Length	0.013	0.003	**<0.001**	0.004	0.004	0.237	-0.006	0.004	0.206	0.007	0.002	**0.001**
Lexical Diversity	0.032	0.568	0.955	-0.158	0.646	0.806	-3.661	0.949	**<0.001**	-0.534	0.423	0.207
Age*Speech Rate	-0.006	0.003	0.081	0.002	0.004	0.730	0.004	0.006	0.469	-0.002	0.003	0.415

Note. P-values in bold are significant at the 0.05 level. Final model structures:

(1) Filled pause ratio ~ Age*Speech rate + Word frequency + Sentence length + Lexical diversity + (1 | Interviewee) + (0 + Age | Interviewee) + (0 + Speech rate | Interviewee)

(2) Repeats ratio ~ Age*Speech rate + Word frequency + Sentence length + Lexical diversity + (1 | Interviewee) + (0 + Age | Interviewee) + (0 + Speech rate | Interviewee)

(3) Repairs ratio ~ Age*Speech rate + Word frequency + Sentence length + Lexical diversity + (1 | Interviewee) + (0 + Age | Interviewee)

(4) All disfluencies ratio ~ Age*Speech rate + Word frequency + Sentence length + Lexical diversity + (1 | Interviewee) + (0 + Age | Interviewee) + (0 + Speech rate | Interviewee)

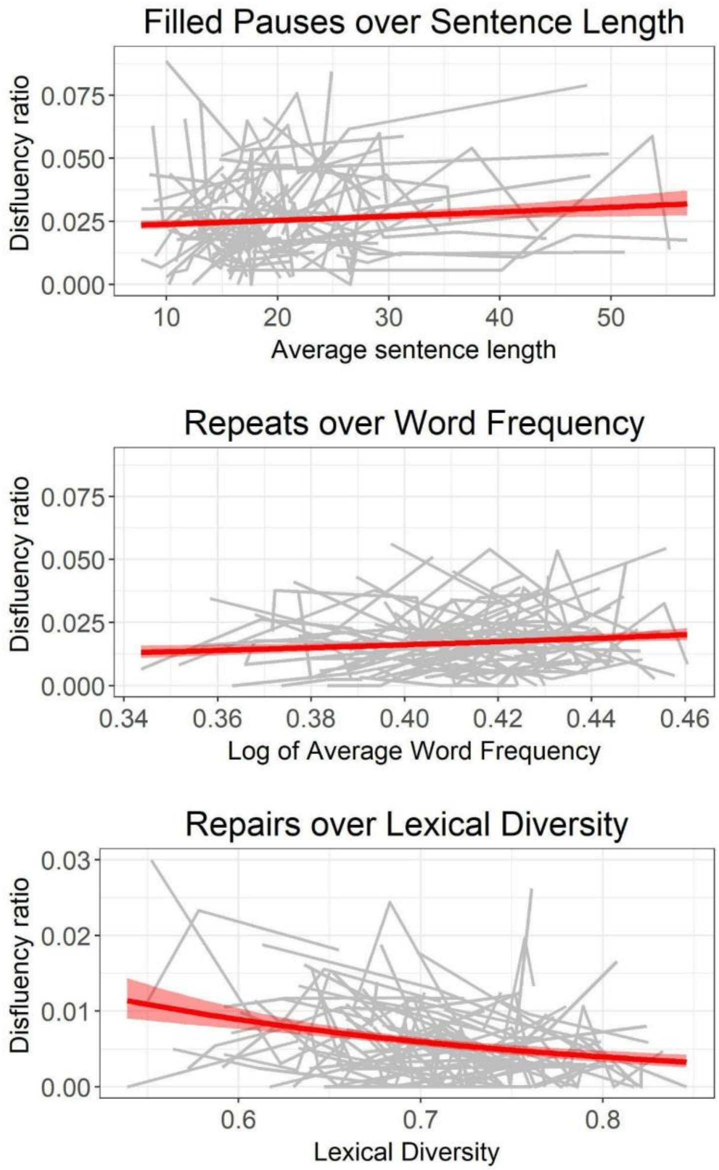

Figure 3. Effects of sentence length, word frequency and lexical diversity on disfluencies. Y-axes represent the disfluency ratio separately calculated for each disfluency type (filled pauses, repeats, and repairs). Gray lines represent data for each individual. Red lines represent best fit. Shaded red areas represent Standard Error.

