

Subjective rating scale for discourse: Evidence from the efficacy of subjective rating scale in amnesic mild cognitive impairments

JungWan Kim, PhD^a, Jihye Shim, MS^b, Ji Hye Yoon, PhD^{c,d,*}

Abstract

In clinical settings, the language ability of patients with neurologic communication disorders have been measured by quantitative parameters such as the total number of words in dialogue and picture description tasks. However, this quantitative analysis requires a long period of time in order to analyze the quantitative parameters, and results can differ according to discourse tasks. The purposes of this study are to explore whether SR-D may predict the quantitative measures of discourse tasks. Forty patients with amnesic MCI and 40 normal elderly participated in the study. We gathered responses to 10 items regarding SR-D and analyzed the quantitative measures of narrative discourse through 3 discourse tasks (i.e, picture description, dialogue, procedural discourse). We found significant differences in MLT_w, CIU, and SR-D scores between the 2 groups. In particular, 4 items were significantly correlated with the performance of MLT_w and CIU. Sensitivity and specificity of these 4 items were 100% and 75%, respectively. In terms of economic opportunity costs, objective measures cannot be evaluated to be practical, since it is used in research rather than clinical diagnosis in general. Therefore, evaluation of discourse using a few items proven in its sensitivity and specificity could allow a wide use of such measure in not only research but also in clinical diagnosis. These findings suggest that subjective measures of narrative discourse may be valid with objective language tests to predict individual discourse performance.

Abbreviations: CIU = corrected information unit, MCI = mild cognitive impairment, MLT_w = mean number of words per T-unit, SR-D = subjective ratings of discourse.

Keywords: aMCI, discourse, elderly, narrative, subjective rating

1. Introduction

Discourse is action in a communicative setting to maintain a suitable context, gather necessary information, and make an intended response in not less than 2 sentences, and is related to a person's efficient use of cognitive resources.^[1] A discourse production task provides information on strengths and weaknesses of one's communicative abilities and evaluates semantic, syntactic and pragmatic information of language, thus making it possible to assess one's comprehensive communicative ability.^[2,3] Therefore, relevant studies have used discourse tasks in order to identify the comprehensive cognitive abilities of normal individ-

uals and to distinguish the cognitive-communication impairments of individuals. It is thought that because discourse incorporates every aspect of language it is formed through complex interchanges between attention, memory, language and cognitive abilities,^[3] and a number of symptoms of individuals with low cognitive ability are identified in discourse.

Individuals with advanced Alzheimer's disease (AD), a typical symptom of degeneration dementia, showed a low rate of corrected information unit (CIU, %) and a low type-token ratio (TTR, %) in a picture description task.^[4] In addition, when quantitative measures of discourse regarding a picture description task utilizing the Boston Diagnostic Aphasia Examination (BDAE) were compared with those of a normal group, it was found out that a discourse task is effective in identifying the decreased semantic memory of an individual with AD.^[5] In a discourse task, normal elderly people showed a limited ability in a semantic aspect,^[6,7] but a relatively well-preserved ability in a syntactic aspect.^[8] However, patients with AD exhibited deficiencies in both memory and semantic aspects of language ability, even from an early stage of AD, due to temporoparietal dysfunctions, which often led to little or no information of discourse, described as "empty speech".^[9,10]

Mild cognitive impairment (MCI) is characterized by memory deficits in clinical terms. Although MCI can exhibit with a variety of symptoms, when memory loss is the predominant symptom it is termed "amnesic type of MCI (aMCI)" and is frequently seen as a preclinical stage of AD.^[11] Studies suggest that these individuals tend to progress to AD at a rate of approximately 10% to 15% per year.^[11] MCI patients also exhibited symptoms of language impairment.^[11] Some studies have held that confrontation naming and semantic verbal fluency tasks are effective in discriminating those with aMCI from the normal elderly,^[12] while others have argued that those tasks are not

Editor: Massimo Tusconi.

The MEC protocol in this study was used with the permission of the original author. We would like to thank Dr. Yves Joannette for allowing us to publish it.

The authors report no conflicts of interest.

^a Department of Speech and Language Pathology, College of Rehabilitation Sciences, ^b Rehabilitation & Science Graduate Program, Daegu University, Gyeongsan, ^c Division of Speech Pathology and Audiology, College of Natural Sciences, Hallym University, ^d Research Institute of Audiology and Speech Pathology, Hallym University, Chuncheon, Republic of Korea.

