



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Giffard, J;Li, R;Roccati, E;Vogel, AP;Bindoff, AD;Goldberg, LR;Bai, Q;Alty, J

Title:

Rapid repetitive syllable sounds associate with episodic memory, executive function, and working memory in cognitively healthy and subjectively impaired older adults

Date:

2025-12-01

Citation:

Giffard, J., Li, R., Roccati, E., Vogel, A. P., Bindoff, A. D., Goldberg, L. R., Bai, Q. & Alty, J. (2025). Rapid repetitive syllable sounds associate with episodic memory, executive function, and working memory in cognitively healthy and subjectively impaired older adults. *Geroscience*, 47 (6), pp.6683-6698. <https://doi.org/10.1007/s11357-025-01739-x>.

Persistent Link:

<https://hdl.handle.net/11343/362201>

License:

CC BY



Rapid repetitive syllable sounds associate with episodic memory, executive function, and working memory in cognitively healthy and subjectively impaired older adults

Julia Giffard · Renjie Li · Eddy Roccati · Adam P. Vogel · Aidan D. Bindoff · Lynette R. Goldberg · Quan Bai · Jane Alty

Received: 18 March 2025 / Accepted: 4 June 2025 / Published online: 24 June 2025
© The Author(s) 2025

Abstract Non-invasive screening tools for Alzheimer’s disease (AD) risk are needed. Decline in episodic memory and subjective cognitive impairment (SCI) are both associated with elevated AD risk. We investigated associations between three cognitive domains (episodic memory, executive function, and working memory) and motor speech performance in older adults with healthy cognition (HC) or SCI. In total, 1014 community-dwelling participants (cross-sectional sample: 843 HC, mean 66.9 years, 72.8% female; 171 SCI, mean 66.3 years, 70.8% female) remotely completed cognitive tests (Paired Associates Learning and Spatial Working Memory from the

Cambridge Neuropsychological Test Automated Battery) and motor speech tests (Tasmanian [TAS] Test). We evaluated whether acoustic motor speech features extracted from 10-s tests of “pa,” “ta,” or “ka” syllable repetition improved episodic memory, executive function, and working memory regression models over age, sex, education, anxiety, and depression ($\Delta AIC > 2$). Motor speech features improved estimation of cognitive scores over fixed demographic and clinical variables in each group and cognitive domain. In summary, we found associations between motor speech performance and cognition in older adults. Short motor speech tests demonstrate potential for non-invasive, cost-effective AD risk screening. Longitudinal research is needed to investigate relationships between individual motor speech features and

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11357-025-01739-x>.

J. Giffard · R. Li · E. Roccati · A. D. Bindoff · L. R. Goldberg · J. Alty (✉)
Wicking Dementia Research and Education Centre,
College of Health and Medicine, Medical Science
Precinct, Medical Sciences 1, University of Tasmania,
Level 4, 17 Liverpool St, Hobart, TAS 7000, Australia
e-mail: jane.alty@utas.edu.au

A. P. Vogel
Audiology and Speech Pathology, The University
of Melbourne, 550 Swanston St, Parkville, VIC 3010,
Australia

A. P. Vogel
Redenlab Inc. Australia, 585 Little Collins St, Suite 669,
Melbourne, VIC 3000, Australia

A. P. Vogel · J. Alty
Tasmanian School of Medicine, Medical Science 1,
University of Tasmania, Level 1, 17 Liverpool St, Hobart,
TAS 7000, Australia

Q. Bai
School of Information and Communication Technology,
Centenary Building, University of Tasmania, Grosvenor
St, Sandy Bay, TAS 7005, Australia

J. Alty
Neurology Department, Royal Hobart Hospital, 48
Liverpool St, Hobart, TAS 7000, Australia

cognitive changes confirmed to result from the progression of AD.

Keywords Alzheimer’s disease · Subjective cognitive impairment · Diadochokinesis · Executive function · Episodic memory · Working memory

Introduction

Alzheimer’s disease (AD) is the most common cause of dementia, accounting for up to 70% of cases [1]. AD is a complex neurodegenerative condition characterized by the accumulation of amyloid beta plaques and tau proteins in the brain followed by cognitive decline [2]. AD symptomatology progresses over 15–25 years, from healthy cognition (HC) with underlying covert pathology, through mild cognitive impairment (MCI), which may be preceded by subjective cognitive impairment (SCI), to dementia [3]. Fourteen modifiable risk factors, such as hypertension and smoking, have been determined to account for up to 45% of dementia cases [4]. Early intervention options are increasing with the development of new disease-modifying therapies [5, 6].

Accessible, cost-effective, and non-invasive methods are needed to detect AD during its preclinical HC and SCI stages to facilitate early risk modification and treatment [2, 7, 8]. Current early detection techniques and assessments include blood- and CSF-based biomarkers, neuropsychological assessments, and PET brain imaging. These methods, however, are not scalable to the population level due to dependencies including specialist equipment, skilled administration, and trained interpretation. Cognitive assessments such as the Montreal Cognitive Assessment (MoCA) may be more scalable but lack sensitivity to earlier preclinical and SCI stages, which are characterized by a lack of objective cognitive impairment.

Motor function tests show promise for early detection of AD [9]. Gait deterioration begins in the preclinical stage about 10 years before dementia diagnosis, leading to slower gait and greater stride variability in people with dementia compared to healthy older adults [10, 11]. Interest is growing in other motor tests such as finger tapping (FT) which is an established method of assessing upper limb motor function but has been used in few studies of cognitive decline and discrimination of preclinical AD, MCI, and dementia

[12, 13]. Since AD pathology typically begins in the hippocampus, a known flag of elevated risk of covert AD is subtle decline in episodic memory [14–17]. FT shows links to cognitive performance, including the domain of episodic memory [13, 18].

Motor function is also a crucial component of speech production. Difficulties with sequencing and repetition are known to occur in people with dementia [19–21]. Diadochokinetic (DDK) motor speech tests involve participants rapidly repeating syllables such as “pa”, “ta”, or “ka” individually (as in “papapa”), or sequentially (as in “pataka”) [22]. In younger adults, performance on DDK tests is associated with executive function, specifically cognitive switching [23, 24]. Higher cognitive switching and updating of working memory are associated with higher accuracy on sequential DDK tests, and higher cognitive switching is associated with slower rates on monosyllabic and sequential DDK tests [24].

DDK tests may complement FT- and gait-based assessments of cognition by adding a focus on the functional daily task of speaking. DDK tests do not require specialist equipment such as gait mats, and they remain accessible for people with reduced mobility, ambulation, and balance. We drew upon the findings in DDK and cognition in younger adults [23, 24], and the emerging evidence in older adults of FT tests aiding the estimation of episodic memory, executive function, and working memory [13, 18, 25], to investigate whether fast, monosyllabic DDK tests elicited similar associations in older adults.

Focusing first on episodic memory due to its sensitivity to AD [14–17], we hypothesized that performance features derived from three monosyllabic DDK tests (“pa”, “ta”, and “ka”) would improve estimation of episodic memory scores in a cognitively healthy group of older adults over null models accounting for readily obtainable demographic and clinical details (age, sex, education, anxiety, and depression). Second, we hypothesized that features from each DDK test would, likewise, improve model fit for executive function and working memory scores.

Since SCI itself is a risk factor for dementia (an estimated 2.3% of affected older adults progress to dementia annually; double the rate of those without SCI) [26], we undertook a secondary analysis in participants with self-reported SCI to examine whether the same set of motor speech features improved estimation of episodic memory, executive function, and working memory scores.

