

15 **Key words:** Alzheimer’s Disease pathology; Alzheimer’s Disease; Logopenic variant primary
16 progressive aphasia; speech characteristics; automatic speech processing

17 **Abstract**

18 **INTRODUCTION:** In this study, we compared digital speech features of AD and lvPPA
19 patients in a biologically confirmed cohort and related them to specific neuropsychiatric test
20 scores and CSF proteins.

21 **METHODS:** We extracted language variables with automated lexical and acoustic pipelines
22 from oral picture descriptions of 44 AD and 21 lvPPA patients with autopsy or CSF confirmation
23 of AD pathology. We correlated distinct speech features with MMSE and BNT test scores and
24 CSF p-tau levels.

25 **RESULTS:** LvPPA patients produced fewer verbs, adjectives, and more fillers with lower
26 lexical diversity and higher pause rate than AD. Both groups showed some shared language
27 impairments compared with normal speakers.

28 **DISCUSSION:** Our speech measures captured differences in speech between the two
29 phenotypes. Also, shared speech markers were linked to the common underlying pathology. This
30 work demonstrates the potential of natural speech analysis in detecting underlying AD
31 pathology.

32

33 **Acknowledgements**

34 The authors thank the patients and caregivers who participated in this study. Funding: This study
35 was funded by grants from the National Institute of Health (AG066597, AG054519, NS109260,
36 P30 AG072979), Alzheimer's Association (AACSF-18-567131, AARF-D-619473, AARF-D-
37 619473-RAPID), Department of Defense (PR192041).

38 **Declaration of interest**

39 Dr. Grossman participates in clinical trials sponsored by Alector, Eisai and Biogen that are
40 unrelated to this study. He also receives research support from Biogen and Avid that is unrelated
41 to this study, and research support from NIH. Dr. Mark Liberman serves on the Scientific
42 Advisory Board for Baidu Research, USA, and is a co-editor of the Annual Review of
43 Linguistics. All other authors have nothing to disclose.

44

45 **1. Background**

46 Speech production is a complex behavior, involving coordinated activation of multiple brain
47 regions. Thus, examining speech production provides potential opportunities to identify
48 neurodegenerative disease markers that are sensitive to specific phenotypes. Since Alzheimer's
49 disease (AD) accounts up to 80% of patients with dementia,¹ much attention has been paid to
50 cognitive and linguistic profiling of AD. Language produced by amnesic AD patients was found
51 to be "empty" with abundance of nonspecific words, circumlocutions, and sparse content.²

52 Logopenic variant primary progressive aphasia (lvPPA) is a recently identified PPA variant,^{3,4}
53 which is an atypical, non-amnesic manifestation of AD,^{3,5-9} with over 50% of autopsied cases
54 associated with underlying AD pathology.^{9,10} Since the identification of this PPA variant, many
55 studies have been dedicated to characterizing its linguistic features. Previous studies have shown
56 that lvPPA patients speak slowly with impaired lexical access,¹¹ and have poor phonemic
57 discrimination¹² with limited auditory-verbal short-term memory, naming impairment,^{4,5} and
58 dysfluencies.¹³

59 Previous studies of neurodegenerative patients suggest that language features are useful as a
60 screening tool¹⁴⁻¹⁸, because speech is easy to collect non-invasively, and it is sensitive to
61 cognitive impairments. Despite the shared pathology of AD and lvPPA, previous studies have
62 focused on the linguistic profiling of these two syndromes separately. With few comparative
63 studies, this leaves an important gap in the literature. Most previous quantitative work has
64 focused on measures such as the Boston Naming Test (BNT) that assess lexical retrieval during
65 confrontation naming of an object. However, descriptions of natural, connected speech in AD
66 and lvPPA are frequently informal. In this study, we identified similarities and differences

67 between AD and lvPPA patients with biological confirmation of underlying AD pathology, by
68 analyzing digitized, natural speech samples with reliable and reproducible automated methods.
69 Based on previous studies, we hypothesized that lvPPA patients would produce more dysfluent
70 speech with more limited lexical content than AD patients. We also hypothesized that AD and
71 lvPPA patients would share linguistic similarities, including decreased speech production. We
72 associated language variables with clinical test scores and CSF analytes for additional validation
73 and specific mechanistic clarification.

74 **2. Methods**

75 **2.1 Participants**

76 We examined oral picture descriptions that were collected from 93 participants in the
77 Department of Neurology at the Hospital of the University of Pennsylvania. Forty-four
78 participants had amnesic AD, and 21 were lvPPA patients with confirmed underlying AD
79 pathology, based on autopsy (n=15; 4 lvPPA and 11 AD patients) or on cerebrospinal fluid
80 (CSF) analytes levels (n=50; 17 lvPPA and 33 AD patients; phosphorylated Tau (p-Tau)/beta-
81 amyloid 42 ($A\beta$) ≥ 0.09 ¹⁹ and total Tau/ $A\beta$ ≥ 0.34 ²⁰). Twenty-eight matched elderly healthy
82 controls (HC) were included as a control group. Participants with other neurological, psychiatric,
83 or medical conditions that could affect cognition were excluded.

84 Table 1 shows the demographic and clinical characteristics of the participants. The groups did
85 not differ in age, sex, or education level. The patient groups did not differ from each other in
86 disease duration ($p=0.45$), the Mini-Mental State Exam (MMSE; $p=0.12$), BNT ($p=0.32$), CSF p-
87 Tau ($p=0.81$), $A\beta$ ($p=0.36$), or p-Tau/ $A\beta$ ratio levels ($p=0.66$).

88 Table 1: Demographic and clinical characteristics of participants.

	AD (N=44)	lvPPA (N=21)	HC (N=28)	p value
Age (years)	62.0 (8.0)	64.1 (8.2)	65.9 (5.9)	0.098
Education (years)	16.0 (2.6)	16.1 (3.2)	15.9 (2.6)	0.965
Sex				0.986
Female	24 (54.5%)	11 (52.4%)	15 (53.6%)	
Male	20 (45.5%)	10 (47.6%)	13 (46.4%)	
Disease Duration (years)	3.7 (2.6)	3.4 (1.5)	NA	
MMSE*	21.1 (4.8)	23.0 (4.2)	29.2 (1.0)	< 0.001
BNT*	20.1 (8.4)	18.6 (8.6)	27.9 (2.7)	< 0.001
CSF p-Tau	38.5 (20.1)	38.4 (16.0)	NA	
CSF Aβ	139.9 (36.2)	135.0 (26.2)	415.0 (NA)	< 0.001
CSF p-Tau/Aβ	0.3 (0.2)	0.3 (0.2)	NA	

89 NOTE: *AD and lvPPA patients did not differ in MMSE or BNT.

90 Abbreviations: AD = Alzheimer's disease, lvPPA = logopenic variant primary progressive

91 aphasia, HC = healthy controls; MMSE = Mini-mental State Exam; BNT = Boston Naming Test;

92 CSF = Cerebrospinal fluid; p-Tau = phosphorylated Tau; A β = beta-amyloid 42.

93

94 **2.2 Data collection**

95 To optimize reliability and reproducibility, we digitally recorded the participants' descriptions of
96 the Cookie Theft picture from the Boston Diagnostic Aphasia Examination.²¹ Descriptions were
97 about 1 minute long. Recordings were orthographically transcribed. The earliest recording of
98 each participant was analyzed with our lexical and acoustic pipelines, as described below.

