Temporal analysis of spontaneous speech for early screening of cognitive impairment among the elderly

Summary of Ph.D. Thesis

Nóra Imre, M.A.

Doctoral School of Clinical Medicine
Department of Psychiatry
Albert Szent-Györgyi Medical School
University of Szeged

Supervisors:

János Kálmán, M.D., Ph.D., D.Sc.
professor and head of department
Department of Psychiatry, Albert Szent-Györgyi Medical School, University of Szeged

Magdolna Pákáski, M.D., Ph.D.
senior research fellow
Department of Psychiatry, Albert Szent-Györgyi Medical School, University of Szeged

Szeged
2022

Original research articles related to the thesis (cumulative impact factor: 5.397):

I. Kálmán, J., Devanand, D. P., Gosztolya, G., Balogh, R., **Imre, N.**, Tóth, L., Hoffmann, I., Kovács, I., Vincze, V., & Pákáski, M. (2022). Temporal speech parameters detect mild cognitive impairment in different languages: validation and comparison of the Speech-GAP Test® in English and Hungarian. *Current Alzheimer Research, 19*(5), 373-386. DOI: 10.2174/1567205019666220418155130.

IF (2021): 3.040; **SJR (2021):** Q2 (Neurology (clinical))

II. **Imre, N.**, Balogh, R., Gosztolya, G., Tóth, L., Hoffmann, I., Várkonyi, T., Lengyel, Cs., Pákáski, M., & Kálmán, J. (2022). Temporal speech parameters indicate early cognitive decline in elderly patients with type 2 diabetes mellitus. *Alzheimer Disease and Associated Disorders, 36*(2), 148-155. DOI: 10.1097/WAD.0000000000000492.

IF (2021): 2.357; **SJR (2021):** Q1 (Clinical Psychology, Gerontology)

Conference abstracts related to the thesis:

I. **Imre, N.**, Balogh, R., Gosztolya, G., Tóth, L., Hoffmann, I., Várkonyi, T., Lengyel, Cs., Pákáski, M., & Kálmán, J. (2021). Temporal speech characteristics in elderly type 2 diabetes mellitus patients – examining cognitive state through speech. *The International Online Workshop on Language in Healthy and Pathological Aging*, 29-30 April (online).

II. **Imre, N.,** Balogh, R., Gosztolya, G., Tóth, L., Hoffmann, I., Várkonyi, T., Lengyel, Cs., Pákáski, M., & Kálmán, J. (2021). Spontán beszéd temporális vizsgálata II. típusú diabétesszel élő idősek körében. *Magyar Pszichiátriai Társaság XXIV. Vándorgyűlése,* 9-12 June (online); *Psychiatria Hungarica, XXXVI. Supplementum I.,* 48-49.

III. Pákáski, M., Davangere, D. P., Gosztolya, G., Balogh, R., **Imre, N.**, Tóth, L., Hoffmann, I., Kovács, I., Vincze, V., & Kálmán, J. (2021). Beszédes jelek kognitív zavarban: Nem kell tolmács. *Magyar Pszichiátriai Társaság XXIV. Vándorgyűlése,* 9-12 June (online); *Psychiatria Hungarica, XXXVI. Supplementum I.,* 86.

I. INTRODUCTION

1. Scope and goals of the work

Neurocognitive disorders are among the most frequent global health problems that affect the life quality of the elderly. On the more severe end of the spectrum lies dementia (in the majority of the cases due to Alzheimer's disease (AD)), with its often devastating consequences to those affected and their families and caregivers. However, there is a less obvious side to cognitive deterioration which, although thoroughly studied by researchers, is only getting acknowledged by the general public: the prodrome of dementia, mild cognitive impairment (MCI).

Therapy of the neuropathological processes underlying dementia has most potential in the introductory stages, therefore early diagnosis is crucial. Due to this, MCI has become center of attention for researchers as screening and diagnosing patients at this stage facilitates closer medical attention, and risk reduction for those with more chance of progressing to dementia.

The studies presented in this thesis focus on a telemedicine-based way of screening for MCI (the so-called *Speech-Gap Test*, or *S-GAP Test* for short), as traditional, pen-and-paper cognitive tests are rarely used in primary care, mainly due to lack of time. This method is based on the evidence that speech problems are telltale symptoms of cognitive deterioration (besides the more well-known memory loss), as language is a sensitive indicator of cortical functioning. Word-finding difficulties and memory retrieval problems often manifest as disfluencies in speech. Analysis of temporal (time-based) characteristics of speech (such as the number and duration of pauses, and also the speed of speech) thus offers valuable information on an elderly person's cognitive state, aiding screening and targeting those most at risk of MCI or dementia.

