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

Nóra Imre

Ph.D. Thesis

Szeged

2022

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

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.

Original research articles not closely related (cumulative impact factor: 12.27):

I. 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,* 1-13. DOI: 10.1017/S1355617721001454.

IF (2021): 3.114; **SJR (2021):** Q2 (Clinical Psychology, Psychiatry and Mental Health, Neurology, Neuroscience)

II. Egas-López, J. V., Balogh, R., **Imre, N.**, Hoffmann, I., Szabó, M. K., Tóth, L., Pákáski, M., Kálmán, J., & Gosztolya, G. (2022). Automatic screening of mild cognitive impairment and Alzheimer's disease by means of posterior-thresholding hesitation representation. *Computer Speech & Language, 75*, 101377. DOI: 10.1016/j.csl.2022.101377.

IF (2021): 3.252; **SJR (2021):** Q2 (Human-Computer Interaction, Software, Theoretical Computer Science)

III. Gosztolya, G., Balogh, R., **Imre, N.**, Egas-López, J. V., Hoffmann, I., Vincze, V., Tóth, L., Devanand, D. P., Pákáski, M., & Kálmán, J. (2021). Cross-lingual detection of mild cognitive impairment based on temporal parameters of spontaneous speech. *Computer Speech & Language, 69*, 101215. DOI: 10.1016/j.csl.2021.101215.

IF (2021): 3.252; **SJR (2021):** Q2 (Human-Computer Interaction, Software, Theoretical Computer Science)

IV. **Imre, N.**, Balogh, R., Papp, E., Kovács, I., Heim, Sz., Karádi, K., Hajnal, F., Kálmán, J., & Pákáski, M. (2019). Knowledge of general practitioners on dementia and mild cognitive impairment: a cross-sectional, questionnaire study from Hungary. *Educational Gerontology, 45*(8), 495-505. DOI: 10.1080/03601277.2019.1660137.

IF (2019): 0.748; **SJR (2019):** Q3 (Geriatrics and Gerontology, Education)

V. Balogh, R., **Imre, N.**, Papp, E., Kovács, I., Heim, S., Karádi, K., Hajnal, F., Kálmán, J., & Pákáski, M. (2019). Dementia in Hungary: General practitioners' routines and perspectives regarding early recognition. *European Journal of General Practice, 26*(1), 7-13. DOI: 10.1080/13814788.2019.1673723.

IF (2019): 1.904; **SJR (2019):** Q1 (Family Practice)

Total impact factor: 17.667

TABLE OF CONTENTS

I. ABBREVIATIONS

AD	Alzheimer's disease
ASR	automatic speech recognition
AUC	area under the curve
BEA	'Beszélt Nyelvi Adatbázis' (Speech Database)
CDT	Clock Drawing Test
COVID-19	coronavirus disease 2019
CT	computerized tomography
CUDA	Compute Unified Device Architecture
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th ed.
DNN	Deep Neural Network
GDS	Geriatric Depression Scale
GDS-15	15-item Geriatric Depression Scale
GDS-30	30-item Geriatric Depression Scale
HC	healthy cognition
HMM	Hidden Markov Model
HTK	Hidden Markov Model Toolkit
IT	information technology
M	mean
MCI	mild cognitive impairment
MRI	magnetic resonance imaging
MMSE	Mini-Mental State Examination
N	number
PCM	Pulse-Code Modulation
ReLU	rectified linear activation unit
ROC	receiver operating characteristic
SD	standard deviation
S-GAP Test	Speech-Gap Test
T2DM	type 2 diabetes mellitus
VD	vascular dementia

II. SCOPE AND GOALS OF THE WORK

Neurocognitive disorders have become one of the most frequent health problems that affect the quality of life of the elderly population all around the world. On the more severe end of the spectrum of cognitive deterioration lies dementia (in the majority of the cases due to Alzheimer's disease (AD)), with its often devastating consequences to those affected and no less to their families and caregivers. However, there is a less obvious side to cognitive deterioration which, although thoroughly studied by researchers, is only becoming acknowledged by the general public: the prodrome of dementia, mild cognitive impairment (MCI).

Therapy of the neurocognitive processes underlying dementia has the most potential in the introductory stages, therefore early diagnosis is crucial for patients and is of utmost interest to clinicians and researchers. However, dementia symptoms are often minimized and considered part of natural aging, and thus are overlooked until the condition gets more severe and less modifiable by treatment. Due to this reason, MCI has become the center of attention for researchers of neurocognitive disorders, as screening and diagnosing patients at this stage facilitates closer medical attention, ideally including regular check-ups, neuropsychological testing, and starting risk reduction for those with more chance of progressing to dementia later.

The studies presented in this thesis focus on a telemedicine-based way of screening for MCI (the so-called Speech-Gap Test), as traditional pen-and-paper cognitive screening tests are rarely used in primary care, mainly due to lack of time. This method is inspired by research and also clinical observations stating that speech problems are telltale symptoms of cognitive deterioration (besides the more well-known memory loss). 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 the speech) thus might offer valuable information on an elderly person's cognitive state, aiding the screening of deterioration and targeting those most at risk of MCI or later, dementia.

This thesis incorporates two original research articles, both using the same methodology of the Speech-Gap Test, which comprises a spontaneous speech task, followed by automatized speech recognition (ASR) and statistical analysis. The main goals of the two studies included the following:

I. To investigate and differentiate temporal speech characteristics among elderly individuals, who (based on traditional screening) are either cognitively healthy controls (HC) or are considered to have MCI.

II. To compare these 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 1***).

III. To explore temporal speech characteristics in elderly people with and 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 (***Study 2***).

III. BACKGROUND

1. From dementia to mild cognitive impairment: epidemiology, definitions, and symptoms

Worldwide, an estimated number of 55 million people live with *dementia* (Gauthier *et al.*, 2021) and this is expected to reach 131 – 152 million by the year 2050 (Prince *et al.*, 2015; GBD 2019 Dementia Forecasting Collaborators, 2022). In Hungary, epidemiological studies are scarce and the available data show great variation, however in 2008, the number of patients living with dementia was estimated to be 530,000 – 917,000 (Érsek *et al.*, 2010). Age-standardized prevalence rate of diagnosed dementia was calculated to be 570 for every 100,000 residents (Balázs *et al.,* 2021), which compared to international estimates is considerably low – supposedly due to possible underdiagnoses (Balogh *et al.*, 2019). As most of Europe, Hungary is also facing the demographic consequences of an aging population – as of 2022, 20.5% of all residents (more than 1,990,000 people) are above the age of 65 (Hungarian Central Statistical Office, 2022), and thus have an elevated risk of dementia – a risk that doubles with every 5 years of age (McCullagh *et al.*, 2018). In Hungary, the median survival after the first diagnosis of dementia was found to be 3.01 years (Balázs *et al.*, 2021).

The concept of dementia has been around for several centuries – moreover, it was referenced even in ancient times. One of the first mentions of cognitive decline in the elderly is attributed

to the Greek physician, Pythagoras from the 7th century BC: he divided the human life cycle into 5 stages, of which the last two (starting above the ages of 63 and 80) was described as the 'senium', when the human body and mind returns to the imbecility of infancy (Berchtold & Cotman, 1998). The exact term and its meaning has gone through profound changes, however: the phrase of 'dementia' stems from the Latin word 'demens', meaning 'from outside', alluding to being outside of one's mind (Vatanabe *et al.*, 2020). The etymology also pinpoints the important fact that the early descriptions did not differentiate between the different sources and background of the decline, but rather meant a generic mental illness (Vatanabe *et al.*, 2020).

Nowadays, dementia is defined as the (usually progressive) decline or loss of cognitive functions, resulting in increasing difficulties in personal, social, educational or occupational functioning, beyond what might be expected from the usual consequences of biological aging (World Health Organization, 2019, 2021a). The eleventh edition of the *International Statistical Classification of Diseases and Related Health Problems* (*ICD-11*; World Health Organization, 2019) categorizes dementia under neurocognitive disorders and essentially requires the impairment to be present in two or more cognitive domains (not necessarily restricted to memory), the evidence to be based on both subjective information (concern from the individual/informant/clinician) and objective data (quantified clinical/neuropsychological assessment), and the symptoms to not be better explained by other chronic mental disorders or temporary conditions (like current substance intoxication, delirium, or recent head trauma).

Regarding etiology, dementia or major neurocognitive disorder can be further specified to occur (probably or possibly) due to a series of conditions: Alzheimer's (AD; about 60%-80% of dementia cases; Alzheimer's Association, 2022a), vascular disease (VD), Lewy-body disease, frontotemporal lobar degeneration, Parkinson's disease, Huntington's disease, prion disease, HIV-infection, traumatic brain injury, substance/medication use, another medical condition, or multiple etiologies (American Psychiatric Association, 2013a). There is research evidence that the pathological processes underlying dementia (specifically AD as well) usually start decades before the symptoms are apparent (Trejo-Lopez *et al.*, 2022). Clinicopathological studies suggest that what is perceived as dementia syndrome is actually caused by multiple, often coexisting and overlapping pathologic processes, even in the cases where a probable etiological diagnosis (*e.g.* AD) is given (Rabinovici *et al.*, 2017). In AD, two representative neuropathologies are described most often: 1) beta-amyloid (Aβ) plaques, affecting over time

the neocortex, the limbic structures, the diencephalon, the basal ganglia, and finally the cerebellum and the brainstem; and 2) neurofibrillary tangles ranging from the transentorhinal region to the limbic system and the neocortex (Trejo-Lopez *et al.*, 2022). However, it is becoming more apparent that in the majority of the cases, additional pathological changes (*e.g.* vascular) contribute to the clinical presentation (Trejo-Lopez *et al.*, 2022).

As of today, our definitions and symptomatology regarding neurocognitive deterioration stems from a large body of research and clinical practice which has grown incomparably since the first mentions of the phenomena in the ancient times – and is continuously evolving. Previous definitions of dementia specified two or more cognitive domains to be significantly impaired for a diagnosis (Arvanitakis *et al.*, 2019), however the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-5*; American Psychiatric Association, 2013a) broadened the definition by including that dementia can be present in a single domain as well. It also changed the nomenclature that was used in the previous version of the manual, *DSM-IV* (American Psychiatric Association, 2000), by substituting the term ’dementia’ with ‘major neurocognitive disorder’.

The DSM-5 definition for *major neurocognitive disorder*, similarly to ICD-11, prescribes evidence of significant cognitive decline from a previous level of performance, however in one or more of 6 specified cognitive domains (***Table 1***), based on 1) concern of the individual or a knowledgeable informant or a clinician, and 2) substantial impairment in cognitive performance, preferably documented by standardized neuropsychological testing or quantified clinical assessment (American Psychiatric Association, 2013a). It is also described that the cognitive deficits 3) interfere with independence in everyday activities; 4) they do not occur exclusively in the context of delirium; and 5) they are not better explained by another mental disorder (*e.g.* major depressive disorder, or schizophrenia) (American Psychiatric Association, 2013a).

As a clinical syndrome, it can be divided into severe, moderate, and mild stages, based on the symptoms and the disturbances in everyday life. Severity can be described based on how much assistance is required with activities of daily living: in the mild stage, instrumental activities are affected (*e.g.* paying bills, managing medications or money), in the moderate stage, basic activities are also impaired (*e.g.* feeding, dressing), while in the severe stage, the patient is fully dependent (American Psychiatric Association, 2013a). As a consequence,

dementia seriously affects not only the patients' overall health and independence, but also the life quality and well-being of their immediate family and/or caregivers; not to mention the huge economic burden that this condition imposes on the health-care system and the society (Burns, 2000; Wimo *et al.*, 2013).

Since available pharmacological therapies are only capable of slowing down the progression of the disease (Alzheimer's Association, 2022a), the focus of clinical interest gradually broadened from examining manifest dementia to screening in the earliest possible stage. With the growing volume of research, finally a prodrome was described that precedes dementia, and was named *mild cognitive impairment* (MCI) (Petersen *et al.*, 2014). The first formal set of features described for the diagnosis of MCI (often referred to as the 'Petersen-criteria') included 1) subjective complaint of memory decline, 2) objective cognitive impairment, 3) preserved general cognition, 4) intact daily functioning, and 5) the individual does not meet criteria for the diagnosis of dementia (Petersen *et al.*, 1999). Although more or less the same condition has been previously referred to under several different terms (*e.g.* questionable dementia, minimal dementia, cognitive impairment without dementia, age-associated memory impairment, age associated cognitive decline, benign senescent forgetfulness), mild cognitive impairment became the most widespread denomination (Taler & Phillips, 2008).

To put MCI into perspective, the spectrum of cognition in the elderly population ranges from normal cognitive decline associated with aging (*e.g.* regarding reaction time, dividing one's attention, learning new information, or verbal fluency), through subjective cognitive impairment (complaints without deficits in screening tests) to MCI, and finally dementia (Hazzard & Halter, 2009; Jongsiriyanyong & Limpawattana, 2018). Subjective memory complaints, although often accompany symptoms of memory loss, are not a definitive precursor of either dementia or MCI − *e.g.* a typical phenomenon, difficulty with recalling names is considered common in aging (Arvanitakis *et al*., 2019).

The condition of MCI, although less severe than dementia, is nonetheless a great focus of concern. It might appear at any point across the lifespan; its risk is dependent on the underlying etiology, however it is increasing by age as the possible causes are getting more prevalent among the elderly (World Health Organization, 2019). According to the estimations of the Alzheimer's Association, approximately 12-18% of people above the age of 60 live with MCI (Alzheimer's Association, 2022b); other sources presented that while prevalence is 8.4%

between the ages of 65 and 69, it increases to 25.2% between the ages of 80 and 84, a quarter of all the elderly in this age range (Petersen *et al.*, 2018). According to research, 10-15% of individuals living with MCI develop dementia (Petersen *et al.*, 1999, 2014), and about 30% of all MCI cases convert to AD within 5 years (Ward *et al.*, 2013). However, a recent study found that about half of the examined MCI cases did not progress to dementia, but had reverted back to normal cognition, as registered 2.4 years after diagnosis (Angevaare *et al.*, 2022). The course of MCI may be static, progressive, or even reversible (it either completely resolves or partially improves) (World Health Organization, 2019), thus early screening for MCI gives clinicians an opportunity to either find (and treat) reversible etiological causes, or to involve the individuals in further assessment and initiate follow-up examinations to recognize signs of progression (Alzheimer's Association, 2022b).

Regarding DSM-5 and ICD-11, they classify MCI as *mild neurocognitive disorder*, and describe it with similar or the same bullet points as major neurocognitive disorder (or dementia), with two key changes: 1) instead of 'significant', it is characterized by 'modest' cognitive decline; and 2) these deficits do not interfere with capacity for independence in everyday activities (i.e. complex instrumental activities are preserved, but they require either greater effort, or compensatory strategies, or accommodation) (American Psychiatric Association, 2013a; World Health Organization, 2019). Furthermore, MCI can be subcategorized based on its symptomatology, into single- or multi-domain (with one or more affected cognitive components), and amnestic or non-amnestic types (with or without impaired memory) (Arvanitakis *et al*., 2019).