99 **2.3 Lexical pipeline**

100 We automatically tagged the part-of-speech (POS) category of all tokens, using spaCy²² with its
101 large language model ('en_core_web_lg') for English. The number of tense-inflected verbs was
102 calculated by summing the number of modal auxiliaries, past-tense, and present-tense verbs.
103 Dysfluency markers, including fillers, repetitions, and partial words, were counted separately.
104 The count of each POS category, tense-inflected verbs, and dysfluency markers were converted
105 to counts per 100 words, controlling for the total number of words per participant.

106 We rated each word for concreteness,²³ semantic ambiguity,²⁴ frequency,²⁵ age of acquisition
107 (AoA),²⁶ familiarity,²⁶ and word length by the number of phonemes,²⁷ based on published norms.
108 We calculated the mean scores of these measures for content words (nouns, verbs, adjectives,
109 and adverbs) per participant.

110 Lastly, we measured lexical diversity, using the moving-average type-token ratio (MATTR),²⁸
111 which has been described as one of the most reliable measures for calculating lexical diversity.²⁹
112 The window length was set at 15 words. Detailed description of the lexical pipeline and
113 validation of the POS tagging accuracy has been previously published.³⁰

114 **2.4 Acoustic pipeline**

115 We employed an in-house speech activity detector (SAD) to segment audio recordings into
116 speech segments and silent pauses. We visually reviewed the segments, and non-speech
117 segments at the beginning and end of each recording and interviewer's prompts were excluded
118 from the analysis.

119 Using the SAD output, we calculated five duration-related measurements, including mean speech
120 segment duration, mean pause segment duration, total speech time, percent of speech, and pause
121 rate per minute. We summed the number of syllables of all words from the published norm²⁷ and
122 computed articulation rate as the number of syllables per second.

123 Additionally, we pitch-tracked all speech segments with Praat.³¹ To normalize physiological
124 differences in pitch (f_0), we converted the pitch values from Hz to semitones (st) using the 10th
125 pitch percentile of each participant as a baseline: $st=12*\log_2(f_0/\text{baseline } f_0)$. We used the
126 converted 90th percentile as a measure of the pitch range of each speaker. Detailed description of
127 the acoustic pipeline has been published previously,³² and the list of all analyzed features is
128 included in the Appendix.

129 **2.5 CSF analysis**

130 Forty-two (30 AD and 12 lvPPA) patients had CSF biomarkers collected within one year of the
131 Cookie Theft recording (mean interval = 4.7 months \pm 3.5), including A β and p-Tau. This subset
132 did not differ demographically or clinically from the larger group of patients. CSF was analyzed
133 with two platforms, Luminex xMAP or Innostest ELISA, which was then transformed to the
134 Luminex scale.³³ We previously related CSF p-Tau levels directly to cerebral tau burden in our

135 autopsy cohort.³³ The AD and lvPPA subset groups did not differ in age ($p=0.64$), sex ($p=0.99$),
136 education ($p=0.82$), disease duration ($p=0.75$), or the time difference between CSF sample
137 collection and the Cookie Theft recording ($p=0.43$). To determine the association of language
138 features with in vivo measures of pathology, we examined the relationship between our language
139 variables and the two CSF biomarkers of A β and p-Tau.

140 **2.6 Statistical considerations**

141 To compare the groups, we tested if requirements for parametric tests were met with a Levene's
142 test. If the data met the requirements for parametric tests, we performed an Analysis of Variance.
143 If not, we performed a Kruskal-Wallis test. We visually assessed residuals of the models to make
144 sure the data was suitable for linear modelling. When a group difference was significant, we
145 additionally performed a posthoc pairwise t-test or pairwise Wilcoxon rank sum test for pairwise
146 group comparisons, adjusting p -values for multiple group comparisons ($n=3$) with the false
147 discovery rate (fdr). We reported the effect size of each group comparison using Cohen's d .

148 Patients' language variables were z-scored using HC's mean and standard deviation. These z-
149 scores were used for visualization and linear regressions to estimate relations of our language
150 variables and clinical ratings. We did not use z-scores to determine significant group differences
151 with Z-tests, since the test statistic in some variables did not follow a normal distribution.

152 The z-scored language variables that showed significant group differences were associated with
153 patients' MMSE and BNT scores to investigate the relations of our language features to clinical
154 ratings of cognitive and language impairment. To examine potential interactions, we included
155 phenotype as an interaction term (MMSE/BNT ~ language variable*phenotype). P -values were
156 adjusted with fdr.