Spontaneous speech is a subtype of connected or self-generated speech, and contains naturally occurring errors and disfluencies. It is suitable for analysis, as 1) it requires ongoing interactions among diverse cognitive processes (*e.g.* semantic storage and retrieval, working memory, executive functions) and is therefore more complex than isolated linguistic tasks; 2) it is produced frequently in everyday context and is highly relevant to real life functioning; and 3) it does not impose high burden on participants and is less likely to be influenced by confounding factors (*e.g.* test-induced anxiety).

This thesis incorporates two original studies, both using the same methodology of the S-GAP Test, which comprises a spontaneous speech task, followed by *automatized speech recognition (ASR)* and statistical analysis. Both studies intended to investigate temporal speech characteristics among the elderly, who (based on traditional screening) are either cognitively healthy controls (HC) or have MCI. Specific goals of the studies were: ***Study 1:*** To compare the same temporal speech characteristics among elderly native speakers of Hungarian and of English, and thus to evaluate the possible similarities and differences in these languages regarding the associations between speech and cognitive state. ***Study 2:*** To explore temporal speech characteristics in elderly people with vs. without type 2 diabetes mellitus (T2DM), as this condition is a major risk factor for cognitive deterioration, and is also associated with reduced performance in a number of cognitive domains, previously only investigated by traditional neuropsychological tests.

2. Temporal speech characteristics in dementia and MCI (Study 1)

The temporal (or time-based) characteristics of speech belong to the language domain of phonetics and phonology, and are examined by a range of informative features, including: the tempo of speech/articulation, the number/length of pauses, and other features based on the ratio of speech/pause. Pauses (or hesitations) are defined as the absence of speech within an utterance for at least 30 ms, and can be categorized into two types: silent pauses (silence) and filled pauses (vocalizations of meaningless filler words like ‘uhm’ or ‘er’).

Temporal analysis of (spontaneous) speech offers particularly informative measures on language skills and sensitive biomarkers for cognitive decline, as the organization of speech reflects and requires the functioning of several underlying cognitive processes: working memory, access to the mental lexicon, planning of speech production, and (depending on the specific task) even episodic memory. The number and/or duration of speech pauses also reflect 1) the time needed for word-retrieval, and 2) the cognitive load of maintaining one’s train of thought – the more/longer the pauses, the slower it is to find the right word, and the harder it is to focus on the message the individual intends to deliver with their speech. Accordingly, increased signs of disfluencies and decreased tempo of speech have been repeatedly detected in the speech of cognitively impaired individuals, either with dementia/AD or with MCI.

However, a recent systematic review highlighted the fact that the methodology of speech-based studies is very heterogeneous; they apply different tasks and diversely calculated speech parameters, which is in contrast with the ideal that procedures used for the screening or the detection of neurocognitive disorders, such as MCI, should be internationally applicable. Therefore, a highly relevant focus of neurolinguistic studies should be to apply and explore the same methods in different language environments − which is what our research group intended to pilot and execute simultaneously in the English and in the Hungarian language, in ***Study 1***.

3. Type 2 diabetes mellitus: a major risk factor to cognitive deterioration (Study 2)

Of all the risk factors for dementia and MCI, type 2 diabetes mellitus (T2DM) should be highlighted, as T2DM is a just as serious global health concern as neurocognitive disorders, given its huge worldwide prevalence. A growing number of evidence confirms increased risk of cognitive decline in elderly diabetic patients compared to nondiabetic individuals. Patients with T2DM are more prone to developing vascular pathology, which in itself can lead to VD or combined with other pathologies to AD, thus doubling the odds of dementia. A quantitative meta-analysis showed that T2DM patients had higher risk for AD, any dementia, and MCI as well. Although this shared pathophysiology is still investigated, it is suggested that diabetes accelerates the aging process in the brain via altered metabolism of glucose, insulin, and amyloid, which all act as serious biological risk factors for dementia. Because of the intertwined pathophysiology, an approach has emerged in which AD is viewed as a neuroendocrine disorder resembling T2DM and thus termed 'type 3 diabetes'. Growing evidence suggests that the metabolic disturbance that is characteristic of T2DM, directly contributes to biochemical, molecular, structural, and functional abnormalities that are associated with AD (*e.g.* neuronal loss, synaptic disconnection, accumulation of beta-amyloid). The role of glucose is all the more prominent: studies reported that AD is characterized by reduced glucose utilization, and the treatment of T2DM improves memory. Insulin also plays a role in the formation of amyloid plaques, and is also indirectly involved in the phosphorylation of tau and thus contributes to the formation of neurofibrillary tangles − which are neuropathological changes described in AD.

