



## OPEN ACCESS

## EDITED BY

Andreea Oliviana Diaconescu,  
University of Toronto, Canada

## REVIEWED BY

Maksymilian Aleksander Brzezicki,  
University of Oxford, United Kingdom  
Ciming Pan,  
Yunnan University of Traditional Chinese  
Medicine, China

## \*CORRESPONDENCE

Varsha D. Badal  
✉ vbadal@health.ucsd.edu

RECEIVED 19 March 2025

ACCEPTED 03 June 2025

PUBLISHED 01 July 2025

## CITATION

Badal VD, Tran C, Brown H, Glorioso DK,  
Daly R, Molina AJA, Moore AA, Bilal E, Lee EE  
and Depp CA (2025) Audio and linguistic  
prediction of objective and subjective  
cognition in older adults: what is the role of  
different prompts?  
*Front. Psychiatry* 16:1596132.  
doi: 10.3389/fpsy.2025.1596132

## COPYRIGHT

© 2025 Badal, Tran, Brown, Glorioso, Daly,  
Molina, Moore, Bilal, Lee and Depp. This is an  
open-access article distributed under the terms  
of the [Creative Commons Attribution License](#)  
(CC BY). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Audio and linguistic prediction of objective and subjective cognition in older adults: what is the role of different prompts?

Varsha D. Badal<sup>1,2\*</sup>, Caitlyn Tran<sup>3</sup>, Haze Brown<sup>3</sup>,  
Danielle K. Glorioso<sup>1,2</sup>, Rebecca Daly<sup>1,2</sup>, Anthony J. A. Molina<sup>2,4</sup>,  
Alison A. Moore<sup>2,4</sup>, Erhan Bilal<sup>5</sup>, Ellen E. Lee<sup>1,2,6</sup>  
and Colin A. Depp<sup>1,2,6</sup>

<sup>1</sup>Department of Psychiatry, University of California San Diego, San Diego, CA, United States, <sup>2</sup>Stein Institute for Research on Aging, University of California San Diego, San Diego, CA, United States, <sup>3</sup>University of California San Diego, San Diego, CA, United States, <sup>4</sup>Department of Medicine, University of California San Diego, La Jolla, CA, United States, <sup>5</sup>International Business Machines Corporation (IBM) Research, Yorktown, NY, United States, <sup>6</sup>Veterans Affairs (VA) San Diego Healthcare System, La Jolla, CA, United States

**Background:** Psycho-linguistic and audio data derived from speech may be useful in screening and monitoring cognitive aging. However, there are gaps in understanding the predictive value of different prompts (e.g., open ended or structured) and the relationship of features to subjective versus objective cognition.

**Objective:** To advance understanding of method variation in speech-analysis based psychometry, we evaluated targeted prompts for classification of impaired cognition and cognitive complaints.

**Method:** A sample of 49 older participants (mean age: 76.9, SD: 8.5) completed short interview questions and cognitive assessments. Acoustic and Linguistic Inquiry through Word Counting i.e., LIWC (verbal content-based) features were derived from answers to open ended questions about aging (AG) and the Cookie Theft task (CT). Outcomes were objective cognitive ability measured using Telephone Interview for Cognitive Status (TICS-m), and subjective cognition using Cognitive Failures Questionnaire (CFQ).

**Results:** A combined feature set including acoustic and LIWC (verbal content) yielded excellent classification results for both CFQ and TICS-m. The F1, precision and recall for CFQ elevation was 0.83, 0.85 and 0.82, and for TICS-m cutoff was 0.92, 0.92 and 0.92 respectively (using single learners). Features derived from CT task were of greater relevance to TICS-m classification, while the features from the AG task were of greater relevance to the CFQ classification.

**Conclusion:** Acoustic and psycholinguistic features are relevant to assessment of cognition and subjective cognitive complaints, with combined features performing best. However, subjective and objective cognitions were predicted to differing extents by the different tasks, and the feature sets.

## KEYWORDS

acoustic, psycholinguistic, cognitive impairment, dementia, machine learning, NLP, Alzheimer's

## 1 Introduction

It is well established that age-related cognitive decline co-occurs with changes detectable in speech (1). Changes in speech appear also to be associated with risk of Alzheimer's disease (2). Speech analysis has primarily derived from samples of responses to structured prompts, but audio and psycho-linguistic analysis of verbal responses in open ended conversation may also predict objective cognitive performance (3). In addition to prediction of objective cognitive impairments, the application of speech analysis to prediction about subjective cognitive complaints or SCC (e.g. concerns about memory or slow thinking) has received little study (4). SCC are a key component of screening and diagnosis of Mild Cognitive Impairment (MCI). SCC remain unexplored through automatic voice and speech-analysis based techniques, leaving a gap in how speech task variants correlate with subjective cognition.

Although both relevant and common, SCC (5–9) are distinct from cognitive impairments (10). SCC is positively correlated with depressive symptoms (11), more so than are objective measures of cognition (12–15). Identifying which features from audio-based samples predict subjective, objective cognition, and both, could be helpful in understanding the potential utility of speech analysis.

The type of conducted dialogue (e.g., unstructured interviews vs. directed instructions) and the topic may influence not only the sentiment and non-verbal vocalizations, but also the content and framing of responses (16–18). Cognitive impairment has been explored through automated speech analysis using several kinds of dialogues with humans or software agents. Some of these dialogues are everyday conversation with humanoid robot (19), computer avatar based conversations (20), casual conversation (21–23), story retelling (24), recalling content of film (25), picture description task (26) and directed questions such as birthplace, name of elementary school, time orientation and backward recitation of three digit numbers (27). Among these, the Cookie Theft task (28, 29), which is a directed picture description task, has been a popular choice (30–34). To our knowledge, few or no studies have evaluated differences in prediction from different speech data sources within the same sample.

Recorded conversational speech offers a variety of features: acoustic, linguistic and verbal content; each offering a different insight (18). These layers often intertwine and influence each other; for example, a speaker's voice acoustics may betray underlying emotions that can significantly impact the interpretation of the information (content) as well as speaker's age and health. Acoustic feature sets are often large, openSMILE (35–37) comprises a set of 88 features, some of which have been related to psychological processes. In speech analysis, “shimmer” is a measured acoustic feature that quantifies the cycle-to-cycle variation in the amplitude (volume) of a voice signal, as to how much the loudness fluctuates between each vocal fold vibration. Shimmer features studies (38, 39) suggest a link with emotions, and as indicators of cognition decline (40). More recently, formant frequencies were shown to undergo a predictable change under cognitive load (41). Linguistic Inquiry and Word Count (LIWC) (42) counts words which are assigned into various psychological and linguistic categories (43). Speech transcribed through text can be effectively processed using LIWC

for content analysis (44) for mood (45) as well as objective cognition (46) and other constructs (47). Although we could find no studies on subjective cognition using LIWC, it is a strong candidate feature set for such analyses.

A recent comprehensive review of NLP and audio based studies on detection of cognitive impairment (48) summarized that most prior work included both NLP and audio analyses in the same sample. Among studies reviewed, speech elicitation methods varied from spontaneous speech, clinical interviews, and conversations with virtual agents. The analysis for the CT task mostly relied upon NLP based techniques using n-grams, BERT-embeddings, Transformer encodings, GPT encodings; only three combined both NLP and acoustic features. Another review focused only on automated speech recognition based methods (49) included only three studies combining NLP and audio features, one included immediate and delayed recall of a short film and two using cookie theft task. Techniques combining NLP and speech performed generally better than either one separately. No study to our knowledge addressed both objective cognition and subjective cognitive complaints while combining NLP and audio features.

In this study, we contrast two approaches to speech data collection: the Cookie theft picture description task that invokes cognitive processing, and the other more open-ended prompting to describe individual experiences of aging. The choice of prompts (Cookie Theft Task, successful aging questions) in this pilot study were partially dictated by the prevalent norms, especially the cookie theft prompt. The Cookie Theft Task (28) has a long history in aphasia diagnosis (29), with as well as in research on Alzheimer's (50). The task was also selected because it has been extensively evaluated using speech analysis (31, 48, 49, 51, 52). The aging questions were chosen as a complementary task because they were previously evaluated in other studies of healthy aging (53) and has been subjected to linguistic and voice analyses (3, 54, 55). Finally, both the Cookie Theft task and successful aging questions can be delivered by remote means making them scalable.

We also evaluated the association of these modalities with differential prediction of subjective cognitive ability and objective impairment. We hypothesized (based on prior literature) that both speech elicitation prompts would yield data that would result in reasonable levels of accuracy in discriminating individuals above the cutoffs from those below, for subjective as well as objective measures. We also hypothesized that acoustic and linguistic features could attain good performance ( $F1 > 0.75$ ) in predicting both subjective and objective cognitive abilities. We explored the variation in contribution of acoustic versus linguistic features to integrated models, and then contrasted features derived from the different speech elicitation methods in predicting subjective versus objective cognitive impairments.