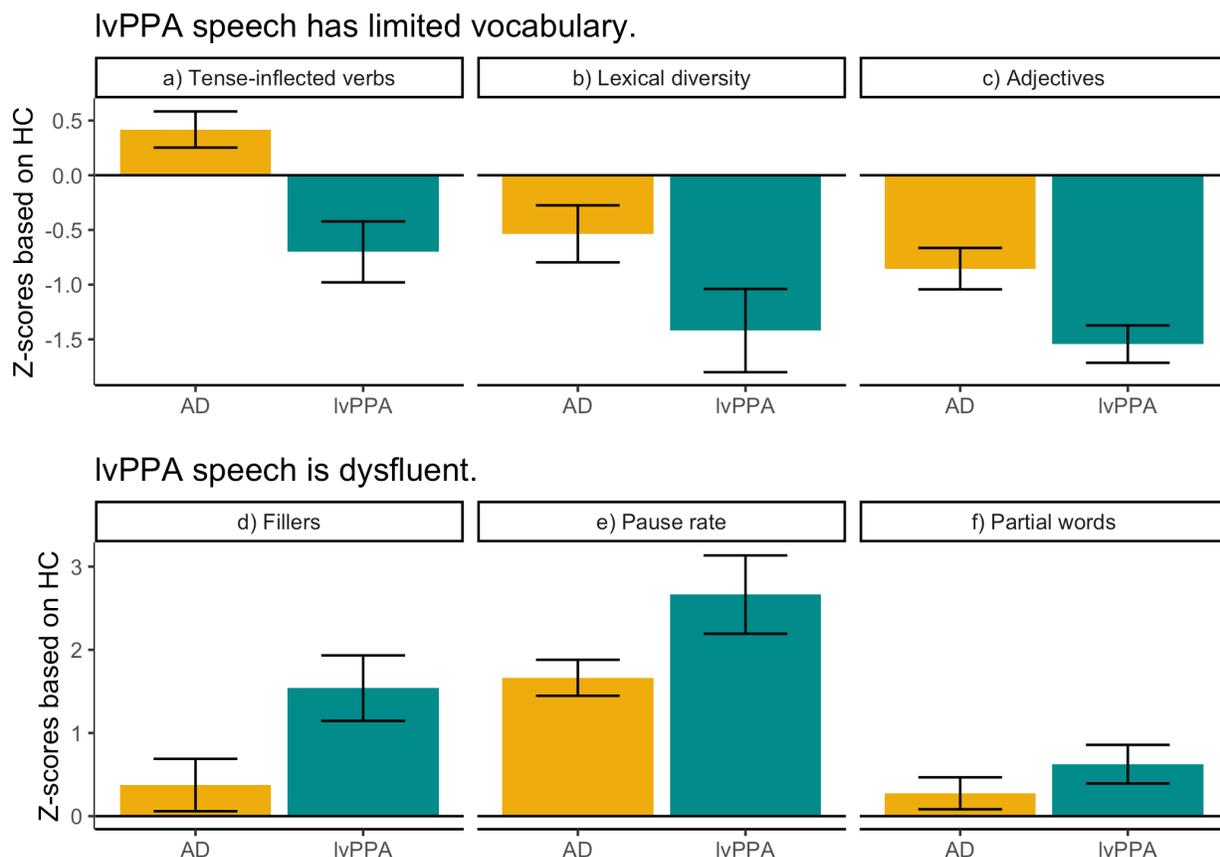
157 To validate our findings with levels of specific CSF biomarkers, we correlated patients' CSF
158 analyte levels to the language variables using Pearson correlation tests. CSF p-Tau levels were
159 log-transformed to normalize the data. We also checked if patients' clinical phenotype and the
160 time difference between Cookie Theft recording and CSF sample collection were significant
161 factors with linear regression models. Since the two factors were not significant, we only
162 reported the results of simple correlations to simplify the models. All statistical analyses were
163 performed with R version 4.1.0³⁴ and RStudio version 1.4.1717.³⁵

164 **2.7 Ethics**

165 The Institutional Review Board of the Hospital of the University of Pennsylvania approved the
166 study of human subjects, and all participants agreed to participate in the study by written
167 consent. All digital data was stored in secured HIPAA-compliant servers and handled by
168 personnel trained in PPI protection.

169 3. Results

170 3.1 Differences between amnesic AD and lvPPA patients



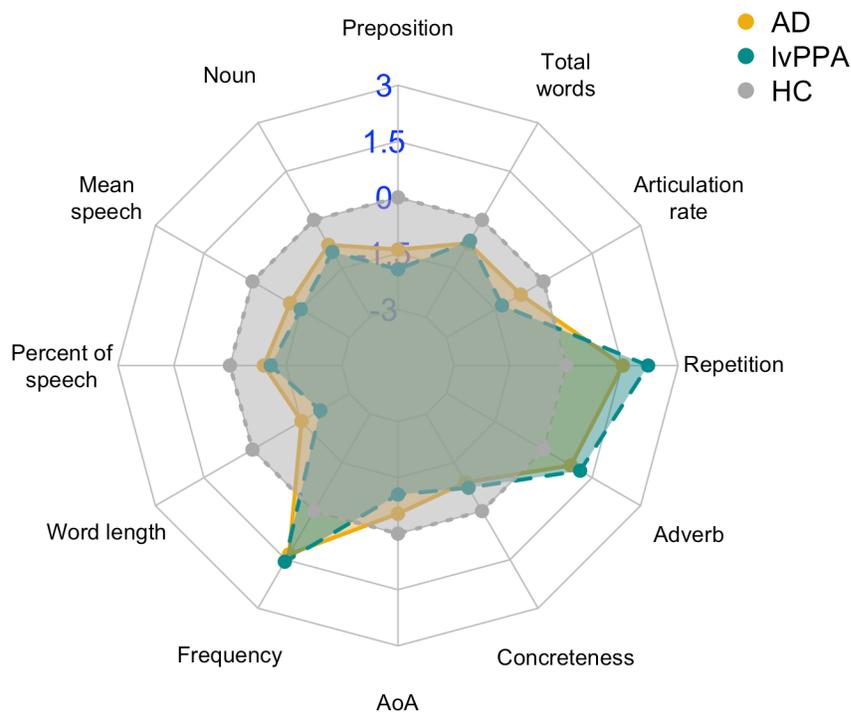
171
172 Figure 1: Speech differences between lvPPA and amnesic AD. For ease of visualization, z-
173 scored values compared to HC mean and standard deviation are plotted. Horizontal black line
174 indicates HC's mean. Abbreviations: AD = Alzheimer's disease, lvPPA = logopenic variant
175 primary progressive aphasia, HC = healthy controls

176 LvPPA patients produced fewer tense-inflected verbs compared to AD ($p=0.001$, $|d|=0.94$) and
177 HC ($p=0.048$, $|d|=0.61$; Fig.1A). AD patients' tense-inflected verb counts did not significantly
178 differ from HC ($p=0.124$, $|d|=0.4$). Patients with lvPPA showed lower lexical diversity than AD

179 ($p=0.05$, $|d|=0.52$) and HC ($p=0.005$, $|d|=1.02$), yet AD patients did not differ from HC ($p=0.149$,
180 $|d|=0.39$; Fig.1B). LvPPA patients produced fewer adjectives than AD patients ($p=0.019$,
181 $|d|=0.66$) and HC ($p<0.001$, $|d|=1.72$); AD patients also produced fewer adjectives than HC
182 ($p=0.003$, $|d|=0.75$; Fig.1C). Thus, both patient groups were impaired in their adjective
183 production, but lvPPA was more severely impaired compared with AD.