Cognition in T2DM has been found to be impaired in several domains, including learning, verbal memory, attention, processing speed, executive functions, psychomotor functions, and language. However, language functions have usually been investigated using the same few neuropsychological tests (in most cases verbal fluency and naming tests), resulting in mixed outcomes. Several studies found no baseline difference between diabetic vs. nondiabetic subjects regarding verbal fluency, naming, or vocabulary tests, while contrastingly, longitudinal follow-up studies have detected a greater decline in the fluency of T2DM patients later in life.

Given the fact that speech features provide highly valuable information regarding cognition, and there is a strong association between cognitive deficits and T2DM, the exploration of temporal speech characteristics would have great clinical importance in this high risk group: elderly individuals with T2DM. However, to the best of our knowledge, no study has investigated this topic previously – therefore, this was our main objective in ***Study 2***.

II. AIMS AND HYPOTHESES

In ***Study 1***, our main aim was to explore and compare temporal speech parameters in elderly native speakers of the English or Hungarian language, using the same methodology, for the purpose of MCI-detection. Until now, phonetic-phonological analyses of speech for the assessment of cognitive impairment have been only independently performed on native speakers of different languages. Based on previous studies, the following hypotheses were proposed:

- H_1) Temporal speech parameters will be able to differentiate the MCI and the HC group, in both the English- and the Hungarian-speaking samples.
- H_2) Classification abilities of the temporal speech parameters will be similar in both languages, as we expect that some temporal speech deficits are language-independent and thus are present in the speech of MCI patients regardless of the native language.

In ***Study 2***, the main objective was to investigate the temporal speech characteristics in diabetic patients and to compare them with nondiabetic, age- and education-matched participants, both in HC and in MCI. To the best of our knowledge, this was the first linguistic study aimed at the cognitive changes manifested in the phonetic/phonological changes in the speech of diabetic patients. Based on the scientific literature, the following hypotheses were formed:

- H_3) Temporal speech deficits (involving the number/duration of pauses and/or tempo) will characterize the speech of diabetic individuals, which are manifest signs of subtle cognitive deficits, based on a shared neuropathology of T2DM with major and mild neurocognitive disorders.
- H_4) Temporal speech parameters will be able to differentiate the diabetic (‘with T2DM’) and nondiabetic (‘without T2DM’) groups, both when cognition is intact (HC), and when impairment is already detected (MCI) based on conventional neuropsychological tests.

III. METHODS AND MATERIALS

1. Participants and study design

1.1 Study 1

Participant recruitment and examination was executed at two institutions, one in the USA, and one in Hungary: 1) at the *Memory Disorders Center of the Department of Psychiatry, New York State Psychiatric Institute and Columbia University* (New York, NY, USA), and 2) at the *Memory Clinic, Department of Psychiatry, University of Szeged* (Szeged, Hungary). In total, 88 individuals were recruited from the outpatient clinics at the two research locations, of whom 66 were found eligible for inclusion in the study. The English-speaking and the Hungarian-speaking samples were of equal sizes (n = 33). All participants were classified as either MCI or as HC based on Petersen's criteria, for which the Mini-Mental State Examination (MMSE) served as the objective measure for cognitive impairment (30-28 points: HC; 27-24 points: MCI). After classification in both the English-speaking and the Hungarian-speaking samples, 4 groups emerged: English-speaking HC (n = 19), English-speaking MCI (n = 14), Hungarian-speaking HC (n = 20), and Hungarian-speaking MCI (n = 13). ***Inclusion/exclusion criteria*** were identical at both research locations and determined the age (a minimum of 60 years), the level of formal education (a minimum of 8 years), and the native language of the potential participants (English in the USA, and Hungarian in Hungary, corresponding to the country of recruitment). ***Exclusion criteria*** included major speech/hearing problems, substance abuse, previous stroke, dementia and current depression. ***Study protocol*** consisted of the following elements: an initial eligibility interview and anamnesis (focused on demographic features and medical history), and a brief neuropsychological test and screening battery including 1) the MMSE, 2) the Clock Drawing Test (CDT), 3) the Geriatric Depression Scale (GDS); and finally, 4) the speech task.