## 2 Materials and methods

### 2.1 Participants and procedures

The participants were drawn from a previously engaged large sample of 1300 community -dwelling residents of San Diego

County for the parent study, the Successful Aging Evaluation (SAGE) (56). That project is detailed elsewhere, and briefly, used random digit dialing to recruit a sample of 1006 persons. Participants completed a baseline assessment consisting of a set of survey instruments and thereafter participants were followed on an annual basis with some exceptional years. A subsample ( $n=311$ ) that had expressed interest in future studies on aging were contacted via mail with a pamphlet describing the goals of this study.

We augmented the SAGE survey (56) to include brief telephone or Zoom interviews of SAGE participants. The SAGE study (56) had the following inclusion criteria: (1) age 50–99 years, (2) having a (landline) telephone at home, (3) physical and mental ability to participate in a telephone interview and to complete a paper and pencil mail survey, (4) informed consent for study participation, and (5) English fluency and the exclusion criteria w: (1) residence in a nursing home, or requiring daily skilled nursing care, and (2) self-reported prior diagnosis of dementia, (3) terminal illness, or requiring hospice care. The study protocol was approved by the IRB of the University of California San Diego.

A subsample ( $n=311$ ) that had expressed interest in future studies on aging were contacted via mail with a pamphlet describing the goals of this study of which 49 participated in the study. The severe attrition of the sample was attributed to several

reasons, some of which are detailed as follows (Figure 1). Some of the phones were out of service or incorrect numbers (68), some did not pick up the phone and no voicemail could be left (35). One or more voicemails was left for some individuals (100). Many of those contacted had lost interest due to illness, scheduling or age (52). Six interviews were cancelled or withdrawn. No contact could be established with the remainder. Of the 49 who consented and interviewed, 40 were interviewed over Zoom and 9 over the phone. All recordings were of acceptable audio quality and were transcribed for use in analysis.

The interview contained three parts, 1) the Cookie Theft task (CT) from the Boston Diagnostic Aphasia Examination (28, 29) (Supplementary, Appendix A). 2) three open ended questions about aging (AG) (Supplementary, Appendix A) and 3) two structured questionnaires, the 12 -item Modified Telephone Interview for Cognitive Status (TICS-m) (57) and the 25-item Cognitive Failures Questionnaire (CFQ) (58, 59). The interviewer was trained to administer the TICS-m and CFQ tasks by a licensed staff psychologist who was also available to answer scoring related questions. Interviews were conducted over Zoom or phone between June 2024 and December 2024 and study data were managed and collected using REDCap electronic data capture tools hosted at UC San Diego (60–62).

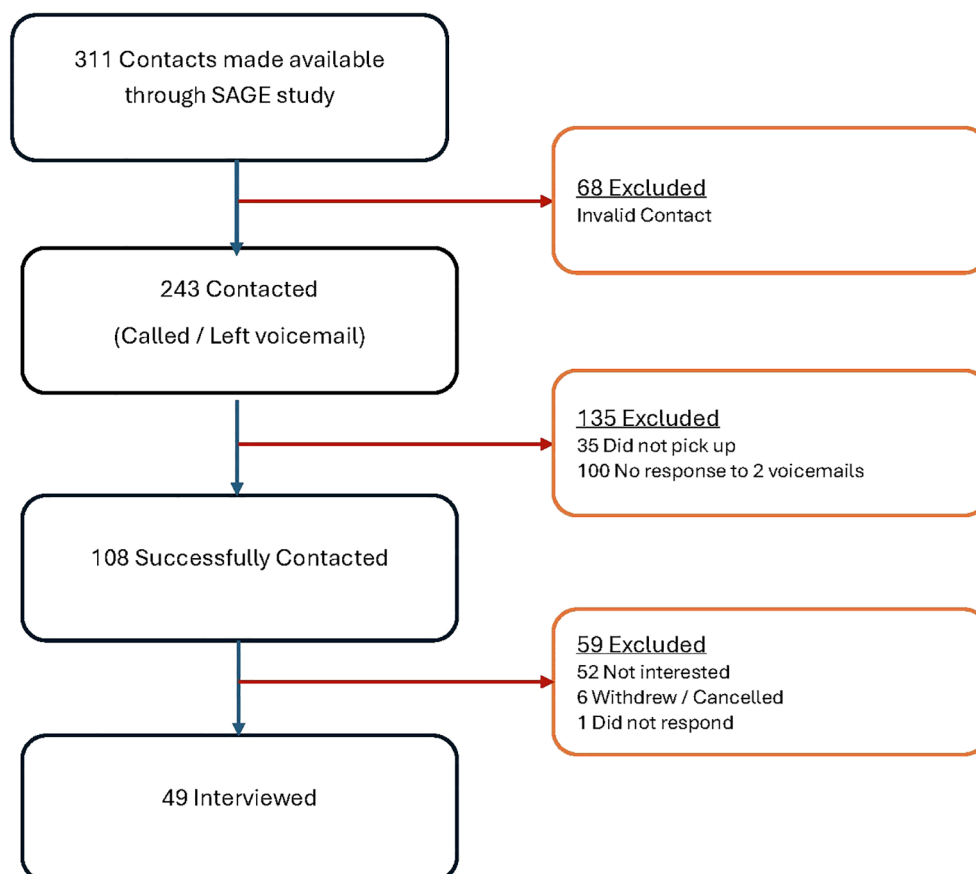


FIGURE 1  
Flowchart of participant retention through the study.

## 2.2 Sociodemographic and clinical neuropsychological measures

**Socio Demographic:** Sociodemographic information was made available from participant details in the parent survey (56) which included age, sex, race, marital-status and education.

**Cognitive Failures Questionnaire (CFQ)** (58) is a 25-item questionnaire of self-reported failures in perception, memory, and motor function. Responses are stable over a long period, tend to show positive correlation among questions, and positively correlated with the number of psychiatric symptoms reported on the Mental Health Quotient (MHQ). A high CFQ score was defined as greater than or equal to 43, which was associated previously with neurosarcoidosis (63).

The Modified Telephone Interview for Cognitive Status (TICS-m) (57, 64) is a concise questionnaire adapted to be used over the phone for screening dementia or mild cognitive impairment (MCI). The questions on TICS-m target attention, orientation, language, and learning and memory like the Mini-Mental Status Exam (MMSE). The modified version includes delayed recall for better detection of memory deficits compared to the original. We administered the 12-item Telephone Interview of Cognitive Status (TICS-m) (57) which is a modified version of 11-item Telephone Interview of Cognitive Status (TICS) (65). Item 10 of TICS and TICS-m which is “With your finger tap five times on the part of the phone you speak into”, was replaced by “Clap your hands five times” where the interviewer could see or hear the participant clapping on the Zoom/phone. Two studies offer detailed comparison of various versions of TICS to assess consistency of cutoffs (66, 67). A meta-analysis recommended a cut-off score of <31 on the TICS, providing 92% sensitivity and 66% specificity for detecting dementia (68). A cut-off score of 30/31 with 85% sensitivity and 83% specificity was suggested for the TICS-m assessment (57) and goes on to suggest that cutoff of 31/32 produces similar discrimination. This is also supported by (69).

## 2.3 Audio preprocessing

The audio recordings were converted to wav format. These recordings were used in entirety to extract acoustic features with an assumption that interviewer utterances comprised of only a small part of the recording and were consistent. Digital recordings of Zoom or phone-based interviews were obtained in m4a or mp3 formats respectively and then converted to wav format using *ffmpeg* (70). The audio recordings included the interviewer’s prompts which were short and generally uniform across the sample.

## 2.4 Features

**Acoustic Features:** We used the concise and curated feature set “eGeMAPSv02” suitable for clinical speech analysis, and described in Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing (71). The GeMAPS is a minimal

set of voice acoustic features that are deemed suitable for both voice (trait) and mood (state) related research (3) and made accessible through the Python openSMILE library (audeERING GmbH), and validated for this purpose (35–37) relevant among these are features of successive formants, F0, F1, and F2, the successive peaks in the frequency spectrum, and voice shimmer. Details on the acoustic features can be found in (71, 72), (Supplementary, Appendix B).

**Psycholinguistic Features:** The recordings were transcribed using whisper (<https://whisperapi.com/speech-to-text-free-tool>). The transcribed text was then manually tagged with “Q:” tags for interviewer utterances and “A:” for the participant utterances. These tags were used to extract participant utterances for further LIWC analysis. LIWC uses a word spotting paradigm as used in Linguistic Inquiry through Word Counting (LIWC) (43), considered to be the gold standard in NLP for psychology applications. The approach emphasizes content over syntax. The technique typically uses a handcrafted dictionary, that has assigned words to categories, to count words in the text that fall in each category. We extracted the full set of 119 LIWC 2022 features described by (43) for each transcript in our dataset. Transformer based approaches such as BERT (73) or large language models (LLMs) require large to huge amounts of data while offering little insights into the relevance of features. Further, they use only textual data and do not incorporate audio features. Finally, the performance as assessed through F1-score was low (74). Therefore, these approaches were not investigated.

