Early recognition of neurocognitive disorders: Dementia screening in primary care and the detection of mild cognitive impairment via verbal fluency tests

Summary of Ph.D. Dissertation

Réka Balogh, M.A.

Doctoral School of Clinical Medicine

Department of Psychiatry, Albert Szent-Györgyi Health Centre

Faculty of Medicine, University of Szeged

Supervisors:

János Kálmán, M.D., Ph.D., D.Sc.

Professor

Department of Psychiatry, Albert Szent-Györgyi Health Centre

Faculty of Medicine, University of Szeged

Gábor Gosztolya, Ph.D.

Senior Research Fellow

ELKH-SZTE Research Group on Artificial Intelligence

Szeged

2022

Papers the thesis is based on:

I. **Balogh, R.**, Imre, N., Papp, E., Kovács, I., Heim, S., Karádi, K., Hajnal, F., Pákáski, M., & Kálmán, J. (2019). Dementia in Hungary: General practitioners' routines and perspectives regarding early recognition. *European Journal of General Practice*, *26*(1), 1–7.

SJR Indicator: D1
IF: 1.904

II. **Balogh, R.**, Imre, N., Gosztolya, G., Hoffmann, I., Pákáski, M., & Kálmán, J., (2022). The Role of Silence in Verbal Fluency Tasks - A New Approach for the Detection of Mild Cognitive Impairment. *Journal of the International Neuropsychological Society*, 24;1-13.

SJR Indicator: Q2
IF: 3.114

Cumulative impact factor of publications related to this thesis: **5.018**

I. INTRODUCTION

1. Mild cognitive impairment and the role of primary care

Neurocognitive disorders represent a worldwide problem: at present, more than 55 million people live with dementia, and the number of cases is estimated to reach around 150 million by the year 2050. Mild cognitive impairment (MCI) is a heterogenous syndrome, characterized by the subtle deficit of memory, language, and executive skills, and is often considered the prodromal stage of dementia. At present, there are no available disease-modifying therapeutic options for MCI or dementia: the existing treatments offer only symptomatic effects on the diseases. However, early recognition and timely diagnosis are crucial: they can provide an opportunity to reduce the rate of cognitive decline, as interventions applied at the early stage of the disease are more likely to be effective. Early detection allows better patient follow-up and helps to observe the disease mechanism as well, while also offering a chance for the patients and their relatives to start planning for the future. Despite its importance, the difficulty of early dementia recognition is a global problem: research suggests that a substantial amount of dementia cases (up to 60%) might be missed in community settings or residential care.

General practitioners (GPs) are greatly involved in the early stages of the dementia recognition process, as they treat the general somatic problems of the elderly and most patients also visit them first to have their initial cognitive examination. Besides the complaints reported by the individual or their family members, GPs' concerns about signs of dementia during patient consultation, targeted case-finding, and population screening can also be potential pathways to the early identification of cognitive disorders. As prior studies suggest that a significant obstacle to effective dementia detection is the low use of cognitive assessment tools, it is crucial to offer healthcare providers cognitive screening instruments that are quick, simple, and can yield objective and reliable results. In addition to this, the evaluation of GPs' routines and views regarding dementia screening could also help to enhance the low recognition rates.

2. Verbal fluency tests

Verbal fluency (or word generation) tests are some of the most common neuropsychological tests, administered both in research and clinical settings. While their asset requirement is minimal, their significant advantage is that they can be administered to individuals of various ages, levels of education, and socioeconomic status. Verbal fluency tests can be divided into two subtypes: *phonemic* (PVF) and *semantic* (SVF) verbal fluency, also known as letter fluency and category fluency, respectively. In the standard versions of the tests, subjects are given 60

seconds to recall as many words as they can, which begin with a given letter (PVF) or belong to a given semantic category (SVF). Besides the above-mentioned, which are the most commonly used types of verbal fluency tasks, action (or verb) fluency also has significant value. When performing action fluency, the participants have to produce as many verbs ("things that people do") as they can. In this thesis, action fluency will be regarded as an SVF task because both semantic fluencies and action fluency are content-oriented speech tasks.