Methods

Participant recruitment and eligibility

A cross-sectional community sample of older adults was recruited from the Island Study Linking Aging and Neurodegenerative Disease (ISLAND), for which a full protocol has been previously published [27]. ISLAND was launched in Tasmania, Australia, in 2019 as a 10-year public health initiative to educate residents aged 50 years and over on reducing their modifiable dementia risk factors [27].

ISLAND participants were invited to complete the TAS Test protocol, an online battery of speech and motor-cognitive tests (full protocol previously published) [2]. In brief, TAS Test comprises five sections, each focusing on different tests and abilities: (1) video hand movement and FT tests, (2) keyboard tapping tests, (3) visuomotor tests for visuoperceptual deficits and reaction times, (4) visuospatial tests of ability and working memory, and (5) tests of motor speech (DDK) and language abilities (verbal picture description).

The participants who elected to be involved self-administered TAS Test remotely and without researcher supervision at their convenience between October and December 2022 by logging into the TAS Test website through a personalized link, providing consent, and following the on-screen text and audiovisual instructions. If participants needed to resume testing after an accidental exit or interruption, they could log in again through their personal link and resume from their most recent incomplete test. Participants were not able to navigate backwards to retake completed test items.

HC and SCI group criteria

Because this study targeted adults without overt symptoms, we excluded any participants who reported conditions associated with cognitive or movement difficulties. Those who responded “yes” to any of the following ISLAND background survey questions were excluded: “Have you been told by a doctor that you have dementia?”; “Have you been told by your doctor that you have a memory impairment, but they were uncertain if you have dementia?”; “Have you been diagnosed with delirium?”; and “Have you been diagnosed with a central nervous system degenerative disease, e.g. Parkinson’s,

Huntington’s, Multiple Sclerosis?” TAS Test survey questions were also considered, and participants were excluded if they reported any prior diagnoses of “Parkinson’s disease,” “Multiple Sclerosis,” or “Mild Cognitive Impairment” in these items.

Following this screening, participants were assigned to our HC or SCI groups depending on their response to the ISLAND background survey question, “Have you noticed a substantial change in your memory and mental function in recent years?” Participants who responded “no” were assigned to the HC group; those who responded “yes” were assigned to the SCI group. Figure 1 illustrates the participant recruitment and classification process.

Demographic and clinical information

Participants’ demographic and clinical details were surveyed at the time of their recruitment to ISLAND (between 2019 and 2020). From these baseline data, we included the following details in our models: age in years (at the median TAS Test login date), sex (self-report response options: male, female, other, prefer not to say), highest level of education attained, anxiety and depression scores as reported using the Hospital Anxiety and Depression Scale (HADS), and “yes” or “no” responses to the question “Have you been diagnosed with a psychiatric disorder e.g. depression, psychosis, bipolar disorder, anxiety disorder?” The HADS and psychiatric items were included to adjust for effects of anxiety and depression on test performance and were not applied as exclusion criteria. Further characteristics of the ISLAND cohort are available in the protocol and interim results [27, 28].

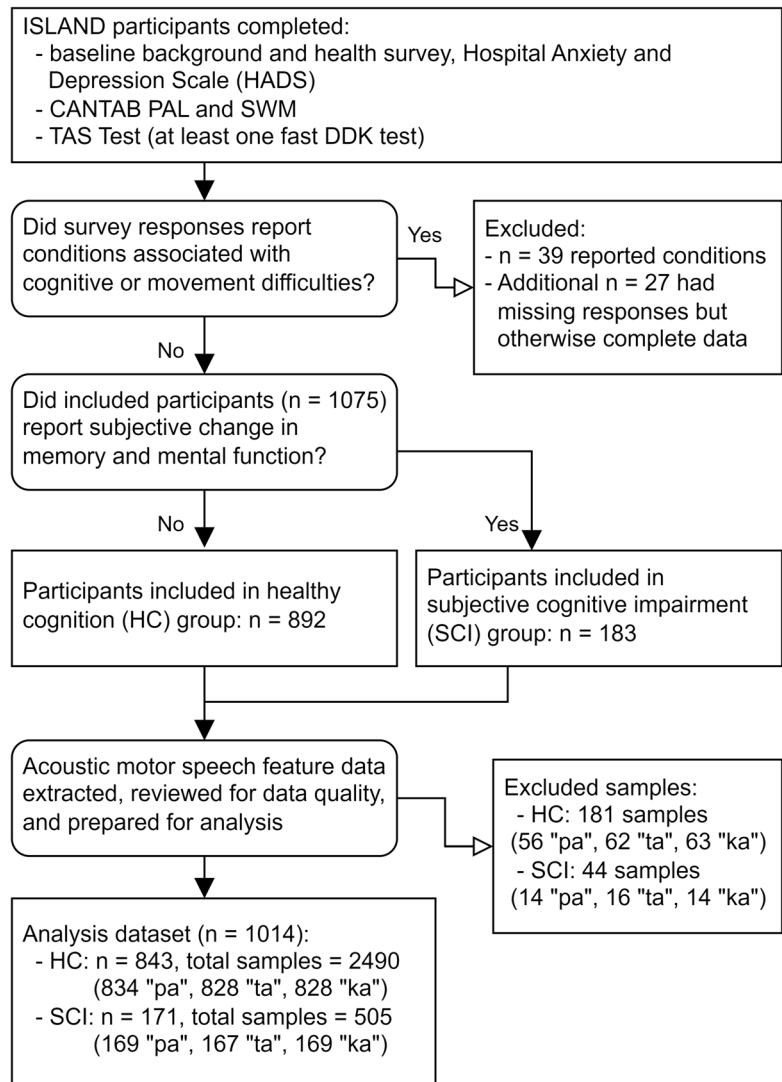
Cognitive assessment

ISLAND participants remotely completed two tests selected from the Cambridge Neuropsychological Test Automated Battery (CANTAB) [29] in August 2021.

Paired associates learning (PAL)

The PAL test assesses visual episodic memory through the memorization and location of increasing numbers of hidden patterns concealed in boxes on-screen. The test typically takes 8 min to administer. The score used in this study is the PALTEA6 score, which represents

Fig. 1 Flow diagram of participant recruitment, cognitive group classification criteria, and motor speech samples included in the final analysis dataset



participants' total errors on the six-pattern sequence, adjusted for incomplete or failed trials.

Spatial working memory (SWM)

The SWM test suite assesses working memory and executive function (planning and decision-making). Participants memorize and identify the locations of hidden pattern sequences and search for predetermined numbers of hidden tokens. The test typically takes 15 min to administer.

The scores used in this study are SWMBE6 for spatial working memory and SWMS for executive function. SWMBE6 measures the number of times

during a six-token search trial that a participant revisits a box in which they had previously found a token. SWMS measures search strategy through the number of unique boxes a participant uses to commence their search processes over the course of all their trials.

Motor speech data collection

Within the motor speech section of the TAS Test protocol, participants completed three fast, 10-s monosyllabic DDK tests ("pa", "ta", and "ka") in fixed order. Participants were presented with the following instructions in text and video, with the video including a demonstration by a speech pathologist ("pa"

example described): “Please repeat the sound ‘pa’ as fast as you can, like this: ‘pa-pa-pa...’. Keep going until you see the ‘Well done’ sign.” Once ready, participants commenced a 5-s countdown to each test by clicking a button labelled “Start the Test.” Fig. 1 in the Online Resource shows the steps of the “pa” test.

No standardized equipment or physical configuration of recording space was used, as participants were instructed to use their personal desktop or laptop computers to access the web-based TAS Test battery at home.

Acoustic feature extraction pipeline

The fast “pa”, “ta”, and “ka” speech recordings were stored in single-channel WAV format at a variety of sample rates from 8 to 48 kHz since capture parameters varied depending on participants’ computer hardware and settings. All recordings with 8 kHz sample rates were excluded upon review due to distorted quality (three total; one “pa,” one “ta,” and one “ka”). Recordings of negligible size or length (likely due to issues during capture or upload) and those containing no utterances were also excluded. We applied non-stationary noise reduction using the Python package “noisereducer” (v3.0.2) [30] to reduce the influence of ambient noises.

