



Using Linguistic Indicators of Difficulty to Identify Mild Cognitive Impairment

Rebecca Lunsford, Peter A. Heeman

Center for Spoken Language Understanding
 Oregon Health & Science University
 Portland, OR, USA
 lunsforr@ohsu.edu

Abstract

Speaking is a complex task, and it is to be expected that speech will be effected when a speaker is faced with cognitive difficulties. To explore how cognitive impairment is manifested in a persons' speech, we compared the speech of elders diagnosed with Mild Cognitive Impairment (MCI) to others who are cognitively intact, while the speakers attempt to retell a story they just heard. We found that the speakers with impairment, as compared to those who are cognitively intact, spent more time engaged in verbalized hesitations (e.g., "and um ...") prior to speaking story content, and that these verbalized hesitations accounted for a larger ratio of the time spent retelling. In addition, we found that a higher percentage of the impaired speakers used phrases such as "I guess" and "I can't recall" to qualify content they were unsure of, or to replace details they couldn't recall. These results provide insight into how speakers manage cognitive impairment, suggesting that these indicators of difficulty could be used to assist in early diagnosis of MCI.

Index Terms: cognitive impairment, linguistic indicators of difficulty, hesitations, disfluencies

1. Introduction

One goal of current clinical research is to improve early identification of Mild Cognitive Impairment (MCI). MCI is a term used to describe the earliest clinically detectable phase of incipient dementia [1]. Although some diagnosed with MCI will remain stable or return to normal functioning others will progress to dementia over time [2, 3]. In fact, among people over age 65, the prevalence rate of dementia is 6-8%, and the prevalence increases with age [4].

Early diagnosis is important from a research perspective, to better understand the disease progression. Although early diagnoses of MCI is important, common screening tools such as the Mini-Mental State Exam (MMSE) are not reliable during the earliest stages of MCI, when the cognitive impairments can be subtle. Yet, a comprehensive assessment can be difficult and costly, requiring a battery of neurological and neuropsychological tests performed by skilled clinicians. For example, at the Layton Aging and Alzheimer's Disease Center at Oregon Health & Science University, research subjects are assessed during two sessions, each lasting up to an hour.

Unfortunately, from a clinical perspective, lengthy diagnostic procedures are impractical, especially given the increasing number of senior patients we will see in the coming decades. Because of the time constraints of clinicians, and despite the importance of early diagnosis, assessment may be postponed until the patient displays clear indicators of impairment and is more easily assessed using tools that do not require substantial

time and effort. Going forward, clinicians and researchers will need tools that speed the identification of people either with existing impairment, or those at increased risk of impairment.

It is well known that MCI and dementia impact a person's speech. Previous research has aimed to determine what specific spoken language characteristics could be used in a classifier to differentiate those with MCI or dementia from those who are cognitively intact. Singh et al. explored temporal aspects of speech, such as the speaker's rate of speech and the duration and rate of pausing, finding that those with dementia paused longer and spoke more slowly than their peers without dementia [5]. Roark et al. expanded on this work, including different measures of syntactic complexity in their analyses, finding that the those with MCI produced shorter utterances, and that their speech was less complex [6].

Other researchers have examined whether the use of non-specific nouns (i.e., words including "thing",) and fillers increases over time in persons who develop dementia. Le, et al., [7] compared the writings of three life-long authors, one who developed dementia, one who is suspected of having developed dementia, and one who remained cognitively intact. They found that the two authors with dementia did have an increasing rate of fillers over time, in contrast to the cognitively intact author whose rate remained low. For non-specific nouns they found no consistent pattern of change. Yet, in a case study conducted by Berisha et al. [8], it was found that Ronald Reagan (who was diagnosed with Alzheimer's disease after his presidency) used an increasing number of combined non-specific nouns and fillers in his press conferences over the course of his presidency, but that George H. W. Bush did not.