Performing verbal fluency tasks requires the simultaneous activation of multiple cognitive processes. Both PVF and SVF tests assess executive functions (such as divergent thinking for generating category examples, and flexibility while searching subcategories) and engage the working memory (the subjects need to keep the exact instruction and prior responses in mind). Cognitive control processes also play a major role in the execution of verbal fluency tests, as during the test one must repress the repetitions and any potentially incorrect or irrelevant responses. While both PVF and SVF require rapid associative exploration, SVF relies more on semantic associations. Thus, SVF reflects more on the integrity of semantic memory, while PVF is more dependent on search strategies based on lexical representation. Since executive functions, among others, are proven to be altered in dementia and other forms of neurocognitive impairments, fluency tests have great potential to become effective screening tools.

The most traditional and widely-used method used to assess fluency performance requires the clinician to score the number of unique and correct words that are produced by the participant while also counting the repetitions, perseverations, and intrusive (incorrect) words. This method can be refined by comparing the scores measured in different time intervals or by only considering the number of produced words regarding one interval. These approaches have the advantage of enabling us to gain information about the temporal dynamics of word production. A more elaborate method, the so-called *cluster-analysis* or *clustering*, is based on grouping the consecutive words that are similar in some respects, e.g., rhyming words or homonyms in case of any PVF, and pets or farm animals in case of the animal fluency task.

In the past years, aiming to address the limitations of manual methods and to achieve large-scale analysis with objective and quick results, there have been multiple attempts to automatize the application and analysis of verbal fluency tests. Most of these attempts are focused on the automatization of word scoring or cluster analysis. A major difficulty regarding the fully automatic end-to-end analysis of audio fluency recordings stems from the characteristics of the general 1-minute response. Voice recordings of fluency test performances often comprise more than solely a sequence of task-relevant words. They usually also contain

extraneous speech, such as filler words or hesitations, irrelevant comments, questions directed at the experimenter, utterances expressing loud thinking, or other parts of speech, like conjunctions. To be able to automatically analyze the relevant parts, fluency recordings need to go through an often time-consuming preparation process prior to analysis: the words irrelevant to the tasks need to be removed from the recording/transcription and, in some cases, words need to be lemmatized (i.e., to be converted to their stem). Given the substantial amount of task-irrelevant content in most fluency recordings, the question arises whether the analysis of these segments could provide valuable information regarding the overall performance of the patient.

II. AIMS OF THE STUDIES

Study 1:

The first study focuses on dementia screening in Hungarian primary care. Therefore, in this exploratory study, we aimed to:

(I) examine GPs' views regarding the early recognition and the current recognition rates of dementia

(II) identify the methods Hungarian GPs use for dementia screening

(III) evaluate GPs' satisfaction with the available and most widespread neurocognitive and dementia screening tests

(IV) explore GPs' ideas about an ideal test for early recognition and those optimal circumstances that could contribute to more effective dementia identification in Hungarian primary care

Study 2:

The focus of the second exploratory study of the thesis was to examine PVF and SVF audio recordings by moving beyond the words listed by the participants and thus by exploring the additional, previously unharvested data present in the fluency recordings. Our aims were to:

(I) examine whether the derived temporal parameters differ between participants classified as healthy control (HC) and as MCI

(II) compare the traditional, word count-based method and the temporal parameters regarding their ability to detect differences in the performance of the HC and MCI groups

(III) compare the different (phonemic and semantic) types of fluency tasks investigating their sensitivity to the presence of MCI

III. METHODS AND MATERIALS

1. Study 1

1.1. Participants

The participants of this study were Hungarian GPs, who were investigated within the framework of a national project in collaboration between the University of Szeged and the University of Pécs. To reach as many GPs as possible from every region of the country, questionnaires were distributed onsite at six major mandatory training courses and three national conferences between February and November 2014. Participation was entirely voluntary and anonymous. Ethical approval was obtained from the Regional and Institutional Research Ethics Committee of the University of Pécs (reference number: 5244).