**Demographic Features:** Age, sex, race and years of education were considered in the feature set.

## 2.5 Feature ranking

Gini ranking (75) was used to rank top 20 features. The contribution of the features in discriminating the cognitively impaired (TICS-m ≤ 31) from those who were not, and those with significant self-assessed cognitive complaints (CFQ >42) from those without, was assessed by first limiting the feature set to include only the top 20 features. The vast total number of features (119 linguistic features, 88 acoustic features from derived from each of the two tasks (Supplementary Figure A1) and 5 demographic features) far exceed the sample size (n=49), creating a strong likelihood of overfitting.

## 2.6 Machine learning models

We used ANN with ReLU, logistic and tanh activation functions, Support vector machine (SVM) and k neighbors (kNN) based models with specified hyperparameters (Supplementary, Appendix C). The aim was to assess features and tasks in their utility, rather than to obtain the best fine-tuned classification model. Transformer based approaches such as BERT (73) or large language models (LLMs) require larger data sets and are also limited in respect to interpretability of features. Further, they use only textual data and do not incorporate audio features. Therefore, these approaches were not applied in the current study.

Several model performance metrics were evaluated such as Area Under the Curve (AUC), F1 score, precision, recall and specificity and the details on how these measures are computed are available in the supplementary material (Supplementary, Appendix D). AUC provides an overall picture of model's ability to classify beyond randomness on a range on operating points, precision and recall may provide a direct measure for comparison in specific applications (e.g. higher recall may be desirable over precision in medical screening applications). We use the harmonic mean of precision and recall, the F1-score, to rate feature-set and model performances.

### 3 Results

The sample ranged from 61 to 93 years in age at the time of interviews (Table 1). Participants were mostly white 42(85.7%), female 26(53.1%) and married or cohabitating (n=33, 67.3%), with high education (mean years 15.6, SD 2.2). Cognitive functioning varied among participants as indicated by TICS-m scores in the range 27–45 with a mean score 35.1 (SD=4.1), with 10 (or 20.4%) falling below the cutoff ( $\leq 31$ ). Subjective failures of cognition as reported in the CFQ were in the range 31–81, with a mean score 55.5 (SD=11.5), with 43 (or 87.8%) reporting above the cutoff ( $>42$ ). The Zoom recordings were compared with phone recordings and no discernible differences in quality were observed. No recording got excluded from analysis due to poor quality. The CT task duration was unimodal, lasting a minute in most cases (Mean= 1.0 minutes, SD=0.5), while the AG task duration was bimodal with one peak a little over a minute and the other about two and a half minutes (Mean=2.2 minutes, SD=1.0) (Figure 2). Supplementary Figure B1 shows effect of computing features on longer time scales may cause

some loss of momentary information. CFQ and TICS-m scores for individuals were not correlated (Figure 3).

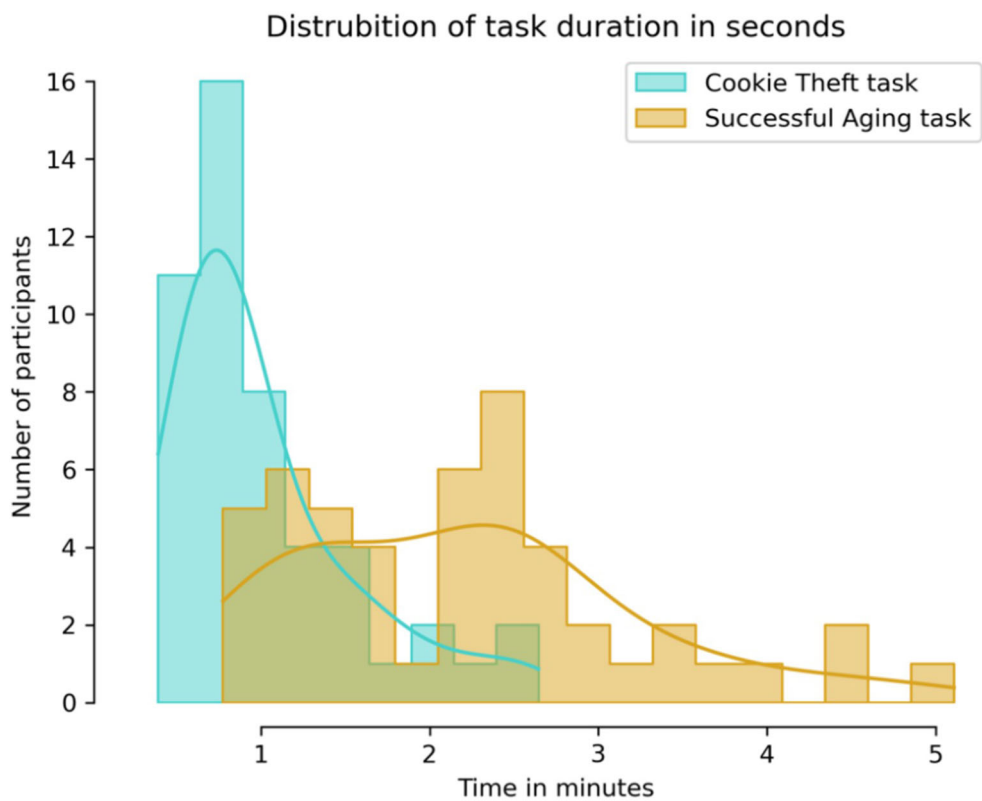
When predicting subjective cognition based on elevated CFQ score, ML models performed best with all combined features including acoustic and content-based features were combined from CT and AG tasks. Using single-learner models an F1-score of 0.83, precision of 0.85 with an AUC of 0.88 was achieved (Table 2A). Similarly, when predicting objective cognition based on TICS-m cutoff, ML models performed best with all combined features including acoustic and content-based features from both CT and AG tasks. We achieved an F1-score and precision of 0.92 with an AUC of 0.90 (Table 2A). Performance improved for some targets when ensemble methods were used (Table 2B). All models and targets yielded AUC equal to or above 0.76, with TICS-m classification approaching 0.90; these values are sufficiently higher than 0.5 (random) suggesting the identified features have great value in classification. Precision, recall and their harmonic mean, the F1 score, too, support our claim of excellent classification. Tables 3, 4 show the best features identified through GINI-index that were used for classification.

The AG and CT tasks, however, contributed differently to discrimination of elevation in subjective cognitive complaints as well as low scores on objective cognition measure. In predicting CFQ, seven of the top 10 features were derived from the AG instead of the CT task. In contrast, when predicting TICS-m scores, seven of the top 10 features were derived from the CT instead of the AG task. Further, the specificity for CFQ target was poor with features derived only from the CT task, suggesting limited suitability of the task for the CFQ classification. The generally lower specificity for the CFQ target, we believe, stems from the fact that 87.76% of our sample had CFQ above the cutoff (Table 1), and the models were eager to classify samples into the category. As expected, using

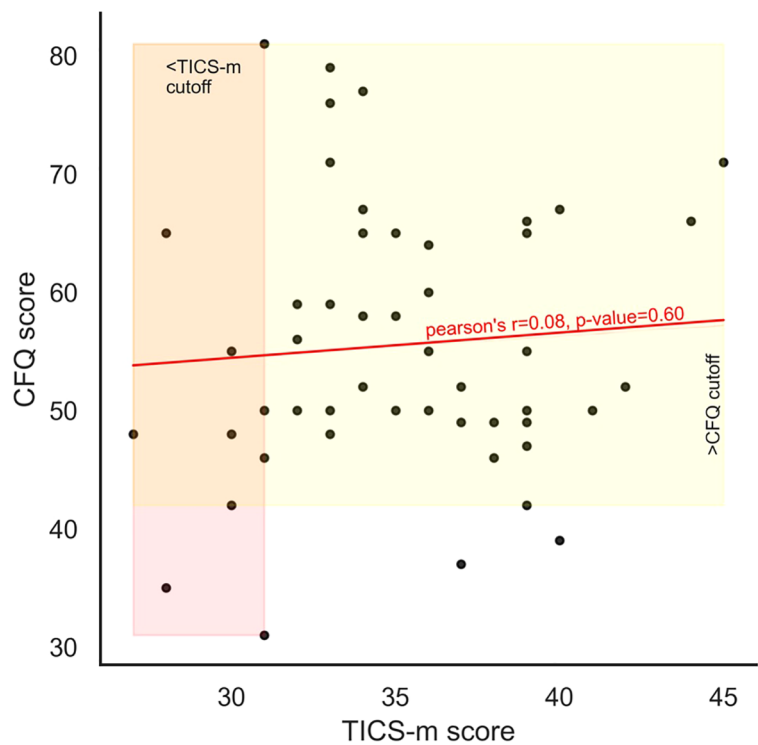
TABLE 1 Demographic and clinical characteristics.