Changes in sentence length, word frequency, and lexical diversity

We examined whether these speech characteristics varied with age and speech rate, irrespective of the disfluency ratio. We used mixed effects models with age and speech rate as fixed effects and random slopes, and interviewee as a random intercept separately for each speech characteristic. Slower speech rate was significantly associated with lower word frequency (see Figure 4 and Table 7 for model structure and output) and with higher lexical diversity. For sentence length, neither age nor speech rate was a significant predictor. The models using syllable rate led to the same results, with the exception that the effect of speech rate on word frequency did not reach significance.

Table 7

Model output for exploratory models on aspects of production as a function of Age and Speech Rate

	Word Frequency			**Sentence Length**			**Lexical Diversity**		
Predictors	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Age	0.000	0.000	0.681	0.003	0.032	0.936	0.000	0.000	0.267
Speech Rate	0.009	0.002	**<0.001**	1.845	1.263	0.145	-0.045	0.006	**<0.001**
Age* Speech Rate	0.000	0.000	0.910	0.046	0.065	0.471	-0.000	0.000	0.570

Note. P-values in bold are significant at the 0.05 level. Final model structures:

(1) Word frequency ~ Age*Speech rate + (1 | Interviewee) + (0 + Age | Interviewee)
(2) Sentence length ~ Age*Speech rate + (1 | Interviewee) + (0 + Speech rate | Interviewee)
(3) Lexical diversity ~ Age*Speech rate + (1 | Interviewee) + (0 + Age | Interviewee)

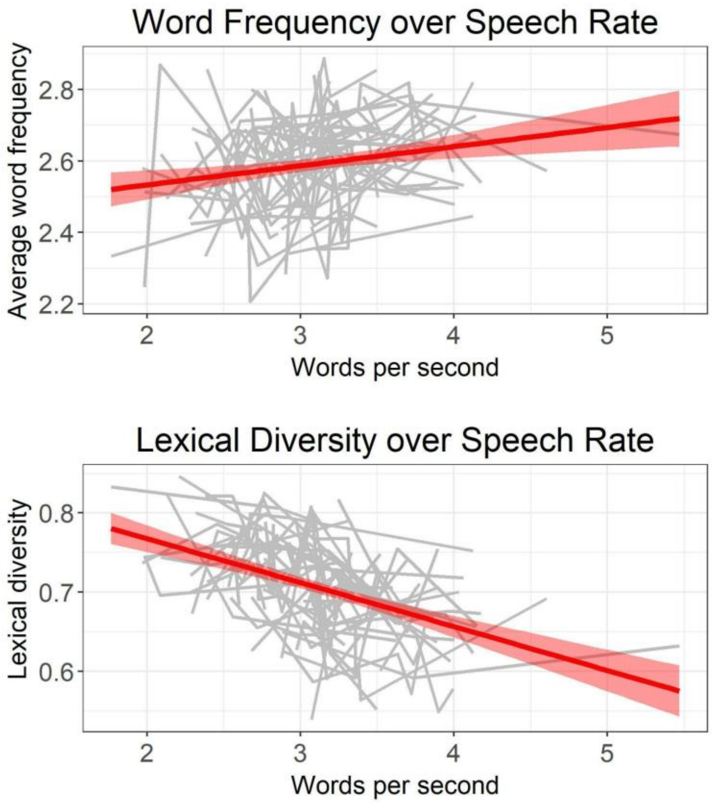

Figure 4. Effects of speech rate on word frequency and lexical diversity. Gray lines represent data for each individual. Red lines represent best fit. Shaded red areas represent Standard Error.

Analysis of interviewees over 60

All analyses reported in the Results were conducted again, on a subset of data excluding interviewees whose final interview age was under 60 (see Method). All of the results were consistent with the analyses conducted on the original sample, with three exceptions: the effect of age on repeats was no longer significant (β = 0.004, *SE* = 0.002, *p* = .077); there was a significant interaction of age and speech rate for filled pauses (β = -0.007, *SE* = 0.004, *p* = .042) in the model that included the Coh-Metrix variables; and the relationship between speech rate and word frequency did not reach significance (β = 0.006, *SE* = 0.003, *p* = .052). These findings indicate that a few of the significant effects found in the full sample have become marginal, potentially due to lower statistical power of this smaller sample. Overall, the findings on this subset of older adults support our conclusions that speech rate and other speech characteristics are better predictors of disfluencies than age, with little evidence for language decline.