* Correspondence: Ji Hye Yoon, Division of Speech Pathology and Audiology, Hallym University, 1 Hallymdaehak-gil, Chuncheon, Kangwon, 24252, Republic of Korea (e-mail: j.yoon@hallym.ac.kr).

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Medicine (2019) 98:2(e14041)

Received: 6 September 2018 / Received in final form: 9 November 2018 /

Accepted: 14 December 2018

<http://dx.doi.org/10.1097/MD.0000000000014041>

suitable for discriminating such groups.^[13,14] In a picture description task, aMCI exhibited low discourse abilities, including shorter discourse and more difficulties in maintaining a discourse topic, than those of normal elderly people, and had difficulty in understanding comprehensive discourse.^[3,15]

To summarize the findings from the relevant studies above, not only one's memory and common cognitive abilities but also one's ability to perform discourse declines from the MCI stage. Studies have identified the importance of research into discourse in the aMCI stage as a prodromal stage of AD^[16] on the assumption that there is a close relationship between typical symptoms of dementia and discourse abilities.^[17] In addition, since it is impossible to explain a reduced cohesion of discourse only by considering one's difficulty in the lexical access at the word level, it has limitations to evaluate the language abilities of patients utilizing MCI with naming tasks only.

Concerning discourse analysis, existing studies have analyzed quantitative measures, such as content units, speech efficiency, the number of words, and anaphora, in order to assess discourse impairments. Some researchers have computed particular discourse behaviors that normally had not been considered in discourse analysis so as to compare and analyze ratios of modalizing discourse to referential discourse (M/R ratio).^[18] In spite of these various approaches for analyzing discourse by normal people or patients with neurologic communication disorders, limitations have arisen in applying discourse analysis in clinical settings because discourse analysis tends to be time-consuming. More specifically, analysis of discourse generally involves the recording of all utterances (about 10–15 minutes) followed by transcription. These complex and lengthy assessments are typically not feasible in clinical settings and particularly within the context of a hospital cognitive-communication screening. Furthermore, it is not simple to compare quantitative measures of discourse due to differences in the types of conducted discourse tasks, linguistic units (i.e., syllable, word, utterance) and time units (i.e., seconds, minutes) used in the analysis, and different languages.

In terms of diagnosis and intervention for diseases, early detection may alter the disease course. The main purpose of the present study is to explore whether subjective ratings of discourse (SR-D) may predict the quantitative measures of discourse tasks. In addition, we investigate the effectiveness of SR-D to identify and distinguish communicative decline in patients with aMCI.

2. Methods

2.1. Participants

Eighty community-dwelling elderly adults (40 aMCI patients and 40 normal elderly) residing in South Korea were recruited from May to December 2016. All participants were subjected to pre-evaluation questions in order to screen their hearing and vision. Participants had no history of neurological or psychiatric disorders. All participants gave informed written consent prior to participating in the study. This research complied with all applicable ethical rules and regulations and was approved by the Institutional Review Board of Daegu University (IRB#: 1040621-201405-HR-004-02).

2.2. Procedure and scoring

All participants were administered the following tests in order to more fully assess the characteristics of the population sample:

Table 1

Subject characteristics.

	NE (n = 40)	aMCI (n = 40)
Age	74.10 (±4.53)	78.50 (±4.76)
Level of education	4.90 (±3.47)	4.50 (±3.99)
MMSE	27.72 (±2.23)	26.29 (±1.82)
VLT-IR	17.17 (±2.93)	11.18 (±4.42)
VLT-DR	6.33 (±1.03)	2.44 (±2.08)
RCFT (copy)	29.86 (±5.01)	28.67 (±8.36)
SGDS	2.20 (±1.82)	3.83 (±2.86)
K-IADL	.12 (±.10)	.15 (±.18)

aMCI = Amnesic mild cognitive impairment, K-IADL = Korean Instrumental Activities Daily Living (cut-off score: 43), MMSE = Mini-Mental Status Examination, NE = Normal elderly, RCFT = Rey Complex Figure Test, SGDS = Short-form Geriatric Depression Test (cut-off score: 8), VLT-DR = verbal learning test (delayed recall), VLT-IR = verbal learning test (immediate recall).