1.2. Instrument

As part of the extensive research project, a self-administered questionnaire was designed specifically to explore a broad range of aspects regarding GPs' views on and their role in dementia detection and management in Hungary. In the survey, several significant topics were investigated, including a subtopic on GPs' routines and perspectives regarding dementia screening and detection, which is addressed in this thesis. The questions analyzed were fixed-response (single or multiple choice) and Likert-type questions; open-ended questions were not applied.

1.3. Statistical analysis

Data were analyzed using the SPSS v.24 statistical analysis software package. Descriptive statistics were applied for all items on the questionnaire. Comparative analysis was executed for one question using the Wilcoxon signed ranks test. The level of significance was set at 0.05.

2. Study 2

2.1. Participants

Participants (patients and their relatives) were recruited at the Memory Clinic of the Department of Psychiatry, University of Szeged. Participation was entirely voluntary and anonymous; the subjects did not receive any remuneration for their participation. Ethical approval was obtained by the Regional Human Biomedical Research Ethics Committee of the University of Szeged (Reference No. 231/2017-SZTE).

The inclusion criteria were: at least 50 years of age, a minimum of 8 years of formal education, and Hungarian as the native language. The two main exclusion criteria were the

presence of dementia or major cognitive deficits (Mini-Mental State Examination (MMSE) ≤ 24) and acute depressive symptoms (15-item Geriatric Depression Scale (GDS-15) ≥ 7). In addition, individuals were excluded from the study if they had any past or present neuropsychological, psychotic, or affective disorders, head injuries, stroke, substance abuse disorders, major (uncorrected) hearing loss, or language problems (e.g., stutter), based on patient history and medical records. Participants with MRI or CT records showing evidence of microhemorrhages, lacunar or other infarctions, cerebral contusion, encephalomalacia, aneurysm, vascular malformation, or space-occupying lesions were also excluded. Out of 79 initially recruited persons, after reviewing and evaluating the criteria, 50 subjects were considered eligible for inclusion in the study.

Participants were split into two groups based on their MMSE scores. MMSE cut-off scores were determined based on the results of previous studies conducted by our research group. In the research described in this thesis, participants achieving a score of 29 to 30 points were considered healthy control (HC) subjects, while participants achieving a score of 25 to 28 points formed the MCI group.

2.2. Study protocol

Each participant performed an extensive neuropsychological test battery: six verbal fluency tasks, the Forward and Backward Digit Span Test, the Non-Word Repetition Test, the Listening Span Test, the Clock Drawing Test (CDT), and the Alzheimer's Disease Assessment Scale – Cognitive Subscale. The fluency tasks were implemented in a fixed order, separated by the five shorter cognitive tests. We also ensured that tasks assessing the same cognitive domain did not follow each other directly.

In the three PVF tasks, the participants were asked to list as many words as they can, starting with the letters "k", "t", and "a", respectively, while avoiding proper nouns. The starting letters in this study were chosen based on previous studies conducted with Hungarian-speaking population. For the SVF tasks, participants had to name as many animals, food items, and verbs as they could. Regarding the SVF tasks, the participants were instructed to avoid saying variations of the same word stem (e.g., horse, horses; go, goes). For all six verbal fluency tasks, participants had 1 minute to perform the task. Every fluency task was recorded using an Olympus Digital Voice Recorder (16 kHz sampling rate, 16-bit resolution).

2.3. Analysis method based on temporal parameters

Voice recordings of all fluency tasks were manually transcribed in Praat, a free language software enabling speech analysis. Besides the task-relevant answers of the participants, the

transcriptions also contained the silent pauses, the hesitation sounds (filled pauses, like *hmm* and *er*), and the irrelevant utterances, such as comments or loud thinking. For each recording, they were annotated based on their boundaries (i.e., their exact start and end times), providing their duration measures.