Discussion

In this study, we employed a novel approach to the study of age-related changes in speech fluency by capitalizing on publicly available interviews to measure these changes longitudinally through a sequential design, rather than cross-sectionally. Our primary goal was to determine whether disfluencies become more common with increasing age and to resolve inconsistencies found in previous cross-sectional designs, some of which reported an increase in speech disfluencies with age, especially filled pauses (James et al., 2018; Manning & Monte, 1981; Spieler & Griffin, 2007; Kemper, 1992), while others found that some aspects of fluency remain stable over the lifespan (Metz & James, 2019; Schmitter-Edgecombe et al., 2000; Cooper, 1990; Duchin & Mysak, 1987; Shewan & Henderson, 1988). We focused on four common measures of fluency: overall speech rate, and the presence of disfluencies including filled pauses, repeated words and phrases, and repairs. We found that older age was associated with slower speech rate, replicating previous findings (Castro & James, 2014; Horton et al., 2010; Gordon et al., 2019). We also observed a tendency for people to produce more repeats as they aged, consistent with some previous research (Bortfeld et al., 2001; Dennis & Hess, 2016; Castro & James, 2014), and slower speech rate was associated with a greater number of filled pauses. However, neither filled pauses nor repairs were predicted by age, contrary to previous studies which have found an age-related increase in filled pauses (Bortfeld et al., 2001; Horton et al., 2010; Metz & James, 2019) and repairs (Dennis & Hess, 2016; Bortfeld et al., 2001; Schmitter-Edgecombe et al., 2000; Spieler & Griffin, 2007). There was also no effect of age on the combined count of all three types of disfluencies.

These findings suggest that determining how age impacts fluency relies heavily on the particular measures of fluency that are used. When considering fluency in terms of speech rate, most studies, including ours, tend to converge on the conclusion that older adults produce slower speech (Duchin & Mysak, 1987; Horton et al., 2010; Castro & James, 2014). However, when focusing on individual types of disfluencies, the effects are more nuanced. In our sample, only repeats increased with age, but this effect was small, and it depended on the inclusion of speech rate as a covariate, as well as whether the sample was restricted to older adults only. Our findings help explain the inconsistencies found in this literature, as previous studies did not always control for other potential predictors such as speech rate. Note that, in our sample, there was no significant relationship between age and aggregated disfluencies. The fact that the effect of age is different when considering different types of disfluencies suggests that rather than asking whether older adults become overall more disfluent, it would be more productive to think of each type of disfluency as potentially arising from separate age-related cognitive changes and then to investigate each distinct type. This view is supported by our exploratory analyses, which found that longer utterances corresponded to more filled pauses, greater lexical diversity corresponded to fewer repairs, and lower word frequency corresponded to fewer repeats. In turn, slower speech corresponded to higher lexical diversity and lower word frequency.

Overall, our results show that age is most strongly associated with decreased speech rate, which in turn is related to disfluencies and indicators of lexical and syntactic complexity. This suggests that the slower speech rate associated with age may reflect the retrieval of less frequent words drawn from a vocabulary characterized by high lexical diversity, and these characteristics are themselves associated with the use of repeats and repairs. On the other hand, slower speech is also related to more filled pauses, which are in turn associated with the production of longer sentences. Although age does not appear to predict these factors in our sample, the direction of these relationships is consistent with the effects that have been reported in a previous cross-sectional corpus study of age-related changes in speech production (Horton et al., 2010). Similarly to our findings, Horton et al. (2010) found that older age was linked to slower speech rate, greater lexical diversity, longer sentence length, and more filled pauses. Additionally, they found a positive relationship between sentence length and filled pauses. Future studies should further explore the possibility that speech rate mediates the effect of age on disfluencies and other speech characteristics.

Why is there little evidence of language decline with age?