- Mini-Mental Status Examination,^[19,20]
- verbal learning test (in Seoul Neuropsychological Screening Battery, SNSB),^[20] This measure involves studying a list of 12 words which is repeated 3-times. The total number of words recalled after each trial and the number of words recalled following a 20-minute delay.
- Rey Complex Figure Test-copy (in SNSB),^[20]
- Short-form Geriatric Depression Test (SGDS),^[21] and
- Korean Instrumental Activities Daily Living (K-IADL).^[22]

We followed the criteria of the modified Petersen criteria for the diagnosis of aMCI^[23] and considered the diagnosis of aMCI if memory loss is the predominant symptom (below -1.0 SD on the verbal learning test). Participants whose MMSE and Rey Complex Figure Test (copy) scores deviated at least 1 SDs from this estimate were excluded from the study.^[20] Subjects showed depression were excluded from the study based on the SGDS score. Subjects with history of head trauma, epilepsy, stroke, or diagnosis of acute psychiatric disorder were also excluded from the study. There were no significant differences between the normal elderly and aMCI with regard to age ($P=.06$), level of education ($P=.17$), MMSE score ($P=.31$), RCFT (copy) ($P=.56$), SGDS ($P=.10$) and K-IADL score ($P=.48$), except in VLT-IR ($P<.001$) and VLT-DR ($P<.01$) (see Table 1).

With respect to the discourse, each participant was required to produce a spoken discourse sample. There are many types of discourse tasks (i.e., conversational, narrative, procedural). In clinical or research settings, picture description is widely used for eliciting narrative discourse samples from adults. Picture description has the advantages of providing a standardized approach to language sampling^[24] and allowing for performance comparison within and across groups.^[18,25] Stimuli included single pictures so that picture description tasks pose fewer demands on memory. However, in contrast to conversation and procedure discourse, the picture description is not a naturalistic task for participants. All discourse tasks do not elicit comparable levels or types of performance.^[26] Since diverse tasks for discourse complexity might detect subtle changes in language ability,^[27] we used 3 different types of tasks regarding discourse. The testing order was counterbalanced to reduce order effects.

2.2.1. Conversation. In order to assess the ability of conversational discourse, stimuli of a sub-test contained in "Protocole Montreal d'évaluation de la communication" (MEC)^[28] which developed to test for patients with traumatic brain injury was used. In the original version of the conversational discourse sub-test in MEC,^[28] the test procedure involved a ten-minute

spontaneous conversation with the examiner on 2 different topics. In accordance with the original version, our examiner initiated 2 different topics of conversation of interest to the subject such as family, hobbies, or current events.

2.2.2. Picture narratives. Spontaneous speech samples in response to the “Beach” picture of the Western Aphasia Battery-revised version (WAB-R),^[29] were elicited in verbal modality. Subjects were instructed to tell the examiner about everything happening in the picture. Verbal responses were tape-recorded and transcribed.

2.2.3. Procedural discourse. Subjects were instructed to tell the examiner “how to make a noodle”. In order to elicit more verbal output in subjects whose narratives were insufficient or incomplete, prompts that did not reveal the content of the narratives such as “do you have anything else to add?” were offered. Discourse samples were audio-taped and transcribed verbatim.

2.3. Discourse rating

2.3.1. Subjective rating. Regarding the subjective rating, rating items for the conversational discourse in MEC^[28] were used. Among the 17 rating items in the original version of MEC, items related to cognitive dysfunction and language problem in aMCI were selected for use in this study. According to content validity, the 4th item (production of inappropriate or unexpected comments), 10th item (misunderstands what is said), 11th item (misunderstands indirect language), 12th item (remains indifferent to humorous comments), 14th item (use of monotone voice), 16th item (lack of facial expression), and the 17th item (inconsistent or absent visual contact) were excluded and the other 10 items were included in SR-D. Following the responses on the conversational discourse task, 2 speech pathologists rated the performance of participants utilizing a 3-point scale, with 0 indicating many; 1, often; and 2, none (Table 2).

2.3.2. Objective analysis. Concerning objective analysis, the speech samples for the 3 discourse tasks were recorded and transcribed. The transcriptions were systematically checked by a second examiner. Data were analyzed for complexity, speech rate, and quality. Complexity was determined by the mean number of words per T-unit (MLT_w) and TTR. Speech rate was determined by the mean number of syllables per second. Quality was determined by the CIU.^[30] CIU (%) was obtained by totaling the number of “relevant, truthful, non-redundant” facts^[30] for 3 discourse tasks.