Characteristics	Specification	Mean (Std. Dev); Min-Max	N(%)
Age		76.9(8.5); 61.2-93.8	
Sex	Female		26(53.1%)
Race	White		42(85.7%)
	Hispanic		4(8.16%)
	Other		3(6.12%)
Education	Number of years	15.6(2.2); 11.0-18.0	
Marital Status	Currently Married/Cohabitating		33(67.3%)
	Never married/divorced/separated		8(16.33%)
	Widowed		7(14.29%)
	Other		1(2.04%)
CFQ score		55.5(11.5); 31.0-81.0	
	CFQ > 42		43(87.8%)
TICS-m score		35.1(4.1); 27.0-45.0	
	TICS-m $\leq 31$		10(20.4%)





**FIGURE 2**  
Distribution of task durations in minutes. CT task was generally completed within a minute, while AG task durations were bi-modal, approximately one half taking a little over a minute, and the other about twice as much.



**FIGURE 3**  
Objective cognition as measured using TICS-m and subjective cognition as measured using CFQ show no correlation.

TABLE 2A Best performing single estimators/learners.

Target	Features	Model	AUC	F1	Precision	Recall	Specificity
CFQ	All	ANN	0.88	0.83	0.85	0.82	0.54
CFQ	AG and CT	ANN	0.86	0.82	0.83	0.82	0.40
CFQ	AG	ANN	0.76	0.82	0.83	0.82	0.40
CFQ	CT	kNN	0.80	0.82	0.77	0.88	0.12
Tics-m	All	ANN	0.90	0.92	0.92	0.92	0.83
Tics-m	AG and CT	ANN	0.90	0.90	0.90	0.90	0.82
Tics-m	AG	ANN	0.76	0.76	0.76	0.80	0.35
Tics-m	CT	ANN	0.88	0.80	0.81	0.80	0.65

TABLE 2B Best performing models when ensemble methods were included.

Target	Features	Model	AUC	F1	Precision	Recall	Specificity
CFQ	All	AdaBoost	0.81	0.92	0.92	0.92	0.70
CFQ	AG and CT	AdaBoost	0.71	0.88	0.88	0.88	0.55
CFQ	AG	AdaBoost	0.71	0.88	0.88	0.88	0.55
CFQ	CT	Random Forest	0.80	0.89	0.89	0.89	0.56
Tics-m	All	ANN	0.90	0.92	0.92	0.92	0.83
Tics-m	AG and CT	ANN	0.90	0.90	0.90	0.90	0.82
Tics-m	AG	XGBoost	0.83	0.81	0.82	0.84	0.44
Tics-m	CT	ANN	0.88	0.80	0.81	0.80	0.65

AG- Acoustic, linguistic features related to aging related questions.  
CT- Acoustic, linguistic features related to Cookie Theft picture description.  
ALL- All combined features (AG, CT and Demographics).  
AUC- Area under the curve.  
F1- F score.  
ANN- Artificial neural network, ReLu.  
kNN- k nearest neighbour.  
XGBoost- Gradient Boosting.

the top ranked features from combined set yielded the best classification; F1 of 0.92 for the TICS-m target and 0.83 for the CFQ (Table 2A). Demographic features (age, sex, race, education and marital status) were of little consequence (Tables 3, 4). CT derived features were more relevant to TICS-m classification while AG derived features were of greater value to the CFQ classification. Importance by the type of features, openSMILE (acoustic) vs. LIWC (content) as represented among top 10, in prediction of the two targets were split evenly. When predicting CFQ based subjective cognition, five were derived from openSMILE. Of these five, three acoustic features were loudness and shimmer (changes in loudness) related (Table 3). The LIWC features derived from AG task encoded tentativeness and leisure in the reminiscence. The CT task derived LIWC features encoded tone and negation (Supplementary, Figure B2). When predicting TICS-m based objective cognition, four features were derived from openSMILE, three of which encode aspects of formant frequencies, F0 (lowest) to F2 (highest). The AG task derived LIWC features encoded work and affiliation (Supplementary, Figure B3). The remaining features encoded negative emotion and anxiety.

## 4 Discussion

Notwithstanding limitations, we found several potentially important aspects of different speech collection modalities and predictive accuracy across subjective and objective cognition in older adults. Our hypotheses were supported; acoustic and linguistic markers derived from either the cookie theft and open-ended modalities achieved acceptable accuracy in predicting objective and subjective cognition. However, the features derived from the cookie theft task were more predictive of objective measure of cognition, whereas the more open ended successful aging questions derived features were more predictive of subjective complaints. In all models, there was a relatively balanced proportion of acoustic versus linguistic markers in prediction of both objective and subjective cognition, with little overlap in top features across prediction of subjective or objective cognition. Therefore, different speech elicitation modalities (cookie theft, open-ended etc.) may have different strengths in predicting objective and subjective cognition, and the combination of acoustic and linguistic markers may be optimal in predicting either outcome.

TABLE 3 Top features for predicting Cognitive Failures Questionnaire (CFQ) score > 42.

CFQ (Subjective experience of cognitive aging)						
Feature rank	Task type	Feature name - description	Information gain	Gini	Student's t	P value
1	AG OpenSmile	<b>shimmerLocaldB_sma3nz_amean</b> – Loudness (Variation)	0.195	0.058	[+] 3.062	0.004*
2	AG OpenSmile	<b>spectralFluxV_sma3nz_amean</b>	0.195	0.058	[+] 2.973	0.005*
3	AG LIWC	<b>Tentat</b> - Tentative (E.g. if, or, any, something)	0.195	0.058	[+] 1.919	0.063
4	CT LIWC	<b>Tone</b> - Emotional tone (Degree of positive (negative) tone)	0.144	0.043	[-] 0.895	0.388
5	CT LIWC	<b>Comm</b> - Communication (E.g. said, say, tell, thank)	0.116	0.043	[-] 1.665	0.150
6	CT LIWC	<b>Negate</b> - Negations (E.g. not, no, never, nothing)	0.167	0.040	[-] 1.796	0.125
7	AG LIWC	<b>Leisure</b> - (E.g. game, fun, play, party)	0.114	0.040	[-] 0.997	0.348
8	AG OpenSmile	<b>loudness_sma3_percentile20.0</b> – Loudness (Baseline)	0.152	0.038	[+] 3.084	0.003*
9	AG OpenSmile	<b>loudness_sma3_stddevFallingSlope</b> – Loudness (Rolloff)	0.152	0.038	[+] 2.810	0.009*
10	AG OpenSmile	<b>mfcc2_sma3_stddevNorm</b>	0.152	0.038	[-] 0.743	0.462
11	CT OpenSmile	<b>F0semitoneFrom27.5Hz_sma3nz_percentile20.0</b> – Formant 0 (Baseline frequency)	0.152	0.038	[+] 0.414	0.688
12	CT OpenSmile	<b>shimmerLocaldB_sma3nz_amean</b> –Loudness (Variation)	0.152	0.038	[+] 1.089	0.291
13	CT OpenSmile	<b>MeanVoicedSegmentLengthSec</b> – Voiced Speech length (mean)	0.152	0.038	[+] 0.631	0.533
14	CT LIWC	<b>Analytic</b> - Analytical thinking (Metric of logical, formal thinking)	0.152	0.038	[+] 2.578	0.026*
15	CT OpenSmile	<b>jitterLocal_sma3nz_amean</b> – Frequency (Shifts)	0.147	0.037	[-] 0.728	0.488
16	CT LIWC	<b>Health</b> - (E.g. medic, patients, physician, health)	0.101	0.034	[-] 0.129	0.900
17	AG OpenSmile	<b>mfcc2_sma3_amean</b> – 2 <sup>nd</sup> Mel Cepstrum (Voice Timber)	0.141	0.034	[-] 4.946	0.001*
18	AG OpenSmile	<b>mfcc2V_sma3nz_amean</b> – 2 <sup>nd</sup> Mel Cepstrum (Voice Timber)	0.141	0.034	[-] 0.4531	0.001*
19	AG LIWC	<b>Allure</b> - words commonly used in successful ads and persuasive communications (76).	0.141	0.034	[-] 1.200	0.267
20	CT OpenSmile	<b>mfcc2_sma3_amean</b> – 2 <sup>nd</sup> Mel Cepstrum (Voice Timber)	0.141	0.034	[-] 4.783	0.001*

AG: Aging questions.  
CT: Cookie Theft picture description.  
[+]: positively related to CFQ score > 42.  
[-]: negatively related to CFQ score > 42.  
Feature names are in bold and a short description is provided in normal text.  
\*: Significant p-values < .05.

Our study contributes to a growing literature evaluating the linguistic and audio features derived specifically from the cookie theft picture description task as well as other brief structured cognitive tasks (77). Our study was different from many in that included people who were randomly selected from a population and evaluated the link to task performance rather than diagnostic characterization (e.g., MCI). A focused review on cookie theft task studies suggested richness of content words, conciseness of expression, and quantity of expression were greater among the control (31). A recent review proceeded to harmonize the linguistic feature nomenclature that abounds (into several 100s) in the literature, found that linguistic feature categories such as phonetic-prosody (breaks and repetitions in connected speech), lexical-semantic (meaning and grammar), speed, coherence and cohesion were very relevant in screening (78), these reviews did not include acoustic

features. We found that language suggestive of negative tone, use of numbers from the CT task, and work and affiliation from the AG task were relevant linguistic features. Acoustic features that captured formant frequencies were also discriminative. Our study evaluated the prediction of a global cognitive screening measures across domains (3), and future studies might employ a comprehensive neuropsychological battery to evaluate which acoustic and linguistic features align with different cognitive domains.