184 LvPPA patients also produced more fillers than AD ($p=0.022$, $|d|=0.6$) and HC ($p=0.01$, $|d|=1.06$;
185 Fig.1D), while AD did not significantly differ from HC ($p=0.383$, $|d|=0.23$). Patients with lvPPA
186 showed a higher pause rate than AD ($p=0.015$, $|d|=0.55$) and HC ($p<0.001$, $|d|=1.58$; Fig.1E); AD
187 patients' pause rate was also higher than HC ($p<0.001$, $|d|=1.34$). Lastly, lvPPA patients
188 produced more partial words than HC ($p=0.042$, $|d|=0.61$), but lvPPA patients did not
189 significantly differ from AD speakers($p=0.147$, $|d|=0.3$); AD patients did not differ from HC
190 ($p=0.313$, $|d|=0.24$). Thus, lvPPA speakers produced an abnormal number of partial words, while
191 AD speakers did not.

192 3.2 Impaired speech features in both AD and lvPPA



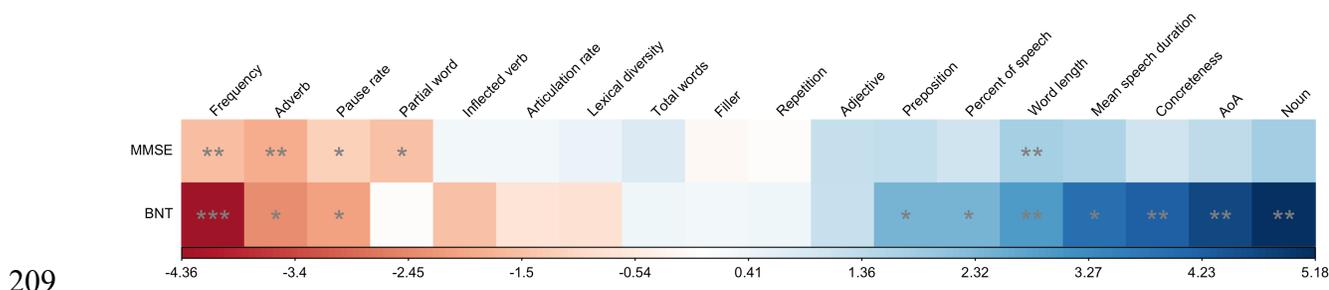
193

194 Figure 2: Impaired speech measures in both lvPPA (green) and AD (yellow). Only speech
195 measures that were significantly different from HC (gray) are plotted, and measures were
196 standardized based on the HC mean and standard deviation. The numbers in blue indicate z-
197 scored values based on HC. Abbreviations: AD = Alzheimer's disease, lvPPA = logopenic
198 variant primary progressive aphasia, HC = healthy controls

199 AD and lvPPA patients produced fewer prepositions and nouns than HC; patients produced
200 shorter speech segments than HC, and their percent of speech time out of the total time was also
201 lower than HC ($p < 0.001$, $|d| > 0.8$ for all comparisons). Both groups' content words were shorter,
202 more frequent (length & frequency: $p < 0.001$, $|d| > 0.8$), earlier-acquired (lvPPA: $p < 0.001$,
203 $|d| = 0.95$; AD: $p < 0.001$, $|d| = 0.56$), and more concrete (lvPPA: $p = 0.028$, $|d| = 0.7$, AD: $p = 0.002$,

204 $|d|=0.86$) than HC's. Both patient groups produced more adverbs (lvPPA: $p=0.029$, $|d|=0.84$; AD:
 205 $p=0.029$, $|d|=0.73$) and repetitions (lvPPA: $p<0.001$, $|d|=1.39$; AD: $p<0.001$, $|d|=0.79$) than HC.
 206 Patients also spoke slower (lvPPA: $p<0.001$, $|d|=1.01$; AD: $p<0.001$, $|d|=0.57$), and they
 207 produced fewer words in total than HC (lvPPA: $p=0.032$, $|d|=0.7$; AD: $p=0.007$, $|d|=0.73$).

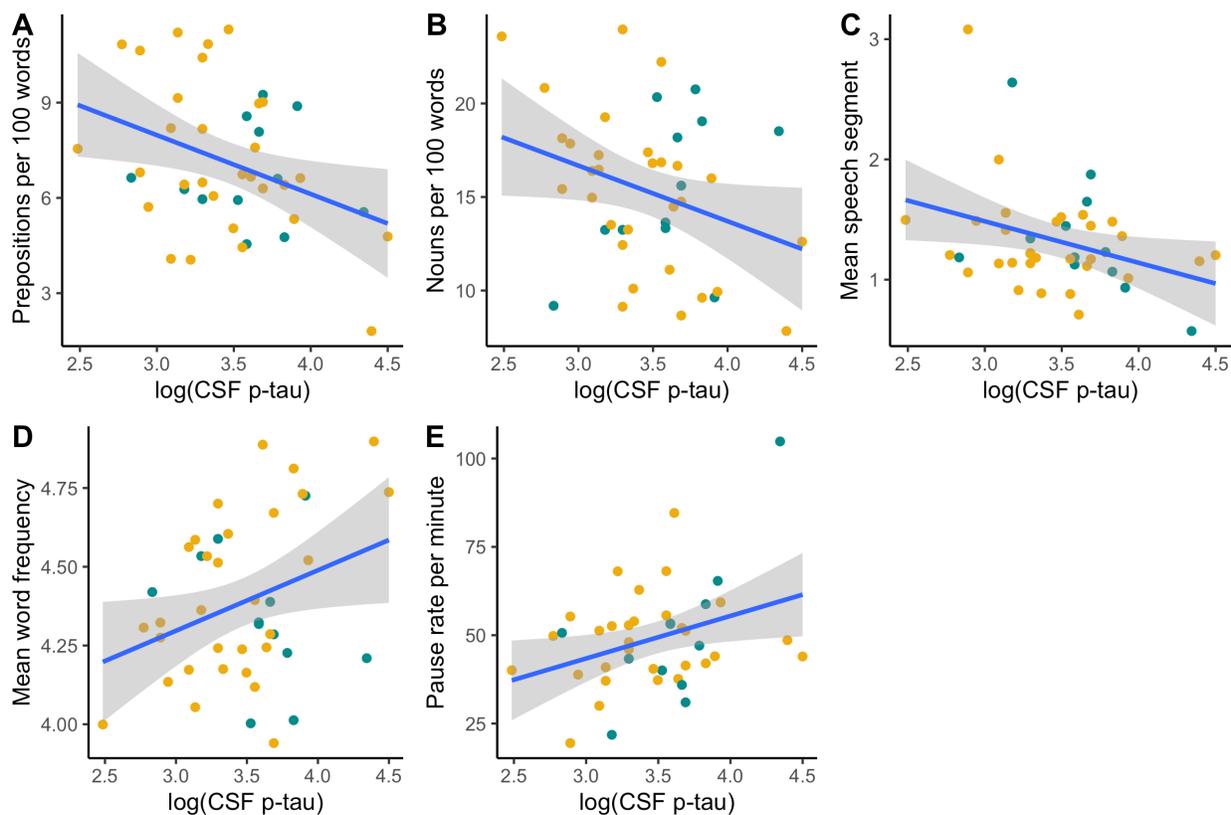
208 3.3 Relationships to clinical measures