The 6 neurocognitive domains and subdomains affected in both major and mild neurocognitive disorder (dementia and MCI) are the following: *language* (*e.g.* expressive language including naming, word-finding, fluency, grammar and syntax; receptive language), *complex attention* (*e.g.* sustained/divided/selective attention; processing speed), *executive function* (*e.g.* planning; decision making; working memory; feedback/error utilization; overriding habits/inhibition; mental/cognitive flexibility), *learning and memory* (*e.g.* immediate memory; recent memory including free/cued recall, recognition; very long term memory including semantic/autobiographical memory; implicit memory and learning), *perceptual-motor function* (*e.g.* visual perception; visuoconstructional ability; praxis; gnosis),

Table 1*. Affected neurocognitive domains, and possible symptoms in major/mild neurocognitive disorder according to DSM-5 (based on: American Psychiatric Association, 2013a).*

Dementia (major neurocognitive disorder)	**Mild cognitive impairment (mild neurocognitive disorder)**
Language	
▪ Significant difficulty with expressive/receptive language ▪ Preference of using general terms/pronouns (*e.g.* 'that thing', 'you know what I mean') ▪ Inability to recall names ▪ Grammatical errors, idiosyncratic word usage ▪ Spontaneity of output, automatic speech ▪ Stereotypy of speech, echolalia ▪ Economy of utterances, mutism	▪ Noticeable word-finding difficulty ▪ Substituting general words for specific terms ▪ Avoiding the use of specific names of acquaintances ▪ Subtle grammatical errors (*e.g.* omission/incorrect use of articles, prepositions, auxiliary verbs)
Complex attention	
▪ Increased difficulty in environments with multiple stimuli (*e.g.* TV, radio, conversation, events) ▪ Taking more time to think than usual ▪ Difficulty with holding new information in mind (*e.g.* freshly recalling phone numbers, addresses) ▪ Inability to perform mental calculations ▪ Information must be simplified for understanding	▪ Taking more time for normal tasks than previously ▪ Making errors in routine tasks, needing more double-checking ▪ Thinking is easier when not competing between stimuli (*e.g.* TV, radio, conversations, cell phone, driving)
Executive function	
▪ Abandoning complex projects ▪ Needing to focus on one task at a time ▪ Needing to rely on others to make decisions ▪ Needing to rely on others to plan instrumental activities of daily living	▪ Increased difficulty with multitasking, and resuming an interrupted task (*e.g.* by a phone call or visitor) ▪ Increased effort to complete multistage projects ▪ Increased effort to follow shifting conversations ▪ Increased fatigue from the extra effort required to organize, plan or make decisions
Learning and memory	
▪ Repeating oneself in conversation (often within the same one) ▪ Inability to keep track of short lists (*e.g.* shopping items, daily plans) ▪ Requiring frequent reminders to orient to the task at hand	▪ Occasionally repeating oneself in conversation (over a few weeks) ▪ Increased reliance on lists/calendars ▪ Difficulty with recalling recent events ▪ Needing occasional reminders to keep track of movie/novel characters (*e.g.* re-reading, re-watching) ▪ Losing track of whether bills have already been paid
Perceptual-motor function	
▪ Significant difficulties with previously familiar activities (*e.g.* using tools, driving) ▪ Difficulty with navigating in familiar environments ▪ More confusion at dusk due to changed perceptions (*e.g.* shadows, lowering levels of light)	▪ Needing to rely more on maps or asking for directions ▪ Needing notes/following others to get to a new place ▪ Finding oneself lost/turned around when not concentrating on the task ▪ Needing to expend greater effort for spatial tasks (*e.g.* parking, sewing, knitting, carpentry, assembly)
Social cognition	
▪ Behaving clearly out of acceptable social range ▪ Insensitivity to social standards of modesty (*e.g.* dressing, political/religious/sexual topics) ▪ Excessive focusing on a topic despite others' disinterest or direct feedback ▪ Decision making without regard to safety ▪ Having little insight to these changes	▪ Subtle changes in behavior/attitude ▪ Subtle change in personality ▪ Subtle/episodic apathy or restlessness ▪ Less ability to recognize social cues or read facial expressions ▪ Increased extraversion/introversion ▪ Decreased empathy, decreased inhibition

and *social cognition* (*e.g.* recognition of emotions; theory of mind) (American Psychiatric Association, 2013a). Possible symptoms or observations in everyday activities regarding all 6 neurocognitive domains are listed in ***Table 1***.

It is worth noting that the range of possible manifestations of cognitive impairment expand well beyond the most stereotypical dementia symptoms stigmatized by society (partly due to depictions in popular culture): like forgetfulness, word-finding difficulty, disorientation, deviant behavior, and depressive state or even suicidal ideation (Low & Purwaningrum, 2020). In certain cases, comorbid behavioral disturbances might also accompany dementia and MCI, like personality change, disinhibition, psychotic symptoms, mood disturbance, irritability, agitation, apathy, or wandering (American Psychiatric Association, 2013a; World Health Organization, 2019).

2. Language domain manifestations of neurocognitive deterioration

The Roman poet Lucretius, in the 1st century BC, wrote that with old age "the intellect grows dim / The tongue talks nonsense and the mind gives way" (cited by McMenemey, 1963). The association between language and neurocognitive deterioration, although already referenced in ancient literature and anecdotally observed in daily life, only received more attention (both scientific and public) in the last few decades, along with the publication of a number of intriguing studies. One of the most well-known example was the longitudinal 'Nun Study', in which researchers demonstrated the predictive power of written language features regarding the development of dementia in old age, with the participation of nuns born at the beginning of the 20th century. Lower idea density and lower grammatical complexity were detected in written autobiographies of nuns which, 58 years later, were associated with poor cognitive functions, and AD confirmed by post-mortem neuropathological examination (Snowdon *et al.*, 1996; Riley *et al.*, 2005). Another study analyzed the novels of Iris Murdoch, a celebrated and critically acclaimed British writer of the 1950s-1980s who was diagnosed with AD, and found both syntactic and lexical decline in her final work compared to her earlier publications (*e.g.* less complex sentences, impoverished lexical diversity, increased word repetition) (Garrard *et al.*, 2005; Pakhomov *et al.*, 2011). Besides written language, spoken language (especially spontaneous speech) became the center of attention after 1984, when a cognitive neuropsychologist analyzed the public utterances given by Ronald Reagan, then president of the United States of America, and revealed signs of linguistic impairment characteristic of

cognitive deterioration (*e.g.* many para-grammatical errors, and according to recent analysis, increase in conversational fillers and non-specific nouns 10 years before his medical diagnosis of AD was established (Venneri *et al.*, 2005; Forbes-McKay & Venneri, 2005; Berisha *et al.*, 2015).

Deficits in both written and spoken language are suggested to be one of the earliest signs of cognitive decline (Mueller *et al.*, 2018; Szatloczki *et al.*, 2015; Taler & Phillips, 2008), as language is a sensitive indicator of cortical functioning (Braaten *et al.*, 2006; Meilán *et al.*, 2012; Boschi *et al.*, 2017). These disruptions might be characteristic in various neurocognitive disorders, including dementia in all stages of severity (Forbes *et al.*, 2002; Laske *et al.*, 2015) and MCI as well (Taler & Phillips, 2008). Language deficits can manifest in almost all major linguistic domains, including phonetics (with the temporal subdomain), semantics, grammatics, and pragmatics (***Table 2***) (Mueller *et al.*, 2018; Szatloczki *et al.*, 2015; Taler & Phillips, 2008).

Spontaneous speech is a subtype of connected or self-generated speech, and is defined opposite to prepared or scripted speech (which are closer to written documents due to intentionally constructed, well-formed sentences) (Hoffmann *et al.*, 2010). Spontaneous speech contains naturally occurring errors and disfluencies, and is especially intriguing from the viewpoint of 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 like picture naming; 2) it is produced very frequently in everyday context, and thus a task involving spontaneous speech is highly relevant to real life functioning; and 3) it imposes a relatively low burden on the individual and is less likely to be influenced by iatrogenic confounding factors (*e.g.* test-induced performance anxiety) (Mueller *et al.*, 2018).

Evidence suggests that impairment on multiple cognitive domains (including language) has more predictive power regarding converting from MCI to dementia than a pure memory impairment (Taler & Phillips, 2008), and that language assessment has advantages in screening at-risk groups like the elderly, patients with MCI, or individuals with a confirmed genetic predisposition to AD (homozygous apolipoprotein ε4) (Venneri *et al.*, 2005). As the recognition of the first, early signs of cognitive deterioration is challenging due to the patients often minimizing their deficits or accounting them for normal concomitants of aging, screening at the prodromal stage would be of utmost importance. Since subtle linguistic changes might manifest

Table 2. *Manifestations of language deficits in dementia and MCI, categorized by domain (based on: Mueller et al., 2018; Szatloczki et al., 2015; Taler & Phillips, 2008)*

Dementia (major neurocognitive disorder)	**Mild cognitive impairment (mild neurocognitive disorder)**
Phonetics-phonology (e.g. temporal characteristics)	
▪ Acoustic differences ▪ Decreased speed of speech/articulation ▪ Decreased fluency (semantic, phonological) ▪ Decreased number of total syllables ▪ Decreased number of words/minute ▪ Decreased duration of sample ▪ Increased phonological errors ▪ Increased hesitation markers ▪ Delays in word-finding	▪ Phonemic paraphasia ▪ Increased number of pauses/hesitations ▪ Increased duration of pauses/hesitations
Semantics (e.g. lexical characteristics)	
▪ Decreased number of content words ▪ Decreased number of nouns ▪ Decreased number of verbs ▪ Decreased idea density ▪ Decreased information content/units ▪ Decreased total number of words, and words/clause ▪ Decreased number of unique words ▪ Increased word-finding difficulties ▪ Increased number of pronouns ▪ Increased number of circumlocutory comments ▪ Increased number of semantic errors ▪ Increased proportion of uninformative utterances ▪ More redundant and deictic words/phrases ▪ Higher proportion of closed-class words ▪ Semantic paraphasia ▪ Less pictorial themes ▪ Less relevant observations ▪ Decreased performance in picture naming	▪ Word-finding difficulties ▪ Word-retrieval difficulties ▪ Naming difficulties ▪ Impairment in vocabulary ▪ Impairment in the recognition of the name/function of objects ▪ Decreased verbal fluency (semantic, phonemic) ▪ Less information units and conciseness ▪ Increased reaction time in comprehension ▪ Impairment in repetition ▪ Impaired receptive language processing (*e.g.* slower single-word identification) ▪ Reduced semantic memory
Syntax (e.g. grammatical characteristics)	
▪ Reduced syntactic complexity (simplification of grammatical structure) ▪ Decreased grammatical form ▪ Decreased error monitoring, and repairs/revisions ▪ Decreased number of subordinate clauses ▪ Increased number of pronouns without antecedents ▪ Increased number of undetected errors	▪ Decreased reaction time in syntax comprehension
Pragmatics (e.g. discourse characteristics)	
▪ Decreased discourse coherence ▪ Impaired response to word-finding errors ▪ Communication breakdowns ▪ Empty speech ▪ Lack/ineffectivity of verbal communication ▪ Mutism (in severe stage)	▪ Mild reduction in productive/receptive discourse-level processing (*e.g.* impaired gist-level processing of texts)

in the early course of disease progression (Ahmed *et al.*, 2013), speech analysis could serve as a viable means for screening (Martínez-Nicolás *et al.*, 2021; Petti *et al.*, 2020; Vigo *et al.*, 2022). Advantages regarding the analysis of spontaneous speech for detecting signs of MCI include that 1) speech can be easily recorded and thus can permit analysis similar to biological samples, 2) it allows a completely noninvasive and usually quickly administrable procedure, and 3) it is cost-effective (Laske *et al.*, 2015).

3. Temporal speech characteristics in dementia and MCI: international and Hungarian antecedents (STUDY 1)

The temporal (or time-based) characteristics of speech belong to the language domain of phonetics and phonology, and are usually investigated via spontaneous speech, reading aloud, or other spoken tasks (Szatloczki *et al.*, 2015). The temporal organization of speech is defined by three main variables: time, pauses, and speech; from which a range of informative features can be calculated, including the tempo of speech, the tempo of articulation, the number of pauses, the 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, and can be categorized into two types: silent pauses and filled pauses. Silent pauses might incorporate any silence that is not solely attributable to articulation constraints that are naturally necessary for the pronunciation of words; while filled pauses are actually not silent, but are vocalizations of meaningless filler words like ‘uhm’ or ‘er’. Regarding tempo, speech tempo can be calculated by counting the number of phonemes per second including pauses/hesitations (it reflects the overall fluency of speech), while articulation tempo is also calculated the same way, but excluding pauses/hesitations (thus representing the pure articulation differences between individual speakers) (Hoffmann *et al.*, 2010).

Temporal analysis of speech, especially of spontaneous speech, offers particularly informative measures on language skills and sensitive biomarkers for cognitive deterioration, 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 and topic) even episodic memory (Mortensen *et al.*, 2006). The number and/or duration of speech pauses also reflects 1) the time needed for word-retrieval, and 2) the cognitive load regarding 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 would like to deliver with their speech (König *et al.*, 2015; Szatloczki *et al.*, 2015). 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 (*e.g.* Hoffmann *et al.*, 2010; Meilán *et al.*, 2012) or with MCI (*e.g.* Roark *et al.*, 2011; Meilán *et al.*, 2020).

A number of phonetic-phonological studies (including the ones using temporal analysis) have been executed for the detection of dementia or MCI based on speech characteristics – mainly among native speakers of Indo-European languages, especially English (***Table 3***).