Our study is consistent with recent studies on speech analysis and objective cognition. A study (77) used audio features extracted using openSMILE and Wave2Vec2.0 (79) which is an alternative audio feature representation. The highest accuracies reported (84.8%) were from the interference and the number reading task while the interview and reading task provided lower accuracies in the 67%-78% range. While the performance of openSMILE and Wave2Vec



TABLE 4 Top features for Modified Telephone Interview for Cognitive Status (Tics-m) score  $\leq 31$ .

TICS-m (Objective assessment of cognitive aging)						
Feature rank	Task type	Feature name-description	Information gain	Gini	Student's t	P value
1	CT LIWC	<b>Number</b> - (E.g. one, two, first, once)	0.154	0.074	[+] 0.980	0.338
2	AG LIWC	<b>Affiliation</b> - (E.g. we, our, us, help)	0.183	0.073	[-] 2.311	0.032*
3	CT LIWC	<b>Feeling</b> - (E.g. feel, hard, cool, felt)	0.126	0.067	[+] 1.671	0.125
4	CT LIWC	<b>emo_neg</b> - negative emotion	0.132	0.066	[+] 0.747	0.467
5	CT OpenSmile	<b>F2frequency_sma3nz_amean</b> - Formant 2 (frequency)	0.166	0.064	[+] 1.935	0.073
6	CT OpenSmile	<b>F0semitoneFrom27.5Hz_sma3nz_stddevFallingSlope</b> - Formant 0 (rolloff variation)	0.166	0.064	[+] 2.957	0.013*
7	CT OpenSmile	<b>F0semitoneFrom27.5Hz_sma3nz_meanFallingSlope</b> - Formant 0 (rolloff mean)	0.166	0.064	[+] 2.729	0.021*
8	AG LIWC	<b>emo_anx</b> - anxiety	0.120	0.061	[+] 1.744	0.111
9	AG LIWC	<b>Work</b> - (E.g. work, school, working, class)	0.155	0.060	[-] 3.267	0.003*
10	CT OpenSmile	<b>slopeV0-500_sma3nz_amean</b>	0.164	0.060	[-] 1.600	0.134
11	CT OpenSmile	<b>F0semitoneFrom27.5Hz_sma3nz_pctrange0-2</b> - Formant 0 (range)	0.161	0.059	[+] 0.498	0.629
12	AG OpenSmile	<b>F1bandwidth_sma3nz_stddevNorm</b> - Formant 1 (bandwidth)	0.118	0.059	[-] 0.384	0.706
13	AG LIWC	<b>Perception</b> - (E.g. in, out, up, there)	0.152	0.055	[-] 0.214	0.835
14	AG OpenSmile	<b>mfcc2V_sma3nz_amean</b> - 2 <sup>nd</sup> Mel Cepstrum (Voice Timber)	0.149	0.053	[+] 0.876	0.399
15	AG OpenSmile	<b>F3frequency_sma3nz_amean</b> - Formant 3 (frequency)	0.149	0.053	[+] 1.066	0.304
16	CT LIWC	<b>Space</b> - E.g. in, out, up, there)	0.132	0.046	[-] 3.105	0.005*
17	CT LIWC	<b>BigWords</b> - Percent words 7 letters or longer	0.132	0.046	[-] 2.663	0.020*
18	CT LIWC	<b>Allnone</b> - (E.g. all, no, never, always)	0.110	0.045	[+] 1.893	0.082
19	AG OpenSmile	<b>mfcc2_sma3_stddevNorm</b> - 2 <sup>nd</sup> Mel Cepstrum (Voice Timber variation)	0.127	0.045	[+] 0.337	0.738
20	AG LIWC	<b>Conj</b> - (E.g. and, but, so, as)	0.124	0.044	[-] 1.069	0.307

AG: Aging questions.

CT: Cookie Theft picture description.

[+]: positively related to Tics-m score  $\leq 31$ .[-]: negatively related to Tics-m score  $\leq 31$ .

Feature names are in bold and a short description is provided in normal text.

\*: Significant p-values &lt; .05.

derived features were identical for the best case of interference task. These accuracies are about 5% lower than our best performances which we attribute to their simpler model choice of Support Vector Machine (SVM), and a lack of feature selection. The feature relevance was not examined, but the study reinforces our finding on different speech elicitation modalities where features derived from tasks of cognition are better predictors of objective cognition. Accuracies like ours were achieved in Chinese language Cookie Theft task with simpler audio features that encoded pauses and hesitation but included visual facial features (80). BERT based models that used transcriptions (only) of the cookie theft task achieved lower accuracies of about 84.8% for non-controls (81), suggesting acoustic features have additional and relevant information besides transcriptions only processing by BERT, a notion also embraced by a recently proposed dementia screening system (82).

Our study was novel in evaluating and applying speech analysis to the prediction of subjective cognition. Acoustic and linguistic markers were able to predict subjective cognition. Notably, the markers were generally different from those of objective cognition, with some overlap of linguistic markers of negative emotions. Recent reviews of longitudinal studies suggest a higher symptom burden on subjective cognition has predictive value for mild cognitive impairment (MCI) and dementia (83), while the symptoms themselves were associated with quality of life (84). Conversely, a younger subjective age was related to higher cognitive performance, and reduced depressive symptoms (85) (86), suggesting subjective cognition, quality of life, subjective age, depressive symptoms and longer term cognitive outcomes remain enmeshed (87). Other studies provide evidence that subjective cognition and depressive symptomology may be directly linked as higher cognitive failure

scores are associated with greater perceived psychological distress and affective disorders (13, 58, 88, 89), and momentary affect among healthy individuals (90). The association of negative emotions with subjective cognition is therefore not surprising. Subjective cognitive complaints are a component of MCI diagnoses, but a challenge is in potentially understanding the specificity of these experiences beyond affective symptoms. Furthermore, since our open-ended question likely elicited more affectively linked content, it is perhaps not surprising that open ended questions content was more linked to subjective compared to objective cognition. In the future, sentiment from NLP and audio features that encode emotions, such as shimmer, could play a role in disentangling symptoms from subjective complaints. The use of multimodal speech elicitation paradigms may help tease apart the subjective complaints tied to objective decline from that tied to affective symptoms. In the future, it would be important to understand the within person trajectories of acoustic and linguistic features and how they might change with subjective and objective cognition. Other linguistic features such as sentence complexity, vocabulary richness and attributes of grammar might be more stable and linked to crystallized knowledge, whereas features that are related to vocalization and sentiment may vary within people, and perhaps in conjunction with affective states.

Acoustic features implicated in subjective cognitive complaints in our study were shimmer, spectral flux and loudness, all derived from the AG task (among top 10, Table 3). In speech analysis, “shimmer” is an acoustic feature that quantifies the cycle-to-cycle variation in the amplitude (loudness) of a voice signal, as how much the loudness fluctuates between each vocal fold vibration. This is in accordance with other studies that have linked shimmer and loudness with emotions (38, 39, 91); and emotions having established link to SCC is in alignment with our previous finding that such complaints are mood dependent (90).

In contrast, key audio features implicated in objective cognition were all related to base and higher formant frequencies derived from the CT task. Formant frequencies were shown to undergo a predictable change under cognitive load (41). Such formant shifts (at a gross level) are manifested as a shift in pitch and Mel-frequency cepstral coefficients (mfcc), which was described as an invariant pattern of cognitive decline (92). There is a greater body of evidence supporting this finding (3, 93, 94). The probable explanation for fundamental frequency (F0) and resonant frequencies (Formants) to encode information about an individual's cognition stems from the mechanics of phonological motor planning and control of vocal speech production apparatus (95). Inclusion of F0, F1, and F2 formant features in analysis of interview prompts that require cognitive processing can be helpful in assessing individual cognitive capacities and as indicators of cognition decline (40).

Our study had several strengths including the focus on multiple modes of speech elicitation, prediction of both objective and subjective cognition, and inclusion of both acoustic and linguistic markers. There are some important limitations and, as such, this study's findings should be considered preliminary and require replication. For one, the sample size was small, and the demographic make-up of the sample was skewed toward white and

persons with high education. The sampling approach employed random-digital dialing (56) but we note that this is a subset of the original sample. Our study's outcomes included brief global screenings of cognitive ability and subjective cognition and so does not speak to the prediction of specific cognitive impairments or diagnoses (e.g., MCI). There is a myriad of potential prompts for elicitation of speech. In the future, data from a larger more diverse population (or integrable data sets from different populations) alongside a wider variety of prompts that are parameterized for variation in subjective or objective cognitive levels derived from normed data would help to specify prompts that produce audio and linguistic patterns linked to either subjective, objective cognition or both. The study is also cross sectional and does not speak to the stability of these findings, and did not include independent validation. In our current survey we did not have questions about subjective cognition from the perspective of caregivers. As such replication would be required to understand the robustness of these findings. Finally, NLP models used supervised approaches and generative or transformer models could provide additional accuracy.