For feature extraction, we used two methods to create a combined set of complementary features from the same set of preprocessed recordings. The first method was our custom Python script, which used the package “librosa” (v0.10.1) [31] to load recordings, normalize signals, evaluate relative changes in energy over time, and extract onset event times, and the “signal” submodule of “SciPy” (v1.13.0) [32] to locate local waveform peaks. Individual features were then computed using our directly specified formulae. The second method was developed by Redenlab® and has been previously published [33–35]. To facilitate the description of our full feature set, Fig. 2 annotates key events and classifications on a waveform excerpt adapted from a DDK recording.

As Fig. 2 shows, the main events and zones in the DDK tests are the onset of speech in each repetition, voice onset (vibration of the vocal folds) at the commencement of the vowel sound, and the distinction between the spoken “syllable” and silent “pause” zones within each repetition. Table 1 presents our full list of features, grouped by their underlying qualities

to reflect the key events or zones used in their calculation. Features derived from the qualities of “onset,” “IOI,” and “energy” were extracted using our custom script, and those derived from the qualities “syllable,” “pause,” and “voice onset” were extracted using the methods developed by Redenlab®.

Statistical analysis

The goal of this analysis was to determine whether any of the DDK features improved cognitive score model fit over demographic and clinical variables, rather than exploring individual feature significance and relationships. To this end, regression models were fitted to 18 combinations of factors from the two cognitive groups (HC or SCI), three cognitive score outcomes (PALTEA6 [episodic memory], SWMS [executive function], and SWMBE6 [working memory]), and three DDK tests. Null model variables were participant age, sex, highest level of education, presence of a psychiatric disorder diagnosis, and baseline HADS anxiety and depression scores. In R, cognitive scores were modelled using generalized linear models with outcome distributions suitable for count data: Poisson for executive function (“glm” function, base R v4.4.0), and negative binomial for episodic and working memory to account for overdispersion (“glm.nb” function, “MASS” v7.3–60.2) [36]. The “dredge” function of the “MuMIn” package (v1.47.5) [37] was used to evaluate whether DDK features made significant contributions to model fit over the null variables in any of the 18 test cases. This was achieved by using “dredge” to fit, evaluate, and rank models in each case, starting with the fixed set of null variables, followed by the null variables plus every linear combination of at least one feature. No interaction terms were included. To reduce computational cost and pairwise collinearity between DDK features, we prohibited the simultaneous inclusion of DDK features in any single model if their pairwise correlation coefficient exceeded 0.6 in all three DDK tests.

All model selection was performed using the corrected Akaike Information Criterion (AICc) [38]. Models containing DDK features were deemed to explain significantly more variance than null models if $\Delta AICc$ (computed as $AICc_{null} - AICc_{speech}$) exceeded the rule of thumb threshold of two [39]. In each test case, if at least one model containing at least one feature was found to be superior to the null

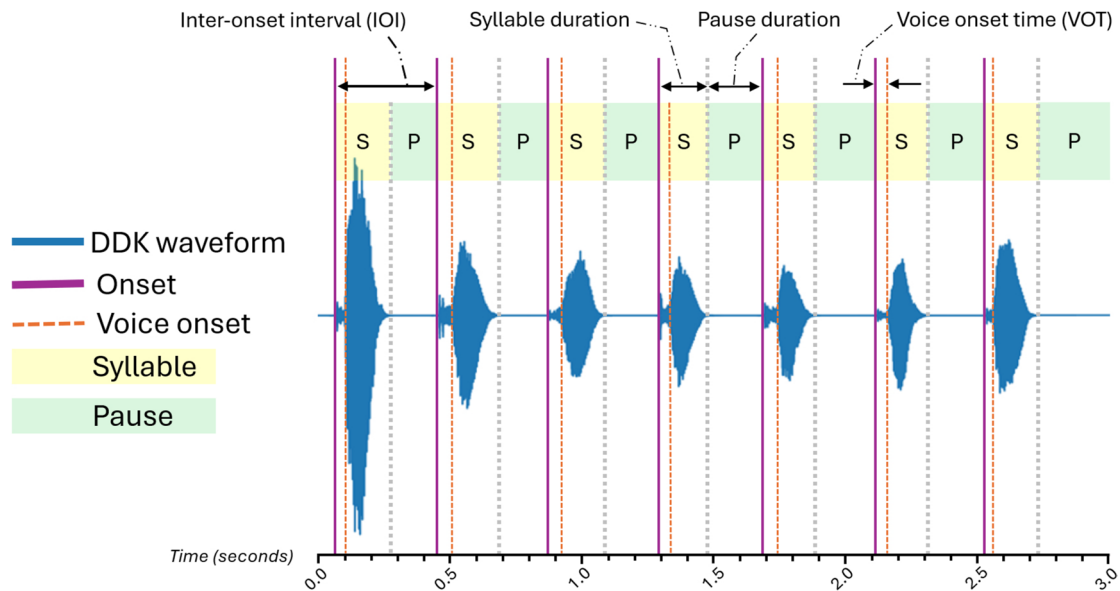


Fig. 2 Annotated amplitude view adapted from a DDK recording showing seven syllable repetitions. The onset events mark the commencement of the syllable in each repetition; the inter-onset intervals (IOIs) are the time between sequential syllable onsets (i.e., time to complete each repetition cycle). The “pa”, “ta”, and “ka” syllables all begin with voiceless plosives (“p”, “t”, and “k”) and the voice onset events mark when the

production of the voiced vowel (“a”) begins in each repetition. In this study, voice onset time (VOT) is measured as the duration between the speech onset event and the voice onset event. Each repetition is comprised of two zones: “syllable” during active speech production and “pause” for the segment between the end of the utterance and the start of the next repetition

model, this was interpreted as supporting evidence that the DDK test improved estimation of the applicable cognitive outcome. Nagelkerke’s pseudo- R^2 (R^2_N) is reported as an estimate of variance explained.

As a supplemental analysis into whether motor speech features improved prediction of HC vs. SCI, we fitted logistic regression models to each syllable-based subset of our data (i.e., “pa”, “ta”, and “ka”). Using the R package “pROC” [40], we compared receiver operating characteristic (ROC) curves for models fitted with null variables only to models fitted with all null model variables plus all motor speech features.

Results

Participants

We included data from 1014 participants without missing values in their eligibility, demographic, clinical, or cognitive variables, and for whom features from at least one DDK recording were available. Twenty-seven participants were excluded due to

incomplete survey data. HC group criteria were met by 843 participants, and 171 participants met criteria for the SCI group. Summary details for the HC, SCI, and excluded participant groups are shown in Table 2.

The number of DDK recordings included in each group varied across test types due to the results of data preprocessing and feature extraction. The counts reported in Table 2 were the totals remaining after preparing the data for analysis. From the HC group, 56 “pa” samples (6% of the total) were excluded, as were 62 “ta” samples (7%), and 63 “ka” samples (7%). From the SCI group, 14 “pa” samples (8%), 16 “ta” samples (9%), and 14 “ka” samples (8%) were excluded. Issues included variable speaking volumes which fell too quiet for microphones and differences in home internet connections, hardware, or configurations affecting file upload or sound quality.