Table 2
Selected 10 items of SR-D.

No.	Items	Many	Often	None
1	Word-finding problems/incorrect choice of words	0	1	2
2	No self-correction of word errors	0	1	2
3	Imprecise expression of ideas	0	1	2
5	Inappropriate topic switches	0	1	2
6	Lack of verbal initiative	0	1	2
7	Excessive talking	0	1	2
8	Repetitiveness	0	1	2
9	Interruptions	0	1	2
13	Loses track of conversation	0	1	2
15	Speech rate too slow	0	1	2

min 0~max 20 point.
SR-D= subjective ratings of discourse.

2.4. Reliability

As a quality control measure in our study, 2 speech and language pathologists trained for 10 different communication behaviors checklist. They were blind to the participants’ clinical status and reevaluated 10% of the patients’ samples subjectively and objectively. The rates of agreement between the evaluators were 91.1% and 97.5%, respectively. A consensus was used in cases of discrepancy.

2.5. Statistical analyses

To identify the objective and subjective differences in performance for the 3 discourse tasks between 2 groups, the level of education and age were controlled through analysis of covariance (ANCOVA). With respect to the relationship between the SR-D and quantitative measures of narrative discourse, and the MMSE score was assessed utilizing multiple linear regression analysis, controlling for age, education, gender, and depressive symptoms. Partial correlation analysis was conducted to define the sub-items of SR-D which correlated with the objective measures. Following the result of the partial correlation analysis, ROC analysis was conducted in order to provide the cut-off score, sensitivity, and specificity of the selected sub-items of SR-D. Finally, the relationship between the selected 4 items of SR-D and quantitative measures of narrative discourse and MMSE score was assessed employing multiple linear regression analysis, controlling for age, education, gender, and depressive symptoms.

3. Results

Of the 80 participants, 4 aMCI subjects (5%) were found to exhibit depressive symptoms. The demographic information is shown in Table 2. An ANCOVA was conducted to confirm the differences in the objective and subjective measures between the normal elderly (NE) and aMCI groups after adjusting for age and education. In comparing the mean number of the objective and subjective measures, there were significant differences found in MLT_w ($F=8.245, P<.01$), CIU ($F=27.202, P<.001$), and SR-D score ($F=192.114, P<.001$) (Table 3). The multiple regression model with variables such as the MMSE score and objective measures accounted for 73.3% of the variance in the SR-D ($F=20.131, P<.001$). Results indicated that only the CIU ($B=.075, P\leq.001$) and the MMSE ($B=.339, P\leq.001$) were significant predictors (Table 4).

To evaluate the relationship between MLT_w, CIU, and sub-items of SR-D, partial correlation analysis was conducted after adjusting for age, education, and MMSE score. Among the 10 items of the SR-D, items 3 (“imprecise expression of

Table 3
Comparison of narrative discourse between 2 groups.

	NE Mean (SD)	aMCI Mean (SD)	F	ANCOVA P value < .05
MLT _w	4.77 (.86)	2.73 (.64)	8.245	$P<.01^{**}$
TTR	77.92 (8.63)	76.93 (4.85)	1.576	$P=.215$
Speech rate	2.55 (.61)	2.37 (.55)	.005	$P=.975$
CIU	80.42 (6.94)	53.25 (12.52)	27.202	$P<.001^{***}$
SR-D score	19.63 (.81)	13.68 (1.06)	192.114	$P<.001^{***}$

aMCI=Amnesic Mild Cognitive Impairment, CIU=corrected information unit (%), MLT_w=mean number of words per T-unit, n.s. = not significant, NE=Normal elderly, speech rate=mean number of syllables per sec, SR-D= subjective ratings of discourse (total score: 20points), TTR=type token ratio. $^{*}P<.05, ^{**}P<.01, ^{***}P<.001$.

Table 4
Relationship between MMSE score and quantitative measures of narrative discourse and SR-D.

	CIU			MMSE		
	B	SE	P	B	SE	P
SR-D	.075	.017	.000***	.339	.077	.000***

* $P < .05$, ** $P < .01$, *** $P < .001$, CIU=corrected information unit (%), MMSE=the Mini-Mental State Examination, n.s.=not significant.

Table 5
Relationship between MMSE score and quantitative measures of narrative discourse and short-form SR-D.