209 Figure 3: Results of linear regression models of speech measures in patients with MMSE and
 210 BNT. Standardized estimated coefficients (x-axis) are plotted with colors: positive coefficients
 211 are in blue and negative ones in red. P-values are adjusted with fdr. *: $p<0.05$, **: $p<0.01$, ***:
 212 $p<0.001$. Abbreviations: AD = Alzheimer's disease, lvPPA = logopenic variant primary
 213 progressive aphasia, HC = healthy controls, MMSE = Mini-mental State Exam; BNT = Boston
 214 Naming Test
 215
 216 MMSE was significantly related to five language variables. Patients with lower MMSE produced
 217 more frequent ($\beta=-1.6$, $p=0.009$) and shorter content words ($\beta=1.75$, $p=0.001$), paused more
 218 frequently ($\beta=-1.21$, $p=0.041$), and produced more adverbs ($\beta=-1.9$, $p=0.003$) and partial words
 219 ($\beta=-1.52$, $p=0.021$).
 220 BNT was significantly associated with ten variables. Patients with lower BNT produced frequent
 221 words ($\beta=-4.36$, $p<0.001$), many adverbs ($\beta=-2.41$, $p=0.041$), and had a high pause rate ($\beta=-$

222 2.09, $p=0.041$). Patients with lower BNT also produced fewer prepositions ($\beta=2.42$, $p=0.048$)
223 and nouns ($\beta=5.18$, $p=0.003$). Patients with lower BNT showed a low percent of speech
224 produced during the picture description ($\beta=2.4$, $p=0.045$), and they produced shorter ($\beta=2.94$,
225 $p=0.003$), less concrete ($\beta=4.24$, $p=0.003$), and earlier-acquired ($\beta=4.69$, $p=0.003$) content
226 words, with shorter speech segments ($\beta=3.89$, $p=0.041$).
227 Only one variable showed a significant interaction with phenotype. AD patients with lower
228 MMSE scores produced more adverbs ($\beta=-1.9$, $p=0.003$), yet lvPPA patients with lower MMSE
229 scores produced fewer adverbs ($\beta=2.67$, $p=0.015$).

230 3.4 CSF results



ClinicalPhenotype1 ● AD ● lvPPA

232 Figure 4: Significant correlations of the speech variables and CSF p-Tau levels. Abbreviations:
233 AD = Alzheimer's disease, lvPPA = logopenic variant primary progressive aphasia, HC =
234 healthy controls, CSF = Cerebrospinal fluid, p-Tau = phosphorylated Tau, A β = beta-amyloid 42
235 Patients' CSF p-Tau level correlated with lower preposition counts ($r=-0.36$, $p=0.019$), lower
236 noun counts ($r=-0.31$, $p=0.047$), and shorter mean speech segment durations ($r=-0.33$, $p=0.032$;
237 Fig.4A-C). Also, patients' p-Tau levels were inversely correlated with a higher pause rate
238 ($r=0.34$, $p=0.026$) and higher word frequency ($r=0.33$, $p=0.036$; Fig.4D-E). The other measures
239 (adverbs, partial words, tense-inflected verbs, articulation rate, lexical diversity, total words,
240 fillers, repetitions, adjectives, percent of speech, word length, concreteness, AoA) were not
241 related to patients' p-Tau levels. A β alone did not significantly correlate with any of the
242 language measures.

243 **4. Discussion**

244 LvPPA is most frequently associated with underlying AD pathology, but direct comparison of
245 lvPPA with amnesic AD patients has been reported rarely. The current study focuses on
246 characterizing the language similarities and differences between lvPPA and AD in a biologically
247 confirmed cohort. To optimize reliability and reproducibility, we used fully automated lexical
248 and acoustic analyses to characterize linguistic markers of AD pathology. We expected that
249 lvPPA patients with non-amnesic AD would produce more dysfluent speech with limited
250 vocabulary than amnesic AD patients because of lvPPA's phenotypic characteristics. Results
251 confirmed that lvPPA patients produced fewer adjectives and tense-inflected verbs with lower
252 lexical diversity than amnesic AD patients and HC. lvPPA patients also paused more frequently
253 and produced more fillers and partial words than HC and/or amnesic AD patients. However, we

254 also found that both patient groups shared impairments in some speech features relative to HC.
255 For example, both patient groups produced more adverbs, including words like “*there*” and
256 “*here*,” but fewer prepositions and nouns than HC. Also, patients’ content words were earlier-
257 acquired, shorter, more frequent, and less concrete than those of HC. Patients produced more
258 repetitions and fewer total words with a slower articulation rate and a shorter mean speech
259 duration than HC. Some of these language variables were significantly related to clinical test
260 scores and p-Tau levels in CSF. We discuss these findings below.

261 The patient groups significantly differed on six language measures: pause rate, partial words,
262 fillers, adjectives, tense-inflected verbs, and lexical diversity. Fillers, adjectives, tense-inflected
263 verbs, and lexical diversity were significantly more impaired in lvPPA than AD patients,
264 emphasizing the deficits in lexical retrieval and poor fluency in these patients. However, these
265 linguistic features distinguishing between lvPPA and amnesic AD were not related to patients’
266 MMSE or BNT. It is important to note that the mechanism thought to subserve retrieval of a
267 single word in response to a stimulus picture appears to differ from lexical retrieval during
268 natural connected speech.³⁶ It is thus critical to monitor these speech features since they are not
269 easily explained by more general and commonly used measures such as BNT and MMSE. To our
270 knowledge, the result that tense-inflected verb production differed between lvPPA and AD has
271 not been previously reported. This finding seems to suggest that lvPPA patients produced fewer
272 complete sentences, assuming that there was one tense-inflected verb per inflection phrase.³⁷
273 This would be consistent with the observation that lvPPA patients’ limited auditory-verbal short-
274 term memory results in briefer sentences.^{3,4} Frequent fillers in lvPPA were in line with previous
275 observations, and consistent with their limited lexical retrieval.^{11,38,39} Lower adjective counts in
276 amnesic AD patients compared to HC have been previously reported,⁴⁰ yet we further showed