Table 3. *A collection of original research studies aimed at the detection of dementia and/or MCI based on phonetic/phonological analysis of speech, organized by the native language of the participants.*

Native language	**Dementia (major neurocognitive disorder)**	**Mild cognitive impairment (mild neurocognitive disorder)**
	Indo-European	
English	Sajjadi *et al.*, 2012; Jarrold *et al.*, 2014; Guo *et al.*, 2019; Luz *et al.*, 2018; De Looze *et al.*, 2018; Sluis *et al.*, 2020	Roark *et al.*, 2011 De Looze *et al.*, 2018
Spanish	Meilán *et al.*, 2012, 2014; López-de-Ipiña *et al.*, 2013; Gonzalez-Moreira *et al.*, 2015 Martínez-Sánchez *et al.*, 2013, 2018	Meilán *et al.*, 2020; Espinoza-Cuadros *et al.*, 2014;
Bengali	Bose *et al.*, 2022	
French	König *et al.*, 2015, 2018; Mirzaei *et al.*, 2018; Tröger *et al.*, 2019	König *et al.*, 2018; Mirzaei *et al.*, 2018
Persian	Nasrolahzadeh *et al.*, 2018	
Italian	Beltrami *et al.*, 2018	Beltrami *et al.*, 2018
Greek	Satt *et al.*, 2013	Satt *et al.*, 2013
Swedish		Fraser *et al.*, 2019; Themistocleus *et al.*, 2020
	Sino-Tibetan	
Chinese	Chien *et al.*, 2018	
	Japonic	
Japanese	Kato *et al.*, 2013; Tanaka *et al.*, 2017;	Kato *et al.*, 2013; Kobayashi *et al.*, 2019; Yamada *et al.*, 2021
	Turkic	
Turkish	Khodabakhsh *et al.*, 2015	
	Uralic	
Hungarian	Hoffmann *et al.*, 2010; Gosztolya *et al.* 2016, 2019, 2021; Vincze *et al.*, 2020	Tóth *et al.*, 2015, 2018a; Vincze *et al.*, 2016, 2020; Gosztolya *et al.*, 2019, 2021; Balogh *et al.*, 2022

However, it was a Hungarian research team (the predecessor of the one which executed the two studies delineated in this thesis) that first published significant differences between mild AD patients and healthy control (HC) individuals regarding speech tempo and hesitation ratio (Hoffmann *et al.*, 2010) In this early study, the preparation of the speech samples before statistical analysis (namely transcription and annotation) was performed manually using a linguistic computer software package, *Praat* (Boersma, 2002). As manual preparation and calculation of speech biomarkers is extremely time-consuming, our research team (along with the international trends) started implementing *automatic speech recognition* (*ASR*) techniques for the detection of cognitive impairments.

ASR represents a relatively simple and reliable technique that allows the analysis of large language datasets in a rapid manner, via machine learning. Based on this technology, our research team has developed a method called the *Speech-Gap Test* (or *S-GAP Test* for short) – this procedure consists of recording a spontaneous speech task, on which a set of temporal speech parameters are calculated based on the phonetic-level segmentation produced by ASR. In earlier studies applying the S-GAP Test, MCI patients could be distinguished from HC based on temporal speech parameters (Tóth *et al.*, 2015, 2018a; Gosztolya *et al.*, 2016, 2019; Vincze *et al.*, 2016), which demonstrated that the proposed features indeed carry clinically relevant information (Tóth *et al.*, 2015). A novelty of these studies was that besides the more well-known silent pauses, filled pauses were also taken into account, as well as the two different types of tempos (speech and articulation, respectively). As the sample size gradually expanded, machine learning techniques were also exploited, substituting traditional statistical analysis, with which differentiation between MCI and HC became more accurate (Tóth *et al.*, 2018a).

A recent systematic review highlighted the fact that the methodology of speech-based studies is very heterogeneous, applying different tasks and diversely calculated speech parameters (de la Fuente Garcia *et al.*, 2020), which is in contrast with the basic requirement that procedures used for the screening or detection of neurocognitive disorders, such as MCI should be internationally applicable (Solomon *et al.*, 2014). 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 in ***Study 1***.

4. Type 2 diabetes mellitus: a major risk factor to cognitive deterioration (STUDY 2)

Risk factors for dementia can be categorized into two major types based on reversibility: immutable and potentially modifiable factors (Patterson *et al.*, 2007). Immutable risk factors are mainly the following: higher age, female gender, lower levels of education, and (in the case of AD) genetic predisposition associated with apolipoprotein ε4 (APOE ε4 allele) (Chen *et al.*, 2009). However, there are numerous risk factors which can be modified by either converting to a healthier lifestyle or in case of comorbid diseases, by preventing or treating them. These include *type 2 diabetes mellitus (T2DM)*, hypertension (high blood pressure), hypercholesterolemia (high blood cholesterol), hyperthyroidism (overactive thyroid), higher estrogen levels (in females), lower free testosterone levels (in males), stroke, head trauma, infections (*e.g.* human immunodeficiency virus, hepatitis C), high body mass index (obesity), depression, and unhealthy lifestyle (high fat intake, low level of physical and mental activities, smoking, heavy drinking) (Chen *et al.*, 2009; Patterson *et al.*, 2007). Environmental risk factors (*e.g.* long-term, excessive exposure to air pollution, toxic metals and chemicals, electric and magnetic fields) can also contribute to higher risk of developing dementia (Killin *et al.*, 2016), although these factors are not eliminable in every case. Risk factors for MCI are fundamentally the same because of the shared pathophysiology with dementia, however it is highly important to note that factors apart from neurodegenerative processes account for the reversibility of MCI in many cases (*e.g.* education, vascular function, hormonal changes, use of anticholinergic drugs, lifestyle) (Gauthier *et al.*, 2006).

Of all the risk factors, T2DM is a just as serious global health concern as dementia, given that its worldwide prevalence is approximately 9.3% of all adults (463 million people), and is projected to rise as high as 10.2% by 2030 and 10.9% by 2050 (Saeedi *et al.*, 2019). In Hungary, the prevalence was estimated to be 9.9% in the adult population, with over 807,885 people; however, taking into account the possible undiagnosed cases, prevalence might even reach 13.4% (Tóth *et al.*, 2018b). It spreads in an almost epidemic manner, as the number of diabetic people in the 1980s was only 108 million globally, and prevalence is rising alarmingly especially in countries with low- or middle-income (World Health Organization, 2021b), including Hungary (Domján *et al.*, 2017).

A growing number of evidence confirms increased risk of cognitive decline in elderly diabetic patients compared to nondiabetic individuals (Cukierman *et al.*, 2005; Ninomiya, 2019;

Sadanand *et al.*, 2016). 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 (Ahtiluoto *et al.*, 2010). A quantitative meta-analysis showed that T2DM patients had higher risk for AD, any dementia, and MCI as well (with relative risks of 1.46, 1.51, and 1.21, respectively) (Cheng *et al.*, 2012). Although this shared pathophysiology is still investigated nowadays, it is suggested that diabetes accelerates the aging process in the brain via altered metabolism of glucose, insulin, and amyloid, which all account for serious biological risk factors for dementia (Biessels *et al.*, 2006). 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, the accumulation of beta-amyloid) (De la Monte, 2014). The role of glucose is all the more prominent as studies reported that AD is characterized by reduced glucose utilization, and the treatment of T2DM improves memory (Leszek *et al.*, 2017). 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 (Kandimalla *et al.*, 2017).

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 (Awad *et al.*, 2004; Degen *et al.*, 2016; Geijselaers *et al.*, 2015; McCrimmon *et al.*, 2012; Wennberg *et al.*, 2014). 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 (Wysokinski *et al.*, 2010). Several studies found no baseline difference between diabetic and nondiabetic subjects regarding verbal fluency, naming, or vocabulary tests (Kumari & Marmot, 2005; Morelli *et al.*, 2017; Palta *et al.*, 2017; Wysokinski *et al.*, 2010), whereas longitudinal follow-up studies have detected a greater decline in the fluency of T2DM patients later in life (Callisaya *et al.*, 2019; Mayeda *et al.*, 2014; Palta *et al.*, 2017; Rawlings *et al.*, 2014).

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***.

IV. AIMS AND HYPOTHESES

In ***Study 1***, our main aim was to explore and compare temporal speech parameters in elderly speakers of both the English and Hungarian language, with the same methodology and with the purpose of MCI detection. Until now, phonetic-phonological analyses of speech for the assessment of cognitive impairment have been independently performed on native speakers of different languages. Based on observations from 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 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 phonetic/phonological study aimed at the research of cognitive changes manifested in the language of diabetic patients. Based on 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, when cognition is intact based on conventional neuropsychological tests (HC), and when impairment is already detected (MCI).

V. METHODS AND MATERIALS

1. Participants and study design

1.1 STUDY 1

Participants. Recruitment and examination of the participants was executed at two institutions, one in the USA, and one in Hungary: at the 1) *Memory Disorders Center of the Department of Psychiatry, New York State Psychiatric Institute and Columbia University* (New York, NY, USA), and at the 2) *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 (***Figure 1***). 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 (Petersen *et al.*, 1999), for which the Mini-Mental State Examination (MMSE) (Folstein *et al.*, 1975) served as the objective measure for cognitive impairment (30-28 points: HC; 27-24 points: MCI).

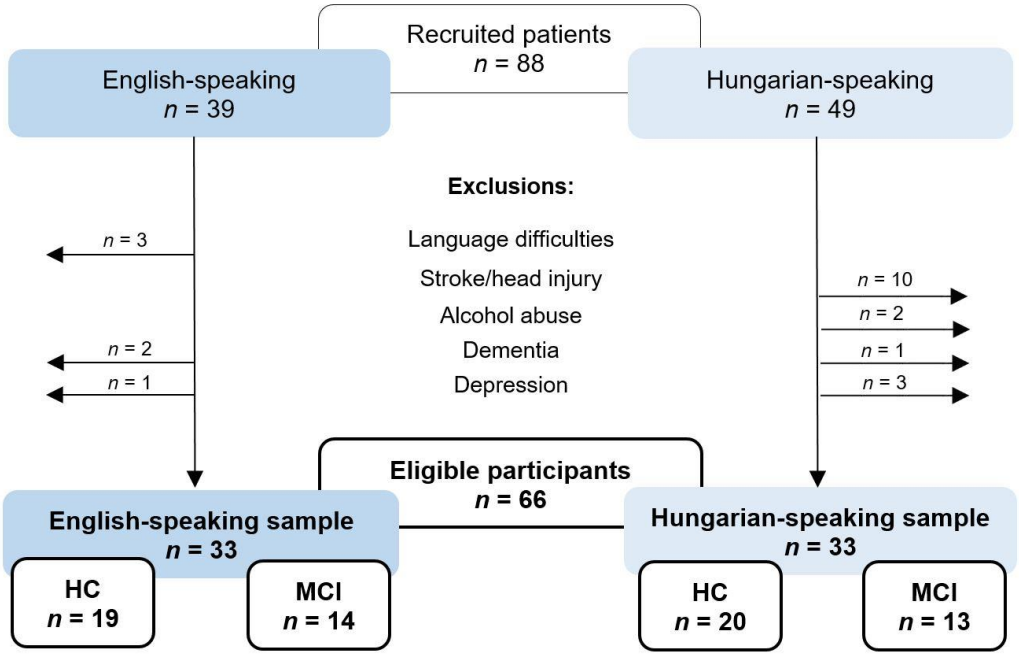

Figure 1. *Flowchart of the inclusion/exclusion process and the final composition of the 4 study groups, namely: English-speaking HC/MCI; Hungarian-speaking HC/MCI.*

Abbreviations: *HC: healthy control; MCI: mild cognitive impairment; n = number*

Ethical approval was granted at both research locations: in the USA, by the Institutional Review Board of the New York State Psychiatric Institute – Columbia University Department of Psychiatry (protocol number: 7611); and in Hungary, by the Regional Human Biomedical Research Ethics Committee of the University of Szeged, Hungary (reference number: 231/2017-SZTE). The study was conducted in compliance with the principles of the Declaration of Helsinki. Written informed consents were obtained from every participant.

Inclusion criteria. The criteria for inclusion and exclusion were identical at both research locations. Initial inclusion criteria determined 1) the age (a minimum of 60 years), 2) the level of formal education (a minimum of 8 years), and 3) the native language of the potential participants (English in the USA, and Hungarian in Hungary, corresponding to the country of recruitment; bilingualism was not taken into account). The ethnical composition of the samples were not defined by inclusion/exclusion criteria, but are hereby reported: at Columbia University, New York, USA the participants were either Caucasian (69.7%), African-American (24.2%), or Hispanic (6.1%), while at the University of Szeged all participants were Caucasian.

Exclusion criteria. Exclusion criteria included the following list of confounding medical conditions: significantly impaired speech (*e.g.* any form of aphasia) or articulation (*e.g.* stutter), major hearing problems (*e.g.* uncorrected hearing loss), evidence of substance use disorder, stroke, or severe head trauma, previous CT/MRI (when available) and/or history of clinically significant cerebral abnormality suggesting another potential etiology for cognitive deficits (*e.g.* lacunar/single large infarct, micro- or macrohemorrhages, cerebral contusion, encephalomalacia, aneurysm, vascular malformations, or clinically significant space-occupying lesions).

Dementia and depression were also exclusion criteria and were therefore evaluated on site, using screening tests. For excluding individuals with signs of possible dementia, the MMSE was applied: those individuals with a score under 24 were excluded. Depressive symptoms were screened using the Geriatric Depression Scale (GDS), corresponding to institutional protocols: patients who scored above 10 on the 30-item version (GDS-30; Yesavage *et al.*, 1983), or above 5 on the 15-item version of GDS (GDS-15; Sheikh & Yesavage, 1986) were not involved in further participation (in the English-speaking/Hungarian-speaking sample, respectively).

Study protocol. The study protocol consisted of the following elements: following an initial eligibility interview and anamnesis (focused on demographic features and medical history), a

brief neuropsychological test and screening battery including 1) the MMSE, 2) the Clock Drawing Test (CDT) (Manos *et al.*, 1994), 3) the GDS; and finally, 4) the speech task.

1.2 STUDY 2

Participants. Recruitment and data collection took place at two departments of the *Albert Szent-Györgyi Health Center, University of Szeged*, Hungary: 1) in the case of the diabetic patients ('with T2DM') at the *Division of Diabetology of the Department of Internal Medicine*, while 2) for nondiabetic subjects ('without T2DM'), at the *Memory Clinic of the Department of Psychiatry*.

Ethical approval was granted by the Regional Human Biomedical Research Ethics Committee of the University of Szeged, Hungary (231/2017-SZTE). The study was conducted in compliance with the principles of the Declaration of Helsinki. Written informed consents were obtained from every participant.