1.2 Study 2

Participants were recruited at two departments of the *Albert Szent-Györgyi Health Center, University of Szeged*, Hungary: 1) diabetic patients ('with T2DM') at the *Division of Diabetology of the Department of Internal Medicine*, while 2) nondiabetic subjects ('without T2DM'), at the *Memory Clinic of the Department of Psychiatry*. A total of 160 individuals were recruited at the two institutions combined; after completing the exclusion process, 100 of them were found eligible for participation. Participants were evaluated by a neuropsychological test battery, including the MMSE serving as the measure for objective cognitive status (30-28 points: HC; 27-25 points: MCI). After classification in both the diabetic and nondiabetic samples, 4 groups emerged: HC with T2DM (*n* = 39), HC without T2DM (*n* = 34), MCI with T2DM (*n* = 12), and MCI without T2DM (*n* = 15). ***Inclusion criteria*** included a minimum age of 50 years, a minimum of 8 years of formal education, and Hungarian as native language. ***Exclusion criteria*** included major speech/hearing problems, substance abuse, previous stroke, dementia and current depression. ***Diabetes-related criterion*** was a medical diagnosis of type 2 diabetes mellitus (insulin-related diseases other than T2DM were excluded). The 1-hour long ***study protocol*** consisted of the following elements: an initial eligibility interview and anamnesis (focused on demographic features and medical history); the speech task described below, and a neuropsychological test sequence. This comprised of 8 instruments: three assessing cognitive state: 1) the MMSE, 2) the CDT, and 3) the Alzheimer's Disease Assessment Scale – Cognitive Subscale (ADAS-Cog); four tests measuring working memory and executive functions: 4) the Digit Span Test Forward and 5) Backward, 6) the Non-Word Repetition Test, and 7) the Listening Span Test; and one scale for screening current depressive symptoms: 8) the GDS.

2. Speech task protocol

The backbone of the S-GAP Test is a speech task that is administered in order to collect spontaneous (unprepared/unplanned) speech recordings, which are later analyzed based on temporal speech characteristics. This task was carefully chosen with the intention to 1) allow remote and repeated testing, 2) to incorporate both working and episodic memory. In both ***Study 1*** and ***Study 2***, the task was identical: an investigator called the participant on a mobile phone, and after a short introduction, asked them to talk about their previous day. The standardized instruction was: *'Please tell me about your previous day in as much detail as you can.'* After this cue, the researchers could not provide verbal prompts or repeat the instructions, but rather they remained silent throughout the call until the participant finished the task. Each participant's monologue was recorded by a call recorder application installed on the mobile phone device.

3. Analysis of speech samples

Preparation of the speech recordings. Recordings were converted into an uncompressed PCM mono, 16-bit wav format with a sampling rate of 8,000 Hz. The beginning and the end of the recordings were manually cut off, so that only the participants' speech remained. ***Building and training of the ASR model.*** ASR is a computerized technology that is used to transform a speech input into a text/phonetic output. The ASR system requires training before utilization: for the Hungarian-speaking recordings in ***Study 1*** and ***Study 2***, the model was trained on a subset of the BEA audio corpus on the speech of 116 speakers, amounting to approximately 44 hours of recordings. The model for the English-speaking recordings in ***Study 1*** was trained on a subset of the TED-LIUM audio corpus, on the speech of 100 speakers and approximately 15 hours of recordings. ***Phoneme-level recognition of speech using ASR.*** Following the preparations, ASR was applied in order to identify pauses (both silent and filled) in each recording. Pauses were defined as the interruption of speech that lasted for more than 30 ms. As language models, simple phone bigrams were used for both languages. The ASR model described above was able to perform phoneme-level recognition, producing a time-aligned phoneme sequence for each recording. ***Extraction and calculation of temporal speech parameters.*** Based on the raw parameters from the ASR output (containing the sequence of phonemes, silent pauses and filled pauses), 15 temporal speech parameters were extracted using simple calculations.

4. Statistical analysis

Descriptive statistics were applied to examine the demographic, neuropsychological, and speech characteristics of participants. Normality of data was tested by the Shapiro-Wilk test of normality. For continuous variables, group comparisons were executed by using either the independent samples *t*-test/Welch's *t*-test (based on equality of variances), or the Mann-Whitney U test (for cases when the normality assumption was not fulfilled). In the cases of categorical variables, the Chi-square test or Fisher's Exact Test was carried out. For correlational analysis, the Kendall-tau correlation was used. With the purpose of determining which temporal speech parameters would be the most sensitive and precise in an automated discrimination of MCI vs. HC, and diabetic vs. nondiabetic participants, receiver operating characteristic (ROC) analysis was applied, and area under the curve (AUC) was calculated. Sensitivity and specificity measures (true positive rate and true negative rate) were calculated using those threshold values that yielded the highest possible sensitivity, while specificity was kept above 50%. Statistical analyses were performed using IBM SPSS 24.0 and MedCalc v.19.4.