Figure 3 shows the pairwise correlation coefficients of all motor speech features (measured from “pa” tests) and cognitive scores. The correlations from the “ta” and “ka” tests are included in Online Resource Fig. 2. The cells marked with crosses indicate the feature pairs which exceeded the correlation

Table 1 Motor speech feature definitions grouped by underlying quality

Quality	Feature name	Definition ^a
Onset	Onset count	Total number of syllable onsets
	Onset rate	Total syllable onset count divided by total sample duration (per second)
Inter-onset interval	IOI mean	Average inter-syllable onset interval (IOI) (seconds)
	IOI SD	Standard deviation (SD) of IOI (seconds)
	Inverse min. IOI	Inverse of minimum IOI (per second)
	Decrement on IOI	Logarithmic ratio derived from first and last IOIs, normalized by total count of IOIs
	Inverse IOI mean	Mean of inverse IOI (per second)
	Inverse IOI COV	Coefficient of variation ^b (COV) of inverse IOI
	Syllable	Articulation rate
Syllable duration mean		Average syllable duration (seconds)
Syllable duration COV		Coefficient of variation of syllable duration
Speech percent		Total speech duration as a percentage of total sample duration
Speech to pause ratio		Ratio of total speech duration to total pause duration
Pause	Pause mean	Average pause duration (seconds)
	Pause COV	Coefficient of variation of pause duration
	Pause percent	Total pause duration as a percentage of total sample duration
Voice onset	VOT mean	Average voice onset time ^c (VOT) (seconds)
	VOT COV	Coefficient of variation of VOT
Energy	Onset strength COV	Coefficient of variation of normalized onset strengths (relative changes in energy over time)

^aUnits provided where applicable

^bCoefficient of variation (COV) = standard deviation/mean

^cSee Fig. 2 for VOT detail

threshold described in “[Methods](#)” and were excluded from simultaneous selection in models.

Stronger associations with executive function in HC participants

The results for both participant groups are displayed in [Table 3](#). In both the HC group and the SCI group, the addition of motor speech features achieved significant improvements in fit (i.e., $\Delta AICc > 2$) over null models for all three cognitive outcomes.

In the HC group, features from all three DDK tests improved models of executive function. The two strongest results found in the HC group were in this domain (“ka” $\Delta AICc = 7.1$, followed by “pa” $\Delta AICc = 6.2$). Significant improvements were also found in episodic memory (“pa” $\Delta AICc = 5.8$ and “ta” $\Delta AICc = 2.1$) and working memory (“pa” $\Delta AICc = 2.5$).

Performance on the “ka” test was associated with executive function only. Features from the “pa” test were the only ones associated with improvements in all three cognitive domains.

Stronger associations with episodic memory in participants with SCI

As [Table 3](#) shows, a total of four significant associations were found in the SCI group. No cognitive domain was found to be associated with performance in all three DDK tests. The two strongest results were both in episodic memory (“pa” $\Delta AICc = 7.4$ and “ka” $\Delta AICc = 7.1$). The next best result was in working memory (“pa” $\Delta AICc = 4.5$), followed by executive function (“pa” $\Delta AICc = 3.8$). Performance on the “ta” test was not found to improve models in any of the cognitive domains.

Table 2 Summary of participant demographics and clinical details, cognitive scores, and DDK recordings

Measure	HC group (<i>n</i> = 843)	SCI group (<i>n</i> = 171)	Excluded ^a (<i>n</i> = 27)
Age			
Mean (SD)	66.9 (7.3)	66.3 (7.3)	68.0 (7.0)
Median (min, max)	67 (51, 92)	66 (52, 82)	68 (54, 79)
Gender (%total)			
Female	614 (72.8%)	121 (70.8%)	21 (77.8%)
Male	229 (27.2%)	50 (29.2%)	6 (22.2%)
Highest level of education (% total)			
High school	83 (9.8%)	13 (7.6%)	4 (14.8%)
Certificate or apprenticeship (including cert 2, 3, or 4)	58 (6.9%)	25 (14.6%)	1 (3.7%)
Diploma/associate degree	170 (20.2%)	30 (17.5%)	5 (18.5%)
Bachelor's degree	194 (23.0%)	39 (22.8%)	5 (18.5%)
Higher university degree (honors, graduate, diploma, masters, or PhD)	338 (40.1%)	64 (37.4%)	11 (40.7%)
Other	0 (0%)	0 (0%)	1 (3.7%)
Psychiatric disorder diagnosis (% total)			
No	715 (84.8%)	119 (69.6%)	20 (74.1%)
Yes	128 (15.2%)	52 (30.4%)	7 (25.9%)
Baseline HADS anxiety score ^b			
Mean (SD)	5.0 (3.4)	6.5 (3.9)	5.3 (3.6)
Median (min, max)	5 (0, 18)	6 (0, 18)	5 (0, 14)
Baseline HADS depression score ^b			
Mean (SD)	2.5 (2.2)	4.7 (3.2)	4.2 (2.5)
Median (min, max)	2 (0, 13)	4 (0, 14)	4 (0, 11)
Hearing impairment (% total) ^c			
No	701 (83.2%)	140 (81.9%)	17 (63.0%)
Yes	141 (16.7%)	31 (18.1%)	10 (37.0%)
Corrected hearing impairment (% of those with hearing impairment) ^c			
No	65 (46.1%)	16 (51.6%)	8 (80.0%)
Yes	74 (52.5%)	15 (48.4%)	2 (20.0%)
Wear corrective glasses (% total) ^c			
No	100 (11.9%)	11 (6.4%)	4 (14.8%)
Yes	740 (87.8%)	160 (93.6%)	23 (85.2%)
Episodic memory: PALTEA6 ^d			
Mean (SD)	4.1 (4.4)	4.5 (4.9)	4.2 (5.9)
Median (min, max)	3 (0, 21)	3 (0, 22)	2 (0, 20)
Executive function: SWMS ^d			
Mean (SD)	7.7 (2.8)	7.9 (2.6)	8.1 (2.9)
Median (min, max)	8 (2, 13)	8 (2, 12)	8 (3, 13)
Working memory: SWMBE6 ^d			
Mean (SD)	3.2 (3.3)	3.4 (3.4)	4.2 (3.7)
Median (min, max)	2 (0, 14)	3 (0, 12)	5 (0, 11)
Number of DDK recordings			
“pa”	834	169	27
“ta”	828	167	27
“ka”	828	169	25

^aParticipants ineligible due to incomplete personal data

^bHADS anxiety and depression scores: normal, 0–7; borderline abnormal, 8–10; abnormal, 11–21. Participants with abnormal scores were offered support service information and contact details

^cMissing responses: hearing impairment (*n* = 1 HC), corrected hearing impairment (*n* = 2 HC), wear corrective glasses (*n* = 3 HC)

^dCANTAB scores: For all three of the selected tests, a lower score represents better performance. The minimum possible score for both PALTEA6 and SWMBE6 is zero; for SWMS, one

In parallel with the HC group, the “pa” test was the only one associated with improvements in each cognitive domain.

Differences in feature prevalence

To help inform future investigations into specific aspects of motor speech performance, Fig. 4 provides a breakdown of individual feature prevalence in each significant test case. The models summarized are those which achieved significant improvements relative to null and were either the best-ranked model or approximately equivalent in fit (i.e., AICc within two units).

Across the full set of results, every feature appeared at least once. The top five features accounted for approximately 55% of all appearances and were syllable duration coefficient of variation (COV; 14.5%), speech to pause ratio (14.3%), onset count (13.4%), inverse of minimum IOI (6.7%), and IOI standard deviation (SD; 6.1%). The three features

which appeared the least were onset strength COV, decrement on IOI, and VOT COV. We note this aggregation disproportionately represents the larger result sets such as working memory in the SCI group.

Comparing across participant groups instead, the feature IOI SD appeared across almost all significant test cases in the HC group, except for working memory and the “pa” test. In the SCI group, however, this feature was only associated with the test case for episodic memory and “ka”. Conversely, the speech to pause ratio was associated with every significant case in the SCI group, but only with one case in the HC group (executive function and “ka”).