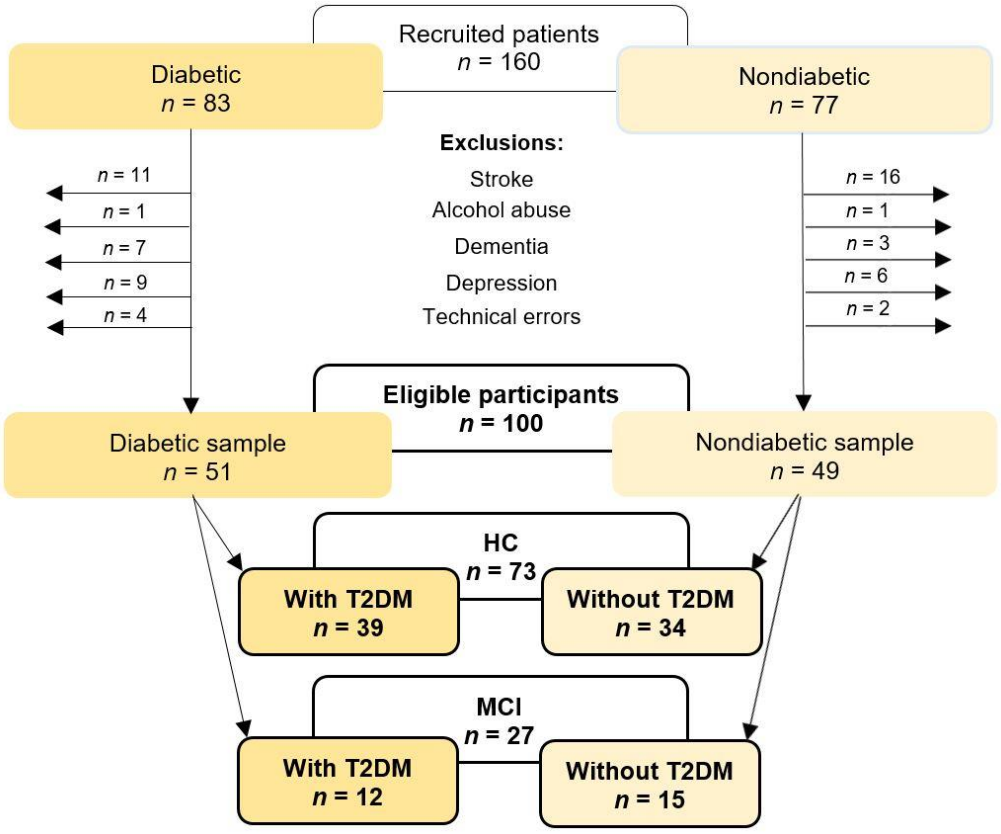

Figure 2. *Flowchart of the inclusion/exclusion process and the final composition of the 4 study groups, namely: HC with/without T2DM; MCI with/without T2DM*

Abbreviations: *T2DM: type 2 diabetes mellitus; HC: healthy control; MCI: mild cognitive impairment; n = number*

Based on the initial inclusion criteria, 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 (***Figure 2***). All participants (both with and without T2DM) were evaluated by means of a neuropsychological test battery (under *Study protocol* in detail), which included 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).

General inclusion/exclusion criteria. The general criteria for inclusion were identical at both research locations, for both diabetic and nondiabetic participants, and included the following: 1) a minimum age of 50 years, 2) a minimum of 8 years of formal education, and 3) Hungarian as native language. Exclusion criteria were the following: significant impairment in speech (*e.g.* any form of aphasia) or in articulation (*e.g.* stutter), major hearing problems (*e.g.* uncorrected hearing loss), evidence of substance use disorder, stroke, or severe head trauma, previous CT/MRI (when available) and/or history of clinically significant cerebral abnormality suggesting another potential etiology for cognitive deficits (*e.g.* lacunar/single large infarct, micro- or macrohemorrhages, cerebral contusion, encephalomalacia, aneurysm, vascular malformations, or clinically significant space-occupying lesions). Finally, those recordings that were below suitable quality (*e.g.* low volume, loud background noises, other technical errors), were not involved in the final analysis (***Figure 2***).

Inclusion/exclusion criteria were checked in an initial patient history interview and from the available medical records. Furthermore, dementia and depression were screened on-site, at the beginning of the protocol, by applying the MMSE and the GDS-15, respectively. Those with a score under 25 on MMSE, or above 5 on GDS-15 were considered as showing signs of dementia or acute depressive symptoms and were therefore excluded from further participation.

Diabetes-related inclusion/exclusion criteria. In the T2DM sample, the initial inclusion criterion was a diagnosis of type 2 diabetes mellitus, verified by medical records. Diagnosis is given following the current international guidelines of the American Diabetes Association (American Diabetes Association, 2014). Exclusion criteria included insulin-related diseases other than T2DM, *e.g.* type 1 diabetes mellitus, prediabetes, or chronic hyperglycemia of any other etiology – patients diagnosed with these conditions were not enrolled. Although

characteristics of diabetes were not specified by inclusion/exclusion criteria, they are hereby reported: average duration of diabetes was 11.4 years (SD = 8.08); treatment was either oral medication (50.9%; *n* = 26), insulin (25.5%; *n* = 13), combined oral medication and insulin (17.6%; *n* = 9), or only diet (5.9%; *n* = 3).

Study protocol. The study protocol consisted of the following elements: initial eligibility interview and anamnesis (focused on demographic features and medical history); and a neuropsychological test sequence, comprising 8 instruments. There were three test batteries with the purpose of taking stock of current cognitive state: 1) the MMSE, 2) the CDT, and 3) the Alzheimer's Disease Assessment Scale – Cognitive Subscale (ADAS-Cog) (Rosen *et al.*, 1984); four tests measuring working memory and executive functions: 4) the Digit Span Test Forward and 5) Backward (Wechsler, 1981), 6) the Non-Word Repetition Test (Gathercole *et al*., 1994), and 7) the Listening Span Test (Daneman & Carpenter, 1980); and one scale for screening current depressive symptoms: 8) the GDS-15. The protocol was about 1-hour long, including 9) the speech task which was administered at approximately the 15-minute mark of the examination, in order to prevent fatigue of the participant. The order of the tests was fixed, and had been assembled in order to ensure that tasks requiring the same cognitive function were separated (*e.g.* tasks based on working memory did not follow each other directly, but were instead alternated with other elements of the neuropsychological test battery).

2. Speech task protocol

The backbone and basis of the S-GAP Test is a speech task that is administered in order to collect spontaneous (unprepared/unplanned) speech samples from the participants, which are recorded for later temporal speech analysis. This task was carefully chosen with the intention to 1) allow remote and repeated testing, 2) to incorporate both working and episodic memory, and 3) because in a previous work of our research group, it was the most sensitive among a range of speech tasks in discriminating between MCI and HC (Vincze *et al.*, 2020). In order to prevent fatigue, this speech task was administered approximately at the 15-minute mark of the protocols of ***Study 1*** and ***Study 2***. In this pilot phase, two researchers took part in the administration of the task (however it is designed in a way that a future mobile application would eliminate the need for assistance and would allow independent use).

In both ***Study 1*** and ***Study 2***, the task was identical and consisted of the following steps: 1) the lead researcher (Investigator 1), who administered the rest of the protocol, explained to the

participant that another researcher from a neighboring room (Investigator 2) will call them on a mobile phone and provide instructions for a new task. Investigator 1 also informed the participant that the conversation will be recorded. Following this cue, Investigator 2 called the mobile phone, and after a short introduction, asked the participant 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 recordings. First, the speech recordings were converted into an uncompressed PCM mono, 16-bit wav format with a sampling rate of 8,000 Hz. Then, the beginning and the end of the recordings (containing the greeting/closing formulas and the instructions) were manually cut off in each case, so that only the participants' speech remained.

Building 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 used here was built on the modified Hidden Markov Model Toolkit (HTK) (Young *et al.*, 2006), in which the acoustic model was replaced by one based on a Deep Neural Network (DNN). This way, a standard HMM/DNN hybrid model was created that is able to outperform traditional HMM models (Hinton *et al.*, 2012). For the DNN acoustic model, a custom DNN implementation (Tóth, 2015) was written in Visual C++ and the CUDA library was used with the purpose of speeding up both model training and evaluation.

As acoustic features, 40 raw Mel-frequency filter bank energy values were used along with the global log-energy which was extended with first- and second-order derivatives ('FBANK + Δ + ΔΔ'), resulting in 123 acoustic features overall. Both training and evaluation was executed on a 150 milliseconds wide sliding window (15 frames), leading to 1,845 input neurons in the actual acoustic models. Finally, the acoustic model DNNs contained 5 fully connected hidden layers with each consisting of 1,024 neurons. The ReLU activation function (Glorot *et al.*, 2011) was employed, having a softmax final layer with the same number of neurons as the phonetic units of the given language.

Training the ASR systems. The ASR system requires training before utilization. In ***Study 1*** and ***Study 2***, for the Hungarian-speaking recordings, the model used by our research group was trained on a subset of the so-called 'Beszélt Nyelvi Adatbázis' audio corpus (BEA) (Neuberger *et al.*, 2014). This corpus was selected because it consists of spontaneous speech and therefore it is expected to contain filled pauses. The training was executed on the speech of 116 speakers, amounting to approximately 44 hours of recordings (for the detailed training of the ASR system, see: Gosztolya *et al.*, 2021). In ***Study 1***, for the English-speaking recordings, another model was trained on a subset of the TED-LIUM audio corpus (Rousseau *et al.*, 2012), on the speech of 100 speakers, and approximately 15 hours of recordings. Previously to training, samples in both corpora were downsampled to a sampling rate of 8,000 Hz in order to match the quality of the recordings in the study.

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 by either 1) complete silence (silent pause) or by 2) filler words/vocalizations like 'er' or 'um' (filled pause) that lasted for more than 30 milliseconds. 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. That is, it provided the corresponding phonetic labels along with their starting and ending time indices, where filled pauses were treated as a special 'phoneme'. The accuracy of this workflow was also tested: silent pauses were detected with high precision (precision: 96.1%, recall: 94.9%, F-measure: 95.5), while filled pauses were also identified with relatively high performance (precision: 83.2%, recall: 69.6%, F-measure: 75.8) (Kálmán *et al.*, 2022). In some cases, filled pauses were confused with prolongations of certain phonemes (*e.g.* m / n / a), which sound similar acoustically and are often uttered by the speakers for similar purposes as filled pauses (Deme & Markó, 2013; Eklund, 2001).

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 (established in previous works of our research group: Tóth *et al.*, 2018a; Gosztolya *et al.*, 2021). The set of 15 parameters are listed and defined in ***Table 4***.

Table 4. *The complete list and definitions of the 15 temporal speech parameters calculated for all spontaneous speech recordings both in* ***Study 1*** *and* ***Study 2****.*

Temporal speech parameters	Definitions
Utterance length (s)	Total length of the utterance (s)
Articulation tempo (1/s)	Total number of phonemes (without hesitations) (count) / total length of the utterance (s)
Speech tempo (1/s)	Total number of phonemes (including hesitations) (count) / total length of the utterance (s)
Occurrence rates of pauses	
Silent pause (%)	Total number of silent pauses (count) x 100 / total number of phonemes (count)
Filled pause (%)	Total number of filled pauses (count) x 100 / total number of phonemes (count)
Total pause (%)	Total number of silent and filled pauses (count) x 100 / total number of phonemes (count)
Duration rates of pauses	
Silent pause (%)	Total length of silent pauses (s) x 100 / total length of the utterance (s)
Filled pause (%)	Total length of filled pauses (s) x 100 / total length of the utterance (s)
Total pause (%)	Total length of silent and filled pauses (s) x 100 / total length of the utterance (s)
Frequency of pauses	
Silent pause frequency (1/s)	Total number of silent pauses (count) / total length of the utterance (s)
Filled pause frequency (1/s)	Total number of filled pauses (count) / total length of the utterance (s)
Total pause frequency (1/s)	Total number of silent and filled pauses (count) / total length of the utterance (s)
Average durations of pauses	
Silent pause average duration (s)	Total length of silent pauses (s) / total number of silent pauses (count)
Filled pause average duration (s)	Total length of filled pauses (s) / total number of filled pauses (count)
Total pause average duration (s)	Total length of silent and filled pauses (s) / total number of silent and filled pauses (count)

4. Statistical analysis

Descriptive statistics were applied to examine the demographic, neuropsychological, and speech characteristics of participants, and were reported by means (*M*) and/or medians (*Md*), and standard deviations (*SD*).

Regarding ***Study 1***, comparisons between the MCI vs. HC groups were executed in both the English- and in the Hungarian-speaking samples. 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 was carried out. For the examination of inter-

language differences (English-speaking HC vs. Hungarian-speaking HC; English-speaking MCI vs. Hungarian-speaking MCI), independent samples *t*-test/Welch's *t*-test or the Mann-Whitney U test was implemented. Regarding ***Study 2***, the Shapiro-Wilk test demonstrated non-normality of data in the case of most continuous/scale variables. Therefore, the Mann-Whitney U test was employed to assess between-group differences on demographic, neuropsychological and temporal speech parameters. In the case of categorical variables, Fisher's Exact Test was applied. For correlational analysis, the Kendall-tau correlation was used.

Receiver operating characteristic (ROC) analysis was applied in both studies to assess which temporal speech parameters have the most classification/identification potential for MCI (***Study 1***) or T2DM patients (***Study 2***), based on their area under the curve (AUC). Sensitivity and specificity measures (also known as 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%.

All statistical analyses were performed using IBM SPSS 24.0 (SPSS Inc., Chicago, IL, USA), except for the inter-language comparison of AUCs (***Study 1***), for which the independent ROC curves module of MedCalc v.19.4 was applied (MedCalc Software Ltd., Ostend, Belgium). For all statistical comparisons, the level of significance was set at the <0.05 level.

VI. RESULTS

1. STUDY 1

1.1 Demographic and neuropsychological characteristics: English- and Hungarian-speaking samples

Demographic characteristics and neuropsychological test scores of all 4 groups (expressed in means and standard deviations) are presented in ***Table 5***. Regarding demographics, no statistically significant differences were detected in gender, age, and years of education between the MCI and the HC groups, neither in the English-speaking nor in the Hungarian-speaking sample. Regarding the neuropsychological tests, the CDT-test demonstrated similar results, however MCI-patients performed significantly poorer than HCs on the MMSE in both language samples (English-speaking sample: $U = 62.500$; $Z = -2.703$; $p = 0.009$; Hungarian-speaking sample: $U = 0.000$; $Z = -4.879$; $p < 0.001$), and MCI-patients also scored higher on the GDS

Table 5. *Demographic characteristics and neuropsychological test scores in the English-speaking and Hungarian-speaking samples (respectively).*

	English-speaking sample		**Hungarian-speaking sample**	
	HC (*n* = 19)	**MCI (*n* = 14)**	**HC (*n* = 20)**	**MCI (*n* = 13)**
	M (SD)	***M (SD)***	***M (SD)***	***M (SD)***
Demographic characteristics				
Sex (male/female)	5/14	6/8	3/17	4/9
Age (years)	74.47 (7.321)	72.36 (6.857)	69.90 (5.609)	73.77 (4.969)
Education (years)	17.84 (3.532)	16.79 (3.118)	13.15 (2.455)	11.77 (2.743)
Neuropsychological test scores				
MMSE	29.16 (1.015)	27.71 (1.773)	28.85 (0.813)	26.31 (0.751)
CDT	8.89 (1.197)	9.21 (1.188)	7.60 (3.152)	7.92 (2.178)
GDS-30 / GDS-15	3.16 (2.853)	5.50 (2.822)	1.65 (1.387)	2.77 (1.013)

Abbreviations: *HC: healthy control; MCI: mild cognitive impairment; MMSE: Mini-Mental State Examination; CDT: Clock Drawing Test; GDS-30: Geriatric Depression Scale (30-item); GDS-15: Geriatric Depression Scale (15-item)*

(indicating increased acute depressive symptoms) in both languages (English-speaking sample: U = 71.000; Z = -2.277; p = 0.024; Hungarian-speaking sample: U = 59.000; Z = -2.736; p = 0.008).