There has also been research on how the speech of normal functioning speakers is impacted when encountering cognitive difficulty, which is relevant to this paper. Smith et al. found that speakers will potentially produce the fillers "um" and "uh", and pause longer, when uncertain how to answer a factual question [9]. Lindstrom et al. found that speakers' rates of disfluencies change when under cognitive load [10]. We refer to speech characteristics such as these as *difficulty indicators* (DIs), and use this term to refer to both symptoms of difficulty (e.g., "uh") and mechanisms used to manage difficulty (e.g., hesitations or qualifying phrases).

We hypothesize that DIs are also present in the speech of people with MCI. However, people with MCI might not exhibit DIs in the same manner, or at the same rate, and this change could happen very early in the progression of MCI. To illustrate, those with MCI may exhibit an increased rate of some DIs due to impairments in planning and producing speech. Alternatively, those with MCI may produce a decreased rate of those DIs that require the speaker to monitor their own speech

10.21437/Interspeech.2015-235

Story A – Immediate	Score
Anna / Thompson / of South / Boston /, employed / as a cook /	___ (0-6)
in a school / cafeteria /, reported / at the City Hall /Station /	___ (0-5)
that she had been held up / on State Street / the night before /	___ (0-3)
and robbed / of fifty-six dollars /. She had four /	___ (0-3)
small children /, the rent was due /, and they had not eaten /	___ (0-3)
for two days /. The police /, touched by the woman’s story /,	___ (0-3)
took up a collection / for her /.	___ (0-2)

Figure 1: Wechsler Logical Memory test narrative. Story elements are separated by “/”.

and processing, as this is a secondary cognitive task that may be impaired.

In this paper we explore differences in how speakers exhibit DIs, comparing elders with MCI to elders who are cognitively intact. Specifically we focus on verbal hesitations, disfluencies, and phrases used to express uncertainty.

2. Subjects and Recordings

For this research, we used existing audio recordings made during neuropsychological examinations administered at the Layton Aging & Alzheimers Disease Center. From these recordings, we use only the Wechsler Logical Memory (LMI) test [11]. In this test, subjects engage in an immediate narrative recall task, in which they retell a story they just heard.¹ The narrative is a detail-rich, short, three sentence story, shown in Figure 1. During this test, the clinician notes which of 25 story elements the participant recalls, and the participant’s score is based on the number of elements correctly recalled.

The LMI test is well-suited to our research for several reasons. First, it is challenging enough that we can expect participants’ speech to be effected. Second, as the participants are engaged in a monologue retelling, our analyses do not need to account for any effects due to the complexities of dialogue. Third, we can directly compare measures between participants, because they are all engaged in the same task, e.g., that of retelling the story.

In this study we included two groups; those diagnosed with MCI (impaired) and those who have no dementia (intact). Participants were placed in each group based on their Clinical Dementia Rating (CDR), which is calculated using a participant’s scores in six cognitive categories: Memory; Orientation; Judgment and Problem Solving; Community Affairs; Home and Hobbies; and Personal Care. Using these scores, an algorithm is applied to produce a summary score of 0.0 (healthy), 0.5 (MCI), 1.0 (dementia), and so on. As we are focused on early signs of impairment, our analyses include only participants with a CDR of 0.0 or 0.5. For more information on the CDR, see Morris, et al. [12]. An added advantage of the CDR for this work, is that it provides an assessment that is independent from the LMI data we use herein.

¹The Wechsler Logical Memory test also includes a delayed recall component (LMI), in which the subject again attempts to retell the story after a minimum of 15 minutes. We did not include this second retelling in these analyses.

3. Transcriptions and Data Treatment

Transcriptions and word time alignments for the LMI recordings were produced as part of a previous study [6]. We then segmented and selected data for analyses as described below.

C-unit Segmentation: We used the guidelines for segmentation previously defined by the Center for Spoken Language Understanding for Systematic Analysis of Language Transcriptions (CSLU/SALT) to segment the speech. The CSLU/SALT guidelines defines a Communication Unit (c-unit) as “an independent clause with its modifiers”, which includes one main clause, with all subordinate clauses attached to it. In essence, a c-unit is the smallest unit that can stand alone as a complete utterance. We choose to use the CSLU/SALT guidelines as a starting point as they are well suited to non-fluent speech.