Based on this, the total number, the average length, and the total length of *silent pauses*, the total number, the average length, and the total length of *hesitations*, and the total number, the average length, and the total length of *irrelevant utterances* were calculated. Besides these parameters, the mean time between two consecutive task-relevant words (*average word transition time*) was also calculated. Not only correct words but also errors and repetitions were considered task-relevant words. The average word transition time (irrelevant of its content, such as silent pause, hesitation, or irrelevant utterance) provided information about the average time the participant needed to produce a new task-relevant word.

2.4. Traditional fluency analysis based on word count

In the traditional scoring method, we calculated the number of correct words, the number of errors, and the number of repetitions (or perseverations). In the case of animal fluency, when a participant recalled synonymous words (e.g., cat and kitten), variations in gender (e.g., hen and rooster), or an animal and its offspring (e.g., horse and foal), words were only scored as one. The participants did not receive points for naming a subcategory if they also gave specific examples of it.

2.5. Statistical analysis

Descriptive statistical analysis was used to examine the demographic features, the neuropsychological test scores, and the fluency measures of the participants. Comparisons were executed using the Mann–Whitney U test; categorical variables were compared using the Chi-square test. Effect sizes were calculated using the Pearson correlation coefficient. Receiver operating characteristic (ROC) analysis was applied to assess the classification abilities. For the comparison of classification abilities, the area under the curve (AUC) variables were compared based on the method of DeLong, DeLong, and Clarke-Pearson. For all statistical comparisons, the level of significance was set at 0.05. All analyses were performed using *SPSS* v.24, except for the comparison of AUCs, for which the *MedCalc Statistical Software* v.19.6. was utilized.

IV. RESULTS

This thesis summary reports the most relevant results and statistical analyses of the two studies.

1. Study 1

Altogether, 402 GPs handed back their completed questionnaire, which is more than 8% of all 4,850 GPs practicing in Hungary in 2014. Regarding the questionnaire, the completion rate varied for each question, therefore, the number of responses is indicated in brackets for each question.

Regarding dementia screening, the vast majority of GPs reported that they ask the patient general questions (91%; *n* = 355) or they gather information from relatives (64%; *n* = 253). Only a quarter of them (24%; *n* = 95) indicated that they utilize cognitive tests for this purpose and some did not perform any examination at all to test for the possible occurrence of dementia (5%; *n* = 22). Two of the most widely used tests for dementia evaluation, the MMSE and the CDT, are fairly well-known among respondents: most GPs reported that they knew CDT (89%; *n* = 307) and slightly fewer people stated familiarity with MMSE (76%; *n* = 265). More than two-thirds of respondents indicated they were (completely or mostly) satisfied with the CDT (69%; *n* = 152) while a slightly lower percentage of them expressed satisfaction with the MMSE (65%; *n* = 98).

Supporting the importance of dementia recognition in the early stages, the vast majority (90%; *n* = 352) believed that early therapy could slow down symptom progression. GPs (97%; *n* = 374) also held the view that early detection enhanced both the patients' and their relatives' well-being. Regarding their views on dementia testing and managing, three-fourths (75%; *n* = 290) of the GPs believed that managing dementia patients and their caregivers took more time than they could afford in their practice. Provided that conditions were suitable, the majority (79%; *n* = 298) would implement standardized cognitive tests for early detection.

From a list of five potential factors contributing to a more effective dementia examination routine, GPs marked the items as necessary with the following percentages: more time for patients (81%; *n* = 311), up-to-date tests (with a maximum of 5 minutes needed for administration and evaluation) (77%; *n* = 297), help from assistants (50%; *n* = 192), more staff (44%; *n* = 170), and, lastly, more examination rooms (26%; *n* = 103). Regarding an optimal, up-to-date instrument, GPs preferred a pen-and-paper test that could be administered by an assistant or the patients themselves and would include information from the patient's caregivers.

2. Study 2

2.1. Demographic characteristics and test scores of the participants

The two participant groups showed no significant differences in gender and years of education (M_{HC} = 13.48 (SD = 2.632); M_{MCI} = 12.36 (SD = 2.827)). However, the mean age of participants was significantly higher (U = 187.000, Z = -2.440, p = 0.015) in the MCI group (M_{MCI} = 71.72 (SD = 5.435)) in comparison with those in the HC group (M_{HC} = 67.32 (SD = 8.300)).