IV. RESULTS

1. Study 1

Regarding *demographics*, no statistically significant differences were detected in gender, age, and years of education between the MCI vs. the HC groups, neither in the English-speaking nor in the Hungarian-speaking sample. Regarding temporal speech characteristics, in the ***English-speaking sample***, the MCI vs. the HC group demonstrated statistically significant differences in 7 of the analyzed 15 temporal speech parameters. MCI-patients showed significantly lower articulation tempo ($t(31) = 4.732$; $p < 0.001$) and lower speech tempo ($t(31) = 4.810$; $p < 0.001$), while on the other hand, a higher occurrence rate of total pauses ($t(31) = -2.736$; $p = 0.010$), duration rate of silent pauses ($t(31) = -3.927$; $p < 0.001$) and total pauses ($t(31) = -4.228$; $p < 0.001$), and average duration of silent pauses ($t(15.802) = -3.108$; $p = 0.007$) and total pauses ($t(15.968) = -3.007$; $p = 0.008$) characterized their spontaneous speech. Furthermore, with the purpose of determining which temporal speech parameters would be the most sensitive in an automated discrimination of MCI vs. HC, ROC analysis was executed. As a result, 8 of the 15 parameters showed significant discrimination abilities (starting from the highest AUC): speech tempo ($AUC = 0.891$; $p < 0.001$), articulation tempo ($AUC = 0.891$; $p < 0.001$), total pause duration rate ($AUC = 0.846$; $p = 0.001$), silent pause duration rate ($AUC = 0.835$; $p = 0.001$), silent pause average duration ($AUC = 0.808$; $p = 0.003$), total pause average duration ($AUC = 0.782$; $p = 0.006$), total pause occurrence rate ($AUC = 0.748$; $p < 0.016$), and filled pause occurrence rate ($AUC = 0.703$; $p = 0.049$). Regarding sensitivity, two parameters achieved a value above 90%, namely speech tempo (sensitivity: 100%; specificity: 63.2%) and articulation tempo (sensitivity: 100%; specificity: 57.9%).

Regarding the ***Hungarian-speaking sample***, the MCI vs. the HC group demonstrated statistically significant differences in 5 of the analyzed 15 temporal speech parameters. The MCI-patients' utterance length ($U = 66.000$; $Z = -2.358$; $p = 0.018$) was significantly shorter, while on the other hand a higher duration rate of silent pauses ($t(31) = -2.750$; $p = 0.010$) and total pauses ($t(31) = -2.367$; $p = 0.024$), as well as higher average duration of silent pauses ($U = 70.000$; $Z = -2.211$; $p = 0.027$) and total pauses ($U = 73.000$; $Z = -2.100$; $p = 0.036$) characterized their spontaneous speech. According to the subsequent ROC-analysis, 5 of the 15 parameters showed statistically significant classification abilities (starting from highest AUC): silent pause duration rate ($AUC = 0.746$; $p = 0.018$), utterance length ($AUC = 0.746$; $p = 0.018$), total pause duration rate (*AUC* = 0.742; *p* = 0.020), silent pause average duration (*AUC* = 0.731; *p* = 0.027), and total pause average duration (*AUC* = 0.719; *p* = 0.036). Regarding sensitivity, three parameters achieved a value above 90%, namely silent pause duration rate (sensitivity: 92.3%; specificity: 60.0%), total pause duration rate (sensitivity: 92.3%; specificity: 55.0%), and total pause average duration (sensitivity: 92.3%; specificity: 55.0%).

Furthermore, ***comparisons of AUCs*** using *z*-statistics were executed between the English- and Hungarian-speaking samples. Based on the results of the analysis, AUCs did not differ statistically significantly between the two languages regarding any of the 15 temporal speech parameters, indicating that the S-GAP Test had similar screening potential in English and Hungarian native language environments.

Additionally to our main objective of investigating temporal speech characteristics separately for English and for Hungarian, ***inter-language comparisons*** were also carried out as supplementary analyses. The purpose of this analysis was to explore if individuals with the same cognitive status (either HC or MCI) demonstrate an alternative temporal speech pattern depending on their native language. These comparisons were executed within the same cognitive status, so that the pure effect of the languages themselves could be contrasted, *i.e.* the English-speaking vs. Hungarian-speaking HC groups, and the English-speaking vs. Hungarian-speaking MCI groups were compared. In the HC groups, 8 parameters, while in the MCI groups, 9 parameters differed between the English- and Hungarian-speaking samples, highlighting natural differences in the two languages: to sum up, English-speaking individuals in our present sample produced longer monologues, while they talked slower and their speech contained more pauses on average, compared to their Hungarian-speaking counterparts.