An important question regarding our pattern of results is why we did not observe an effect of age on disfluencies and speech characteristics other than speech rate and repeats, and why we found little evidence of age-related decline in speech production, contrary to many previous studies (e.g., Bortfeld et al., 2001; James et al., 2018; Metz & James, 2019; Dennis & Hess, 2016; Castro & James, 2014). One reason may be that individuals in our corpus were relatively high-functioning middle-aged and older adults even in the latter interviews at which they were assessed, and they may also have had disproportionately more linguistic experience than a typical individual from the population due to their occupations (e.g., actors, politicians) and their experience with public speaking and interviews. However, the fact that our results closely match the effects reported by another corpus study that used a more typical sample (Horton et al., 2010) suggests that our results do generalize, despite the selective nature of our corpus.

In past studies, effects of age have emerged by aggregating data from multiple individuals across different cohorts. However, the effect of age on speech production may be highly variable from person to person. For example, a previous longitudinal study found that age-related decline in grammatical complexity and semantic content was predicted by individual differences in initial working memory capacity and vocabulary size, respectively (Kemper et. al., 2001). Similarly, in our corpus, some individuals showed drastic decreases in speech fluency with age, while others showed no decline in speech fluency at all. These individual differences may be due to a range of environmental and genetic factors, affecting how the tradeoff between increased language experience and slower processing speed is expressed in each individual. Therefore, age itself is not as good a predictor of speech fluency as other changes that tend to correlate with age. Some of these changes can be detected in speech, such as cognitive slowing (reflected in slower speech rate) and greater vocabulary (reflected in the use of less frequent words and greater lexical diversity). But the degree to which individuals display these age-related changes is highly variable. Some individuals may experience greater cognitive slowing and thus may produce slower speech than other individuals. Those older adults who slow down their speech to compensate for a decrease in speed of processing may then produce more filled pauses. Slower speech rate may itself result from greater difficulty in retrieving words due to larger vocabulary and knowledge accumulated across the lifespan (Ramscar et al., 2014). On the other hand, other individuals do not slow down their speech rate with age, and show little change in speech characteristics and fluency. Future studies should further explore the role of individual differences in age-related changes in language and cognition more broadly. Ideally, a longitudinal study that also includes tests of cognitive function at different ages (e.g., working memory) would be well-suited to address this question, although the practical challenges are considerable. Another avenue of exploration is to assess vocabulary size and language experience more directly (e.g., through the author recognition task; Stanovich & West, 1989; Mar et al., 2006), which would permit an explicit assessment of the idea that at least some types of disfluencies arise from the tradeoff between word retrieval and greater knowledge rather than from cognitive decline (Ramscar et al., 2014).

Implications on the cognitive mechanisms that support fluency

Our study contributes to the broader question of how the speech production system changes over the lifespan, and how it is affected by other changes in cognitive function. The most significant finding is that, as people age, they produce more repeated words and phrases. Based on the framework of repeats proposed by Clark and Wasow (1988), we can speculate that older adults are motivated to avoid filled and unfilled pauses while still accommodating the moment-to-moment availability of concepts as the processes that underlie speaking unfold. In contrast, we did not observe an increased tendency to produce more repairs, contrary to what has been reported previously (e.g., Bortfeld et al., 2001; Dennis & Hess, 2016; Engelhardt et al., 2010, 2012). This pattern might imply that changes in inhibition associated with increasing age (Healey et al., 2013; Zacks et al., 1996; Hartman & Hasher, 1991; Daigneault et al., 1992) are not linked to language processing mechanisms, highlighting the possibility of dissociable inhibitory systems.

Our results can also be interpreted in light of the Transmission Deficit Hypothesis (Burke & Shafto, 2008; Burke, Mackay & James, 2000; Mackay & Burke, 1990), which would attribute disfluencies to age-related deficits in the speed and efficacy of information transmission, rather than an increase in vocabulary size and world knowledge (Ramscar et al., 2014). This hypothesis predicts that the ability to retrieve lexical and phonological information declines with age because the connections between representations in the lexical-semantic network have weakened. Results from our exploratory analyses are consistent with this view, under the assumption that some types of disfluencies (e.g., repeats) potentially reflect strategies to gain time for word retrieval (Clark & Wasow, 1998). To compensate for age-related deficits in information transmission, older adults may slow down their speech to gain more time to retrieve low-frequency words from a diverse set of vocabulary items. Similarly, the robust effect of slower speech rate across age may be explained by age-related slowing in the speed of transmission between representations during speech planning.