As a basis for future work, next steps would include replication in larger sample and designing of prompts that elicit content predictive of objective versus subjective cognition. It would be helpful also to contrast people with and without objective cognitive impairments on the acoustic and linguistic predictors of subjective complaints. Further, the influence of mood and other factors on the stability of speech features would be useful, in particular via longitudinal study that might evaluate MCI conversion as an endpoint. A larger study sample on casual conversation would facilitate topic-modelling and clustering approaches for thematic analysis. Furthermore, understanding how speech markers evolve over time in concert with subjective and objective cognitive, as well as brain and other biological markers, would be highly informative. Analyzing such conversations using LLMs with ingrained reasoning and BERT based classification are natural next steps. Together, these findings are consistent with recent reviews indicating the potential for speech analysis in understanding cognitive aging, with our study indicating that this also extends to subjective cognitive decline.

## Data availability statement

The study/data is governed by University of California San Diego Human Research Protections Program (HRPP) rules and other contract. Clinical data or the code is not publicly available due to privacy concerns, including HIPAA regulations. For machine learning parameters access, qualified researchers may contact the corresponding author. Requests to access the datasets should be directed to vbadal@health.ucsd.edu.

## Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board (IRB) University of

California San Diego. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

VB: Supervision, Methodology, Software, Conceptualization, Investigation, Validation, Writing – review & editing, Formal analysis, Resources, Visualization, Writing – original draft, Funding acquisition, Project administration. CT: Writing – review & editing, Data curation. HB: Data curation, Writing – review & editing. DG: Resources, Project administration, Writing – review & editing. RD: Writing – review & editing, Data curation, Resources. AM: Writing – review & editing. AM: Writing – review & editing. EB: Writing – review & editing. EL: Investigation, Funding acquisition, Writing – review & editing. CD: Writing – review & editing, Funding acquisition, Supervision, Conceptualization, Investigation.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work is supported by Center for Healthy Aging and the Stein Institute for Research on Aging at UC San Diego 2023 Pilot Grant Program award to Varsha D. Badal.

## Acknowledgments

We thank the following persons for suggestions related to administrative, RedCap, and SAGE study, Stein Institute for

Research on Aging at UC San Diego: Vanessa L. Scott and Paula Smith.

## Conflict of interest

EB is an employee of International Business Machines Corporation (IBM).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsy.2025.1596132/full#supplementary-material>

## References

- Clarke N, Foltz P, Garrard P. How to do things with (thousands of) words: Computational approaches to discourse analysis in Alzheimer's disease. *Cortex*. (2020) 129:446–63. doi: 10.1016/j.cortex.2020.05.001
- Eyigöz E, Mathur S, Santamaria M, Cecchi G, Naylor M. Linguistic markers predict onset of Alzheimer's disease. *EClinicalMedicine*. (2020) 28:100583. doi: 10.1016/j.eclinm.2020.100583
- Badal VD, Reinen JM, Twamley EW, Lee EE, Fellows RP, Bilal E, et al. Investigating acoustic and psycholinguistic predictors of cognitive impairment in older adults: modeling study. *JMIR Aging*. (2024) 7:e54655. doi: 10.2196/54655
- 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 Psychiatrica Scandinavica*. (2014) 130:439–51. doi: 10.1111/acps.2014.130.issue-6
- Hong JY, Lee PH. Subjective cognitive complaints in cognitively normal patients with parkinson's disease: A systematic review. *J Movement Disord*. (2023) 16:1. doi: 10.14802/jmd.22059
- Mark RE, Sitskoorn MM. Are subjective cognitive complaints relevant in preclinical Alzheimer's disease? A review and guidelines for healthcare professionals. *Rev Clin Gerontology*. (2013) 23:61–74. doi: 10.1017/S0959259812000172
- Jacob L, Haro JM, Koyanagi A. Physical multimorbidity and subjective cognitive complaints among adults in the United Kingdom: a cross-sectional community-based study. *Sci Rep*. (2019) 9:12417. doi: 10.1038/s41598-019-48894-8
- Ávila-Villanueva M, Rebollo-Vázquez A, Ruiz-Sánchez De León JM, Valenti M, Medina M, Fernández-Blázquez MA. Clinical relevance of specific cognitive complaints in determining mild cognitive impairment from cognitively normal states in a study of healthy elderly controls. *Front Aging Neurosci*. (2016) 8:233. doi: 10.3389/fnagi.2016.00233
- Molinuevo JL, Rabin LA, Amariglio R, Buckley R, Dubois B, Ellis KA, et al. Implementation of subjective cognitive decline criteria in research studies. *Alzheimer's Dementia*. (2017) 13:296–311. doi: 10.1016/j.jalz.2016.09.012
- Bassett SS, Folstein MF. Memory complaint, memory performance, and psychiatric diagnosis: a community study. *J geriatric Psychiatry Neurol*. (1993) 6:105–11. doi: 10.1177/089198879300600207
- Pfund GN, Spears I, Norton SA, Bogdan R, Oltmanns TF, Hill PL. Sense of purpose as a potential buffer between mental health and subjective cognitive decline. *Int Psychogeriatrics*. (2022) 34:1045–55. doi: 10.1017/S1041610222000680
- Hill NL, Mogle J, Wion R, Munoz E, Depasquale N, Yevchak AM, et al. Subjective cognitive impairment and affective symptoms: a systematic review. *Gerontologist*. (2016) 56:e109–27. doi: 10.1093/geront/gnw091
- Markova H, Andel R, Stepankova H, Kopecek M, Nikolai T, Hort J, et al. Subjective cognitive complaints in cognitively healthy older adults and their relationship to cognitive performance and depressive symptoms. *J Alzheimer's Dis*. (2017) 59:871–81. doi: 10.3233/JAD-160970