277 that the adjective production was more impaired in non-amnestic lvPPA than amnestic AD.
278 Lexical diversity has been frequently examined in the AD literature,^{14,15,41,42} where previous
279 studies have found that AD patients' lexical diversity was lower than that of HC. We showed
280 that lexical diversity was even lower in lvPPA than in AD. The fact that lvPPA and AD patients
281 significantly differed on these measures suggests that our language variables may capture subtle
282 but unique phenotypic differences between lvPPA and AD. Additionally, traditional clinical
283 ratings such as the MMSE and BNT are relatively insensitive to these linguistic features. Also,
284 none of the six variables, except pause rate, correlated with CSF p-tau, and regression analyses
285 failed to find an interaction of p-Tau level with phenotype in the assessment of pause rate. This
286 suggests that these speech markers are related more narrowly to the phenotype and not
287 necessarily to the underlying AD pathology burden. Further studies, including the anatomic
288 distribution of pathology in an autopsy cohort with quantitative measures of pathological burden,
289 may help shed light on this issue.

290 Some speech measures were related to MMSE and BNT. Pause rate showed more impairment in
291 lvPPA than AD. This might indicate word-finding difficulty in lvPPA which could provoke
292 frequent pausing to recall an appropriate word from their lexicon. It could also be that lvPPA
293 patients spoke slowly – patients' articulation rate was lower than that of HC – due to their
294 difficulty in retrieving words to generate utterances. Pause rate was significantly related to both
295 MMSE, an indicator of general cognitive impairment, and BNT, a measure of confrontation
296 naming. Elevated pause rate, therefore, may reflect in part both patients' word-finding
297 difficulties and general cognitive impairments. Partial word count, which was impaired only in
298 lvPPA, only correlated with MMSE, suggesting that it reflected in part lvPPA patients' disease
299 severity and general cognitive impairments, but not impaired object naming.

300 Word-finding difficulty in AD and lvPPA has been previously noted,^{11,43–48} where studies have
301 shown that patients had impairments in auditory-verbal short-term memory and could not recall
302 the phonological form of a word. However, comparative studies have not been reported to
303 examine if both amnesic AD and non-amnesic lvPPA patients would show word-finding
304 difficulty to a similar degree during natural speech. In our study, both patient groups produced
305 content words that were more abstract, earlier-acquired, more frequent and shorter than those of
306 HC, suggesting that they had difficulties in retrieving the full spectrum of lexical items needed to
307 describe the picture. Word frequency and length were significantly related to both MMSE and
308 BNT, which suggests that these lexical measures reflect in part patients' disease severity and
309 difficulties in lexical retrieval during confrontation naming. On the other hand, concreteness and
310 AoA were significantly associated only with BNT, indicating that these may be more sensitive to
311 word-finding difficulty in patients. AD and lvPPA patients did not significantly differ in these
312 measures, confirming that some degree of word-finding difficulty is present in both amnesic and
313 non-amnesic AD.

314 Adverb counts were greater in patients compared to HC. This may be related to patients
315 frequently using “pro-adverbs,” including “*here*” and “*there*,” which replaced locational
316 prepositional phrases. Patients typically produced utterances like “Mom is standing *here*,” for
317 example, when HC produced “Mom is standing *in front of the sink*.” Elevated adverb counts
318 were associated with low BNT scores in AD, suggesting that greater adverb use reflected AD
319 patients' difficulties in naming specific locations during natural speech. Patients' difficulty in
320 producing locational phrases were also partly reflected in the decreased preposition counts
321 compared to HC, which was also related to BNT scores. Additional work is needed to determine

322 if decrease preposition counts and increase adverb use are related to temporal propositions and
323 other features in connected speech.

324 Some of our language variables correlated with CSF p-Tau levels, but not with A β . This finding
325 is in line with previous findings that patients' cognitive impairment is generally not related to A β
326 levels but to accumulation of p-Tau.⁴⁹⁻⁵¹ Language production is one of the most essential daily
327 functions of humans, which needs to be taken into consideration in AD clinical trials and may
328 serve in monitoring response to treatment. Since our automated procedures for collecting speech
329 features is highly reliable and reproducible, investigations of speech variables as secondary
330 outcome measures should be considered in disease-modifying trials targeting tau.

331 The strengths of our study include that we directly compared speech features in AD and lvPPA
332 patients with biological evidence of underlying AD pathology. We examined natural connected
333 speech quantitatively using automated analyses of digitized speech samples. We inspected
334 language differences and similarities in these groups and showed that our language variables
335 could capture subtle linguistic differences between the two phenotypes, and traditional cognitive
336 measures such as MMSE and BNT appear to be insensitive to some of the features distinguishing
337 between amnesic AD and lvPPA. These methods may be useful in screening for underlying AD
338 pathology and monitoring disease progression and response to therapeutic interventions, because
339 collecting one-minute speech samples is easy, highly reproducible, and less costly and can be
340 collected remotely compared with other biomarkers. Future studies, testing the value of these
341 speech features in longitudinal datasets and developing machine learning classifiers for patient
342 stratification of AD pathology would be valuable. Shortcomings include the assessment of
343 relatively small samples, the difficulty obtaining train-test generalizability data in rare lvPPA and
344 early-onset AD cases such as these, the use of a single stimulus picture to elicit the speech

345 sample, and the absence of high resolution MRI data to assess the anatomic associates of these
346 linguistic features.

347 **5. Conclusion**

348 We implemented automated methods in analyzing acoustic and lexical characteristics of
349 amnesic AD and lvPPA patients' natural speech. We found speech markers that differed between
350 the groups. These features seemed to relate to the clinical phenotype and less so to the
351 underlying pathology. We also found speech markers that were shared between these two AD
352 phenotypes and linked to the common underlying pathology. This work demonstrates the
353 potential of natural speech to reflect underlying AD pathology while distinguishing between
354 specific phenotypes with the same pathology. Considering the cost effectiveness and reliability
355 of speech data, such markers could contribute to screening in AD clinical trials and longitudinal
356 monitoring in a more precise and inclusive way.