1.2 Temporal speech characteristics and sensitivity measures: English-speaking sample

Regarding the English-speaking sample, the MCI vs. the HC group demonstrated statistically significant differences in 7 of the analyzed 15 temporal speech parameters. These were namely the following: 1) articulation tempo, 2) speech tempo, 3) total pause occurrence rate, 4) silent pause duration rate, 5) total pause duration rate, 6) silent pause average duration, and 7) total pause average duration. In practical terms, the MCI-patients showed significantly lower articulation tempo and lower speech tempo, while on the other hand, a higher occurrence rate of total pauses, duration rate of silent and total pauses, and average duration of silent and total pauses characterized their spontaneous speech (***Table 6***).

Furthermore, with the purpose of determining which temporal speech parameters would be the most sensitive and precise in an automated discrimination of MCI vs. HC individuals, ROC analysis was executed. Based on the ROC analysis, 8 of the 15 parameters had statistically significant classification abilities. These were the following (starting with the highest AUC): 1) speech tempo, 2) articulation tempo, 3) total pause duration rate, 4) silent pause duration rate,

5) silent pause average duration, 6) total pause average duration, 7) total pause occurrence rate, and 8) filled pause occurrence rate. 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%). Sensitivity and specificity measures were calculated using threshold values tailored for early screening, and are detailed for each statistically significant temporal speech parameters in ***Table 8***.

Table 6. *Descriptive statistics and group comparisons in the English-speaking sample using the independent samples t-test / Mann-Whitney U test. (The p-values that indicate statistically significant differences at the <0.05 level are in* ***bold****.)*

	English-speaking HC (*n* = 19)	**English-speaking MCI (*n* = 14)**	***t*-test / Mann-Whitney U test**	
Temporal speech parameters	***M (SD)***	***M (SD)***	**Test statistics**	***p***
Utterance length (s)	275.33 (120.02)	201.94 (135.07)	U = 82.000; Z = -1.858	0.065
Articulation tempo (1/s)	8.88 (1.21)	6.78 (1.32)	$t(31)$ = 4.732	**<0.001**
Speech tempo (1/s)	10.07 (1.10)	8.02 (1.34)	$t(31)$ = 4.810	**<0.001**
Occurrence rates of pauses				
Silent pause (%)	9.43 (3.17)	12.11 (4.35)	U = 85.000; Z = -1.748	0.084
Filled pause (%)	2.55 (1.08)	3.63 (1.73)	U = 79.000; Z = -1.967	0.050
Total pause (%)	11.98 (3.55)	15.75 (4.34)	$t(31)$ = -2.736	**0.010**
Duration rates of pauses				
Silent pause (%)	31.43 (8.72)	45.61 (12.05)	$t(31)$ = -3.927	**<0.001**
Filled pause (%)	5.64 (3.23)	6.56 (5.22)	U = 126.000; Z = -0.255	0.815
Total pause (%)	37.07 (9.27)	52.17 (11.23)	$t(31)$ = -4.228	**<0.001**
Frequency of pauses				
Silent pause (1/s)	0.93 (0.30)	0.95 (0.28)	$t(31)$ = -0.139	0.890
Filled pause (1/s)	0.25 (0.09)	0.28 (0.14)	U = 122.000; Z = -0.401	0.706
Total pause (1/s)	1.18 (0.33)	1.24 (0.30)	$t(31)$ = -0.453	0.653
Average durations of pauses				
Silent pause (s)	0.34 (0.07)	0.51 (0.18)	$t(15.802)$ = -3.108	**0.007**
Filled pause (s)	0.21 (0.05)	0.21 (0.09)	U = 105.000;Z = -1.020	0.321
Total pause (s)	0.31 (0.05)	0.44 (0.14)	$t(15.968)$ = -3.007	**0.008**

Abbreviations: *HC: healthy control; MCI: mild cognitive impairment; M: mean; SD: standard deviation*

1.3 Temporal speech characteristics and sensitivity measures: Hungarian-speaking sample

Regarding the Hungarian-speaking sample, the MCI vs. the HC group demonstrated statistically significant differences in 5 of the analyzed 15 temporal speech parameters. These were namely the following: 1) utterance length, 2) silent pause duration rate, 3) total pause duration rate, 4) silent pause average duration, and 5) total pause average duration. In practical terms, the MCI-patients' utterance length was significantly shorter, while on the other hand a higher duration rate of silent and total pauses, as well as higher average duration of silent and total pauses characterized their spontaneous speech (***Table 7***).

Table 7. *Descriptive statistics and group comparisons in the Hungarian-speaking sample using the independent samples t-test / Mann-Whitney U test. (The p-values that indicate statistically significant differences at the <0.05 level are in* ***bold****.)*

	Hungarian-speaking HC (*n* = 20)	**Hungarian-speaking MCI (*n* = 13)**	***t*-test / Mann-Whitney U test**	
Temporal speech parameters	***M (SD)***	***M (SD)***	**Test statistics**	***p***
Utterance length (s)	155.06 (70.21)	107.82 (87.65)	U = 66.000; Z = -2.358	**0.018**
Articulation tempo (1/s)	9.90 (1.97)	8.63 (1.75)	$t(31)$ = 1.878	0.070
Speech tempo (1/s)	10.67 (1.87)	9.47 (1.62)	$t(31)$ = 1.894	0.068
Occurrence rates of pauses				
Silent pause (%)	4.88 (1.64)	5.91 (1.83)	$t(31)$ = -1.678	0.103
Filled pause (%)	2.69 (1.83)	3.28 (2.10)	U = 112.000; Z = -0.663	0.524
Total pause (%)	7.58 (3.13)	9.20 (3.37)	U = 94.500; Z = -1.308	0.194
Duration rates of pauses				
Silent pause (%)	23.49 (9.72)	32.46 (8.16)	$t(31)$ = -2.750	**0.010**
Filled pause (%)	6.26 (4.10)	7.03 (4.68)	$t(31)$ = -0.494	0.625
Total pause (%)	29.76 (11.81)	39.49 (11.07)	$t(31)$ = -2.367	**0.024**
Frequency of pauses				
Silent pause (1/s)	0.49 (0.11)	0.54 (0.13)	$t(31)$ = -1.008	0.321
Filled pause (1/s)	0.26 (0.14)	0.28 (0.15)	$t(31)$ = -0.336	0.739
Total pause (1/s)	0.76 (0.21)	0.83 (0.22)	U = 108.000; Z = -0.811	0.434
Average durations of pauses				
Silent pause (s)	0.47 (0.18)	0.62 (0.17)	U = 70.000; Z = -2.211	**0.027**
Filled pause (s)	0.21 (0.06)	0.24 (0.10)	U = 123.000; Z = -0.258	0.813
Total pause (s)	0.39 (0.14)	0.48 (0.10)	U = 73.000; Z = -2.100	**0.036**

Abbreviations: *HC: healthy control; MCI: mild cognitive impairment; M: mean; SD: standard deviation*

Similarly as in the English-speaking sample, a subsequent ROC analysis was executed on the Hungarian-speaking sample as well, regarding the MCI vs. HC discrimination abilities of the analyzed temporal speech parameters. As a result, 5 of the 15 parameters showed statistically significant classification abilities. These were the following (starting with the highest AUC): 1) silent pause duration rate, 2) utterance length, 3) total pause duration rate, 4) silent pause average duration, and 5) total pause average duration. 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%). Sensitivity and specificity measures were calculated using threshold values tailored for early screening, and are detailed for each statistically significant temporal speech parameters in ***Table 8***.

Table 8. *Accuracy measures of temporal speech parameters with statistically significant classification ability in the English-speaking sample and the Hungarian-speaking sample (respectively) using ROC analysis. (The p-values that indicate statistically significant differences at the <0.05 level are in* ***bold****.)*

English-speaking sample	**Accuracy measures**						
Temporal speech parameters	***p***	**AUC**	**95% CI-**	**95% CI+**	**Threshold value**	**Sensitivity (%)**	**Specificity (%)**
Speech tempo (1/s)	**0.000**	0.891	0.784	0.998	9.843	100	63.2
Articulation tempo (1/s)	**0.000**	0.891	0.779	1.000	8.772	100	57.9
Total pause duration rate (%)	**0.001**	0.846	0.711	0.980	36.689	85.7	52.6
Silent pause duration rate (%)	**0.001**	0.835	0.695	0.974	32.398	85.7	63.2
Silent pause average duration (s)	**0.003**	0.808	0.654	0.963	0.346	85.7	52.6
Total pause average duration (s)	**0.006**	0.782	0.614	0.950	0.329	78.6	57.9
Total pause occurrence rate (%)	**0.016**	0.748	0.578	0.918	12.078	78.6	52.6
Filled pause occurrence rate (%)	**0.049**	0.703	0.524	0.882	2.567	78.6	52.6
Hungarian-speaking sample	**Accuracy measures**						
Temporal speech parameters	***p***	**AUC**	**95% CI-**	**95% CI+**	**Threshold value**	**Sensitivity (%)**	**Specificity (%)**
Silent pause duration rate (%)	**0.018**	0.746	0.579	0.914	24.191	92.3	60.0
Utterance length (s)	**0.018**	0.746	0.558	0.934	132.345	76.9	60.0
Total pause duration rate (%)	**0.020**	0.742	0.573	0.912	27.280	92.3	55.0
Silent pause average duration (s)	**0.027**	0.731	0.551	0.910	0.438	84.6	55.0
Total pause average duration (s)	**0.036**	0.719	0.537	0.902	0.349	92.3	55.0

Abbreviations: *ROC: receiver operating characteristics; AUC: area under the curve; CI: confidence interval*

1.4 Classification potential of temporal speech characteristics: English-speaking vs. Hungarian-speaking accuracy measures

Furthermore, in order to investigate whether the analyzed temporal speech parameters are characterized by different classification abilities in the two languages, pairwise comparisons of AUCs 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 (***Table 9***).

Table 9. *Pairwise comparison of the AUCs in the English- and the Hungarian-speaking samples regarding the 15 temporal speech parameters. (The p-values that indicate statistically significant differences at the <0.05 level are in* ***bold****.)*

	English-speaking HC vs. MCI	**Hungarian-speaking HC vs. MCI**	***Pairwise statistics***	
Temporal speech parameters	**AUC**		***z- statistic***	***p***
Utterance length (s)	0.692	0.746	0.384	0.701
Articulation tempo (1/s)	0.891	0.692	1.741	0.082
Speech tempo (1/s)	0.891	0.685	1.828	0.068
Occurrence rates of pauses				
Silent pause (%)	0.680	0.658	0.163	0.871
Filled pause (%)	0.703	0.569	0.931	0.352
Total pause (%)	0.748	0.637	0.827	0.408
Duration rates of pauses				
Silent pause (%)	0.835	0.746	0.784	0.433
Filled pause (%)	0.528	0.508	0.120	0.904
Total pause (%)	0.846	0.743	0.927	0.354
Frequency of pauses				
Silent pause (1/s)	0.541	0.631	0.600	0.548
Filled pause (1/s)	0.541	0.523	0.119	0.905
Total pause (1/s)	0.560	0.585	0.169	0.866
Average durations of pauses				
Silent pause (s)	0.808	0.731	0.630	0.529
Filled pause (s)	0.605	0.527	0.492	0.623
Total pause (s)	0.782	0.719	0.486	0.627

Abbreviations: *HC: healthy control; MCI: mild cognitive impairment; AUC: area under the curve*

1.5 Inter-language comparisons of temporal speech parameters

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 (***Table 10***). 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 (E-HC vs. H-HC), and the English-speaking vs. Hungarian-speaking MCI groups (E-MCI vs. H-MCI) were compared.

Table 10. *Inter-language comparisons of the temporal speech parameters of speech using the independent samples t-test / Mann-Whitney U test. (The p-values that indicate statistically significant differences at the <0.05 level are in* ***bold****.)*

	English- vs. Hungarian-speaking HC (E-HC vs. H-HC)		**English- vs. Hungarian-speaking MCI (E-MCI vs. H-MCI)**	
Temporal speech parameters	***t*-test / Mann-Whitney U test**	***p***	***t*-test / Mann-Whitney U test**	***p***
Utterance length (s)	*t*(28.729) = 3.794	**0.001**	*U* = 47.000; *Z* = -2.135	**0.033**
Articulation tempo (1/s)	*t*(31.801) = -1.949	0.060	*t*(25) = -3.120	**0.005**
Speech tempo (1/s)	*t*(31.081) = -1.219	0.232	*t*(25) = -2.529	**0.018**
Occurrence rates of pauses				
Silent pause (%)	*t*(26.715) = 5.570	<**0.001**	*U* = 8.000; *Z* = -4.028	<**0.001**
Filled pause (%)	*U* = 179.000; *Z* = -0.309	0.771	*U* = 74.000; *Z* = -0.825	0.430
Total pause (%)	*U* = 65.000; *Z* = -3.512	<**0.001**	*t*(25) = 4.347	<**0.001**
Duration rates of pauses				
Silent pause (%)	*t*(37) = 2.678	**0.011**	*t*(25) = 3.293	**0.003**
Filled pause (%)	*U* = 174.000; *Z* = -0.450	0.667	*U* = 85.000; *Z* = -0.291	0.793
Total pause (%)	*t*(37) = 2.142	**0.039**	*t*(25) = 2.951	**0.007**
Frequency of pauses				
Silent pause (1/s)	*t*(23.309) = 5.898	<**0.001**	*t*(25) = 4.652	<**0.001**
Filled pause (1/s)	*t*(37) = -0.400	0.691	*U* = 89.000; *Z* = -0.0.97	0.943
Total pause (1/s)	*U* = 55.000; *Z* = -3.793	<**0.001**	*U* = 17.000; *Z* = -3.591	<**0.001**
Average durations of pauses				
Silent pause (s)	*U* = 110.000; *Z* = -2.248	**0.024**	*t*(25) = -1.537	0.137
Filled pause (s)	*U* = 148.000; *Z* = -1.180	0.247	*U* = 73.000; *Z* = -0.873	0.402
Total pause (s)	*U* = 141.000; *Z* = -1.377	0.175	*t*(25) = -0.789	0.437

Abbreviations: *E-HC: English-speaking sample - healthy control; H-HC: Hungarian-speaking sample - healthy control; E-MCI: English-speaking sample - mild cognitive impairment; H-MCI: Hungarian-speaking sample - mild cognitive impairment*

Regarding the HC groups, 8 temporal speech parameters differed in a statistically significant way between the English- and Hungarian-speaking samples. These were the following: 1) utterance length (E-HC > H-HC), 2) silent pause occurrence rate (E-HC > H-HC), 3) total pause occurrence rate (E-HC > H-HC), 4) silent pause duration rate (E-HC > H-HC), 5) total pause duration rate (E-HC > H-HC), 6) silent pause frequency (E-HC > H-HC), 7) total pause frequency (E-HC > H-HC), and 8) silent pause average duration (H-HC > E-HC).