14. Mogle J, Hill NL, Bhargava S, Bell TR, Bhang I. Memory complaints and depressive symptoms over time: A construct-level replication analysis. *BMC geriatrics*. (2020) 20:1–10. doi: 10.1186/s12877-020-1451-1
15. Bell TR, Beck A, Gillespie NA, Reynolds CA, Elman JA, Williams ME, et al. A traitlike dimension of subjective memory concern over 30 years among adult male twins. *JAMA Psychiatry*. (2023) 80(7):718–727. doi: 10.1001/jamapsychiatry.2023.1004
16. Fauconnier G. *Mappings in thought and language*. Cambridge UP: Cambridge University Press (1997). doi: 10.1017/CBO9781139174220
17. Mehrabian A. *Nonverbal communication*. Routledge (2017).
18. Huang C-F, Akagi M. A three-layered model for expressive speech perception. *Speech Communication*. (2008) 50:810–28. doi: 10.1016/j.specom.2008.05.017
19. Yoshii K, Kimura D, Kosugi A, Shinkawa K, Takase T, Kobayashi M, et al. Screening of mild cognitive impairment through conversations with humanoid robots: Exploratory pilot study. *JMIR Formative Res*. (2023) 7:e42792. doi: 10.2196/42792
20. Tanaka H, Adachi H, Ukita N, Ikeda M, Kazui H, Kudo T, et al. Detecting dementia through interactive computer avatars. *IEEE J Trans Eng Health Med*. (2017) 5:1–11. doi: 10.1109/JTEHM.2017.2752152
21. Sumali B, Mitsukura Y, Liang K-C, Yoshimura M, Kitazawa M, Takamiya A, et al. Speech quality feature analysis for classification of depression and dementia patients. *Sensors*. (2020) 20:3599. doi: 10.3390/s20123599
22. Yamada Y, Shinkawa K, Shimmei K. Atypical repetition in daily conversation on different days for detecting alzheimer disease: evaluation of phone-call data from a regular monitoring service. *JMIR Ment Health*. (2020) 7:e16790. doi: 10.2196/16790
23. Khodabakhsh A, Yesil F, Guner E, Demiroglu C. Evaluation of linguistic and prosodic features for detection of Alzheimer's disease in Turkish conversational speech. *EURASIP J Audio Speech Music Process*. (2015) 2015:1–15. doi: 10.1186/s13636-015-0052-y
24. Roark B, Mitchell M, Hosom J-P, Hollingshead K, Kaye J. Spoken language derived measures for detecting mild cognitive impairment. *IEEE Trans audio speech Lang Process*. (2011) 19:2081–90. doi: 10.1109/TASL.2011.2112351
25. Tóth L, Hoffmann I, Gosztolya G, Vincze V, Szatlóczi G, Bánréti Z, et al. A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech. *Curr Alzheimer Res*. (2018) 15:130–8. doi: 10.2174/1567205014666171121114930
26. Pastoriza-Dominguez P, Torre IG, Dieguez-Vide F, Gómez-Ruiz I, Geladó S, Bello-López J, et al. Speech pause distribution as an early marker for Alzheimer's disease. *Speech Communication*. (2022) 136:107–17. doi: 10.1016/j.specom.2021.11.009
27. Kato S, Homma A, Sakuma T, Nakamura M. (2015). Detection of mild Alzheimer's disease and mild cognitive impairment from elderly speech: Binary discrimination using logistic regression, in: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, 25–29 August 2015. pp. 5569–72. IEEE. doi: 10.1109/EMBC.2015.7319654
28. Goodglass H, Kaplan E. The assessment of aphasia and related disorders. (1983).
29. Cummings L. Describing the cookie theft picture: Sources of breakdown in Alzheimer's dementia. *Pragmatics Soc*. (2019) 10:153–76. doi: 10.1075/ps.17011.cum
30. Kokkinakis D, Fors KL, Fraser KC, Nordlund A. (2018). A Swedish cookie-theft corpus, in: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki (Japan), 7–12 May 2018.
31. Mueller KD, Hermann B, Mecollari J, Turkstra LS. Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *J Clin Exp Neuropsychol*. (2018) 40:917–39. doi: 10.1080/13803395.2018.1446513
32. Keator LM, Faria A, Kim T. Cookie theft picture description: linguistic and neural correlates. *Acad Aphasia 56th Annu Meeting*. (2018) 10:21–3. doi: 10.3389/conf.fnhum.2018.228.00097
33. Berube SK, Goldberg E, Sheppard SM, Durfee AZ, Ubellacker D, Walker A, et al. An analysis of right hemisphere stroke discourse in the modern cookie theft picture. *Am J speech-language Pathol*. (2022) 31:2301–12. doi: 10.1044/2022\_AJSLP-21-00294
34. Williams C, Thwaites A, Buttery P, Geertzen J, Randall B, Shafto MA, et al. The Cambridge cookie-theft corpus: A corpus of directed and spontaneous speech of brain-damaged patients and healthy individuals. *LREC*. (2010), 2824–30. <https://pure.qub.ac.uk/en/publications/the-cambridge-cookie-theft-corpus-a-corpus-of-directed-and-spontaneous-speech>
35. Eyben F, Wöllmer M, Schuller B. (2009). OpenEAR—introducing the Munich open-source emotion and affect recognition toolkit, in: *2009 3rd international conference on affective computing and intelligent interaction and workshops*, Amsterdam, Netherlands, 10–12 September 2009. pp. 1–6. IEEE. doi: 10.1109/ACII.2009.5349350
36. Schuller B, Eyben F, Rigoll G. (2007). Fast and robust meter and tempo recognition for the automatic discrimination of ballroom dance styles, in: *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, Honolulu, HI, USA, 15–20 April 2007. pp. I-217–I-220. IEEE. doi: 10.1109/ICASSP.2007.366655
37. Schuller B, Steidl S, Batliner A. The interspeech 2009 emotion challenge. (2009). doi: 10.21437/Interspeech.2009
38. Brockmann M, Drinnan MJ, Storck C, Carding PN. Reliable jitter and shimmer measurements in voice clinics: the relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *J voice*. (2011) 25:44–53. doi: 10.1016/j.jvoice.2009.07.002
39. Li X, Tao J, Johnson MT, Soltis J, Savage A, Leong KM, et al. (2007). Stress and emotion classification using jitter and shimmer features, in: *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, Honolulu, Hawaii, April 15–20, 2007. pp. IV-1081–IV-1084. IEEE.
40. Mahon E, Lachman ME. Voice biomarkers as indicators of cognitive changes in middle and later adulthood. *Neurobiol Aging*. (2022) 119:22–35. doi: 10.1016/j.neurobiolaging.2022.06.010
41. Yap TF, Epps J, Ambikairajah E, Choi EH. Formant frequencies under cognitive load: Effects and classification. *EURASIP J Adv Signal Process*. (2011) 2011:1–11. doi: 10.1155/2011/219253
42. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. *J Lang Soc Psychol*. (2010) 29:24–54. doi: 10.1177/0261927X09351676
43. Boyd R, Ashokkumar A, Seraj S, Pennebaker J. *The Development and Psychometric Properties of LIWC-22*. Austin, TX: University of Texas at Austin (2022).
44. Eichstaedt JC, Kern ML, Yaden DB, Schwartz HA, Giorgi S, Park G, et al. Closed- and open-vocabulary approaches to text analysis: A review, quantitative comparison, and recommendations. *psychol Methods*. (2021) 26:398. doi: 10.1037/met0000349
45. Sun J, Schwartz HA, Son Y, Kern ML, Vazire S. The language of well-being: Tracking fluctuations in emotion experience through everyday speech. *J Pers Soc Psychol*. (2020) 118:364. doi: 10.1037/pspp0000244
46. Asgari M, Kaye J, Dodge H. Predicting mild cognitive impairment from spontaneous spoken utterances. *Alzheimers Dement (N Y)*. (2017) 3:219–28. doi: 10.1016/j.trci.2017.01.006
47. Crossley SA, Kyle K, Mcnamara DS. Sentiment Analysis and Social Cognition Engine (SEANCE): An automatic tool for sentiment, social cognition, and social-order analysis. *Behav Res Methods*. (2017) 49:803–21. doi: 10.3758/s13428-016-0743-z
48. Shankar R, Bunde A, Mukhopadhyay A. A systematic review of natural language processing techniques for early detection of cognitive impairment. *Mayo Clinic Proceedings: Digital Health*. (2025), 100205. doi: 10.1016/j.mcpdig.2025.100205
49. Martínez-Nicolás I, Llorente TE, Martínez-Sánchez F, Meilán JJG. Ten years of research on automatic voice and speech analysis of people with Alzheimer's disease and mild cognitive impairment: a systematic review article. *Front Psychol*. (2021) 12:620251. doi: 10.3389/fpsyg.2021.620251
50. Slegers A, Filiu R-P, Montembeault M, Brambati SM. Connected speech features from picture description in Alzheimer's disease: A systematic review. *J Alzheimer's Dis*. (2018) 65:519–42. doi: 10.3233/JAD-170881
51. Gagliardi G. Natural language processing techniques for studying language in pathological ageing: A scoping review. *Int J Lang Communication Disord*. (2024) 59:110–22. doi: 10.1111/1460-6984.12870
52. Robin J, Harrison JE, Kaufman LD, Rudzicz F, Simpson W, Yancheva M. Evaluation of speech-based digital biomarkers: review and recommendations. *Digital Biomarkers*. (2020) 4:99–108. doi: 10.1159/000510820
53. Reichstadt J, Depp CA, Palinkas LA, Folsom DP, Jeste DV. Building blocks of successful aging: a focus group study of older adults' perceived contributors to successful aging. *Am J Geriatr Psychiatry*. (2007) 15:194–201. doi: 10.1097/JGP.0b013e318030255f
54. Badal VD, Nebeker C, Shinkawa K, Yamada Y, Rentscher KE, Kim H-C, et al. Do words matter? Detecting social isolation and loneliness in older adults using natural language processing. *Front Psychiatry*. (2021) 12. doi: 10.3389/fpsy.2021.728732
55. Badal VD, Graham SA, Depp CA, Shinkawa K, Yamada Y, Palinkas LA, et al. Prediction of loneliness in older adults using natural language processing: exploring sex differences in speech. *Am J Geriatr Psychiatry*. (2021) 29:853–66. doi: 10.1016/j.jagp.2020.09.009
56. Jeste DV, Savla GN, Thompson WK, Vahia IV, Glorioso DK, Martin AS, et al. Association between older age and more successful aging: critical role of resilience and depression. *Am J Psychiatry*. (2013) 170:188–96. doi: 10.1176/appi.ajp.2012.12030386
57. Welsh KA, Breitner JC, Magruder-Habib KM. Detection of dementia in the elderly using telephone screening of cognitive status. *Cogn Behav Neurol*. (1993) 6(2):103–10. <https://scholars.duke.edu/publication/756056>
58. Broadbent DE, Cooper PF, Fitzgerald P, Parkes KR. The cognitive failures questionnaire (CFQ) and its correlates. *Br J Clin Psychol*. (1982) 21:1–16. doi: 10.1111/j.2044-8260.1982.tb01421.x
59. Rast P, Zimprich D, Van Boxtel M, Jolles J. Factor structure and measurement invariance of the cognitive failures questionnaire across the adult life span. *Assessment*. (2009) 16:145–58. doi: 10.1177/1073191108324440
60. Lawrence CE, Dunkel L, McEver M, Israel T, Taylor R, Chiriboga G, et al. A REDCap-based model for electronic consent (eConsent): moving toward a more personalized consent. *J Clin Trans Sci*. (2020) 4:345–53. doi: 10.1017/cts.2020.30
61. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'neal L, et al. The REDCap consortium: building an international community of software platform partners. *J Biomed Inf*. (2019) 95:103208. doi: 10.1016/j.jbi.2019.103208
62. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inf*. (2009) 42:377–81. doi: 10.1016/j.jbi.2008.08.010