357

358 Appendix

359 Table A. List of features used in the analysis.

Feature set	Features	Example or definition
Part-of-speech categories	Nouns	<i>cookie</i>
	Tense-inflected verbs	Sum of auxiliaries, past-tense, present tense verbs
	Adjectives	<i>nice</i>
	Adverbs	<i>only, here</i>
	Conjunctions	<i>and</i>
	Determiners	<i>the</i>
	Pronouns	<i>she</i>
	Prepositions	<i>in, from</i>
	Particles	<i>give up</i>
	Interjections	<i>jeez</i>
Dysfluency markers	Fillers	<i>um, uh</i>
	Partial words	<i>coo- cookie</i>
	Repetitions	<i>II can't</i>
	Concreteness	How concrete a word's meaning is

Feature set	Features	Example or definition
Lexical characteristics	Frequency	Log ₁₀ scaled word frequency per million words
	Ambiguity	Number of a words' meaning in a context
	Age of acquisition	Average age when children acquire a given word
	Familiarity	Number of people who know a given word
	Word length	Number of phonemes in a word
	Lexical diversity	Average type-token ratio of all windows
Acoustic measures	Mean speech segment duration	Average of all speech segment durations
	Mean pause segment duration	Average of all pause segment durations
	Total speech time	Sum of all speech segment duration
	Pause rate	The number of pause segments per minute
	Percent of speech	The percentage of speech time out of total time
	Articulation rate	The number of syllables articulated per second
	Pitch range	90th percentile of individual's pitch values after excluding pitch-tracking errors

360 **References**

- 361 1. Alzheimer's Association. *2020 Alzheimer's disease facts and figures.*; 2020.
362 doi:10.1002/alz.12068
- 363 2. Nicholas M, Obler L, Albert M, Goodglass H. Lexical Retrieval in Healthy Aging.
364 *Cortex*. 1985;21(4):595-606. doi:10.1016/S0010-9452(58)80007-6
- 365 3. Henry ML, Gorno-Tempini ML. The logopenic variant of primary progressive aphasia.
366 *Current opinion in Neurology*. 2010;23(6):633-637. doi:10.1097/WCO.0b013e32833fb93e.The
- 367 4. Gorno-Tempini ML, Brambati SM, Ginex V, et al. The logopenic/phonological variant of
368 primary progressive aphasia. *Neurology*. 2008;71:1227-1234.
- 369 5. Gorno-Tempini ML, Dronkers NF, Rankin KP, et al. Cognition and anatomy in three
370 variants of Primary Progressive Aphasia. *Annals of Neurology*. 2004;55:335-346.
- 371 6. Rohrer JD, Rossor MN, Warren JD. Alzheimer's pathology in primary progressive
372 aphasia. *Neurobiology of Aging*. 2012;33(4):744-752. doi:10.1016/j.neurobiolaging.2010.05.020
- 373 7. Gefen T, Gasho K, Rademaker A, et al. Clinically concordant variations of Alzheimer
374 pathology in aphasic versus amnesic dementia. *Brain*. 2012;135(5):1554-1565.
375 doi:10.1093/brain/aws076
- 376 8. Leyton CE, Britton AK, Hodges JR, Halliday GM, Kril JJ. Distinctive pathological
377 mechanisms involved in primary progressive aphasias. *Neurobiology of Aging*. 2016;38:82-92.
378 doi:10.1016/j.neurobiolaging.2015.10.017

- 379 9. Giannini LAA, Irwin DJ, Mcmillan CT, et al. Clinical marker for Alzheimer disease
380 pathology in logopenic primary progressive aphasia. *Neurology*. 2017;88(24):2276-2284.
381 doi:10.1212/WNL.0000000000004034
- 382 10. Bergeron D, Gorno-Tempini ML, Rabinovici GD, et al. Prevalence of amyloid- β
383 pathology in distinct variants of primary progressive aphasia. *Annals of Neurology*.
384 2018;84(5):729-740. doi:10.1002/ana.25333
- 385 11. Wilson SM, Henry ML, Besbris M, et al. Connected speech production in three variants
386 of primary progressive aphasia. *Brain*. 2010;133:2069-2088. doi:10.1093/brain/awq129
- 387 12. Johnson JCS, Jiang J, Bond RL, et al. Impaired phonemic discrimination in logopenic
388 variant primary progressive aphasia. 2020:1-6. doi:10.1002/acn3.51101
- 389 13. Ash S, Evans E, O'Shea J, et al. Differentiating primary progressive aphasias in a brief
390 sample of connected speech. *Neurology*. 2013;81(4):329-336.
391 doi:10.1212/WNL.0b013e31829c5d0e
- 392 14. Fraser K, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer's disease in
393 narrative speech. *Journal of Alzheimer's Disease*. 2016;49(2):407-422. doi:10.3233/JAD-150520
- 394 15. Bucks RS, Singh S, Cuerden JM, Wilcock GK. Analysis of spontaneous, conversational
395 speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical
396 performance. *Aphasiology*. 2000;14(1):71-91. doi:10.1080/026870300401603
- 397 16. Rentoumi V, Paliouras G, Danasi E, et al. Automatic detection of linguistic indicators as
398 a means of early detection of Alzheimer's disease and of related dementias: A computational

- 399 linguistics analysis. In: *8th IEEE International Conference on Cognitive Infocommunications,*
400 *CogInfoCom 2017.*; 2017:33-38. doi:10.1109/CogInfoCom.2017.8268212
- 401 17. Hernández-Domínguez L, Ratté S, Sierra-Martínez G, Roche-Bergua A. Computer-based
402 evaluation of Alzheimer's disease and mild cognitive impairment patients during a picture
403 description task. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring.*
404 2018;10:260-268. doi:10.1016/j.dadm.2018.02.004
- 405 18. Eyigoz E, Courson M, Sedeño L, et al. From discourse to pathology: Automatic
406 identification of Parkinson's disease patients via morphological measures across three languages.
407 *Cortex.* 2020;132:191-205. doi:10.1016/j.cortex.2020.08.020
- 408 19. Irwin D, McMillan CT, Toledo JB, et al. Comparison of cerebrospinal fluid levels of tau
409 and A β 1-42 in Alzheimer disease and frontotemporal degeneration using 2 analytical platforms.
410 *Archives of Neurology.* 2012;69(8):1018-1025. doi:10.1001/archneurol.2012.26
- 411 20. Lleó A, Irwin DJ, Illán-Gala I, et al. A 2-step cerebrospinal algorithm for the selection of
412 frontotemporal lobar degeneration subtypes. *JAMA Neurology.* 2018;75(6):738-745.
413 doi:10.1001/jamaneurol.2018.0118
- 414 21. Goodglass H, Kaplan E, Weintraub S. *Boston Diagnostic Aphasia Examination.* (Febiger
415 L&, ed.). Philadelphia; 1983.
- 416 22. Honnibal M, Johnson M. An improved non-monotonic transition system for dependency
417 parsing. In: *EMNLP 2015: Conference on Empirical Methods in Natural Language Processing.*;
418 2015:1373-1378. doi:10.18653/v1/d15-1162