To generalize across “pa” cases, pause COV and the structurally related features speech percent and pause percent appeared in most cases for both participant groups. The inverse IOI mean showed associations with the SCI group but not HC.

Focusing on individual cognitive domains, articulation rate was associated with all episodic memory test cases. For executive function, results were

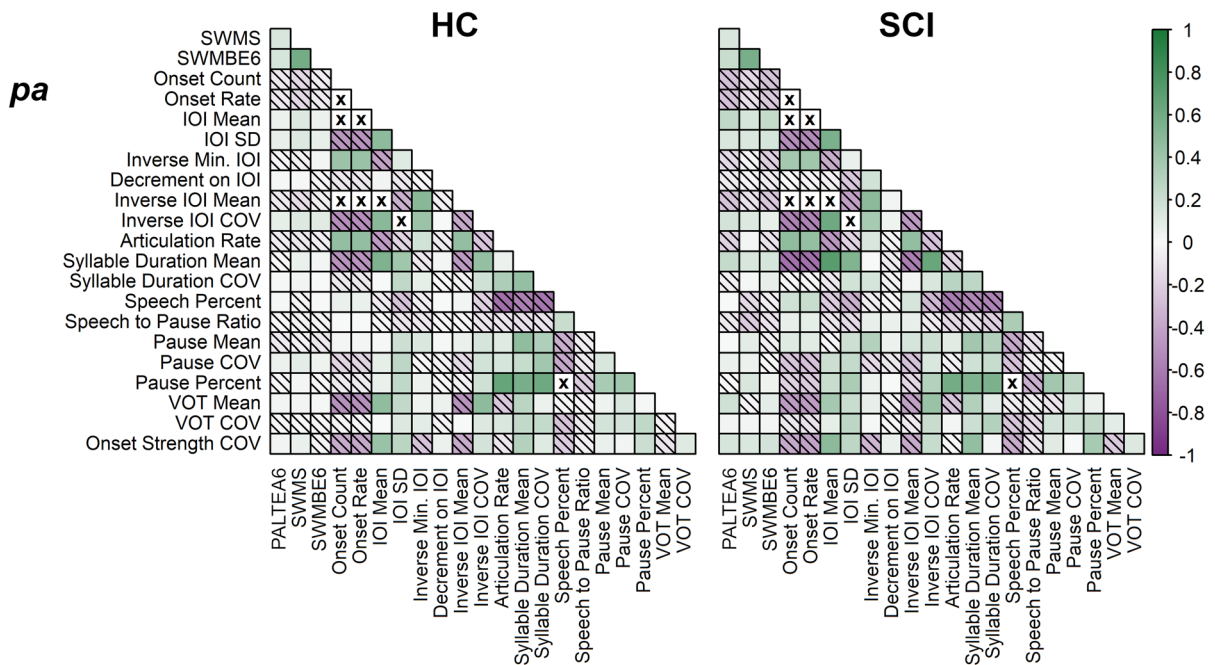


Fig. 3 Pairwise correlation coefficients between cognitive scores (PALTEA6 [visual episodic memory], SWMS [executive function], and SWMBE6 [spatial working memory]) and all motor speech features listed in Table 1, measured from “pa” tests (“ta” and “ka” results are shown in Online Resource Fig. 2). Correlations are displayed separately for each par-

ticipant cognitive group (healthy cognition [HC] or subjective cognitive impairment [SCI]). The feature pairs excluded from simultaneous selection in modelling are indicated by “x” (correlations not shown). Diagonal hatching indicates negative correlations

Table 3 Results in model significance ($\Delta AICc > 2$), fit (R^2_N), and proportion of additional variance explained (ΔR^2_N) by participant group, cognitive domain, and DDK test

Participant group	Cognitive domain (test score)	DDK test	$\Delta AICc^a$	$R^2_{N, \text{null}}$ (%)	$R^2_{N, \text{best}}^b$ (%)	ΔR^2_N (%)
Healthy cognition (HC)	Episodic memory (PALTEA6)	<i>“pa”</i>	5.8	9.0	10.1	1.1
		<i>“ta”</i>	2.1	8.6	9.3	0.7
		<i>“ka”</i>	0.0	8.5	8.5	0.0
	Executive function (SWMS)	<i>“pa”</i>	6.2	11.9	13.0	1.1
		<i>“ta”</i>	2.7	11.7	12.2	0.5
		<i>“ka”</i>	7.1	11.9	13.5	1.6
	Working memory (SWMBE6)	<i>“pa”</i>	2.5	6.4	7.4	1.0
		<i>“ta”</i>	0.0	6.3	6.5	0.2
		<i>“ka”</i>	0.1	5.9	6.2	0.3
Subjective cognitive impairment (SCI)	Episodic memory (PALTEA6)	<i>“pa”</i>	7.4	4.1	9.5	5.4
		<i>“ta”</i>	1.3	3.9	6.0	2.1
		<i>“ka”</i>	7.1	4.2	13.1	8.9
	Executive function (SWMS)	<i>“pa”</i>	3.8	13.6	16.7	3.1
		<i>“ta”</i>	0.0	13.8	13.8	0.0
		<i>“ka”</i>	0.0	13.6	13.6	0.0
	Working memory (SWMBE6)	<i>“pa”</i>	4.5	8.3	16.9	8.6
		<i>“ta”</i>	0.0	7.8	7.8	0.0
		<i>“ka”</i>	0.0	7.5	7.5	0.0

^aItalicized rows indicate the cases in which a model with at least one motor speech feature achieved a significantly improved fit ($\Delta AICc > 2$) over the demographic and clinical variable-only (null) model ($AICc_{\text{null}}$). $\Delta AICc = AICc_{\text{null}} - AICc_{\text{best}}$

^bThe model deemed “best” in each case was the one which returned the lowest AICc (i.e., $AICc_{\text{best}}$). This was either the null model or a model which included at least one motor speech feature

concentrated in the HC group due to minimal results in the SCI group. The HC group cases shared associations with IOI SD, inverse IOI COV, VOT mean, and onset strength COV. The HC and SCI working memory cases both shared inverse IOI COV, pause mean, and onset strength COV.

The results of our supplementary analysis into the classification of HC vs. SCI groups are available in the Online Resource (as Table 1 and Figs. 3, 4, and 5). In brief, inclusion of “pa”- or “ta”-measured motor speech features significantly improved ($p < 0.05$) classification performance over models fitted using null variables only.

Discussion

Performance on short motor speech tests improved estimation of cognitive performance in the domains of episodic memory, executive function, and working

memory. These three tests summed to only 30 s of active speech time per participant and achieved superior results beyond predictions using demographic (age, sex, education) and psychological (depression and anxiety) information alone.

In older adult participants without any cognitive symptoms or prior diagnoses (the HC group, $n = 843$), performance on each of the three DDK tests improved estimation of scores in all three cognitive domains. This partially supported our first and second hypotheses. Our first hypothesis expected features from all three DDK tests would improve models of episodic memory in this group, but only features from the “pa” and “ta” tests showed this association. Further, only one “ta”-based model of borderline significance was found, which may be a consequence of searching over a large variable combination space in the modelling process. Our second hypothesis anticipated that features from all three DDK tests would improve models of executive function and working memory. While features from all three tests did

demonstrate associations with executive function, only the “pa” test was found to be associated with working memory. Across the HC results, model fit improved the most in executive function using “ka” features ($\Delta R^2_N = 1.6\%$, total variance explained 13.5%).