2 Study 2

Regarding *demographic and neuropsychological data* within the HC sample, participants with vs. without T2DM did not differ in either of the demographic or the neuropsychological variables; within the MCI sample, performance was revealed to be lower on the Backwards Digit Span Test among the diabetic MCI-patients (*U* = 49.000; *Z* = -2.161; *p* = 0.047).

Comparison of the diabetic vs. nondiabetic groups was executed first ***within the HC sample***. The HC with T2DM group produced shorter utterances (*U* = 407.000; Z = -2.831; *p* = 0.005), and on the other hand, higher duration rate of silent pauses (*U* = 429.000; *Z* = -2.588; *p* = 0.010) and total pauses (*U* = 474.000; *Z* = -2.090; *p* = 0.037), and also average duration of silent pauses

(U = 453.000; Z = -2.322; p = 0.020) and total pauses (U = 419.000; Z = -2.698; p = 0.007), compared to the HC without T2DM group. The follow-up ROC-analysis showed that discrimination has statistically significant potential in the case of the same 5 temporal speech parameters. The highest AUC was produced with utterance length (AUC = 0.693; p = 0.005), followed by average duration of total pause (AUC = 0.684; p = 0.007), silent pause duration rate (AUC = 0.676; p = 0.010) and silent pause average duration (AUC = 0.658; p = 0.020), and finally total pause duration rate (AUC = 0.643; p = 0.037). Regarding sensitivity, total pause average duration ended up first (79.5%). However, ***within the MCI sample***, no statistically significant difference could be detected between the diabetic vs. nondiabetic groups, regarding neither of the 15 temporal speech parameters. Results of the subsequent ROC analysis revealed a similar tendency, as no parameters had statistically significant abilities to discriminate MCI patients with vs. without T2DM.

Regarding ***associations with demographical characteristics*** across the 4 groups, *age* of the participants correlated with the temporal speech parameters in 4 cases: articulation tempo (HC with T2DM: τ_b = -0.221, p = 0.050), speech tempo (HC with T2DM: τ_b = -0.229, p = 0.042), and silent pause frequency (MCI without T2DM: τ_b = 0.390, p = 0.046). Regarding *education*, weak to moderate correlations were found with 8 parameters: with utterance length (HC without T2DM: τ_b = 0.269, p = 0.035; MCI with T2DM: τ_b = 0.478, p = 0.044), articulation tempo (MCI with T2DM: τ_b = 0.478, p = 0.044), speech tempo (MCI with T2DM: τ_b = 0.546, p = 0.021), filled pause occurrence rate (HC with T2DM: τ_b = 0.274, p = 0.022), filled pause duration rate (HC with T2DM: τ_b = 0.268, p = 0.025; MCI without T2DM: τ_b = 0.596, p = 0.004), silent pause average duration (MCI with T2DM: τ_b = -0.580, p = 0.014), filled pause average duration (MCI without T2DM: τ_b = 0.618, p = 0.003), and total pause average duration (MCI with T2DM: τ_b = -0.615, p = 0.010).

V. DISCUSSION

1. Main findings and general discussion

Both studies that were incorporated in this thesis served the purpose of investigating temporal speech characteristics as sensitive indicators for MCI, in elderly (above the years of 50 and 60) target groups. In ***Study 1***, the research was realized as an international collaboration with the aim to (for the first time) analyze and compare temporal speech characteristics of native speakers using identical methodology, in two different languages: English and Hungarian. In ***Study 2***, the same methodological framework was implemented, in order to investigate (also for the first time) temporal speech characteristics in patients with T2DM, a chronic medical condition that is a major risk factor for cognitive disorders. The data input for both investigations were based on recordings of a spontaneous (unprepared) speech task in which participants retold their previous day. This task was applied specifically because it requires the work of complex cognitive processes and therefore performance might reflect subtle signs of cognitive decline. Based on our results, the studies yielded a number of novel findings:

1) Using the same task and the same methodology, the role of pauses (especially silent pauses) were found to be the most informative temporal speech markers regarding MCI-screening for both English- and Hungarian-speakers (H_1).
2) Discrimination was fundamentally similar in accuracy in the two examined languages, although they did not show an identical pattern regarding MCI-related temporal speech characteristics, which has to be taken into account when comparing the performance of native speakers in different countries (H_2).
3) Silent pauses were more prominent in the spontaneous speech of diabetic patients when compared with age- and education-matched nondiabetic controls, which extends the association between the disturbance of neurocognitive processes and T2DM (H_3).
4) Temporal speech parameters were indicative of subtle cognitive deficits in the case of diabetic patients who did not present other manifest symptoms, suggesting the beginning of underlying pathophysiological processes that are not yet detectable by the most widespread neuropsychological tests. However, no evidence of further deficits could be detected due to T2DM when impairment has already reached the diagnostic threshold for MCI (H_4).