63. Voortman M, De Vries J, Hendriks CM, Elfferich MD, Wijnen PA, Drent M. Everyday cognitive failure in patients suffering from neurosarcoidosis. *Sarcoidosis Vascularitis Diffuse Lung Dis.* (2019) 36:2. doi: 10.36141/svdl.v36i1.7412
64. Cook SE, Marsiske M, McCoy KJ. The use of the Modified Telephone Interview for Cognitive Status (TICS-M) in the detection of amnesic mild cognitive impairment. *J Geriatric Psychiatry Neurol.* (2009) 22:103–9. doi: 10.1177/0891988708328214
65. Espeland MA, Rapp SR, Katula JA, Andrews LA, Felton D, Gaussoin SA, et al. Telephone interview for cognitive status (TICS) screening for clinical trials of physical activity and cognitive training: the seniors health and activity research program pilot (SHARP-P) study. *Int J Geriatr Psychiatry.* (2011) 26:135–43. doi: 10.1002/gps.v26.2
66. Fong TG, Fearing MA, Jones RN, Shi P, Marcantonio ER, Rudolph JL, et al. Telephone interview for cognitive status: Creating a crosswalk with the Mini-Mental State Examination. *Alzheimer's Dementia.* (2009) 5:492–7. doi: 10.1016/j.jalz.2009.02.007
67. Chappelle SD, Gigliotti C, Léger GC, Peavy GM, Jacobs DM, Banks SJ, et al. Comparison of the telephone-Montreal Cognitive Assessment (T-MoCA) and Telephone Interview for Cognitive Status (TICS) as screening tests for early Alzheimer's disease. *Alzheimer's Dementia.* (2023) 19:4599–608. doi: 10.1002/alz.v19.10
68. Elliott E, Green C, Llewellyn DJ, Quinn TJ. Accuracy of telephone-based cognitive screening tests: systematic review and meta-analysis. *Curr Alzheimer Res.* (2020) 17:460–71. doi: 10.2174/1567205017999200626201121
69. Knopman DS, Roberts RO, Geda YE, Pankratz VS, Christianson TJ, Petersen RC, et al. Validation of the telephone interview for cognitive status-modified in subjects with normal cognition, mild cognitive impairment, or dementia. *Neuroepidemiology.* (2010) 34:34–42. doi: 10.1159/000255464
70. Developers, F. ffmpeg tool (Version be1d324)(2016). Available online at: <http://ffmpeg.org> (Accessed February 1, 2025).
71. Eyben F, Scherer KR, Schuller BW, Sundberg J, André E, Busso C, et al. The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Trans Affect computing.* (2015) 7:190–202. doi: 10.1109/TAFFC.2015.2457417
72. Kamiloglu RG, Boateng G, Balabanova A, Cao C, Sauter DA. Superior communication of positive emotions through nonverbal vocalisations compared to speech prosody. *J Nonverbal Behav.* (2021) 45:419–54. doi: 10.1007/s10919-021-00375-1
73. Roshanzamir A, Aghajan H, Soleymani Baghshah M. Transformer-based deep neural network language models for Alzheimer's disease risk assessment from targeted speech. *BMC Med Inf Decision Making.* (2021) 21:1–14. doi: 10.1186/s12911-021-01456-3
74. Mao C, Xu J, Rasmussen L, Li Y, Adekanattu P, Pacheco J, et al. AD-BERT: Using pre-trained language model to predict the progression from mild cognitive impairment to Alzheimer's disease. *J Biomed Inf.* (2023) 144:104442. doi: 10.1016/j.jbi.2023.104442
75. S.D. Brown AJM. *Comprehensive chemometrics* (2009). Available online at: <https://www.sciencedirect.com/topics/mathematics/gini-index> (Accessed February 1, 2025).
76. Kannan R, Tyagi S. Use of language in advertisements. *English specific purposes World.* (2013) 37(13):1–10.
77. Braun F, Erzigkeit A, Lehfeld H, Hillemacher T, Riedhammer K, Bayerl SP. (2022). Going beyond the cookie theft picture test: Detecting cognitive impairments using acoustic features, in: *International Conference on Text, Speech, and Dialogue*, pp. 437–48. Cham: Springer International Publishing. doi: 10.1007/978-3-031-16270-1\_36
78. Richard AB, Lelandais M, Reilly KT, Jacquin-Courtois S. Linguistic markers of subtle cognitive impairment in connected speech: A systematic review. *J Speech Language Hearing Res.* (2024) 67:4714–33. doi: 10.1044/2024\_JSLHR-24-00274
79. Baevski A, Zhou Y, Mohamed A, Auli M. wav2vec 2.0: A framework for self-supervised learning of speech representations. *Adv Neural Inf Process Syst.* (2020) 33:12449–60.
80. Wang J, Gao J, Xiao J, Li J, Li H, Xie X, et al. A new strategy on Early diagnosis of cognitive impairment via novel cross-lingual language markers: a non-invasive description and AI analysis for the cookie theft picture. *medRxiv.* (2024) 2024–06. doi: 10.1101/2024.06.30.24309714
81. Guo Y, Li C, Roan C, Pakhomov S, Cohen T. Crossing the “Cookie Theft” corpus chasm: applying what BERT learns from outside data to the ADReSS challenge dementia detection task. *Front Comput Sci.* (2021) 3:642517. doi: 10.3389/fcomp.2021.642517
82. Zolnoori M, Zolnour A, Topaz M. ADscreen: A speech processing-based screening system for automatic identification of patients with Alzheimer's disease and related dementia. *Artif Intell Med.* (2023) 143:102624. doi: 10.1016/j.artmed.2023.102624
83. Earl Robertson F, Jacova C. A systematic review of subjective cognitive characteristics predictive of longitudinal outcomes in older adults. *Gerontologist.* (2023) 63:700–16. doi: 10.1093/geront/gnac109
84. Hill NL, McDermott C, Mogle J, Munoz E, Depasquale N, Wion R, et al. Subjective cognitive impairment and quality of life: a systematic review. *Int psychogeriatrics.* (2017) 29:1965–77. doi: 10.1017/S1041610217001636
85. Fernández-Ballbé Ó, Martín-Moratinos M, Saiz J, Gallardo-Peralta L, Barrón López De Roda A. The relationship between subjective aging and cognition in elderly people: A systematic review. *Healthcare.* (2023) 11(4):3115. doi: 10.3390/healthcare11243115
86. Alonso Debreczeni F, Bailey PE. A systematic review and meta-analysis of subjective age and the association with cognition, subjective well-being, and depression. *Journals Gerontology: Ser B.* (2021) 4:471–82, 76. doi: 10.1093/geronb/gbaa069
87. Kleineidam L, Wagner M, Guski J, Wolfsgruber S, Miebach L, Bickel H, et al. Disentangling the relationship of subjective cognitive decline and depressive symptoms in the development of cognitive decline and dementia. *Alzheimer's Dementia.* (2023) 19:2056–68. doi: 10.1002/alz.12785
88. Sullivan B, Payne TW. Affective disorders and cognitive failures: a comparison of seasonal and nonseasonal depression. *Am J Psychiatry.* (2007) 164:1663–7. doi: 10.1176/appi.ajp.2007.06111792
89. Payne TW, Schnapp MA. The relationship between negative affect and reported cognitive failures. *Depression Res Treat.* (2014) 2014:396195. doi: 10.1155/2014/396195
90. Badal VD, Campbell LM, Depp CA, Parrish EM, Ackerman RA, Moore RC, et al. Dynamic influence of mood on subjective cognitive complaints in mild cognitive impairment: A time series network analysis approach. *Int Psychogeriatrics.* (2024) 37:100007. doi: 10.1016/j.inpsyc.2024.100007
91. Yanushevskaya I, Gobl C, Ni Chasaide A. Voice quality in affect cueing: does loudness matter? *Front Psychol.* (2013) 4:335. doi: 10.3389/fpsyg.2013.00335
92. Favaro A, Dehak N, Thebaud T, Villalba J, Oh E, Moro-Velázquez L. Discovering invariant patterns of cognitive decline via an automated analysis of the cookie thief picture description task. In: *Proc. The Speaker and Language Recognition Workshop (Odyssey 2024)* (2024). p. 201–8.
93. Ding H, Lister A, Karjadi C, Au R, Lin H, Bischoff B, et al. Detection of mild cognitive impairment from non-semantic, acoustic voice features: the framingham heart study. *JMIR Aging.* (2024) 7:e55126. doi: 10.2196/55126
94. Nishikawa K, Akihiro K, Hirakawa R, Kawano H, Nakatoh Y. Machine learning model for discrimination of mild dementia patients using acoustic features. *Cogn Robotics.* (2022) 2:21–9. doi: 10.1016/j.cogr.2021.12.003
95. Boersma P. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proc institute phonetic Sci.* (1993) 17:97–110.