The mean MMSE score of the HC subjects was 29.44 (SD = 0.507) points, whereas participants in the MCI group scored 26.96 (SD = 1.060) on average. Regarding their GDS-15 score (M_{HC} = 1.84 (SD = 1.724); M_{MCI} = 2.40 (SD = 1.225)), no significant difference was found between the two groups.

2.2. Temporal parameters of verbal fluency performance

Considering the PVF tasks, in the "a" fluency, the average length and the total length of irrelevant utterances were significantly higher in the MCI group, while none of the temporal parameters differed between the two groups in the case of the "k" and "t" PVF.

Regarding the three SVFs, the total number of silent pauses was significantly higher in the HC group in the animal and action fluency tasks, whereas the average length of silent pauses and the average word transition time were significantly higher in the MCI group throughout all three tasks.

2.3. Traditional word count measures of verbal fluency performance

In the three PVF tasks, no statistically significant difference was found between the groups regarding the number of correct words and the number of repetitions or perseverations. However, in the "a" PVF task, participants from the MCI group produced on average more errors than participants from the HC group.

As for the SVF tests, participants from the HC group had a significantly higher number of correct words in the case of all three (animals, food items, and actions) tasks. In the number of errors, repetitions or perseverations, there was no statistically significant difference between the two study groups in any of the SVF tests.

2.4. ROC analysis of the significant temporal parameters and word count measures

ROC analysis of the temporal parameters and word count measures were carried out on all the variables that, based on the previously conducted comparative tests, showed significant differences between the HC and MCI groups. For every ROC analysis, sensitivity and specificity were determined using threshold values optimal for early screening, i.e., maximizing the sensitivity, while keeping specificity greater than or equal to 50%.

Accuracy measures of the traditional and temporal variables are given in *Tables 1* and *2*.

		Accuracy measures					
Fluency tasks	**Traditional measures**	***p***	**AUC**	**95% CI-**	**95% CI+**	**Sensitivity (%)**	**Specificity (%)**
Letter 'a'	Number of errors	0.116	0.630	0.474	0.785	36.0	88.0
Animals	Number of correct words	**0.000**	0.842	0.734	0.949	100.0	56.0
Food items	Number of correct words	**0.002**	0.750	0.616	0.884	76.0	64.0
Actions	Number of correct words	**0.022**	0.689	0.543	0.834	68.0	52.0

Table 1. Accuracy measures of the significant traditional fluency measures. p < 0.05 indicates that the measure is significantly better than chance at discriminating between individuals of the two groups. AUC: area under the curve; CI: confidence interval.

Fluency tasks	**Temporal parameters**	**Accuracy measures**					
		p	**AUC**	**95% CI-**	**95% CI+**	**Sensitivity (%)**	**Specificity (%)**
Letter 'a'	Average length of irrelevant utterances (s)	**0.010**	0.712	0.569	0.855	80.0	52.0
	Total length of irrelevant utterances (s)	**0.035**	0.674	0.523	0.824	80.0	52.0
Animals	Total number of silent pauses (count)	**0.004**	0.740	0.598	0.882	76.0	50.0
	Average length of silent pauses (s)	**0.016**	0.702	0.549	0.855	72.0	50.0
	Average word transition time (s)	**0.001**	0.787	0.651	0.922	96.0	62.5
Food items	Average length of silent pauses (s)	**0.031**	0.678	0.528	0.828	68.0	52.0
	Average word transition time (s)	**0.006**	0.726	0.587	0.866	76.0	52.0
Actions	Total number of silent pauses (count)	**0.013**	0.706	0.562	0.849	72.0	52.0
	Average length of silent pauses (s)	**0.019**	0.693	0.544	0.841	72.0	52.0
	Average word transition time (s)	**0.024**	0.686	0.536	0.837	80.0	52.0