C-unit Selection: During the recall task, participants would sometimes engage in conversational speech unrelated to the story retell task (e.g., “I have six children.”). These c-units were annotated as “Conversation” and excluded from our analyses. In addition, the clinician would, in some cases, re-prompt the participant, asking if they could recall any other details from the story. C-units after the re-prompt were annotated as “Retell” and were also excluded from our analyses. Thus, only spontaneous, first attempt, story-based c-units were included.

4. Annotations and Measures

4.1. Verbal Hesitations:

In this work we use the term *verbal hesitation* and define this measure in a manner similar to the dialogue-act “hold”, as described in the Switchboard SWDB-DAMSL Shallow-Discourse-Function Annotation Coder Manual [13]. Table 1 lists the words and phrases which were annotated as verbal hesitations in our data. More specifically, any combinations of these terms, occurring at the beginning of a c-unit, were annotated as being part of a verbal hesitation. The duration of a verbal hesitation was calculated as the amount of time from the start of the first word in the verbal hesitation, to the start of the first non-hesitation word following the verbal hesitation. A total verbal hesitation time was then calculated for each participant.

Table 1: Verbal Hesitations: words and phrases.

oh	and
then	so
well	yeah
anyway	let’s see
uh	um
hmm	hm
mm	eh
let me think	let’s start with

4.2. Disfluencies:

The reparandum component of word repeats, phrase repeats and revisions were annotated as disfluencies. The reparandum is the component of a speech repair that the speaker intends to be replaced [14, 15, 16]. If a c-unit was abandoned prior to completion, the speech was also annotated as disfluent. C-unit internal fillers, such as “uh”, were also marked as disfluent. The duration of disfluent speech was calculated as the time from the start of the first word in the reparandum, to either the start of

the word following the reparandum or the end of the last word in the reparandum if there was no further speech in the c-unit.

In our annotation scheme, verbal hesitations and disfluencies are mutually exclusive. Thus, any c-unit words that were repeated or revised were annotated as disfluencies unless they: 1) occurred at the beginning of a c-unit and 2) consisted exclusively of words or phrases on the verbal hesitation list shown in Table 1.

4.3. Content Qualifiers and Surrogates:

When faced with difficulty recalling something, speakers may qualify their statements, e.g., “I think it was forty dollars.” or use a surrogate phrase, e.g., “I can’t remember her name, but ...”. We identified content qualifiers and surrogates by annotating c-units that contained phrases with the word “I”. These phrases are shown in Table 2.

Table 2: Content qualifier and detail surrogate phrases found in the data.

I suppose	I don't know
I can't remember	I guess
I'm gonna go with	I forgot
I mean	I don't
I was trying to think	I said
I don't remember	I know it didn't say
I think it was	I've forgotten
I don't recall	I think
I'm not sure	all I remember
I believe	I... (unfinished)

Although measuring use of the word “I” may at first seem odd, it is clear from the list of phrases shown in Table 2 that this measure captures phrases that participants use to express uncertainty. However, it is likely this measure does not capture all of them. We use this measure here purely for expediency and leave it to future work to identify additional phrases.

5. Results

We used data for 43 participants: 15 with impairment and 28 intact. As the group sizes for these analyses is somewhat small, we chose to use non-parametric statistical analyses, specifically Wilcoxon Rank Sum and Fisher Exact tests. As such, we report median (rather than mean) as a measure of central tendency.

5.1. Verbal Hesitations and Disfluencies

First we examine the time participants spent speaking, as this is of interest when comparing to previous work using this data [6]. In addition, we will use speaking time when calculating what ratio of participants’ speaking times was spent in hesitations and disfluencies. Speaking time was calculated as the sum of each participant’s c-unit durations. Silent gaps between c-units were not included in the speaking time measure.

As shown in Figure 2, the two groups differed little in the amount of time they spent speaking, with the impaired group having a median of 29.49 seconds and intact group a median of 28.04 seconds, $p > 0.5$, NS. This result appears to conflict with the results reported by previous researchers, who found significant differences in total speaking time. We expect this is because we exclude silent gaps between c-units from our speaking time measures.