2. Temporal speech characteristics in different languages for the detection of MCI

The aim of this international study (***Study 1***) was to, for the first time, apply temporal speech analysis on two languages with the same study design (from the identical inclusion/exclusion criteria in an age- and education-matched participant pool, through the specific speech task, to the calculation of the same set of temporal speech parameters) and explore if the methodology already tried in Hungarian could be used in English as well, and with what results.

When comparing the significantly different parameters between MCI vs. HC in the two language environments, 4 of them were present in both the English-speaking and in the

Hungarian-speaking samples: MCI patients showed 1) higher silent pause duration rate, 2) total pause duration rate, 3) silent pause average duration, and 4) total pause average duration. Based on this finding, these parameters might serve as sensitive biomarkers of MCI in *both languages*. In the context of previous studies, higher number and/or length of pauses, and decrease of articulation/speech tempo have been described in varying severity of cognitive impairments, however each of these were conducted using different methodologies and diverse tasks including reading aloud, picture description, narrative recall, or spontaneous speech.

Regarding results of the ROC analysis, the English-speaking MCI vs. HC cases were best discriminated based on speech tempo and articulation tempo (with 100% sensitivity) and on further three pause-related parameters with high sensitivity (85.7%). In the Hungarian-speaking sample, ROC analysis showed the highest sensitivity for silent and total pause duration rate and also for total pause average duration (92.3%). These results suggest that the S-GAP Test is applicable in both languages with fair efficiency, but might detect MCI slightly more sensitively in the English-speaking than in the Hungarian-speaking sample.

The prominence of pause-related speech characteristics are indicative of retrieval difficulties which are related to degeneration in hippocampal brain regions. They are also associated with atrophy of gray matter in the frontopolar (or Brodmann) area of the cortex, which plays a role in higher-order cognitive functions like multitasking or memory retrieval. It is hypothesized that the increased number and duration of pauses are manifestations of the increased cognitive load required for maintaining one's train of thought during speech. Even though these slight changes are not necessarily perceptible to the ear, speech analysis suggests that silence is a significant indicator of planning, word-retrieval, and executive difficulties due to cognitive decline. Language functions in general (*e.g.* measured by naming or verbal fluency tasks) also correlate with the volume of gray matter in the left temporal lobe of MCI and AD patients.

Learning from the observed inter-language differences, international application of the S-GAP Test in clinical settings would require thorough preparations as our results emphasize the need for gathering normative data for international adaptations. Linguistic differences (*e.g.* naturally different speed tempo in English vs. Hungarian) would have to be considered individually in different countries when defining screening thresholds, as temporal speech characteristics (even among HCs) have substantially different mean values in each language.

3. Temporal speech characteristics as indicators of early cognitive deficit in T2DM

To the best of our knowledge, ***Study 2*** was the first to investigate the speech of diabetic (or T2DM) patients with the purpose of looking for signs of subtle cognitive deficits manifested in temporal speech characteristics. One of the main findings was that the speech of elderly diabetic patients (with HC cognitive status) compared significantly worse on a number of temporal characteristics than age- and education-matched, also HC nondiabetic individuals.

Our main goal was to investigate the temporal speech characteristics of elderly T2DM patients who have been classified as HC based on conventional neuropsychological screening. Our results demonstrated that their speech indicated subtle, underlying cognitive deficits when compared to HC subjects without T2DM. Specifically, five temporal speech parameters showed statistically significant differences between the diabetic vs. nondiabetic groups: HC with T2DM patients produced decreased utterance length, higher duration rate of silent pauses and total pauses, and also higher average duration of silent pauses and total pauses compared to HC without T2DM participants. It might be intriguing to observe that the temporal speech parameters that showed differentiating power between the HC with/without T2DM groups in ***Study 2*** were also highlighted in the HC/MCI comparison of the Hungarian-speaking sample in ***Study 1***. This might further confirm that from the full 15 temporal speech parameters of the S-GAP Test, these few have possibly the most discriminative potential for future clinical use.

These differences are in tune with previous studies: more and/or longer pauses (interpreted as markers of decreased lexical access and word-finding difficulties) had been observed in the speech of patients with various neurocognitive impairments. These results complemented by the present study confirm that speech pauses offer highly valuable information on language functions and thus cognitive state, especially in the early, non-symptomatic stages of neurocognitive disorders. This is the phase when other cognitive domains have not yet deteriorated in such a magnitude to be detected by conventional neuropsychological tests. In the case of T2DM patients, subtle cognitive changes might be explained by pathophysiological alterations in the brain associated with diabetes − such as inflammation, vascular damage, impaired insulin signaling, neuronal insulin resistance, mitochondrial dysfunction, or disturbances in synaptic plasticity, for all can lead to an onset of cognitive decline.