	CIU			MMSE		
	B	SE	P	B	SE	P
short-form SR-D	-.018	.008	.021*	-.092	.035	.011*

CIU=corrected information unit (%), MMSE=the Mini-Mental State Examination, n.s.=not significant.
 * $P < .05$, ** $P < .01$, *** $P < .001$.

ideas”), 6 (“lack of verbal initiative”), 7 (“excessive talking”), 13 (“loses track of conversation”) were significantly correlated with the performance of MLT_w and CIU. Based on the results of the ROC analysis regarding these 4 items (short-form SR-D), the cut-off score was determined to be 7 points. Sensitivity and specificity were 100% and 75%, respectively (Fig. 1). The multiple regression model with the variables such as MMSE score and objective measures significantly predicted 57.8% of the variance in short-form SR-D ($F=9.791$, $P < .001$). Results indicated that only the CIU ($B = -.018$, $P < .05$) and the MMSE ($B = -.092$, $P < .05$) were significant predictors (Table 5).

4. Discussion

Under the premise that there are some limitations in determining MCI in elderly individuals through language batteries mainly

comprised of naming task, this study intended to examine the correlation between objective measurement and subjective measurement of discourse using a discourse task that reflects comprehensive and complex integrated communication skills, and to examine the clinical utility and significance of subjective measures of discourse.

Results showed that the performance on objective measures, the MLT_w and CIU parameters, declined in the aMCI group compared to the NE group. First, MLT_w refers to the MLT_w , and thus, the decline of MLT_w in the aMCI group reflects a small number of words in a sentence and shorter utterance length. As a consequence, there is the possibility that MLT_w includes simple sentence structures. CIU is defined as “words that are clear in context and provide suitable information on the topic or task”.^[30–32] Therefore, lower CIU reflects insufficient contextual information. The lack of speech rate difference between the