Limitations and future directions

The most important limitation of this study's method is that the selection of publicly available interviews constrained us to include relatively famous individuals who had been interviewed at several points in their lives. As such, all middle-aged and older adults in our corpus were successful, high-functioning individuals who may have had greater experience with language and with the interview format than the average population. Notably, this limitation is shared with other longitudinal corpora that employed much smaller sample sizes (e.g., Harrington et al., 2000; Reubold et al., 2010). We addressed the increased language experience due to occupation-specific training (e.g., for politicians or actors) by selecting individuals from a range of occupations. Our corpus was also limited by the necessity to include only individuals who were high-functioning enough to conduct an interview at an older age, which therefore excluded anyone who was not sufficiently fluent or healthy. However, this issue of "survivor bias"—being limited to studying older adults who "survived" and are relatively high-functioning—is inherent in all aging research, which is why most studies in this field are constrained to investigations of healthy aging. As our pattern of results closely corresponded to another cross-sectional study of age-related changes in speech production (Horton et al., 2010), our conclusions may generalize to a less selective sample.

Another limitation is that the agreement between the disfluency coders was lower than expected. All coders were given the same set of instructions, where the definition of each type of disfluency was provided in as much detail as possible and with several examples. However, this study reveals the high degree of subjectivity in identifying disfluencies even under the same definition. This was particularly true for repairs, being defined as a word that “could still fit with the sentence if swapped with the reparandum”, which involves a linguistic judgment on the part of the coder. Moreover, previous work on naive participants’ ability to detect disfluencies has revealed that people have difficulty recognizing and reporting disfluencies present in spontaneous speech (Lickley & Bard, 1996). Importantly, this low reliability may be one of the reasons for the inconsistent findings that have been reported in the literature on disfluencies. Our findings show that it is essential to employ at least two coders so that analyses can be conducted on disfluencies that were independently detected by both.

Other limitations of our approach relate to the difficulty in controlling features of our data, such as how many interviews were available for each individual and at what exact age they were conducted, the setting of the interview and topics of conversation, and the presence of other interviewees or background music. Previous studies showed that the effect of age on fluency and discourse characteristics can vary as a function of the topic and setting in which speech is produced (e.g., whether the content is familiar or unfamiliar, produced spontaneously or rehearsed; Castro & James, 2014; James et al., 2018; Kemper et al., 2003; Metz & James, 2019). An advantage of our dataset is that all speech samples originated from the somewhat similar setting of live interviews; however, the nature of our corpus meant that we could not constrain the topics that were discussed. Thus, further analyses of our publicly available corpus data may be conducted to explore whether these factors moderate the relationship between age and disfluencies.

Lastly, the occasional presence of multiple talkers, background music, or other sounds prevented us from analyzing other features of speech that may be cues of deteriorated fluency, such as voice quality and silent pauses (Ryan & Burk, 1974). While these were unavoidable limitations due to the limited availability of publicly available interviews, future longitudinal studies may be able to circumvent these issues by recording speech directly from a sample of individuals as they age. This would also allow researchers to directly measure aspects of cognitive function in order to assess how they relate to speech production.

Given the nature of our sample, we took all possible precautions to ensure that our data were sampled and analyzed without bias. The individuals in our corpus and their interviews were all selected, transcribed, and coded for disfluencies by research assistants who were unaware of the study's hypotheses. Thus, although we had little control over which individuals would meet our criteria for inclusion, we ensured that the sampling and coding of all our data were not influenced by our a priori hypotheses.