Time Speaking

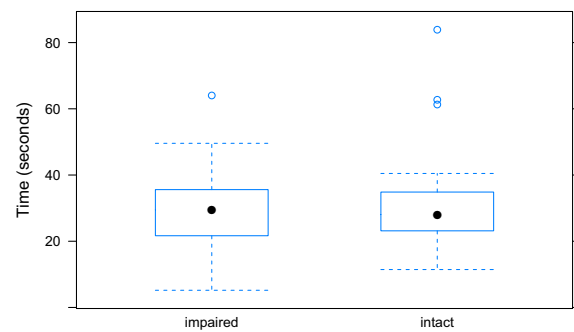


Figure 2: Box and whisker plot showing participants’ time spent speaking.

Disfluencies

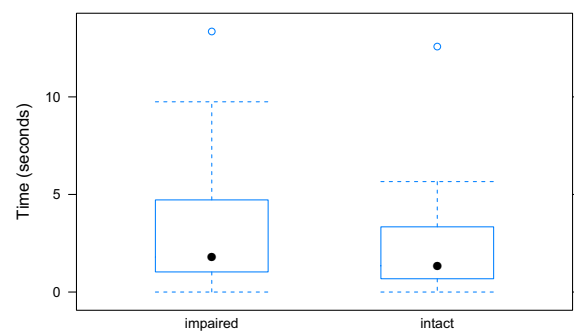


Figure 3: Time spent in disfluencies.

Next we look at the amount of time the participants spent in disfluent speech, as shown in Figure 3. The median for the impaired group was 1.79 seconds vs 1.34 seconds for the intact group. Here also, we do not find a significant difference between the two groups, $p > 0.1$, NS.

In contrast to time spent speaking and time spent in disfluent speech, the two groups do differ in the amount of time they spend engaged in verbal hesitations, as shown in Figure 4. The median verbal hesitation time for the impaired group is 3.92 seconds, and is 1.72 seconds for intact group, and the two groups differed significantly by Wilcoxon Rank Sum test, $W=297$, $p < 0.02$.

Although there is a significant difference between the two groups in terms of time spent in verbal hesitations, it is only a valid comparison because all the participants were performing the same task - that of retelling the same story. However, it is likely that this measure would not be comparable across tasks. For example, speakers may hesitate more when retelling a longer story. For this reason, we also compare the participants’ ratio of time spent in verbal hesitations to the total time they spent speaking. The ratios are shown in Figure 5. Here we see a clearer differentiation of the two groups. The impaired group’s median ratio is 0.14 and the intact group’s median is 0.06, a significant difference by Wilcoxon Rank Sum test, $W=298$, $p < 0.02$.

Next we look at how much time was spent engaging in verbal hesitations and disfluent speech combined, as shown in Figure 6. For this measure, the impaired group had a median ratio of 0.27 and the intact group a median ratio of 0.12, and the

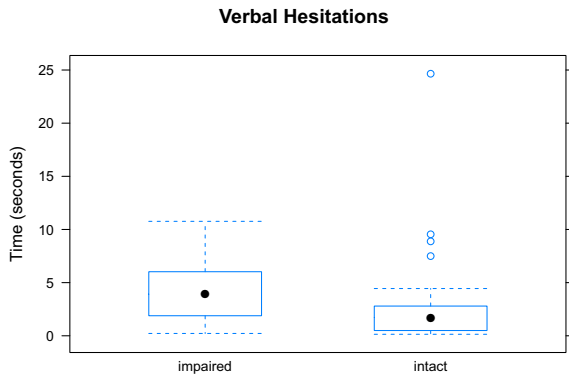


Figure 4: Time spent in verbal hesitations.