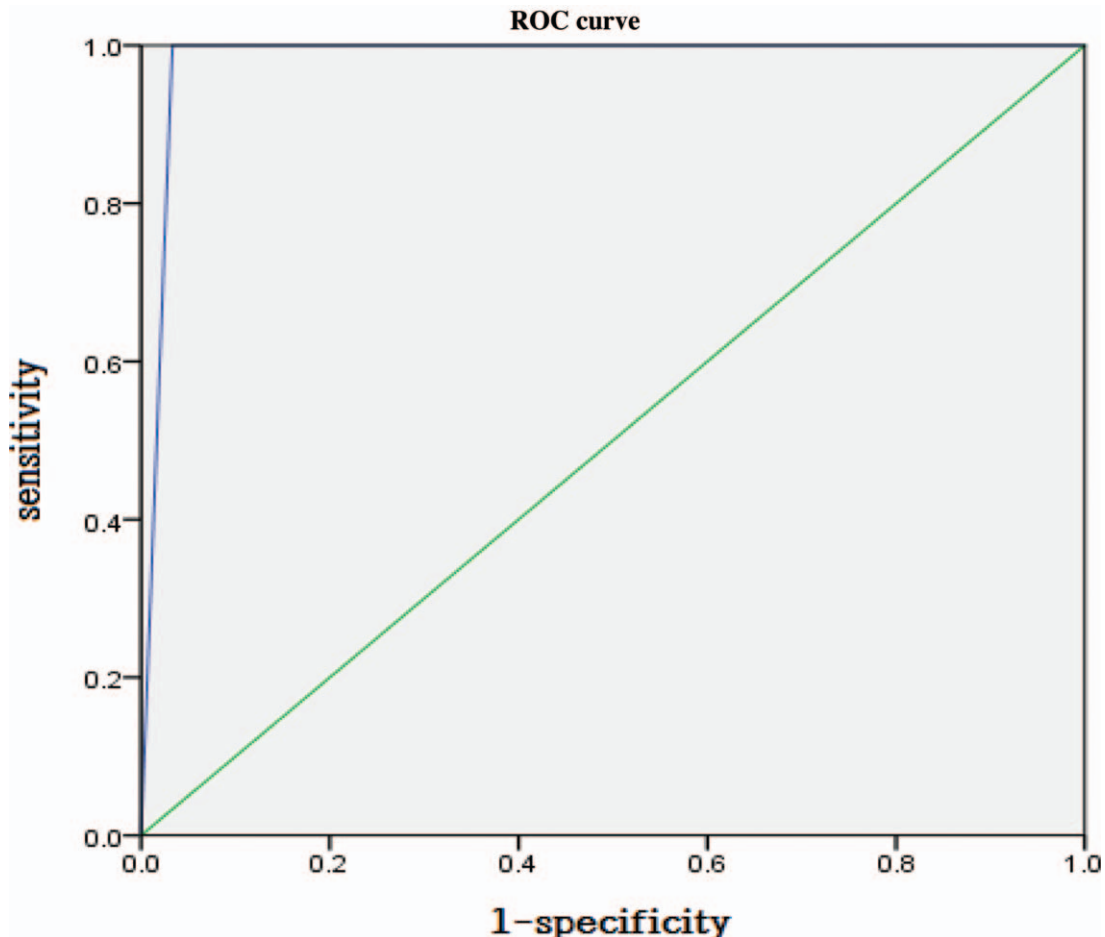


Figure 1. Result of analysis on ROC curve. cut-off score: 7 points (sensitivity: 100%, specificity: 75%, under the area:.983). ROC=receive operating characteristics.

2 groups, despite the decline in MLTw and CIU in the aMCI group, shows that although the 2 groups produce words at a similar speed, the aMCI group used relatively simpler sentence structure and insufficient contents, and produced utterance with declined efficiency and information communication in discourse. Further, a lack of statistically significant differences between aMCI and NE groups regarding TTR could be interpreted as that although the aMCI group produces various types of words during discourse, these words may not contain accurate information. To summarize these results in an objective perspective, although there is no change or increase in the physical amount of discourse produced by the elderly, the amount of contextually relevant information and the amount of communicated topic declined, which is consistent with previous studies.^[7,33,34] Therefore, these results suggest the importance of observing the differences in quality, as well as physical difference, in analyzing discourse.

The second notable finding is that the performance in the subjective measures of SR-D total score in the 2 groups differed. This result shows that subjective measures are sufficient to verify the decline in discourse ability in the aMCI group. To identify the effects of SR-D score and other variables in more detail, multiple regression analysis was conducted, in which MMSE, as well as CIU, could predict SR-D score by approximately 73%. Since CIU is a type of linguistic parameter, it is not surprising that the SR-D score can be predicted. However, MMSE is a screening test to measure overall cognitive function that includes questions assessing language function, as well as sub-areas of cognition such as attention, memory, and orientation. Therefore, the fact that MMSE could predict SR-D total score shows that discourse and various cognitive processes are related. This is in agreement with previous studies that reported that only the activation of the language region in the left hemisphere is required to perform language tasks at the single word or sentence level, while the right hemisphere and frontal cortex activation is required to accompany left hemisphere activation in order to produce discourse requiring not only basic language processing capacity such as semantic, syntax, and phonology but also higher cognitive processing capacity, including memory and executive function.^[35–37]

However, SR-D total score reflects the sum of the 10 questions that assess various aspects of discourse, and thus a selective process was required to identify which question is more closely related to objective measures. Therefore, SR-D items were tested using partial correlation analysis in relation to CIU and MLTw, which showed significant results in objective measures, in which 4 items were found to have high correlation. First, “item 3: imprecise expression of ideas” and “item 6: lack of verbal initiative” could be interpreted as a measure of decline in the ability to access the lexicon and to retrieve lexical entry (i.e., words) to effectively express appropriate information in the aMCI group, instead of the deficit in the lexical-semantic storage itself. Therefore, these items reflect delay in retrieving words in a discourse situation, resulting in difficulty in starting utterance and decline in suitability of content. Further, “item 13: loses track of conversation” could be interpreted as a measure of decline in understanding the context or in executive function to organize and to summarize information in the discourse to understand the overall topic of discourse, due to the decline in attention or working memory. “Item 7: excessive talking” is related to the lack of effective communication despite the large number of utterance, due to lengthy expression to compensate for the defect in retrieving suitable words, and eventually leading to production

of excessive utterances compared with the appropriate number of utterances that should be produced. Based on the results, the short-form SR-D presented in this study compressed into 4 questions is comprised of questions that can evenly evaluate the expression, reception and pragmatics domains of discourse. It can be concluded to show a correlation with efficiency of information delivery and various cognitive processes.