Regarding the MCI groups, 9 significantly different parameters were demonstrated: 1) utterance length (E-MCI > H-MCI), 2) articulation tempo (H-MCI > E-MCI), 3) speech tempo (H-MCI > E-MCI), 4) silent pause occurrence rate (E-MCI > H-MCI), 5) total pause occurrence rate (E-MCI > H-MCI), 6) silent pause duration rate (E-MCI > H-MCI), 7) total pause duration rate (E-MCI > H-MCI), 8) silent pause frequency (E-MCI > H-MCI), and 9) total pause frequency (E-MCI > H-MCI).

2 STUDY 2

2.1 Demographic and neuropsychological characteristics: diabetic and nondiabetic samples

Demographic data and neuropsychological test performances in all 4 groups, namely HC with/without T2DM and MCI with/without T2DM are expressed in means, medians and standard deviations in **Table 11**. Within the HC sample, participants with and without T2DM did not differ statistically significantly in either of the demographic or the neuropsychological variables. However, within the MCI sample, the Backwards version of the Digit Span Test was revealed to be lower among the MCI patients with T2DM compared to those without (MCI-diabetic < MCI-nondiabetic).

2.2 Temporal speech characteristics: HC-sample (diabetic vs. nondiabetic)

Comparison of the diabetic and the nondiabetic groups was executed both within the HC and separately, within the MCI samples. Within the HC sample (***Table 12***), 5 of the 15 parameters differed significantly, as follows: 1) utterance length, higher 2) silent pause duration rate and 3) total pause duration rate, and also higher 4) silent pause average duration and 5) total pause average duration. Taking the direction of the differences into account, this indicates that the HC with T2DM group produced shorter utterances, and on the other hand, higher duration rate and average duration of silent and total pauses.

Table 11. *Demographic characteristics and neuropsychological test scores in the HC with/without T2DM, and the MCI with/without T2DM groups, using the Mann-Whitney U Test or Fisher's Exact Test (in italics). (The p-value that indicates statistically significant differences at the <0.05 level is in* ***bold****.)*

	HC with T2DM (*n* = 39)			**HC without T2DM (*n* = 34)**			**Mann-Whitney U Test / *Fisher's Exact Test***		
	M	***Mdn***	***SD***	***M***	***Mdn***	***SD***	***U***	***Z***	***p***
Demographical data									
Sex (male/female)		13/26			9/25		-	-	*0.613*
Age	65.31	66.00	8.059	67.74	68.00	6.934	548.000	-1.273	0.203
Education (years)	13.03	12.00	2.748	13.29	12.00	2.505	609.500	-0.608	0.543
Neuropsychological tests									
MMSE	28.72	29.00	0.647	29.00	29.00	0.778	531.000	-1.582	0.114
CDT	7.62	9.00	3.159	7.50	9.00	3.077	612.000	-0.584	0.559
ADAS-Cog	7.08	6.15	2.989	6.61	6.95	2.608	607.500	-0.435	0.664
Listening span	2.53	2.60	0.583	2.75	2.85	0.602	504.500	-1.782	0.075
Non-word repetition	5.18	5.00	1.715	4.74	5.00	1.620	552.000	-1.275	0.202
Digit span: forward	5.56	5.00	0.995	5.85	5.50	1.158	579.500	-0.975	0.330
Digit span: backward	4.13	4.00	0.894	4.18	4.00	0.999	642.000	-0.243	0.808
GDS-15	2.00	1.00	1.717	2.00	2.00	1.595	645.000	-0.205	0.838
	MCI with T2DM (*n* = 12)			**MCI without T2DM (*n* = 15)**			**Mann-Whitney U Test / *Fisher's Exact Test***		
	M	***Mdn***	***SD***	***M***	***Mdn***	***SD***	***U***	***Z***	***p***
Demographical data									
Sex (male/female)		2/10			5/10		-	-	*0.408*
Age	70.42	73.50	9.120	72.60	74.00	6.311	83.500	-0.318	0.755
Education (years)	11.17	11.50	2.855	11.73	12.00	2.865	76.000	-0.712	0.516
Neuropsychological tests									
MMSE	26.17	26.00	0.835	26.27	26.00	0.799	84.000	-0.315	0.792
CDT	5.50	4.50	3.529	7.33	8.00	2.870	64.000	-1.281	0.217
ADAS-Cog	9.38	9.00	2.070	10.61	10.60	3.104	64.000	-1.271	0.217
Listening span	2.32	2.15	0.476	2.23	2.30	0.434	87.000	-0.151	0.905
Non-word repetition	3.58	5.00	2.575	3.67	4.00	1.718	81.000	-0.450	0.683
Digit span: forward	5.00	5.00	1.128	5.33	5.00	0.617	60.500	-1.668	0.152
Digit span: backward	3.25	3.00	0.754	3.93	4.00	0.799	49.000	-2.161	**0.047**
GDS-15	1.92	2.00	1.505	2.53	2.00	1.187	62.000	-1.436	0.183

Abbreviations: *HC: healthy cognition; MCI: mild cognitive impairment; T2DM: type 2 diabetes mellitus; M: mean; Mdn: median; SD: standard deviation; MMSE: Mini-Mental State Examination; CDT: Clock Drawing Test; ADAS-Cog: Alzheimer's Disease Assessment Scale – Cognitive Subscale; GDS-15: 15-item Geriatric Depression Scale*

Table 12. *Descriptive and comparative statistics of the HC with and without T2DM groups using the Mann-Whitney U test. (The p-values that indicate statistically significant differences at the <0.05 level are in* ***bold****.)*

	HC with T2DM (*n* = 39)			**HC without T2DM (*n* = 34)**			**Mann-Whitney U test**		
Temporal speech parameters	***M***	***Mdn***	***SD***	***M***	***Mdn***	***SD***	***U***	***Z***	***p***
Utterance length (s)	114.00	93.36	68.274	205.68	151.88	235.281	407.000	-2.831	**0.005**
Articulation tempo (1/s)	9.27	9.49	1.907	9.65	9.68	2.001	602.000	-0.675	0.500
Speech tempo (1/s)	10.05	10.30	1.872	10.46	10.48	1.850	597.000	-0.730	0.465
Occurrence rates of pauses									
Silent pause (%)	5.55	5.35	1.562	5.29	4.83	2.458	536.000	-1.404	0.160
Filled pause (%)	2.57	2.15	1.613	3.09	2.56	2.123	573.000	-0.995	0.320
Total pause (%)	8.11	7.32	2.642	8.38	7.41	4.268	639.000	-0.265	0.791
Duration rates of pauses									
Silent pause (%)	32.16	29.40	10.991	25.79	24.13	10.850	429.000	-2.588	**0.010**
Filled pause (%)	5.81	5.04	4.054	6.92	6.03	3.940	556.000	-1.183	0.237
Total pause (%)	37.97	37.90	11.495	32.71	30.79	12.700	474.000	-2.090	**0.037**
Frequency of pauses									
Silent pause (1/s)	0.53	0.53	0.101	0.52	0.48	0.142	580.000	-0.918	0.359
Filled pause (1/s)	0.24	0.23	0.140	0.30	0.27	0.150	516.000	-1.626	0.104
Total pause (1/s)	0.78	0.74	0.174	0.82	0.78	0.241	620.000	-0.476	0.634
Average durations of pauses									
Silent pause (s)	0.62	0.55	0.248	0.50	0.46	0.169	453.000	-2.322	**0.020**
Filled pause (s)	0.22	0.20	0.072	0.22	0.22	0.056	590.500	-0.802	0.423
Total pause (s)	0.50	0.45	0.164	0.41	0.37	0.128	419.000	-2.698	**0.007**

Abbreviations: *HC = healthy cognition; T2DM: type 2 diabetes mellitus; M: mean; Mdn: median; SD: standard deviation*

Furthermore, in order to investigate which temporal speech parameters would be the most informative and sensitive in the detection of diabetic vs. nondiabetic individuals based solely on their speech, ROC analysis was executed. The results 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 (0.693), while average duration of total pause yielded the highest sensitivity (79.5%). Sensitivity and specificity measures were derived from ROC analysis, using threshold values tailored for early screening, and are detailed for parameters with an AUC above 0.600 in ***Table 14***.

2.3 Temporal speech characteristics: MCI-sample (diabetic vs. nondiabetic)

Comparison of the diabetic and the nondiabetic groups was also executed within the MCI sample. However, based on the statistical analysis, no statistically significant difference could be detected regarding the 15 temporal speech parameters (MCI-diabetic ~ MCI-nondiabetic) (***Table 13***).

Results of the subsequent ROC analysis revealed a similar tendency, as in no parameters had statistically significant abilities to discriminate MCI patients with or without T2DM. Nevertheless, upon examination of the AUCs, parameters concerning the filled pauses produced the highest AUCs. Sensitivity and specificity measures of temporal parameters were derived from ROC analysis; parameters with an AUC above 0.600 are shown in ***Table 14***.

Table 13. *Descriptive and comparative statistics of the MCI with and without T2DM groups using the Mann-Whitney U test.*

	MCI with T2DM (*n* = 12)			**MCI without T2DM (*n* = 15)**			**Mann-Whitney U test**		
Temporal speech parameters	***M***	***Mdn***	***SD***	***M***	***Mdn***	***SD***	***U***	***Z***	***p***
Utterance length (s)	119.50	80.10	93.150	131.70	79.40	39.058	83.000	-0.342	0.755
Articulation tempo (1/s)	9.26	9.64	2.644	8.76	8.20	1.703	76.000	-0.683	0.516
Speech tempo (1/s)	9.99	10.37	2.555	9.57	9.09	1.582	77.000	-0.634	0.548
Occurrence rates of pauses									
Silent pause (%)	5.77	5.74	2.504	5.73	5.47	1.841	88.000	-0.098	0.943
Filled pause (%)	2.19	2.74	1.344	3.13	2.72	2.009	67.000	-1.122	0.277
Total pause (%)	7.97	7.98	3.445	8.85	8.63	3.272	77.000	-0.634	0.548
Duration rates of pauses									
Silent pause (%)	33.94	32.68	16.602	31.93	28.69	7.933	89.000	-0.049	0.981
Filled pause (%)	4.41	5.03	2.883	6.84	7.65	4.474	62.000	-1.366	0.183
Total pause (%)	38.35	36.75	17.231	38.77	36.57	10.476	86.000	-0.195	0.867
Frequency of pauses									
Silent pause (1/s)	0.52	0.53	0.128	0.53	0.54	0.135	88.000	-0.098	0.943
Filled pause (1/s)	0.20	0.19	0.129	0.28	0.27	0.152	62.000	-1.366	0.183
Total pause (1/s)	0.73	0.77	0.204	0.81	0.78	0.214	73.000	-0.830	0.427
Average durations of pauses									
Silent pause (s)	0.64	0.56	0.255	0.62	0.62	0.164	82.000	-0.390	0.719
Filled pause (s)	0.21	0.21	0.041	0.24	0.23	0.097	76.000	-0.683	0.516
Total pause (s)	0.53	0.47	0.210	0.49	0.49	0.099	87.000	-0.146	0.905

Abbreviations: *MCI: mild cognitive impairment; T2DM: type 2 diabetes mellitus; M: mean; Mdn: median; SD: standard deviation*

Table 14. *Accuracy measures of temporal parameters with AUC above 0.600 in the HC and in the MCI samples, respectively (comparing the 'with T2DM' and 'without T2DM' subgroups), using ROC analysis. (The p-values that indicate statistically significant differences at the <0.05 level are in* ***bold****.)*

HC groups (with vs. without T2DM)			**Accuracy measures**				
Temporal speech parameters	***p***	**AUC**	**95% CI-**	**95% CI+**	**Threshold value**	**Sensitivity (%)**	**Specificity (%)**
Utterance length (s)	**0.005**	0.693	0.572	0.815	131.845	74.4	61.8
Total pause average duration (s)	**0.007**	0.684	0.560	0.808	0.374	79.5	55.9
Silent pause duration rate (%)	**0.010**	0.676	0.553	0.800	24.192	74.4	52.9
Silent pause average duration (s)	**0.020**	0.658	0.532	0.785	0.471	74.4	55.9
Total pause duration rate (%)	**0.037**	0.643	0.514	0.771	31.705	66.7	55.9
Filled pause frequency (1/s)	0.104	0.611	0.481	0.740	0.246	61.5	58.8
MCI groups (with vs. without T2DM)			**Accuracy measures**				
Temporal speech parameters	***p***	**AUC**	**95% CI-**	**95% CI+**	**Threshold value**	**Sensitivity (%)**	**Specificity (%)**
Filled pause duration rate (%)	0.172	0.656	0.446	0.865	6.754	83.3	53.3
Filled pause frequency (1/s)	0.172	0.656	0.443	0.868	0.229	66.7	60.0
Filled pause occurrence rate (%)	0.262	0.628	0.408	0.848	2.715	50.0	53.3

Abbreviations: *T2DM = type 2 diabetes mellitus; HC = healthy cognition; MCI: mild cognitive impairment; ROC: receiver operating characteristic; AUC: area under the curve; CI: confidence interval*

2.4 Correlations of temporal speech parameters with age and education

Regarding the relationship between the 15 temporal speech parameters and the *age* of the participants across the 4 groups, correlation was found statistically significant in the following cases: 1) articulation tempo (HC with T2DM: τ_b = -0.221, p = 0.050), 2) speech tempo (HC with T2DM: τ_b = -0.229, p = 0.042), and 3) silent pause frequency (MCI without T2DM: τ_b = 0.390, p = 0.046). Regarding *education*, weak to moderate but statistically significant correlations were found with 1) utterance length (HC without T2DM: τ_b = 0.269, p = 0.035; MCI with T2DM: τ_b = 0.478, p = 0.044), 2) articulation tempo (MCI with T2DM: τ_b = 0.478, p = 0.044), 3) speech tempo (MCI with T2DM: τ_b = 0.546, p = 0.021), 4) filled pause occurrence rate (HC with T2DM: τ_b = 0.274, p = 0.022), 5) filled pause duration rate (HC with T2DM: τ_b = 0.268, p = 0.025; MCI without T2DM: τ_b = 0.596, p = 0.004), 6) silent pause average duration (MCI with T2DM: τ_b = -0.580, p = 0.014), 7) filled pause average duration (MCI without T2DM: τ_b = 0.618, p = 0.003), and 8) total pause average duration (MCI with T2DM: τ_b = -0.615, p = 0.010). The comprehensive table containing all correlations is in ***Table 15***.