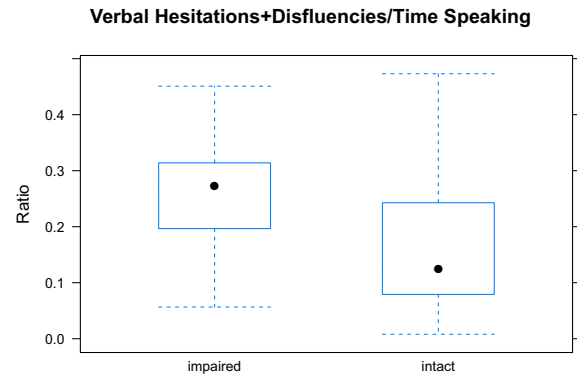


Figure 6: Ratio of participant's time spent in verbal hesitations.

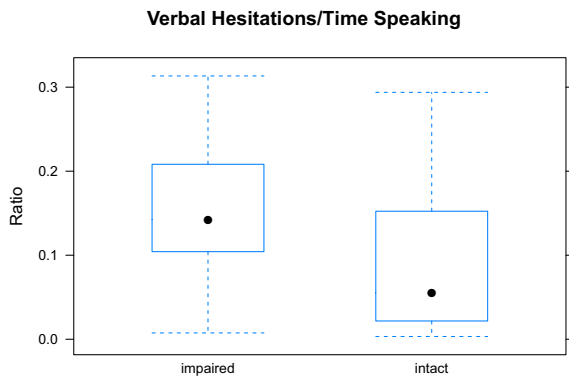


Figure 5: Ratio of participant's time spent in verbal hesitations and disfluencies combined.

two groups differed significantly by Wilcoxon Rank Sum test, $W=307$, $p<0.01$.

It is particularly interesting that, although for speaking time, disfluencies, and verbal hesitations, there were a number of outliers (e.g., the bubbles in Figure 2, 3, and 4), no outliers are present for the two ratio measures (e.g., Figures 5 and 6). This suggests that: 1) speakers who spend more time in verbal hesitations and disfluencies also tend to spend more time speaking, and 2) these ratio metrics would be good candidates for normative measures of impairment.

5.2. Content Qualifiers and Surrogates

We now look at participants' use of content qualifiers and surrogates. Table 3 shows the number of participants in each group for whom these phrases were either present in, or absent from, the story retelling. Nine of the fifteen (60%) participants with impairment used these phrases, as compared to five of the twenty-eight (18%) intact participants, a significant difference by Fisher's Exact test, $p<0.01$.

Table 3: Number of participants that used content qualifying or surrogate phrases while retelling the story.

Diagnosis	Present	Absent
impaired	9	6
intact	5	23

6. Discussion

In this work, we found that the speech of those with MCI does, in terms of certain indicators of difficulty, differ from the speech of those who are cognitively intact. Specifically, although the two groups did not significantly differ in the amount of time they spent retelling the story, the speakers with impairment spent significantly more time in verbal hesitations and their verbal hesitations accounted for a greater ratio of their story retelling time. When we add the time spent in disfluent speech to the time spent hesitating, the difference between the two groups in the ratio of time spent "off-content" is also significant. In some sense, the increased rate of verbal hesitations and disfluencies found in the speakers with impairment can be seen as an indication that they are having difficulty getting the story out.

We also found that speakers with impairment were more likely to use content qualifier and surrogate phrases. By using these phrases, a speaker can communicate that they do remember that the story contained a given element, but don't recall, or are unsure of, the details of that element. For example, "I can't remember the street she was on." communicates that the speaker does remember that a street name was given, but is unable to recall its name. Just as hesitations can be seen as difficulty in getting the story out, this use of content qualifying and surrogate phrases can be seen as indicative of difficulty in getting the story right.

In terms of assessing and understanding cognitive impairment, these content qualifications and replacements also provide insight into the participants' ability to manage memory. Specifically, it appears that those with MCI are often able to capture and maintain the structure of the story, for example recalling that a woman with children was robbed, but are sometimes unable to either store or recall specific details like the woman's name, how many children she has, or how much money was stolen.

In this work we used manually transcribed and time-aligned data to perform our analyses. However, we anticipate that automated analyses of the difficulty indicators could be used as one facet of a toolkit that speeds the diagnoses of MCI.