The third important result is that clinical effectiveness of the subjective measure, which could detect decline in language capacity, was verified through sensitivity and specificity analysis of the SR-D item. Traditionally, objective measures were effective in detecting defects, but its limitations included time-consuming transcription and analysis of discourse. In terms of economic opportunity costs, objective measures cannot be evaluated to be practical, since it is used in research rather than clinical diagnosis in general. Therefore, the evaluation of discourse utilizing a few items proven in its sensitivity and specificity could allow a wide use of such measures in not only research but also in clinical diagnosis.

5. Conclusion

Since lexical-semantic processing is not affected in every case of Alzheimer’s disease,^[12] the discourse task that observes comprehensive communication skills can be utilized further to verify reduction of communication ability due to decline of cognitive ability. This study examined utility of subjective measurement of discourse. Indeed, subjective measures have limitations in producing different results dependent on the examiner’s ability. However, the specialization of the scoring system as seen in the simplified 3-score measure (e.g., 0-1-2) in the SR-D scoring system could address such limitations, and online or offline training courses for examiners would promote the standardization of scoring for examiners and establish the reliability of results. This study can be utilized in examining decline of cognitive ability through quick and easy discourse screening test in patients with cognitive-communication disorders and briefly verify the declined dimensions of discourse, allowing for a more detailed and comprehensive picture on the neuropsychological profile of aMCI regarded as the prodromal stage of Alzheimer’s disease. Further studies will be needed to confirm the external and internal validation for the new evaluation tool with conservative language tests. This study has some limitations in that we classified the participants only according to the amnesic type which has memory deficits. If a sufficient number of subjects are recruited in future studies, we will be able to identify the characteristics of the subtypes of MCI (amnesic type or non-amnesic type). In addition, we did not classify the 2 subtypes of aMCI (aMCI with single domain or multiple domains) in this study. Distinguishing between these 2 subtypes will necessitate checking visuospatial function, language ability, and frontal-executive function, in addition to memory function. However, due to the time constraints of the experiment, other tests besides VLT and RCFT could not be performed. Further studies will be necessary to check the characteristics of the sub-types of aMCI by adding additional tests such as frontal-executive function tests.

Author contributions

Conceptualization: JungWan Kim, Jihye Shim, Ji Hye Yoon.
Data curation: Jihye Shim.
Formal analysis: JungWan Kim, Jihye Shim.
Methodology: JungWan Kim.

Supervision: JungWan Kim, Ji Hye Yoon.

Writing – original draft: JungWan Kim, Ji Hye Yoon.

Writing – review & editing: JungWan Kim, Ji Hye Yoon.

Ji Hye Yoon orcid: 0000-0003-1403-2276.