Our secondary analysis in the SCI group ($n = 171$) found similar patterns. In this group, performance features from the “pa” and “ka” tests were associated with episodic memory. In terms of model fit improvements, the “pa” features doubled, and “ka” features tripled, the proportions of variance explained. This is notable as both SCI and reductions in episodic memory are established proxy measures for AD risk. Features from the “pa” test alone were found to be associated with executive function and working memory. The best fitting model was in working memory using “pa” features ($\Delta R^2_N = 8.6\%$, total variance explained 16.9%).

The full results suggest that associations between FT performance and episodic memory, executive function, and working memory in healthy older adults [18] extend to motor speech performance, and to participants with SCI. In FT, features from the dominant hand outperformed those from the nondominant hand in all three cognitive domains (with only dominant features significant in working memory) [18]. In this study, the “pa” test yielded results in all six combinations of cognitive domain and group, compared to two each for “ta” and “ka”. The “pa” features performed the best across almost all results and were the only ones significant in working memory. This raises questions of how associations between cognitive functions and plosive motor recruitment vary by place of articulation (bilabial “p”, alveolar “t”, or velar “k”), and whether bilabial motor associations are more dominant.

Interest is developing in tongue strength as a possible predictor of cognitive decline since it is affected by sarcopenia, which is, itself, a risk factor [41]. As articulation of the “ta” and “ka” syllables involves tongue movement, tongue strength may influence their associations with cognitive decline. Including measures of tongue strength could offer a localized way to adjust DDK performance for oral frailty and to explore performance differences across different stages of cognitive decline. In our results, “pa” had the broadest range of associations, suggesting that tongue-independent articulation tasks may have wider screening applications, at least in the earliest stages of cognitive decline. Unfortunately, the fixed order of

DDK tests in the TAS Test protocol (1, “pa”; 2, “ta”; 3, “ka”) may have influenced our results, along with fatigue, since DDK tests occurred after approximately 20 min of testing. Future work would benefit from randomizing or counter-balancing test order to validate syllable-specific associations.

Executive function has previously been linked with motor speech performance in younger adults [23, 24] and this study supports this association in older adults. Features from all three DDK tests improved estimation of executive function in the HC group, but only the “pa” test was associated with the SCI group (with only one candidate model identified). The only feature to appear in the SCI model was the speech to pause ratio. The same feature appeared only in the “ka” results for the HC group but in all candidate models. Feature prevalence varied over the HC cases, with the most frequent feature in “pa” models being pause COV. Both “pa” and “ka” shared moderate selection preference for onset-based features, and “ta” and “ka” tended toward IOI SD. The most notable difference between the HC and SCI groups was the overall decrease in results. Possible neural mechanisms for this require additional investigation. In healthy adults undertaking single-repetition DDK tasks, the areas activated in brain networks of motor control and phonological processing regions changed across age groups without a decrease in performance [42]. Our results require validation but may suggest that SCI is associated with a departure from the pattern of network reorganization expected in healthy aging.

A key strength of this study is that our models adjusted for important demographic and psychological covariates, which were set as the benchmark variables for estimating cognitive test performance. Other strengths include the breadth of investigation across cognitive conditions (i.e., asymptomatic HC group, subtle complaints in the SCI group), and the use of participant-collected data without researcher supervision or specialist equipment. We extracted a wide range of acoustic features from three established DDK tests, using complementary methods to capture subtle variations in motor speech performance.

It is important to note that the CANTAB tests studied relate to the specific cognitive domains of episodic memory, executive function, and working memory, and do not represent global cognitive function. Participants could also have had impairments in other domains which were not measured. We also note that

Participants	HC Group						SCI Group			
	Episodic Memory ["pa", "ta"]		Executive Function ["pa", "ta", "ka"]			Working Memory ["pa"]	Episodic Memory ["pa", "ka"]		Executive Function ["pa"]	Working Memory ["pa"]
Onset Count			54%		62%		44%	31%		59%
Onset Rate			38%		19%		30%	10%		9%
IOI Mean	10%		4%		5%			10%		3%
IOI SD	90%	100%	4%	80%	81%			24%		
Inverse Min. IOI			8%		5%		4%	3%		58%
Decrement on IOI			8%		5%		4%			
Inverse IOI Mean					5%		26%	7%		6%
Inverse IOI COV	10%		21%	20%	5%	100%		10%		14%
Articulation Rate	10%	100%			5%		11%	10%		15%
Syllable Duration Mean	10%			20%	100%					3%
Syllable Duration COV	10%				14%		7%	100%		88%
Speech Percent	10%		4%		5%		22%			39%
Speech to Pause Ratio					100%		26%	76%	100%	62%
Pause Mean	100%		8%		5%	100%				18%
Pause COV	10%		83%				4%	10%		6%
Pause Percent	10%		4%		5%		22%			39%
VOT Mean			8%	40%	14%		11%	69%		3%
VOT COV	10%		8%					3%		
Onset Strength COV			4%	20%	5%	50%				3%
<i>No. best performing models</i>	10	1	24	5	21	2	27	29	1	66
Feature appearance rates in the best performing models (%)			0%	> 0%	20%	40%	60%	80%	100%	

◀**Fig. 4** Summary of feature inclusion rates in the set of approximately equivalent best performing models for each participant group, cognitive domain, and DDK test case. In each case, the set of models considered to be these candidates for best fit were those which achieved a significant improvement over their respective null model and were ranked within $\Delta AICc < 2$ of the model with the lowest AICc. A 100% appearance rate means that a feature was included in all the candidates for best performing model in that specific test case. Expanded feature definitions are available in Table 1

changes in associations between motor speech performance and cognition over time cannot be inferred from this study due to its cross-sectional design. We reported associations with episodic memory and SCI, both of which are relevant to AD risk but not exclusive. These data suggest that longitudinal evidence of associations with cognitive decline (for instance, subtle changes in episodic memory, or later diagnoses of MCI and AD) will be valuable for finding and characterizing risk profiles specific to AD. A longitudinal investigation of these associations may be possible through ISLAND. We acknowledge, however, the limited cultural and linguistic diversity of participants recruited from ISLAND, most of whom were female, highly educated, and of White, Northern European ancestry. This limits our findings from generalizing to a broader domestic or international population.

Longitudinal and meta-analytical evidence supports the validity of using a single question to classify SCI in participants [43, 44]. However, the eligibility criteria used did not preclude the possibility that some participants classified as HC may have had presymptomatic neurodegenerative disease. Similarly, it is possible that some participants classified as SCI may have had clinically significant impairment which was undiagnosed at the time of survey completion. Participants in either group may also have had difficulty understanding the medical terminology used in the personal health surveys, incorrect responses to which could have contributed to group misclassification. Since participants were recruited from a highly educated cohort, we expect the risk of misunderstanding to have been low, but we are unable to confirm this as comprehension was not assessed.

The remote collection of speech samples introduced substantial variation between participants in hardware, test environment and its spatial configuration, and internet connection quality. This has likely affected results. Test delivery and data collection

through smartphones, while also subject to variability, may have the potential to offer greater community reach and accessibility compared to desktop and laptop computers [45, 46]. While delivery of the test battery through smartphones may help accommodate participants with lower digital literacy, written survey items and instructions may remain a barrier to participants with lower literacy levels. Although the selected DDK tests were provided with text and audiovisual instructions to facilitate access, and were intended to impose minimal language- and culture-specific requirements, members ultimately required computer and internet access, digital literacy, and reading ability to participate. Further studies are also needed to account for the effects on motor speech performance of factors such as fluency in everyday speech, denture usage, and differences in hearing ability. These factors are relevant across all ages but will be especially important in distinguishing pathological and aging-associated changes in motor speech function. Hearing loss, for example, is a known risk factor for dementia and, although the risk can be mitigated with correction [4], a slim majority of our SCI group participants with hearing impairment reported no use of corrective devices. In adolescents, hearing impairment is associated with DDK performance [47], but further research is needed into these relationships and the effects of corrective device use in older adults.