- 419 23. Brysbaert M, Warriner AB, Kuperman V. Concreteness ratings for 40 thousand generally
420 known English word lemmas. *Behavior Research Methods*. 2014;46(3):904-911.
421 doi:10.3758/s13428-013-0403-5
- 422 24. Hoffman P, Lambon Ralph MA, Rogers TT. Semantic diversity: A measure of semantic
423 ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*.
424 2013;45(3):718-730. doi:10.3758/s13428-012-0278-x
- 425 25. Brysbaert M, New B. Moving beyond Kučera and Francis: A critical evaluation of
426 current word frequency norms and the introduction of a new and improved word frequency
427 measure for American English. *Behavior Research Methods*. 2009;41(4):977-990.
428 doi:10.3758/BRM.41.4.977
- 429 26. Brysbaert M, Mandera P, Keuleers E. Word prevalence norms for 62 ,000 English
430 lemmas. *Behavior Research Methods*. 2018;(July 2018):467-479.
- 431 27. Carnegie Mellon Speech Group. The Carnegie Mellon University Pronouncing
432 Dictionary. 2014. <http://www.speech.cs.cmu.edu/cgi-bin/cmudict>.
- 433 28. Covington MA, McFall JD. Cutting the Gordian knot: The movingaverage type-token
434 ratio (MATTR). *Journal of Quantitative Linguistics*. 2010;17(2):94-100.
- 435 29. Fergadiotis G, Wright HH, Green SB. Psychometric evaluation of lexical diversity
436 indices: Assessing length effects. *Journal of Speech, Language, and Hearing Research*.
437 2015;58:840-852. doi:10.1044/2015
- 438 30. Cho S, Nevler N, Ash S, et al. Automated analysis of lexical features in Frontotemporal
439 Degeneration. *Cortex*. 2021;137:215-231. doi:10.1016/j.cortex.2021.01.012

- 440 31. Boersma P, Weenink D. Praat: doing phonetics by computer. 2019.
- 441 32. Nevler N, Ash S, Jester C, Irwin DJ, Liberman M, Grossman M. Automatic measurement
442 of prosody in behavioral variant FTD. *Neurology*. 2017;89:1-8.
- 443 33. Irwin DJ, Lleó A, Xie SX, et al. Ante mortem cerebrospinal fluid tau levels correlate with
444 postmortem tau pathology in frontotemporal lobar degeneration. *Annals of Neurology*.
445 2017;82(2):247-258. doi:10.1002/ana.24996
- 446 34. R Core Team. R: A language and environment for statistical computing. 2019.
- 447 35. RStudio Team. RStudio: Integrated Development for R. 2020. <http://www.rstudio.com/>.
- 448 36. Levelt WJM. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press;
449 1989.
- 450 37. Pollock J-Y. Verb movement, universal grammar, and the structure of IP. *Linguistic*
451 *Inquiry*. 1989;20(3):365-424.
- 452 38. Ash S, McMillan C, Gross RG, et al. Impairments of speech fluency in Lewy Body
453 Spectrum Disorder. *Brain and Language*. 2013;120(3):290-302.
454 doi:10.1016/j.bandl.2011.09.004.Impairments
- 455 39. Gorno-Tempini ML, Hillis A, Weintraub S, et al. Classification of primary progressive
456 aphasia and its variants. *Neurology*. 2011;76(11):1006-1014.
- 457 40. Croisile B, Ska B, Brabant MJ, et al. Comparative study of oral and written picture
458 description in patients with Alzheimer's disease. *Brain and Language*. 1996;53(1):1-19.
459 doi:10.1006/brln.1996.0033

- 460 41. Kavé G, Goral M. Word retrieval in picture descriptions produced by individuals with
461 Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*. 2016;38(9):958-
462 966. doi:10.1080/13803395.2016.1179266
- 463 42. Kavé G, Dassa A. Severity of Alzheimer's disease and language features in picture
464 descriptions. *Aphasiology*. 2018;32(1):27-40. doi:10.1080/02687038.2017.1303441
- 465 43. Buxbaum LJ, Schwartz MF, Carew TG. The Role of Semantic Memory in Object Use.
466 *Cognitive Neuropsychology*. 1997;14(2):219-254. doi:10.1080/026432997381565
- 467 44. Kempler D. Lexical and pantomime abilities in Alzheimer's disease. *Aphasiology*.
468 1988;2(2):147-159. doi:10.1080/02687038808248900
- 469 45. Ochipa C, Rothi LJG, Heilman KM. Conceptual apraxia in alzheimer's disease. *Brain*.
470 1992;115(4):1061-1071. doi:10.1093/brain/115.4.1061
- 471 46. Adlam ALR, Bozeat S, Arnold R, Watson P, Hodges JR. Semantic knowledge in mild
472 cognitive impairment and mild Alzheimer's disease. *Cortex*. 2006;42(5):675-684.
473 doi:10.1016/S0010-9452(08)70404-0
- 474 47. Giffard B, Desgranges B, Nore-Mary F, et al. The nature of semantic memory deficits in
475 Alzheimer's disease: New insights from hyperpriming effects. *Brain*. 2001;124(8):1522-1532.
476 doi:10.1093/brain/124.8.1522
- 477 48. Venneri A, McGeown WJ, Hietanen HM, Guerrini C, Ellis AW, Shanks MF. The
478 anatomical bases of semantic retrieval deficits in early Alzheimer's disease. *Neuropsychologia*.
479 2008;46(2):497-510. doi:10.1016/j.neuropsychologia.2007.08.026

- 480 49. Huber CM, Yee C, May T, Dhanala A, Mitchell CS. Cognitive Decline in Preclinical
481 Alzheimer's Disease: Amyloid-Beta versus Tauopathy. *Journal of Alzheimer's Disease*.
482 2018;61(1):265-281. doi:10.3233/JAD-170490
- 483 50. Dickson DW, Crystal HA, Mattiace LA, et al. Identification of normal and pathological
484 aging in prospectively studied nondemented elderly humans. *Neurobiology of Aging*.
485 1992;13(1):179-189. doi:10.1016/0197-4580(92)90027-U
- 486 51. Engelborghs S, Maertens K, Vloeberghs E, et al. Neuropsychological and behavioural
487 correlates of CSF biomarkers in dementia. *Neurochemistry International*. 2006;48(4):286-295.
488 doi:10.1016/j.neuint.2005.11.002