Table 2. *Accuracy measures of the significant temporal parameters. p < 0.05 indicates that the measure is significantly better than chance at discriminating between individuals of the two groups. AUC: area under the curve; CI: confidence interval.*

2.5. Comparison of classification abilities

By comparing the AUCs of the significant variables we aimed to compare the classification ability of the traditional and temporal fluency measures. In the animal SVF, the results indicated

no significant differences regarding AUCs between the number of correct words and the total number of silent pauses (Z = 1.433, $p = 0.151$) or the average word transition time (Z = 1.579, $p = 0.114$); however, the classification ability of the average length of silent pauses was smaller (Z = 2.043, $p = 0.041$) compared to the correct word count. In the case of the food item fluency, the AUC of the number of correct words was similar to both the average length of silent pauses (Z = 0.978, $p = 0.328$), and the average word transition time (Z = 0.662, $p = 0.508$). Furthermore, in action fluency, the classification ability of correct word count did not differ from either the total number of silent pauses (Z = 0.267, $p = 0.789$), the average length of silent pauses (Z = 0.056, $p = 0.954$), or the average word transition time (Z = 0.046, $p = 0.962$).

V. DISCUSSION

1. Study 1

Based on the results, Hungarian GPs accept the idea of dementia screening in primary care; furthermore, more than two-thirds of them are satisfied with the most commonly used cognitive screening tests. However, a discrepancy can be discovered between GPs' overall views regarding testing for dementia in primary care versus their actual habits: only a quarter of them apply cognitive tests for dementia detection while the vast majority of GPs apply general questions to the patient or the informant, and a few of them do not perform any examinations at all. Since Hungarian GPs seem to have ambivalent views regarding the effectiveness of the anti-dementia medication, their often insufficient screening habits may reflect therapeutic nihilism.

The findings of the study indicated that the main obstacle to testing for dementia might be short consultation time with patients (which is approximately 6 minutes on average in Hungary). Besides the shortage of time, GPs mentioned the need for quickly administrable cognitive tools and more help from healthcare staff. All of these concerns are reflected by previous studies (e.g., from the UK and the Netherlands). Despite the obvious time restrictions, current views of scientific literature advocate for the integration of targeted case-finding into primary care, prompting early dementia identification.

Cultural differences in attitudes toward age-related memory problems may also affect the success of dementia detection. In Hungary, early dementia symptoms are often overlooked and thus do not prompt taking steps toward recognition. The tolerance for cognitive decline associated with older age may be lower in countries where elderly people live far from their

families and lead a more independent life (e.g., the USA), which would be greatly endangered by a mental illness.

The underuse of standardized instruments and the underdiagnosis of dementia in primary care may also be attributed to the prioritization of somatic diseases over cognitive problems among the elderly. Since more than 65% of people over the age of 65 have multiple chronic conditions, the examination of memory functions might often end up at the bottom of the priority list. Furthermore, the progression of dementia is a slow process and thus is less obvious than the sudden onset of a somatic, sometimes painful symptom requiring urgent examination.

The underutilization of validated cognitive tests might be partly due to the lack of agreements on the most effective ways of dementia recognition, leaving the GPs without an unambiguous source of reference. A crucial way to improve recognition rates would be the regular update of international and national dementia guidelines.

2. Study 2

In the second study of the thesis, according to the assessed needs of GPs, we introduced a novel practical framework for the analysis of verbal fluency tests. The main finding was that in the case of SVF tests, some of the temporal parameters based on silent pauses (total number of silent pauses, average length of silent pauses) and the average word transition time parameter can successfully discriminate between cognitively healthy subjects and individuals with slightly impaired cognition. However, no other temporal measure differed systematically between the two groups.

An inconsistency can be observed concerning the direction of differences regarding the silence-based parameters: the average lengths of the silent pauses and the average word transition times were longer in the MCI group, whereas HC participants had a higher number of silent pauses in the case of the SVF tasks. Since silent pauses were defined as the absence of speech/sound regardless of length, every detectable silent segment found in the recordings was annotated as a silent pause, including even the brief transitions between words. Therefore, the number of silent pauses was increased by the number of words uttered by the participant. Since the HC group produced significantly more correct words in the SVF tasks, the number of silent pauses was also significantly increased in this group.