The promising results of this study facilitate future work investigating the significance and relationships of specific motor speech features with cognition, as this type of interpretation is not supported by the current study design. Additional lines of investigation could include similar analyses using features which represent DDK accuracy, comparisons between self- and maximal-paced tests to represent motor reserve, and extensions into the clinically useful sequential trisyllabic “pataka” test. Further, investigation of associations between motor speech performance and blood-, CSF-, or imaging-based biomarkers of AD pathology may assist with linking at-risk cognitive functions to underlying AD pathology via non-invasive speech measures. Emerging multimodal and artificial intelligence-based methods for early detection of neurodegenerative diseases [48, 49] may offer alternative approaches to feature extraction, analysis, and pooling with other modes of testing such as FT.

In summary, this study demonstrates that performance on short, self-administered motor speech tests is associated with cognition in older adults. Performance features from the “pa” DDK test improved estimation of test scores in episodic memory, executive function, and working memory in both participant groups (HC and SCI). This study establishes the potential of motor speech tests to be part of the solution in the urgent search for non-invasive, cost-effective, and community-scalable screening tools for AD risk.

Acknowledgements The authors are grateful for the support of the ISLAND project team, and the assistance of James Vickers and Anna King in reviewing and editing the manuscript.

Author contribution Drafting or revision of manuscript: JG, RL, ER, APV, ADB, LRG, QB, JA. Study concept or design: ADB, LRG, QB, JA. Acquisition of data: RL, ER, APV, QB, JA. Analysis or interpretation of data: JG, RL, ER, APV, ADB, LRG, JA.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This study was funded through a National Health and Medical Research Council grant (2004051) and the Australian Government’s Medical Research Future Fund (MRFF1170820). A. P. Vogel was funded by an Australian Research Council Fellowship (FT220100253). J. Giffard is supported by an Australian Government Research Training Program Scholarship. The funders had no role in study design, data collection, analysis, or interpretation, in the preparation of the manuscript, or in the decision to publish.

Data availability The authors may share the de-identified data with qualified investigators whose proposal of data use has been approved by an independent review committee.

Declarations

Ethics approval and consent to participate ISLAND and TAS Test were approved by the University of Tasmania Human Research Ethics Committee (H0018264 and H0021660, respectively) in accordance with the guidelines of the National Health and Medical Research Council. Informed consent was provided by all participants.

Consent for publication All authors and named contributors have provided consent for publication.

Competing interests A. P. Vogel is Chief Science Officer of Redenlab Inc. The remaining authors declare no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Fiest KM, Roberts JI, Maxwell CJ, Hogan DB, Smith EE, Frolkis A, et al. The prevalence and incidence of dementia due to Alzheimer’s disease: a systematic review and meta-analysis. *Can J Neurol Sci.* 2016;43:S51–82. <https://doi.org/10.1017/cjn.2016.36>.
2. Alty J, Bai Q, Li R, Lawler K, St George RJ, Hill E, et al. The TAS Test project: a prospective longitudinal validation of new online motor-cognitive tests to detect preclinical Alzheimer’s disease and estimate 5-year risks of cognitive decline and dementia. *BMC Neurol.* 2022;22:266. <https://doi.org/10.1186/s12883-022-02772-5>.
3. Scheltens P, De Strooper B, Kivipelto M, Holstege H, Chételat G, Teunissen CE, et al. Alzheimer’s disease. *The Lancet.* 2021;397:1577–90. [https://doi.org/10.1016/S0140-6736\(20\)32205-4](https://doi.org/10.1016/S0140-6736(20)32205-4).
4. Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *The Lancet.* 2024;404:572–628. [https://doi.org/10.1016/S0140-6736\(24\)01296-0](https://doi.org/10.1016/S0140-6736(24)01296-0).
5. Chu D, Yang X, Wang J, Zhou Y, Gu J-H, Miao J, et al. Tau truncation in the pathogenesis of Alzheimer’s disease: a narrative review. *Neural Regen Res.* 2024;19:1221. <https://doi.org/10.4103/1673-5374.385853>.
6. Singh H, Das A, Khan MM, Pourmotabbed T. New insights into the therapeutic approaches for the treatment of tauopathies. *Neural Regen Res.* 2024;19:1020. <https://doi.org/10.4103/1673-5374.385288>.
7. Bature F, Pappas Y, Pang D, Guinn B-A. Can non-invasive biomarkers lead to an earlier diagnosis of Alzheimer’s disease? *Curr Alzheimer Res.* 2021;18:908–13. <https://doi.org/10.2174/1567205018666211207094630>.
8. Gauthier S, Rosa-Neto P, Morais JA, Webster C. World Alzheimer Report 2021: journey through the diagnosis of dementia. London, England: Alzheimer’s Disease International; 2021.