Table 15. *Correlations between the 15 temporal speech parameters and demographic features (age, education) in all 4 groups (HC with and without T2DM; MCI with and without T2DM) using the Kendall-tau correlation. (The τ_b and p-values that indicate statistically significant differences at the <0.05 level are in* ***bold****.)*

		HC with T2DM (*n* = 39)		**HC without T2DM (*n* = 34)**		**MCI with T2DM (*n* = 12)**		**MCI without T2DM (*n* = 15)**	
Temporal speech parameters		Age	Education	Age	Education	Age	Education	Age	Education
Utterance length (s)	τ_b	-0.161	0.107	0.084	**0.269**	-0.188	**0.478**	-0.117	-0.085
	p	0.153	0.369	0.494	**0.035**	0.406	**0.044**	0.550	0.678
Articulation tempo (1/s)	τ_b	**-0.221**	-0.146	-0.197	0.074	-0.063	**0.478**	0.137	-0.170
	p	**0.050**	0.221	0.108	0.562	0.782	**0.044**	0.486	0.406
Speech tempo (1/s)	τ_b	**-0.229**	-0.164	-0.223	0.078	-0.063	**0.546**	0.215	-0.085
	p	**0.042**	0.170	0.069	0.542	0.782	**0.021**	0.273	0.678
Occurrence rates of pauses									
Silent pause (%)	τ_b	0.139	0.092	0.055	-0.035	-0.031	-0.410	0.234	0.106
	p	0.217	0.439	0.655	0.784	0.890	0.084	0.232	0.604
Filled pause (%)	τ_b	-0.093	**0.274**	-0.055	0.125	0.156	-0.102	-0.176	0.277
	p	0.410	**0.022**	0.655	0.329	0.489	0.666	0.370	0.177
Total pause (%)	τ_b	0.060	0.193	-0.011	0.035	0.000	-0.341	0.020	0.234
	p	0.594	0.105	0.929	0.784	1.000	0.150	0.921	0.254
Duration rates of pauses									
Silent pause (%)	τ_b	0.079	0.015	-0.029	-0.070	-0.063	-0.444	0.039	0.021
	p	0.482	0.901	0.812	0.583	0.782	0.061	0.842	0.917
Filled pause (%)	τ_b	-0.150	**0.268**	-0.004	0.176	0.313	0.000	-0.234	**0.596**
	p	0.183	**0.025**	0.976	0.170	0.166	1.000	0.232	**0.004**
Total pause (%)	τ_b	0.011	0.101	-0.029	-0.027	-0.094	-0.410	-0.020	0.234
	p	0.923	0.396	0.812	0.831	0.678	0.084	0.921	0.254
Frequency of pauses									
Silent pause (1/s)	τ_b	0.033	0.095	-0.007	-0.078	-0.094	-0.341	**0.390**	0.128
	p	0.771	0.424	0.953	0.542	0.678	0.150	**0.046**	0.533
Filled pause (1/s)	τ_b	-0.172	0.229	-0.110	0.144	0.281	0.068	-0.195	0.319
	p	0.127	0.055	0.372	0.259	0.213	0.774	0.319	0.120
Total pause (1/s)	τ_b	-0.134	0.173	-0.055	0.059	0.250	-0.034	0.078	0.256
	p	0.235	0.148	0.655	0.647	0.268	0.886	0.690	0.213
Average durations of pauses									
Silent pause (s)	τ_b	0.074	-0.059	0.022	-0.051	-0.125	**-0.580**	-0.332	-0.106
	p	0.513	0.618	0.858	0.692	0.580	**0.014**	0.090	0.604
Filled pause (s)	τ_b	-0.061	0.191	0.223	0.199	0.094	-0.273	-0.117	**0.618**
	p	0.586	0.110	0.069	0.120	0.678	0.250	0.550	**0.003**
Total pause (s)	τ_b	0.066	-0.098	0.051	-0.066	-0.125	**-0.615**	-0.098	0.106
	p	0.561	0.410	0.677	0.604	0.580	**0.010**	0.619	0.604

Abbreviations: *HC: healthy cognition; MCI: mild cognitive impairment; T2DM: type 2 diabetes mellitus*

VII. 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 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 decline.

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. Extraction of temporal speech parameters from these recordings was carried out in an automated way using ASR, which was then followed by statistical analysis to determine which parameters have the best discrimination abilities for MCI in the different target groups (English-/Hungarian-speakers, diabetic/nondiabetic patients) and are therefore the most informative regarding cognition. 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 for the at-risk individuals themselves.

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 further confirms

and 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 was to, for the first time, apply temporal speech analysis on two languages with the same study design (from the identical inclusion/exclusion criteria of 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 whether the methodology already tried in the Hungarian language could be used in the English language as well, and with what kind of results.

The applied methodology, also termed the S-GAP Test, has been under development for over a decade, and started with the discovery of statistically significant differences between mild AD and HC individuals, based on speech tempo and hesitation ratio (Hoffmann *et al.*, 2010). At this beginning stage, the speech task was different from that of today (it elicited spontaneous speech, however it was less dependent on episodic memory), and only four temporal speech parameters were calculated. The next milestone in the development process was the targeting of the prodromal stage, MCI; the usefulness of temporal speech analysis was demonstrated again in the successful discrimination of MCI vs. HC participants (Tóth *et al.*, 2015). Parallel to the introduction of MCI as a target group, another major step was the implementation of automated analysis (using ASR techniques) instead of relying on labor-intensive and time-consuming manual counting of acoustic features and labeling of the speech recordings. Finally, a machine learning model was constructed (Tóth *et al.*, 2015; Tóth *et al.*, 2018a), with which analysis proved to have similar, or even improved efficiency in MCI-detection than the previous manual model (Gosztolya *et al.*, 2016). Extending applicability to new targets, the S-GAP Test demonstrated successful differentiation not only between MCI and HC, but also between MCI and mild AD patients (Gosztolya *et al.*, 2019, 2021).

The present study was a new stepping stone regarding the evolution of temporal analysis of spontaneous speech, as the S-GAP Test was able to detect MCI-cases not only in native Hungarian-speaking but in a native English-speaking population as well. 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.

Additional to the above-mentioned 4 shared parameters, the English-speaking MCI group also showed lower articulation tempo and speech tempo compared to the HC group. These two features had been previously found to differ between mild AD/MCI vs. HC in the Hungarian language, using both manual and automatic analysis (Hoffmann *et al.*, 2010; Tóth *et al.*, 2015, 2018a; Gosztolya *et al.*, 2019, 2021). Interestingly, Hungarian-speaking MCI/AD patients of previous studies demonstrated a reduction in articulation and speech tempo, while in our present sample, this phenomenon was only tendentious. A possible explanation might be the variation regarding the task that was implemented for speech elicitation: namely in some of our previous studies, a film description task was used (in which the participants had to retell the events of a one-minute long silhouette animation, specifically constructed for the research), instead of the 'previous day' task applied in the present study. While both tasks require memory function, the latter might be more complex in the sense that it mobilizes longer term memory, assuming successful elimination of confounding information (*e.g.* similar memories from another day), and also the decision-making and planning abilities regarding the detailed/simplified reporting of the events. Since all these functions bear more cognitive load, it might manifest more in pauses and less in the general slowing down of 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 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.

In the context of previous, non-Hungarian 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 (Meilán *et al.*, 2012; De Looze *et al.*, 2018; Espinoza-Cuadros *et al.*, 2014), picture description (Satt *et al.*, 2013), narrative recall (Roark *et al.*, 2011), or spontaneous speech (Gayraud *et al.*, 2011; Lee *et al.*, 2011; Singh *et al.*, 2001).

The prominence of pause-related speech characteristics are indicative of retrieval difficulties (Szatloczki *et al.*, 2015) which are related to degeneration in hippocampal brain regions (Sarazin *et al.*, 2010). They are also associated with atrophy of gray matter in the frontopolar (or Brodmann) area of the cortex (Pistono *et al.*, 2016), which plays a role in higher-order cognitive functions like multitasking (Roca *et al.*, 2011) or memory retrieval (Simons *et al.*, 2005). It is hypothesized that increased number and duration of pauses are manifestations of increased cognitive load required for maintaining one's train of thought during speech (König *et al.*, 2015). 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 (Meilán *et al.*, 2012; Gayraud *et al.*, 2011). 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 (Arlt *et al.*, 2013).

It is important to note that the majority of the previous investigations regarding speech analysis in MCI/AD did not focus on reporting accuracy metrics (Roark *et al.*, 2011; Sajjadi *et al.*, 2012). However in the last few years, more relevant data have been shared (Tóth *et al.*, 2018a; Gosztolya *et al.*, 2019, 2021; König *et al.*, 2015, 2018; Hernández-Dominguez *et al.*, 2018). To provide international examples, classification sensitivity of an automated analysis based on linguistic and phonetic features (using a picture description task) was 85% between HC vs. AD/MCI cases in a Canadian-Mexican study (Hernández-Dominguez *et al.*, 2018), while in a French-speaking population, diagnostic utility of automated speech analysis had 79% and 86% classification accuracy between HC vs. MCI (König *et al.*, 2015, 2018). Regarding earlier works of our research group, 75% accuracy was achieved in differentiating MCI from HC (Gosztolya *et al.*, 2019), however it was based on a different language elicitation task (a video description, as mentioned earlier). Compared to these metrics, the present version of the

S-GAP Test applied for English- and Hungarian-speaking populations has relatively fair sensitivity in detecting MCI.

Regarding sensitivity and specificity in the present study, the optimal threshold values for each temporal speech parameter were defined with the goal of maximizing sensitivity. This, as a result of trade-off between the two measures, decreased specificity (although it exceeded 50% in every case). Considering the serious consequences of undiagnosed cognitive decline (mainly the possibility of silently progressing from MCI to dementia (Gauthier *et al.*, 2006)), the primary goal was to construct an especially sensitive MCI-screening tool targeted at high-risk individuals, thus high true positive rate was prioritized even at the expense of lower true negative rate.

Learning from the observed inter-language differences (E-HC vs. H-HC; E-MCI vs. H-MCI), 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. For example, 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. These and similar linguistic differences would have to be considered individually in different countries when defining screening thresholds, as temporal speech characteristics (even among HCs) can 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, the present study 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 nondiabetic individuals (who were also classified as HC).

As a main goal, the intention 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 contained signs of 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 pause and total pause, and also higher average duration of silent pause and total pause 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 Hungarian-speaking sample HC/MCI comparison of ***Study 1*** (***Table 12*** vs. ***Table 7***). This might further confirm that from the full set of 15 temporal speech parameters of the S-GAP Test, these few have possibly the most discriminative potential in future clinical use.

These differences are in tune with the results of 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, *e.g.* due to MCI (Roark *et al.*, 2011; Meilán *et al.*, 2020), to AD (Hoffmann *et al.*, 2010; López-de-Ipiña *et al.*, 2013; Martínez-Sánchez *et al.*, 2013), or even to Parkinson's disease (Hlavnička *et al.*, 2017; Alvar *et al.*, 2019). These results complemented by the present study confirm that speech pauses offer a highly valuable information source on language functions and thus cognitive state, especially in the beginning, early, non-symptomatic stages of neurocognitive disorders. This is the stage when other cognitive domains have not yet deteriorated in such a magnitude to be detected by conventional neuropsychological test batteries. In the case of T2DM patients, these 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 (Biessels *et al.*, 2006; Bello-Chavolla *et al.*, 2019; Stranahan *et al.*, 2008).

Speech of MCI patients with/without T2DM were also explored in terms of temporal characteristics. According to statistical analysis, no significant differences could be detected in any of the investigated temporal speech parameters, suggesting that these two groups performed similarly. A possible explanation for the lack of differences 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

(Petersen *et al.*, 1999, 2001). Nonetheless, the underlying neuropathological deterioration is usually present for a much longer period, more or less without clinical symptoms (Albert *et al.*, 2018). 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 not necessarily aggravates 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.

With regard to the relationship between demographic and temporal speech characteristics, age demonstrated a statistically significant (albeit weak) correlation with three parameters: a negative correlation with articulation tempo and speech tempo, and a positive correlation with silent pause frequency. Education correlated (weakly to moderately) with eight parameters: positively with utterance length, articulation tempo, speech tempo, filled pause occurrence rate, filled pause duration rate, and filled pause average duration; while negatively with silent pause average duration and total pause average duration. After thorough examination of the positive and negative directions of the statistically significant correlations, 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 (Luck *et al.*, 2010; Patterson *et al.*, 2007)). 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 (Luck *et al.*, 2010)).

The implementation of telemedicine in remote diabetes management is a dynamically emerging area, however to this date no such technique has been used for the cognitive examination of diabetic patients. In a possible future application, subtle speech deficit detected by the S-GAP Test could serve as an indication for a thorough medical and neuropsychological follow-up examination to find the possible underlying cause and to monitor the patient more closely (*e.g.* with frequent medical check-ups). Remote assessment is gaining increasing

interest and clinical relevance −mainly in light of the current COVID-19 pandemic which restricts the once usual face-to-face medical appointments. The S-GAP Test would ideally be used in the format of a mobile application which could offer a rapid, non-invasive, no-contact, and cost-effective form of cognitive screening for the elderly – and complemented by the present results, could be utilized for the cognitive monitoring of diabetic patients as well.

4. Limitations and future perspectives

In our effort to aid accurate interpretation and to improve study design, a number of limitations must be considered regarding the studies elaborated in this thesis. Firstly, both studies reckoned MCI as a syndromic, or non-categorized phenomenon, meaning that possible subtypes (*e.g.* amnestic, non-amnestic, single- or multi-domain) were not differentiated. This decision was made purposefully in the planning stage of the research, as our main aim was to screen for individuals most at risk of MCI and thus dementia, regardless of subtypes, and not to thoroughly explore the symptomatology or to establish a diagnosis. For the latter, a detailed clinical picture must be drawn, ideally accompanied by a series of neuropsychological tests and neuroimaging, which belong to the scope of secondary care, while the S-GAP Test is intended for use of the individuals themselves or at primary care (*i.e.* family practices). In the majority of similar neurolinguistic studies, MCI is also often examined as a syndromic disorder with unspecified/unknown cause, and subtyping is not evident. Nevertheless, in future studies, it would be of interest to take into account the heterogeneity of MCI by involving more patients and by creating groups based on the different subtypes. For the same reasons, probable or possible etiology of MCI was also not considered, although it could be hypothesized that the majority of participants who were considered MCI-patients belong to the AD-variant, as 1) Alzheimer-type changes in the neuropathology of the brain are accountable for 60%-80% of all dementia cases (Alzheimer's Association, 2022a), and as 2) major vascular malformations, any evidence of other possible neurodegenerative etiology, and substance abuse were all among the exclusion criteria (as described in *Methods and materials*). Regarding T2DM (***Study 2***), future works could also incorporate a wider spectrum of diabetes-related medical characteristics, or even a full panel blood test, which could allow the creation of subgroups based on variants of T2DM (*e.g.* severity, levels of insulin, *etc.*).

Secondly, in future studies involving the S-GAP Test, the full set of 15 temporal speech parameters could be reduced, as it would improve the robustness of the results. The present

studies however, were intentionally aimed at identifying those few parameters that provide the highest sensitivity and the most promising differentiating potential to be included later in a telemedicine-based assessment only containing the most relevant features for screening (*e.g.* in a mobile application). Thus, multiple correction testing was not applied for the statistical comparisons, which needs to be taken into account during the interpretation of the results.