The significant classification ability of the silent pause parameters in the study suggests that differentiation between HC and MCI patients' SVF performance may be possible by examining only the silent pauses in their speech. Focusing on the disfluencies present in the

recordings has the advantage of making the processing of verbal fluency tasks swifter and more undemanding, as variables based on them do not need to be adapted to different languages and cultures. Based on the preliminary outcomes of the verbal fluency analysis proposed by this study, further projects should be focused on the collection of more and higher quality data in order to define precise reference values for the quantity of silent pauses associated with MCI.

Consistent with earlier studies, our results confirmed the superiority of SVFs over PVFs in the detection of the cognitive changes associated with MCI. The observation that MCI participants are not able to perform at the level of HC participants in either of the SVF tasks can reflect the degradation of semantic networks in MCI patients, causing difficulty to access semantic associations. This observed impairment might indicate that the neuropathology of MCI extends beyond the hippocampal region affecting the cortical areas related to semantic memory. This performance pattern is further underpinned by the results of an earlier cross-sectional study concluding that executive functions may be a key in differentiating MCI patients from cognitively intact individuals.

VI. CONCLUSIONS

Based on the results of the studies, the novel findings of the present thesis are the following:

I. Although Hungarian GPs are aware of and are satisfied with the available cognitive tools, only a quarter of them use standardized cognitive tests for dementia screening.
II. Time pressure and the lack of cost- and time-effective instruments seem to be the most significant barriers to effective dementia screening in primary care.
III. When analyzing SVF tasks, silence-related temporal parameters can differentiate between the HC and MCI groups; however, hesitations and irrelevant utterances showed no systematic differences between the groups.
IV. The discriminatory ability of some silence-related temporal parameters is comparable to the most traditionally applied count-based measures.
V. In line with previous studies, our results confirmed the superiority of SVF over PVF tests regarding the differentiation between the HC and MCI groups, highlighting the important role of the semantic network in the detection of MCI.

Developing widely applicable, computerized diagnostic tools is remarkably important nowadays, seeing the globally heavy burden on the healthcare systems and the growing necessity of telemedicine-based health consultations. A significant advantage of verbal fluency tests in this regard is that they can be self-administered, and, if the appropriate tools are

available, they could be recorded and evaluated remotely. Given their short and easy administration, computerized fluency tests would also meet the needs expressed by GPs in our study, and it would likely lead to an improvement in the detection rates of MCI and dementia.

VII. ACKNOWLEDGEMENTS

Throughout the years of my doctoral studies and during the writing of this dissertation I have received a lot of support from several people who, in one way or another, contributed and extended their valuable assistance. I hereby would like to take the opportunity to thank them all.

First of all, I wish to express my gratitude to my supervisor, Prof. Dr. János Kálmán, for allowing me to conduct my research at the Department of Psychiatry, University of Szeged, as well as for his guidance and support throughout the process. I am especially grateful to my supervisor, Dr. Gábor Gosztolya, for the many phone consultations, and his valuable contribution and constructive suggestions on my manuscripts and the dissertation. I would like to thank Dr. Ildikó Hoffmann and Dr. Magdolna Pákáski for their help and professional advice during my work. I would like to extend my sincere thanks to all my co-authors and I would also like to thank all those who participated in the research.

I am very grateful for the friendship of Nóra Imre, who has stood (sat) beside me for the past ten years, and with whom I have been able to share all the joys and sorrows of the life of a Ph.D. student. I would like to express my deepest gratitude to all my friends who have supported me from near and far, offline and online, who, depending on the stage of this journey, inspired me, consoled me, or celebrated with me.

Last but not least, I would like to thank my family for the endless love, support, encouragement, and patience they have given me, not only in the past years but throughout my entire life.