References

- [1] Foos PW. Working memory resource allocation by young, middle-aged, and old adults. *Exp Aging Res* 1995;21:239–50.
- [2] Burke DM, Shafto MA. Language and aging, in *The Handbook of Aging and Cognition*, eds Craik F. I. M., Salthouse T. A., editors. (Mahwah, NJ: Lawrence Erlbaum), 2007; 2013–2029.
- [3] Fleming VB, Harris JL. Complex discourse production in mild cognitive impairment: detecting subtle changes. *Aphasiology* 2008;22:729–40.
- [4] Kavé G, Dassa A. Severity of Alzheimer's disease and language features in picture descriptions. *Aphasiology* 2018;32:27–40.
- [5] Ahmed S, de Jager CA, Haigh AM, et al. Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease. *Neuropsychol* 2013;27:79–85.
- [6] Albert MS, Heller HS, Milberg W. Changes in naming ability with age. *Psychol Aging* 1988;3:173–8.
- [7] Kemper S, Thompson M, Marquis J. Longitudinal change in language production: effects of aging and dementia on grammatical complexity and propositional content. *Psychol Aging* 2001;16:600–14.
- [8] Obler LK, Nicholas M, Albert ML, et al. On comprehension across the adult life span. *Cortex* 1985;21:273–80.
- [9] Ehrlich JS, Obler LK, Clark L. Ideational and semantic contributions to narrative production in adults with dementia of the Alzheimer's type. *J Commun Disord* 1997;30:79–99.
- [10] Kempler D, Lubinski R. Language changes in dementia of the Alzheimer type. *Dementia and communication*. Decker Inc, Philadelphia: B. C.:1995;98–114.
- [11] Petersen RC, Doody R, Kurz A, et al. Current concepts in mild cognitive impairment. *Arch Neurol* 2001;58:1985–92.
- [12] Taler V, Phillips NA. Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. *J Clin Exp Neuropsychol* 2008;30:501–56.
- [13] Beinhoff U, Hilbert V, Bittner D, et al. Screening for cognitive impairment: a triage for outpatient care. *Dement Geriatr Cogn Disord* 2005;20:278–85.
- [14] Lopez-Higes R, Prados JM, del Rio D, et al. Semantic verbal fluency of animals in amnesia-type mild cognitive impairment. *Revista Neurologia* 2014;58:493–9.
- [15] Chapman SB, Zientz J, Weiner M, et al. Discourse changes in early Alzheimer disease, mild cognitive impairment, and normal aging. *Alzheimer Dis Assoc Disord* 2002;16:177–86.
- [16] Aramaki E, Shikata S, Miyabe M, et al. Vocabulary size in speech may be an early indicator of cognitive impairment. *PLoS One* 2016;11:e0155195.
- [17] Snow P, Douglas J, Ponsford J. Conversational discourse abilities following severe traumatic brain injury: a follow up study. *Brain Inj* 1998;12:911–35.
- [18] Duong A, Tardif A, Ska B. Discourse about discourse: what is it and how does it progress in Alzheimer's disease? *Brain Cogn* 2003;53:177–80.
- [19] Folstein MF, Folstein SE, McHugh PR. Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98.
- [20] Kang YW, Na DL. Seoul Neuropsychological Screening Battery. Incheon South Korea Hum Brain Res Consult 2003.
- [21] Ki BS. A preliminary study for the standardization of geriatric depression scale short form-Korea version. *J Korean Neuropsychiatr Assoc* 1996;35:298–307.
- [22] Kang SJ, Choi SH, Lee BH, et al. The reliability and validity of the Korean instrumental activities of daily living (K-IADL). *J Korean Neurol Assoc* 2002;20:8–14.
- [23] Roberts RO, Geda YE, Knopman DS, et al. The Mayo clinic study of aging: design and sampling, participation, baseline measures and sample characteristics. *Neuroepidemiology* 2008;30:58–69.
- [24] Cooper PV. Discourse production and normal aging: performance on oral picture description tasks. *J Gerontol* 1990;45:210–4.
- [25] Mackenzie C, Brady M, Norrie J, et al. Picture description in neurologically normal adults: concepts and topic coherence. *Aphasiology* 2007;21:340–54.
- [26] Shadden BB, Burnette RB, Eikenberry BR, et al. All discourse tasks are not created equal. *Clin Aphasiology* 1991;20:327–41.
- [27] North AJ, Ulatowska HK, Macaluso-Haynes S, et al. Discourse performance in older adults. *Int J Aging Hum Dev* 1986;23:267–83.
- [28] Joannette Y, Ska B, Côté H. Protocole MEC protocole Montréal d'évaluation de la communication (MEC). Isbergues Ortho Edition 2004.
- [29] Kim H, Na DL. Paradise Korean version-the western aphasia battery (revised). Seoul: Paradise 2012.
- [30] Brookshire RH, Nicholas LE. Speech sample size and test-retest stability of connected speech measures for adults with aphasia. *J Speech Lang Hear Res* 1994;37:399–407.
- [31] Kwon M, Kim H, Choi SS, et al. A study for analyzing spontaneous speech of Korean adults with CIU scoring system. *Korean J Commun Disord* 1998;3:35–49.
- [32] Nicholas LE, Brookshire RH. A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *J Speech Lang Hear Res* 1993;36:338–50.
- [33] Juncos-Rabadán O, Pereiro AX, Rodríguez MS. Narrative speech in aging: quantity, information content, and cohesion. *Brain Lang* 2005;95:423–34.
- [34] Marini A, Boewe A, Caltagirone C, et al. Age-related differences in the production of textual descriptions. *J Psycholinguist Res* 2005;34:439–63.
- [35] Blake ML. Clinical relevance of discourse characteristics after right hemisphere brain damage. *Am J Speech Lang Pathol* 2006;15:255–67.
- [36] Gold DP, Andres D, Arbuckle T, et al. Off-target verbosity and talkativeness in elderly people. *Can J Aging* 1993;12:67–77.
- [37] Shadden BB. The use of discourse analyses and procedures for communication programming in long-term care facilities. *Top Lang Disord* 1995;15:75–86.