Speech of MCI patients with/without T2DM were also explored, resulting in no significant differences suggesting that these two groups performed similarly in terms of temporal characteristics. A possible explanation for the lack of differences (besides the small sample size) could be that the pathophysiological processes in the brain are accelerated and facilitated by T2DM and consequently, cognitive performance gradually declines. Based on the medical protocols currently in effect, a diagnosis of MCI is only given when, besides fulfilling other criteria, cognitive symptoms reach a measurable level and can be confirmed by an objective test, assessment, or evaluation tool. Nonetheless, the underlying neuropathological deterioration is usually present for a much longer period, more or less without clinical symptoms. It could be argued that in the case of diabetic patients, the *onset* of the latent phase of transitioning from HC to MCI might start earlier, therefore speech disfluencies might precede the more robust symptoms by a longer period of time than in the case of nondiabetic subjects. Our results also imply that temporal speech characteristics of diabetic and nondiabetic subjects tend to be become similar when the cognitive deterioration reaches the level of diagnosable MCI, which would suggest that once the transition to MCI has manifested, the presence of T2DM does not necessarily aggravate the already deteriorated temporal speech symptoms. It would be of high clinical interest to further explore the effects of T2DM on cognition from a longitudinal viewpoint and to study whether temporal speech features differ in the next stage of cognitive decay, dementia with T2DM.

Regarding the relationship between demographics and temporal speech characteristics, it can be observed that the increased amount of silent pauses (higher frequency or average length) was aligned with the demographic risk factors of cognitive decline (lower education, higher age). On the other hand, the ability to produce more and faster speech (longer utterance length, higher articulation and speech tempo) was associated with lower dementia-risk (higher education and lower age).

VI. CONCLUSION

To summarize, the results of the S-GAP Test implemented both in the English and Hungarian native speaker populations suggest that similar changes can be observed across different languages in temporal parameters of spontaneous speech. Based on these findings, it could be suggested that pauses are language-independent indicators of word-finding and speech planning difficulties and that the S-GAP Test has the potential to become a useful method for early MCI screening both in English-speaking and Hungarian-speaking populations.

The speech of T2DM patients were explored for the first time, building on the shared pathophysiology of T2DM and neurocognitive disorders, and the strong association between speech deficits and cognitive decline. Results revealed that the speech of diabetic patients, otherwise classified as HC, contained an increased number and length of silent pauses compared to nondiabetic matched individuals. Since these subjects performed similarly well on global cognitive and traditional neuropsychological tests, it could be suggested that temporal speech analysis might offer a more sensitive screening potential in the very early, introductory stages of cognitive impairments and also for identifying those diabetic individuals who have increased risk of developing manifest MCI or even dementia.

Regarding future perspectives, speech analysis might permit both the clinical screening and the research of prodromal stages of different types of dementia via an easy-to-use, interactive smartphone application. This could offer a non-invasive, non-stressful, and low-cost technology that allows rapid, easy, ecologically valid, and remote assessment. A further advantage is that the recording of spontaneous speech is less stressful for elderly patients than a neuropsychiatric test. The major clinical perspective behind both studies was to aid the development of a quickly administrable neuropsychological screening tool that can ease the burden of cognitive screening for primary care or allow self-check for the at-risk individuals themselves.

ACKNOWLEDGEMENTS

My work was supported by the New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund, Hungary (UNKP-21-4-SZTE-114), and also the Ministry of Human Capacities, Hungary (EFOP-3.6.3-VEKOP-16-2017-00009). I would like to thank my supervisors, Prof. Dr. János Kálmán and Dr. Magdolna Pákáski for their guidance and trust towards me during the years of my PhD study, and for all the valuable experiences that they granted me the chance to prove myself in. On the same note, I would like to express my gratitude to all members of our research group, especially to Dr. Gábor Gosztolya and Dr. Ildikó Hoffmann for their well-meaning attitude and honest feedback, and for treating me as a colleague from the very beginning. Most of all, I am deeply grateful to my loved ones and family, especially to my mother and my father, for their unconditional love and continuous support. It makes me truly happy to be able to share this milestone with them as a thank you for their tremendous effort. The same goes for my friends (specifically including Réka Balogh), whom I could always lean on, and who all accompanied me on my journey with empathy, inspiration, and humor.