

II. Data and Annotation Conventions

A. Subjects

The subjects in this study came from existing community cohort studies of brain aging at the Layton Aging & Alzheimer's Disease Center, an NIA-funded Alzheimer's center for research at Oregon Health & Science University (OHSU). The Layton Center defines MCI in two ways: 1) via the Clinical Dementia Rating (CDR) scale [11] and 2) via a psychometrically driven concept of degraded performance on a large set of neuropsychological tests. This paper is investigating speech produced during the neuropsychological tests that play a role in the latter definition of MCI. Thus, to provide an independent unconfounded reference objective for evaluation, our reference classification is based on the CDR scale, defined next. The CDR has been shown to have high expert inter-annotator reliability [17], and, critically, is not deterministically derived from neuropsychological test results.

The CDR is assigned with access to information from clinical examinations, and is calculated by assigning scores for six cognitive categories: Memory; Orientation; Judgment and Problem Solving; Community Affairs; Home and Hobbies; and Personal Care. The score for each category is an assessment of impairment in that category, and can take one of five values: 0 (None); 0.5 (Questionable); 1 (Mild); 2 (Moderate); and 3 (Severe). The information that is used to provide the scores in each of these categories comes from clinical interviews with the subject and an informant, and a clinician-administered Neurobehavioral Cognitive Status Examination. To derive the CDR from these independently scored categories, the score M of the Memory category is taken as primary. The CDR is M if at least three of the remaining five categories also have a score of M . Otherwise, the CDR is calculated by looking at the scores of categories with score other than M . See [11] for specific details.

We collected audio recordings of 74 neuropsychological examinations administered at the Layton Center at OHSU. The two subject groups were: 1) those who were assigned a Clinical Dementia Rating (CDR) of 0 (healthy); and 2) those who were assigned a CDR of 0.5 (Mild Cognitive Impairment; MCI). 37 subjects were in the CDR = 0 group (healthy); and 37 subjects are in the CDR = 0.5 group (MCI).

B. Neuropsychological Tests

Table I presents means and standard deviations for age, years of education and the manually calculated scores of a number of standard neuropsychological tests that were administered during the recorded session. Again, these tests are not used to assign the CDR. These tests include: the Mini Mental State Examination (MMSE) [18]; the CERAD Word List Acquisition (A) and Recall (R) tests [19]; the Wechsler Logical Memory (LM) I (immediate) and II (delayed) narrative recall tests [20]; Category Fluency, Animals (A) and Vegetables (V); and Digit Span (WAIS-R) forward (F) and backward (B) [21]. Test IDs are given in the table for reference later in the paper.

The Wechsler Logical Memory I/II tests are the basis of our study on spoken language derived measures. The original narrative is a short, three-sentence story:

Anna Thompson of South Boston, employed as a cook in a school cafeteria, reported at the police station that she had been held up on State Street the night before and robbed of fifty-six dollars. She had four small children, the rent was due, and they had not eaten for two days. The police, touched by the woman's story, took up a collection for her.