9. Mollica MA, Tort-Merino A, Navarra J, Fernández-Prieto I, Valech N, Olives J, et al. Early detection of subtle motor dysfunction in cognitively normal subjects with amyloid- β positivity. *Cortex*. 2019;121:117–24. <https://doi.org/10.1016/j.cortex.2019.07.021>.
10. Beauchet O, Allali G, Berrut G, Hommet C, Dubost V, Assal F. Gait analysis in demented subjects: interests and perspectives. *Neuropsychiatr Dis Treat*. 2008;4:155–60. <https://doi.org/10.2147/ndt.s2070>.
11. Gras LZ, Kanaan SF, McDowd JM, Colgrove YM, Burns J, Pohl PS. Balance and gait of adults with very mild Alzheimer disease. *J Geriatric Phys Ther*. 2015;38:1. <https://doi.org/10.1519/JPT.000000000000020>.
12. Rudd KD, Lawler K, Callisaya ML, Alty J. Investigating the associations between upper limb motor function and cognitive impairment: a scoping review. *GeroScience*. 2023. <https://doi.org/10.1007/s11357-023-00844-z>.
13. Wang X, St George RJ, Bindoff AD, Noyce AJ, Lawler K, Roccati E, et al. Estimating presymptomatic episodic memory impairment using simple hand movement tests: a cross-sectional study of a large sample of older adults. *Alzheimer's & Dementia*. 2024;20:173–82. <https://doi.org/10.1002/alz.13401>.
14. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*. 2011;7:280–92. <https://doi.org/10.1016/j.jalz.2011.03.003>.
15. Schindler SE, Jasielec MS, Weng H, Hassenstab JJ, Grober E, McCue LM, et al. Neuropsychological measures that detect early impairment and decline in preclinical Alzheimer disease. *Neurobiol Aging*. 2017;56:25–32. <https://doi.org/10.1016/j.neurobiolaging.2017.04.004>.
16. Bäckman L, Small BJ, Fratiglioni L. Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain*. 2001;124:96–102. <https://doi.org/10.1093/brain/124.1.96>.
17. Swainson R, Hodges JR, Galton CJ, Semple J, Michael A, Dunn BD, et al. Early detection and differential diagnosis of Alzheimer's disease and depression with neuropsychological tasks. *Dement Geriatr Cogn Disord*. 2001;12:265–80. <https://doi.org/10.1159/000051269>.
18. Li R, Wang X, Lawler K, Garg S, St George RJ, Bindoff AD, et al. Brief webcam test of hand movements predicts episodic memory, executive function, and working memory in a community sample of cognitively asymptomatic older adults. *Alzheimer's & Dement: Diagnos, Assess Dis Monit*. 2024;16:e12520. <https://doi.org/10.1002/dad2.12520>.
19. Banovic S, Zunic LJ, Sinanovic O. Communication difficulties as a result of dementia. *Mater Sociomed*. 2018;30:221–4. <https://doi.org/10.5455/msm.2018.30.221-224>.
20. Vogel AP, Poole ML, Pemberton H, Caverlé MWJ, Boonstra FMC, Low E, et al. Motor speech signature of behavioral variant frontotemporal dementia: refining the phenotype. *Neurology*. 2017;89:837–44. <https://doi.org/10.1212/WNL.0000000000004248>.
21. Poole ML, Vogel AP. Linking motor speech function and dementia. *Genetics, Neurology, Behavior, and Diet in Dementia*, Elsevier; 2020, 665–76. <https://doi.org/10.1016/B978-0-12-815868-5.00042-6>.
22. Kent RD, Yunjung K, Li-mei C. Oral and laryngeal diadochokinesis across the life span: a scoping review of methods, reference data, and clinical applications. *J Speech, Language Hear Res*. 2022;65:574–623. https://doi.org/10.1044/2021_JSLHR-21-00396.
23. Shen C, Janse E. Maximum speech performance and executive control in young adult speakers. *J Speech Lang Hear Res*. 2020;63:3611–27. https://doi.org/10.1044/2020_JSLHR-19-00257.
24. Janse E, Shen C, De Kerf E. Diadochokinesis performance and its link to cognitive control: alternating versus non-alternating diadochokinesis. *J Speech Lang Hear Res*. 2024;67:4096–106. https://doi.org/10.1044/2023_JSLHR-23-00076.
25. Rudd KD, Lawler K, Callisaya ML, Bindoff AD, Chiranakorn-Costa S, Li R, et al. Hand motor dysfunction is associated with both subjective and objective cognitive impairment across the dementia continuum. *Dement Geriatr Cogn Disord*. 2024. <https://doi.org/10.1159/000540412>.
26. Mitchell AJ, Beaumont H, Ferguson D, Yadegarfar M, Stubbs B. Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: meta-analysis. *Acta Psychiatr Scand*. 2014;130:439–51. <https://doi.org/10.1111/acps.12336>.
27. Bartlett L, Doherty K, Farrow M, Kim S, Hill E, King A, et al. Island Study Linking Aging and Neurodegenerative Disease (ISLAND) targeting dementia risk reduction: protocol for a prospective web-based cohort study. *JMIR Res Protocol*. 2022;11:e34688. <https://doi.org/10.2196/34688>.
28. Bartlett L, Bindoff A, Doherty K, Kim S, Eccleston C, Kitsos A, et al. An online, public health framework supporting behaviour change to reduce dementia risk: interim results from the ISLAND study linking ageing and neurodegenerative disease. *BMC Publ Health*. 2023;23:1886. <https://doi.org/10.1186/s12889-023-16805-2>.
29. Robbins TW, James M, Owen AM, Sahakian BJ, McInnes L, Rabbitt P. Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dement Geriatr Cogn Disord*. 1994;5:266–81. <https://doi.org/10.1159/000106735>.
30. Sainburg T. timsainb/noisereduce: v1.0 2019. <https://doi.org/10.5281/ZENODO.3243139>
31. McFee B, McVicar M, Faronbi D, Roman I, Gover M, Balke S, et al. librosa/librosa: 0.10.1 2023. <https://doi.org/10.5281/ZENODO.8252662>

32. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17:261–72. <https://doi.org/10.1038/s41592-019-0686-2>.
33. van Prooije T, Knuijt S, Oostveen J, Kapteijns K, Vogel AP, van de Warrenburg B. Perceptual and acoustic analysis of speech in spinocerebellar ataxia type 1. *Cerebellum*. 2023. <https://doi.org/10.1007/s12311-023-01513-9>.
34. van Gool R, Golden E, Goodlett B, Zhang F, Vogel AP, Tourville JA, et al. Characterization of central manifestations in patients with Niemann-Pick disease type C. *Genet Med*. 2024;26:101053. <https://doi.org/10.1016/j.gim.2023.101053>.
35. Tamplin J, Haines SJ, Baker FA, Sousa TV, Thompson Z, Crouch H, et al. ParkinSong Online: feasibility of telehealth delivery and remote data collection for a therapeutic group singing study in Parkinson's. *Neurorehabil Neural Repair*. 2024;38:122–33. <https://doi.org/10.1177/15459683231219269>.
36. Ripley B, Venables B, Bates DM, Hornik K, Gebhardt A, Firth D. MASS: support functions and datasets for Venables and Ripley's MASS 2024. <https://doi.org/10.32614/CRAN.package.MASS>.
37. Kamil Bartoń. MuMIn: Multi-Model Inference 2010. <https://doi.org/10.32614/CRAN.package.MuMIn>.
38. Hurvich CM, Tsai CL. Regression and time series model selection in small samples. *Biometrika*. 1989;76:297–307. <https://doi.org/10.1093/biomet/76.2.297>.
39. Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. *Sociol Methods Res*. 2004;33:261–304. <https://doi.org/10.1177/0049124104268644>.
40. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77. <https://doi.org/10.1186/1471-2105-12-77>.
41. Yitbarek GY, Alty J, Lawler K, Goldberg LR. Current evidence on the association of tongue strength with cognitive decline in older adults and the known risk factors of frailty, sarcopenia and nutritional health: a scoping review protocol. *BMJ Open*. 2023;13:e076005. <https://doi.org/10.1136/bmjopen-2023-076005>.
42. Sörös P, Bose A, Sokoloff LG, Graham SJ, Stuss DT. Age-related changes in the functional neuroanatomy of overt speech production. *Neurobiol Aging*. 2011;32:1505–13. <https://doi.org/10.1016/j.neurobiolaging.2009.08.015>.
43. Kang M, Li C, Mahajan A, Spat-Lemus J, Durape S, Chen J, et al. Subjective cognitive decline plus and longitudinal assessment and risk for cognitive impairment. *JAMA Psychiat*. 2024;81:993–1002. <https://doi.org/10.1001/jamapsychiatry.2024.1678>.
44. Pike KE, Cavuoto MG, Li L, Wright BJ, Kinsella GJ. Subjective cognitive decline: level of risk for future dementia and mild cognitive impairment, a meta-analysis of longitudinal studies. *Neuropsychol Rev*. 2022;32:703–35. <https://doi.org/10.1007/s11065-021-09522-3>.
45. Li R, Huang G, Wang X, Lawler K, Goldberg LR, Roccati E, et al. Smartphone automated motor and speech analysis for early detection of Alzheimer's disease and Parkinson's disease: validation of TapTalk across 20 different devices. *Alzheimer's Dement Diagn, Assess Dis Monit*. 2024;16:e70025. <https://doi.org/10.1002/dad2.70025>.
46. Alty J, Goldberg LR, Roccati E, Lawler K, Bai Q, Huang G, et al. Development of a smartphone screening test for preclinical Alzheimer's disease and validation across the dementia continuum. *BMC Neurol*. 2024;24:127. <https://doi.org/10.1186/s12883-024-03609-z>.
47. Robb MP, Hughes MC, Frese DJ. Oral diadochokinesis in hearing-impaired adolescents. *J Commun Disord*. 1985;18:79–89. [https://doi.org/10.1016/0021-9924\(85\)90038-3](https://doi.org/10.1016/0021-9924(85)90038-3).
48. Huang G, Li R, Bai Q, Alty J. Multimodal learning of clinically accessible tests to aid diagnosis of neurodegenerative disorders: a scoping review. *Health Inf Sci Syst*. 2023;11:32. <https://doi.org/10.1007/s13755-023-00231-0>.
49. Li R, Wang X, Lawler K, Garg S, Bai Q, Alty J. Applications of artificial intelligence to aid early detection of dementia: a scoping review on current capabilities and future directions. *J Biomed Inform*. 2022;127:104030. <https://doi.org/10.1016/j.jbi.2022.104030>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.