Conclusions

We have presented results from a corpus study of conversational speech to quantify age-related changes in speech fluency sequentially rather than cross-sectionally. We found that older adults spoke more slowly, consistent with previous reports (Horton et al., 2010; Castro & James, 2014; Gordon et al., 2019), and we observed that they produced more repeated words. At the same time, older adults did not produce more repairs or filled pauses, nor was there an effect of age on all three types of disfluencies combined. Our results show that whether a relationship between age and disfluency is observed depends on the type of disfluency measured, which in turn helps explain previous inconsistent findings. Overall, we suggest that while age is not a strong predictor of fluency measures other than speech rate, there are large individual differences in how other speech characteristics change with age, even in relatively high-functioning older adults, reflecting the tradeoff between slower processing speed (Salthouse, 1996; Salthouse & Meinz, 1995) and accumulated vocabulary and language experience (Ramscar et al., 2014). Thus, some individuals slow down their speech as they age—a change associated with higher lexical diversity and the use of less frequent words, but also more filled pauses—while others do not. Our findings challenge the prevalent assumption that older age leads to more disfluent speech by showing that other changes in speech production (i.e., overall speech rate, word frequency, lexical diversity, and sentence length) are better predictors of disfluencies than age.

Footnotes

1. Throughout this paper, we use the term “speech” to refer to spoken language production and the underlying cognitive processes that include various levels of processing, such as conceptual planning, lexical and phonological retrieval, and articulation.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278.

Bates D., Mächler M., Bolker B., Walker S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. doi: 10.18637/jss.v067.i01.

Beier, E., Chantavarin, S., & Ferreira, F. (2023, January 30). Do disfluencies increase with age? Evidence from a sequential corpus study of disfluencies. https://doi.org/10.17605/OSF.IO/EW26P

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018) "quanteda: An R package for the quantitative analysis of textual data". *Journal of Open Source Software, 3*(30), 774. https://doi.org/10.21105/joss.00774.

Berisha, V., Wang, S., LaCross, A., & Liss, J. (2015). Tracking discourse complexity preceding Alzheimer's disease diagnosis: a case study comparing the press conferences of presidents Ronald Reagan and George Herbert Walker Bush. *Journal of Alzheimer's Disease, 45(*3), 959-963.

Boersma, P., & Weenink, D. (2021). *Praat: Doing phonetics by computer* (6.1.40). http://www.praat.org/

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech, 44*(2), 123-147.

Burke, D. M., MacKay, D. G., & James, L. E. (2000). Theoretical approaches to language and aging. In T. Perfect & E. Maylor (Eds.), *Models of cognitive* aging (pp. 204-237). Oxford University Press.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults?. *Journal of Memory and Language, 30*(5), 542-579.

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (p. 373–443). Psychology Press.

Castro, N., & James, L. E. (2014). Differences between young and older adults' spoken language production in descriptions of negative versus neutral pictures. *Aging, Neuropsychology, and Cognition, 21*(2), 222–238.

Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37*(3), 201-242.

Connor, L. T., Spiro, A., Obler, L. K., & Albert, M. L. (2004). Change in Object Naming Ability During Adulthood. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 59*(5), 203–P209.

Cooper, P. V. (1990). Discourse production and normal aging: Performance on oral picture description tasks. *Journal of Gerontology, 45*(5), 210-P214.

Daigneault, S., Braun, C. M., & Whitaker, H. A. (1992). Early effects of normal aging on perseverative and non-perseverative prefrontal measures. *Developmental Neuropsychology, 8*(1), 99-114.

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., ... & Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin, 92*(1), 135-152.

Dennis, P. A., & Hess, T. M. (2016). Aging-related gains and losses associated with word production in connected speech. *Aging, Neuropsychology, and Cognition, 23*(6), 638–650.

Dodge, H. H., Zhu, J., Hughes, T. F., Snitz, B. E., Chang, C. C. H., Jacobsen, E. P., & Ganguli, M. (2017). Cohort effects in verbal memory function and practice effects: a population-based study. *International Psychogeriatrics, 29*(1), 137-148.

Dodge, H. H., Zhu, J., Lee, C. W., Chang, C. C. H., & Ganguli, M. (2014). Cohort effects in age-associated cognitive trajectories. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 69*(6), 687-694.

Duchin, S. W., & Mysak, E. D. (1987). Disfluency and rate characteristics of young adult, middle-aged, and older males. *Journal of Communication Disorders, 20(*3), 245-257.

Engelhardt, P. E., Corley, M., Nigg, J. T., & Ferreira, F. (2010). The role of inhibition in the production of disfluencies. *Memory & Cognition, 38(*5), 617-628.

