**Supplementary File 1**

**Audio Recording Tasks**

Picture Description Task

The Cookie Theft Scene is a component of the Boston Diagnostic Aphasic Examination. Subjects are asked to describe the scene while viewing a drawing.

Paragraph Reading Task

Patients are asked to read the paragraph below:

“Bamboo walls are getting to be very popular. They are strong, easy to use, and good-looking. They provide a good background and can create a look of a Japanese garden. Bamboo is one of the largest and most rapidly growing grasses all over the world. Many varieties of bamboo are grown in Asia, although it is also grown in America. Last year we bought a new home and have been working on the flower garden. In a few more days, we will be done with the bamboo wall in our garden. We have really enjoyed the project.”

**Supplementary File 2**

*Speech Analysis*

Audio recordings were analyzed using the Winterlight Labs speech analysis platform. The audio recordings were transcribed, and linguistic and acoustic variables were extracted through automated speech analysis. The generation of annotations such as speaker segmentation, transcription, and utterance segmentation was performed by trained raters using custom software. Data processing and feature extraction were performed using the Winterlight Labs pipeline (www.winterlightlabs.com) using Python-based standard acoustic and language processing libraries (e.g., spacy) and custom code. For each speech sample, 546 variables were computed based on the sound file and accompanying transcript for each recording. These variables reflect the acoustic (e.g., properties of the sound wave, speech rate, number of pauses), lexical (e.g., rates and types of words used, and their characteristics such as frequency and imageability, which reflect how commonly words are used and how easy they are to picture, respectively), semantic (relating to the meaning of the words, e.g. semantic relatedness of subsequent utterances, semantic relatedness of utterances to the items in the picture) and syntactic (relating to the grammar of the sentences, e.g., syntactic complexity, use of different syntactic constructions) aspects of the sample.

*Statistical Methods*

Given the lack of prior published data on speech in HE and a large number of speech variables, we first performed an exploratory procedure to identify speech variables that were relevant to HE status. A Mann-Whitney test compared all speech variables derived from the picture description task between those with low PHES (≤ -4) vs. normal PHES (> -4). Given the exploratory nature of this feature selection step, we used a p-value of 0.05 as the threshold to identify candidate speech variables. To reduce redundancy between selected variables, we next examined the correlations between the selected speech variables using Pearson correlations. For any pairs of variables with |correlation coefficient| > 0.9, we selected the variable with the lower p-value based on the Mann-Whitney test, and eliminated the other variable. Selecting a threshold of |correlation coefficient| > 0.95 or 0.99 resulted in selection of the same variables. Next, to account for differences based on demographic factors, we performed logistic regressions on the remaining variables with low PHES vs. normal PHES as the binary outcome, adjusted by factors of age, sex, and years of education. Speech variables with significant effects in these models (p < 0.05) were selected for further analysis. Finally, we tested for significant associations between these speech variables and continuous PHES scores, using non-parametric Spearman partial correlations controlling for age, sex and years of education, and Bonferroni-corrected for multiple comparisons. We performed subsequent analyses on the set of variables found to be associated with PHES scores based on this exploratory procedure, and used Bonferroni-adjusted p-value thresholds for multiple corrections on those analyses.

**Supplementary File 3: Patient Characteristics at Baseline**

|  |  |
| --- | --- |
| Characteristics\* | Values (N = 76) |
| Age, years | 59 $\pm $ 10 |
| Male sex, n (%) | 47 (62%) |
| Years of education | 16 $\pm $ 3 |
| MELD score | 12 $\pm $6  |
| Etiology of cirrhosis, n (%) |  |
| Alcohol | 28 (37%) |
| Non-alcoholic steatohepatitis (NASH) | 16 (21%) |
| Viral | 10 (13%) |
| Mix of above | 9 (12%) |
| Other | 13 (17%) |
| History of overt HE, n (%) | 40 (53%) |
| Current lactulose use, n/n (%) | 25/40 (63%) |
| Current rifaximin use, n/n (%) | 27/40 (68%) |
| Either lactulose or rifaximin use, n/n (%) | 36/40 (90%) |
| Low PHES (≤ -4), n (%) | 23 (30%) |

Data are presented as mean $\pm SD $unless mentioned otherwise.

\*MELD denotes model for end-stage liver disease; HE hepatic encephalopathy, PHES psychometric HE score.

**Supplementary File 4**

**Speech Variables Differing by Psychometric HE Score**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Speech variable category** | **Speech variable** | **Definition** | **Median difference between groups (normal PHES - low PHES)** | **p-value** |
| Acoustic | Average word duration | Average number of seconds per word  | -0.093 | 0.012 |
| Acoustic | Speech rate | Average number of words per minute | 22.2 | 0.013 |
| Acoustic | Power spectrum (MFCC 0) | Mean value of one of the Mel-frequency cepstral coefficients that represent the power spectrum of the audio | -0.032 | 0.014 |
| Lexical/Syntactic | Use of particles | Usage of particles, as a proportion of all syntactic structures | 0.0085 | 0.014 |
| Lexical/Syntactic | Use of personal pronouns | Usage of personal pronouns as noun phrases, as a proportion of all syntactic structures | -0.017 | 0.014 |
| Acoustic | Power spectrum (MFCC 34) | Mean value of one of the Mel-frequency cepstral coefficients that represent the power spectrum of the audio | 0.0011 | 0.015 |
| Lexical/Syntactic | Use of verb phrases including “to” | Usage of verb phrases with the infinitival “to”, as a proportion of all syntactic structures | 0.0086 | 0.017 |
| Acoustic | Kurtosis of power spectrum (MFCC 40) | Kurtosis of one of the Mel-frequency cepstral coefficients that represent the power spectrum of the audio | -40.9 | 0.022 |
| Lexical/Syntactic | Use of particles 2 | Usage of particles, as a proportion of all words | 0.014 | 0.024 |
| Acoustic | Kurtosis of power spectrum (MFCC 23) | Kurtosis of one of the Mel-frequency cepstral coefficients that represent the power spectrum of the audio | -682.5 | 0.029 |
| Acoustic | Skewness of power spectrum (MFCC 16) | Skewness of one of the Mel-frequency cepstral coefficients that represent the power spectrum of the audio | -8.5 | 0.029 |
| Lexical/Syntactic | Use of verb phrases including gerunds and declarative clauses | Usage of verb phrases with gerunds and declarative clauses, as a proportion of all syntactic structures | 0.000031 | 0.033 |

*Speech variables that differed between groups based on low (≤ -4) vs. normal (> -4) PHES scores, when adjusted for age, sex and years of education. Table lists the median difference between groups for each variable and the p-value (uncorrected) for the Mann-Whitney test comparing groups. MFCCs refer to Mel-frequency cepstral coefficients, which are derived from the power spectrum of the recording.*