7. Acknowledgements

This work was supported by NIH grants P30AG008017, P30AG024978, R01AG024059, P30AG024978, M01RR000334, and UL1RR024140, and by funding from Intel Corporation, and the Department of Veterans Affairs.

8. References

- [1] K. Ritchie and J. Touchon, "Mild cognitive impairment: conceptual basis and current nosological status," *The Lancet*, vol. 355, no. 9199, pp. 225–228, Jan. 2000.
- [2] S. Larrieu, L. Letenneur, J. M. Orgogozo, C. Fabrigoule, H. Amieva, N. Le Carret, P. BarbergerGateau, and J. F. Dartigues, "Incidence and outcome of mild cognitive impairment in a population-based prospective cohort," *Neurology*, vol. 59, no. 10, pp. 1594–1599, Nov. 2002.
- [3] D. B. Howieson, R. Camicioli, J. Quinn, L. C. Silbert, B. Care, M. M. Moore, A. Dame, G. Sexton, and J. A. Kaye, "Natural history of cognitive decline in the old old," *Neurology*, vol. 60, no. 9, pp. 1489–1494, May 2003.
- [4] D. S. Knopman, B. F. Boeve, and R. C. Petersen, *Essentials of the Proper Diagnoses of Mild Cognitive Impairment, Dementia, and Major Subtypes of Dementia*. Mayo Foundation for Medical Education and Research, Oct. 2003, vol. 78, no. 10.
- [5] S. Singh, R. S. Bucks, and J. M. Cuerden, "Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech," *Aphasiology*, vol. 15, no. 6, pp. 571–583, 2001.
- [6] B. Roark, M. Mitchell, J.-p. Hosom, K. Hollingshead, and J. Kaye, "Spoken language derived measures for detecting mild cognitive impairment," in *IEEE Transactions on Audio, Speech & Language Processing*, 2011.
- [7] X. Le, I. Lancashire, G. Hirst, and R. Jokel, "Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists," *Literary and Linguistic Computing*, vol. 26, no. 4, pp. 435–461, Dec. 2011.
- [8] V. Berisha, S. Wang, A. LaCross, and J. Liss, "Tracking Discourse Complexity Preceding Alzheimer's Disease Diagnosis: A Case Study Comparing the Press Conferences of Presidents Ronald Reagan and George Herbert Walker Bush," *Journal of Alzheimer's Disease*, vol. 45, no. 3, pp. 959–963, Jan. 2015.
- [9] V. L. Smith and H. H. Clark, "On the course of answering questions," *Journal of Memory and Language*, vol. 32, no. 1, pp. 25–38, 1993.
- [10] A. Lindström, J. Villing, S. Larsson, A. Seward, N. Åberg, and C. Holtelius, "The effect of cognitive load on disfluencies during in-vehicle spoken dialogue," in *InterSpeech*, 2008, pp. 1196–1199.
- [11] D. Wechsler, *Wechsler Memory Scale - Third Edition Manual*, San Antonio, TX, 1997.
- [12] J. C. Morris, "The Clinical Dementia Rating (CDR): current version and scoring rules," *Neurology*, vol. 43, no. 11, pp. 2412–2414, Nov. 1993.
- [13] D. Jurafsky, L. Shriberg, and D. Biasca, "Switchboard: SWBD-DAMSL Coders Manual," 1997.
- [14] P. A. Heeman and J. F. Allen, "Intonational Boundaries, Speech Repairs and Discourse Markers: Modeling Spoken Dialog," in *Proceedings of the Eighth Conference on European Chapter of the Association for Computational Linguistics*, ser. EACL '97. Stroudsburg, PA, USA: Association for Computational Linguistics, 1997, pp. 254–261.
- [15] W. Levelt, "Monitoring and self-repair in speech," *Cognition*, vol. 14, no. 1, pp. 41–104, Jul. 1983.
- [16] E. E. Shriberg, "Preliminaries to a Theory of Speech Disfluencies." Ph.D. dissertation, University of California at Berkeley, 1994.