Engelhardt, P. E., Veld, S. N., Nigg, J. T., & Ferreira, F. (2012). Are language production problems apparent in adults who no longer meet diagnostic criteria for attention-deficit/hyperactivity disorder? *Cognitive Neuropsychology, 29*(3), 275-299.

Fraundorf, S. H., & Watson, D. G. (2014). Alice's adventures in um -derland: Psycholinguistic sources of variation in disfluency production. *Language, Cognition and Neuroscience, 29(*9), 1083–1096.

Gahl, S., & Baayen, R. H. (2019). Twenty-eight years of vowels: Tracking phonetic variation through young to middle age adulthood. *Journal of Phonetics, 74,* 42-54.

Gahl, S., Cibelli, E., Hall, K., & Sprouse, R. (2014). The "Up" corpus: a corpus of speech samples across adulthood. *Corpus Linguistics and Linguistic Theory, 10*(2), 315-328.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36(*2), 193-202.

Gordon, J. K., Andersen, K., Perez, G., & Finnegan, E. (2019). How old do you think I am? Speech-language predictors of perceived age and communicative competence. *Journal of Speech, Language, and Hearing Research, 62*(7), 2455-2472.

Harrington, J., Palethorpe, S., & Watson, C. (2000). Monophthongal vowel changes in Received Pronunciation: an acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association, 30*, 63-78.

Hartig, F. (2022). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.6. http://florianhartig.github.io/DHARMa/

Hartman, M., & Hasher, L. (1991). Aging and suppression: memory for previously relevant information. *Psychology and Aging, 6*(4), 587.

Hartsuiker, R. J., & Notebaert, L. (2010). Lexical Access Problems Lead to Disfluencies in Speech. *Experimental Psychology, 57(*3), 169–177.

Healey, M. K., Hasher, L., & Campbell, K. L. (2013). The role of suppression in resolving interference: evidence for an age-related deficit. *Psychology and Aging, 28*(3), 721.

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.

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1378.

James, L. E., Chambers, B. N., & Placzek, C. L. (2018). How scenes containing visual errors affect speech fluency in young and older adults. *Aging, Neuropsychology, and Cognition*, *25*(4), 520–534.

Kavé, G., & Goral, M. (2017). Do age-related word retrieval difficulties appear (or disappear) in connected speech? *Aging, Neuropsychology, and Cognition, 24*(5), 508–527.

Kavé, G., Knafo, A., & Gilboa, A. (2010). The rise and fall of word retrieval across the lifespan. *Psychology and Aging*, *25*(3), 719–724.

Kemper, S. (1992). Adults' sentence fragments: Who, what, when, where, and why. *Communication Research, 19*(4), 444-458.

Kemper, S., Herman, R. E., & Lian, C. H. (2003). The costs of doing two things at once for young and older adults: Talking while walking, finger tapping, and ignoring speech of noise. *Psychology and Aging*, *18*(2), 181.

Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging, 16*(4), 600-614.

Kemper, S., Schmalzried, R., Hoffman, L., & Herman, R. (2010). Aging and the vulnerability of speech to dual task demands. *Psychology and Aging, 25*(4), 949.

Le Dorze, G., & Bédard, C. (1998). Effects of age and education on the lexico-semantic content of connected speech in adults. *Journal of Communication Disorders, 31(*1), 53-71.

Lickley, R. J., & Bard, E. G. (1996). On not recognizing disfluencies in dialogue. In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP'96*, 3, 1876-1879.

Lu, N., Han, Y., Chen, T., Gunzler, D. D., Xia, Y., Lin, J. Y., & Tu, X. M. (2013). Power analysis for cross-sectional and longitudinal study designs. *Shanghai Archives of Psychiatry, 25*(4), 259–262. https://doi.org/10.3969/j.issn.1002-0829.2013.04.009

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). "performance: An R Package for Assessment, Comparison and Testing of Statistical Models." *Journal of Open Source Software, 6*(60), 3139. doi:10.21105/joss.03139.

MacKay, D. G., & Burke, D. M. (1990). Cognition and Aging: A Theory of New Learning and the Use of Old Connections. In *Advances in Psychology* (Vol. 71, pp. 213–263). Elsevier.

Manning, W. H., & Monte, K. L. (1981). Fluency breaks in older speakers: Implications for a model of stuttering throughout the life cycle. *Journal of Fluency Disorders, 6*(1), 35-48.

Mar, R. A., Oatley, K., Hirsh, J., Dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality, 40*(5), 694-712.

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix.* Cambridge University Press.

Metz, M. J., & James, L. E. (2019). Specific effects of the Trier Social Stress Test on speech fluency in young and older adults. *Aging, Neuropsychology, and Cognition, 26*(4), 558–576.

Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2006). Age-related effects on speech production: A review. *Language and Cognitive Processes, 21*(1-3), 238–290.

Nicholas, M., Obler, L., Albert, M., & Goodglass, H. (1985). Lexical retrieval in healthy aging. *Cortex, 21*(4), 595-606.

R Core Team (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rabaglia, C. D., & Salthouse, T. A. (2011). Natural and constrained language production as a function of age and cognitive abilities. *Language and Cognitive Processes*, *26(*10), 1505-1531.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, *6*(1), 5-42.

Rastle, K. G., & Burke, D. M. (1996). Priming the tip of the tongue: Effects of prior processing on word retrieval in young and older adults. *Journal of Memory and Language, 35*(4), 586-605.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging, 21*(3), 448.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication*, *52*, 638-651.

Rogalski, Y., Peelle, J., & Reilly, J. (2011). Effects of perceptual and contextual enrichment on visual confrontation naming in adult aging. *Journal of Speech, Language, and Hearing Research, 54*, 1349-1360.

Ryan, W. J., & Burk, K. W. (1974). Perceptual and acoustic correlates of aging in the speech of males. *Journal of Communication Disorders, 7*(2), 181-192.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403.

Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory, and speed. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 50*(6), 297-306.

Schmitter-Edgecombe, M., Vesneski, M., & Jones, D. W. R. (2000). Aging and word-finding: A comparison of spontaneous and constrained naming tests. *Archives of Clinical Neuropsychology*, *15*(6), 479-493.

Shafto, M. A., James, L. E., Abrams, L., Tyler, L. K., & Cam-CAN. (2017). Age-related increases in verbal knowledge are not associated with word finding problems in the Cam-CAN cohort: What you know won’t hurt you. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *72(*1), 100–106.

Shewan, C. M., & Henderson, V. L. (1988). Analysis of spontaneous language in the older normal population. *Journal of Communication Disorders, 21*(2), 139-154.

Spieler, D. H., & Griffin, Z. M. (2006). The influence of age on the time course of word preparation in multiword utterances. *Language and Cognitive Processes, 21*, 291-321.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 402-433.

Stine-Morrow, E. A. L., Miller, L. M. S., & Hertzog, C. (2006). Aging and self-regulated language processing. *Psychological Bulletin, 132*(4), 582–606. https://doi.org/10.1037/0033-2909.132.4.582

Warren, D. E., Rubin, R., Shune, S., & Duff, M. C. (2018). Memory and language in aging: How their shared cognitive processes, neural correlates, and supporting mechanisms change with age. In M. Rizzo, S. Anderson, & B. Fritzsch (Eds.), *The Wiley handbook on the aging mind and brain* (p. 270–295). Wiley Blackwell.

Whitbourne, S. K. (2019). Longitudinal, cross-sectional, and sequential designs in lifespan developmental psychology. *Oxford Research Encyclopedia of Psychology.* https://doi.org/10.1093/acrefore/9780190236557.013.344

Xu, Z., Shen, X., & Pan, W. (2014). Longitudinal analysis is more powerful than cross-sectional analysis in detecting genetic association with neuroimaging phenotypes. *PloS One, 9*(8), e102312. https://doi.org/10.1371/journal.pone.0102312

Zacks, R. T., Radvansky, G., & Hasher, L. (1996). Studies of directed forgetting in older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 143.

Zec, R. F., Markwell, S. J., Burkett, N. R., & Larsen, D. L. (2005). A longitudinal study of confrontation naming in the "normal" elderly. *Journal of the International Neuropsychological Society, 11*, 716-726.