Thirdly, on a more technical note: although the sampling rate of the speech recordings could be higher than the one used in the present studies (8,000 Hz), this specification was an intentional choice. The S-GAP Test is specifically designed to be applied either independently (by the user) or by a general practitioner, possibly in the form of a mobile application. Since 8,000 Hz is usually available on most smartphone devices (even on more simple models), this would enable wider adoption of this tool.

Finally, for increasing the statistical power of both studies, greater sample sizes should be aimed for in the future. In particular, a higher number of MCI participants would be ideal, as it could also have contributed to the lack of between-group differences within the MCI-sample (**Study 2**). In our case, moderate sample sizes were a result of 1) the wide range of inclusion and exclusion criteria that were applied to eliminate possible confounding factors; 2) the complexity of the recruitment process due to the cooperation with outside research sites (Department of Internal Medicine and Department of Psychiatry; University of Columbia, New York, USA and University of Szeged, Hungary); and 3) also the sensitive nature or even stigma that surrounds the topic of cognitive decline, causing many possible candidates to be reluctant to participate.

VIII. 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 the S-GAP Test has the potential to become a useful method for early MCI screening both in English-speaking and Hungarian-speaking populations. Early diagnosis of cognitive decline is of much help for patients and their families, both for starting early treatment and for planning the future. Nonetheless, it is important to highlight that this method can only serve as a first step towards the diagnostic process of MCI, as it is not intended (nor suitable) to substitute detailed clinical examination.

The speech of T2DM patients were explored for the first time, building on the shared pathophysiology of T2DM and neurocognitive disorders, and also the strong association between speech deficits and cognitive decline. Results revealed that the speech of diabetic patients, otherwise classified as HC, nevertheless 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, as the situation of a phone call is part of everyday life. Furthermore, this approach might also aid measuring the objective efficacy of pharmacotherapy and drug candidate molecules in various cognitive impairments.

IX. 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.

X. REFERENCES

Ahmed, S., Haigh, A.M., de Jager, C.A., & Garrard, P. (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain, 136*(12), 3727–3737. DOI: 10.1093/brain/awt269.

Ahtiluoto, S., Polvikoski, T., Peltonen, M., Solomon, A., Tuomilehto, J., Winblad, B., Sulkava, R., & Kivipelto, M. (2010). Diabetes, Alzheimer-disease, and vascular dementia: a population-based neuropathologic study. *Neurology, 75*(13), 1195-1202. doi: 10.1212/WNL.0b013e3181f4d7f8.

Albert, M., Zhu, Y., Moghekar, A., Mori, S., Miller, M.I., Soldan, A., Pettigrew, C., Selnes, O., Li, S., & Wang, M.-C. (2018). Predicting progression from normal cognition to mild cognitive impairment for individuals at 5 years. *Brain, 141*(3), 877-887. DOI: 10.1093/brain/awx365.

Alvar, A.M., Lee, J., & Huber, J.E. (2019). Filled pauses as a special case of automatic speech behaviors and the effect of Parkinson's disease. *American Journal of Speech-Language Pathology, 28*(2S), 835-843. DOI: 10.1044/2018_AJSLP-MSC18-18-0111.

Alzheimer's Association (2022a). *2022 Alzheimer's Disease Facts and Figures.* Retrieved from: https://www.alz.org/media/Documents/alzheimers-facts-and-figures.pdf. Last accessed: July 25, 2022.

Alzheimer's Association (2022b). *Special report: More than normal aging: understanding mild cognitive impairment.* Retrieved from: https://www.alz.org/media/Documents/alzheimers-facts-and-figures-special-report.pdf. Last accessed: July 25, 2022.

American Diabetes Association. (2014). Diagnosis and classification of diabetes mellitus. *Diabetes Care, 37*(1), S81-90. DOI: 10.2337/dc14-S081.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th Ed., Text Revision). Washington, DC: American Psychiatric Association.

American Psychiatric Association (2013a). *Diagnostic and Statistical Manual of Mental Disorders* (5th Ed.). Arlington, VA: American Psychiatric Association.

American Psychiatric Association (2013b). *Desk Reference to the Diagnostic Criteria From DSM-5.* Arlington, VA: American Psychiatric Association.

Angevaare, M.J., Vonk, J.M.J., Bertola, L., Zahodne, L., Wei-Ming Watson, C., Boehme, A., Schupf, N., Mayeux, R., Geerlings, M.I., & Manly, J.J. (2022). Predictors of incident mild

cognitive impairment and its course in a diverse community-based population. *Neurology, 98*(1), e15-e26. DOI: 10.1212/WNL.0000000000013017.

Arlt, S., Buchert, R., Spies, L., Eichenlaub, M., Lehmbeck, J.T., & Jahn, H. (2013). Association between fully-automated MRI-based volumetry of different brain regions and neuropsychological test performance in patients with amnestic mild cognitive impairment and Alzheimer's disease. *European Archives of Psychiatry and Clinical Neuroscience, 263*(4), 335-344. DOI: 10.1007/s00406-012-0350-7.

Arvanitakis, Z., Shah, R.C., & Bennett, D.A. (2019). Diagnosis and management of dementia: a review. *JAMA, 322*(16), 1589-1599. DOI: 10.1001/jama.2019.4782.

Awad, N., Gagnon, M., & Messier, C. (2004). The relationship between impaired glucose tolerance, type 2 diabetes, and cognitive function. *Journal of Clinical and Experimental Neuropsychology, 26*(8), 1044-1080. DOI: 10.1080/13803390490514875.

Balázs, N., Ajtay, A., Oberfrank, F., Bereczki, D., & Kovács, T. (2021). Dementia epidemiology in Hungary based on data from neurological and psychiatric specialty services. *Scientific Reports, 11*, 10333. DOI: 10.1038/s41598-021-89179-3.

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), 7-13. DOI: 10.1080/13814788.2019.1673723.

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*, 1-13. DOI: 10.1017/S1355617721001454.

Bello-Chavolla, O.Y., Antonio-Villa, N.E., Vargas-Vázquez, A., Ávila-Funes, J.A., & Aguilar-Salinas, C.A. (2019). Pathophysiological mechanisms linking type 2 diabetes and dementia: review of evidence from clinical, translational and epidemiological research. *Current Diabetes Reviews, 15*(6), 456-470. DOI: 10.2174/1573399815666190129155654.

Beltrami, D., Gagliardi, G., Rossini Favretti, R., Ghidoni, E., Tamburini, F., & Calzà, L. (2018). Speech analysis by natural language processing techniques: a possible tool for very early detection of cognitive decline? *Frontiers in Aging Neuroscience, 10*, 369. DOI: 10.3389/fnagi.2018.00369.

Berchtold, N.C., & Cotman, C.W. (1998). Evolution of the conceptualization of dementia and Alzheimer's disease: Greco-Roman period to the 1960s. *Neurobiology of Aging, 19*(3), 173-189. DOI: 10.1016/S0197-4580(98)00052-9.

Berisha, V., Wang, S., LaCross, A., & Liss, J. (2015). Tracking discourse complexity preceding Alzheimer's disease diagnosis: a case study comparing the press conferences of Presidents Ronald Reagan and George Herbert Walker Bush. *Journal of Alzheimer's Disease, 45*(3), 959-963. DOI: 10.3233/JAD-142763.

Biessels, G.J., Staekenborg, S., Brunner, E., Brayne, C., & Scheltens, P. (2006). Risk of dementia in diabetes mellitus: a systematic review. *The Lancet. Neurology, 5*(1):64-74. DOI: 10.1016/S1474-4422(05)70284-2.

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot International, 5*(9-10), 341–347.

Boschi, V., Catricalà, E., Consonni, M., Chesi, C., Moro, A., & Cappa, S.F. (2017). Connected speech in neurodegenerative language disorders: a review. Frontiers in Psychology, 8(269). DOI: 10.3389/fpsyg.2017.00269.

Bose, A., Dutta, M., Dash, N.S., Nandi, R., Dutt, A., & Ahmed, S. (2022). Importance of task-selection for connected speech analysis in patients with Alzheimer's disease from an ethnically diverse sample. *Journal of Alzheimer's Disease, 87*(4), 1475-1481. DOI: 10.3233/JAD-220166.

Braaten, A.J., Parsons, T.D., McCue, R., Sellers, A., & Burns, W.J. (2006). Neurocognitive differential diagnosis of dementing diseases: Alzheimer's dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. *The International Journal of Neuroscience, 116*(11), 1271–1293. DOI: 10.1080/00207450600920928.

Burns, A. (2000). The burden of Alzheimer's disease. *The International Journal of Neuropsychopharmacology, 3*(7), 31 – 38. DOI: 10.1017/S1461145700001905.

Callisaya, M. L., Beare, R., Moran, C., Phan, T., Wang, W., & Srikanth, V. K. (2019). Type 2 diabetes mellitus, brain atrophy and cognitive decline in older people: a longitudinal study. *Diabetologia, 62*(3), 448-458. DOI: 10.1007/s00125-018-4778-9.

Chen, J.-H., Lin, K.-P., & Chen, Y.-C. (2009). Risk factors for dementia. *Journal of the Formosan Medical Association, 108*(10), 754-764. DOI: 10.1016/S0929-6646(09)60402-2.

Cheng, G., Huang, C., Deng, H., & Wang, H. (2012). Diabetes as a risk factor for dementia and mild cognitive impairment: a meta-analysis of longitudinal studies. *Internal Medicine Journal, 42*(5), 484-491. DOI: 10.1111/j.1445-5994.2012.02758.x.

Chien, Y., Hong, S., Cheah, W., Fu, L., & Chang, Y. (2018). An assessment system for Alzheimer's disease based on speech using a novel feature sequence design and recurrent neural network. *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 3289-3294 DOI: 10.1109/SMC.2018.00557.

Cukierman, T., Gerstein, H.C., & Williamson, J.D. (2005). Cognitive decline and dementia in diabetes – systematic overview of prospective observational studies. *Diabetologia, 48*(12), 2460-2469. DOI: 10.1007/s00125-005-0023-4.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*(4), 450-466. doi: 10.1016/S0022-5371(80)90312-6.

Degen, C., Toro, P., Schönknecht, P., Sattler, C., & Schröder, J. (2016). Diabetes mellitus Type II and cognitive capacity in healthy aging, mild cognitive impairment and Alzheimer's disease. *Psychiatry Research, 240,* 42–46. DOI: 10.1016/j.psychres.2016.04.009.

De la Fuente Garcia, S., Ritchie, C.W., & Luz, S. (2020). Artificial intelligence, speech, and language processing approaches to monitoring Alzheimer's disease: a systematic review. *Journal of Alzheimer's Disease, 78*(4), 1547-1574. DOI: 10.3233/JAD-200888.

De la Monte, S. (2014). Type 3 diabetes is sporadic Alzheimer's disease: mini-review. *European Neuropsychopharmacology, 24*(12), 1954-1960. DOI: 10.1016/j.euroneuro.2014.06.008.

De Looze, C., Kelly, F., Crosby, L., Vourdanou, A., Coen, R.F., Walsh, C., Lawlor, B.A., & Reilly, R.B. (2018). Changes in speech chunking in reading aloud is a marker of mild cognitive impairment and mild-to-moderate Alzheimer's disease. *Current Alzheimer Research, 15*(9), 828–847. DOI: 10.2174/1567205015666180404165017.

Deme, A., & Markó, A. (2013). Lengthenings and filled pauses in Hungarian adults' and children's speech. *Proceedings of Disfluency in Spontaneous Speech (DiSS),* 21-24.

Domján, B.A., Ferencz, V., Tänczer, T., Szili-Janicsek, Z., Barkai, L., Hidvégi, T., Jermendy, G., Kempler, P., Winkler, G., Gerő, L., & Tabák, A.D. (2017). Large increase in the

prevalence of self-reported diabetes based on a nationally representative survey in Hungary. *Primary Care Diabetes, 11*(2), 107-111. DOI: 10.1016/j.pcd.2016.09.001.

Eklund, R. (2001). Prolongations: A dark horse in the disfluency stable. *Proceedings of Disfluency in Spontaneous Speech (DiSS),* 5-8.

Érsek, K., Kárpáti, K., Kovács, T., Csillik, G., Gulácsi, Á.L., & Gulácsi, L. (2010). [Epidemiology of dementia in Hungary.] *Ideggyógyászati Szemle, 63*(5-6), 175-182.

Espinoza-Cuadros, F., Garcia-Zamora, M.A., Torres-Boza, D., Ferrer-Riesgo, C.A., Montero-Benavides, A., Gonzalez-Moreira, E., & Hernandez-Gómez, L.A. (2014). A spoken language database for research on moderate cognitive impairment: Design and preliminary analysis. *Advances in Speech and Language Technologies for Iberian Languages (IberSpeech), 8854*, 219-228. DOI: 10.1007/978-3-319-13623-3_23.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). „Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198. doi: 10.1016/0022-3956(75)90026-6.

Forbes, K.E., Venneri, A., & Shanks, M.F. (2002). Distinct patterns of spontaneous speech deterioration: an early predictor of Alzheimer's disease. *Brain and Cognition, 48*(2-3), 356-361. DOI: 10.1006/brcg.2001.1377.

Forbes-McKay, K.E., & Venneri, A. (2005). Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. *Neurological Sciences, 26*(4), 243-254. DOI: 10.1007/s10072-005-0467-9.

Fraser, K.C., Fors, K.L., Eckerström, M., Öhman, F., & Kokkinakis, D. (2019). Predicting MCI status from multimodal language data using cascaded classifiers. *Frontiers in Aging Neuroscience, 11,* 205. DOI: 10.3389/fnagi.2019.00205.

Garrard, P., Maloney, L.M., Hodges, J.R., & Patterson, K. (2005). The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain, 128*(2), 250-260. DOI: 10.1093/brain/awh341.

Gathercole, S.E., Willis, C.S., Baddeley, A.D., et al. The children's test of Nonword Repetition: a test of phonological working memory. In: Gathercole, S.E., McCarthy, R.A., (eds.) (1994). *Memory Tests and Techniques.* Lawrence Erlbaum Associates, Hove, 103-127. DOI: 10.1080/09658219408258940.

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R.C., Richie, K., & Broich, K. (2006). Mild cognitive impairment. *Lancet, 367*(9518), 1262-1270. DOI: 10.1016/S0140-6736(06)68542-5.

Gauthier, S., Rosa-Neto, P., Morais, J.A., & Webster, C. (2021). *World Alzheimer Report 2021. Journey through the diagnosis of dementia.* London, England: Alzheimer's Disease International. https://www.alzint.org/u/World-Alzheimer-Report-2021.pdf. Last accessed: July 25, 2022.

Gayraud, F., Lee, H.R., & Barkat-Defradas, M. (2011). Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clinical Linguistics and Phonetics, 25*(3), 198-209. DOI: 10.3109/02699206.2010.521612.

GBD 2019 Dementia Forecasting Collaborators (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *The Lancet. Public Health, 7*(2), e105:e125. DOI: 10.1016/S2468-2667(21)00249-8.

Geijselaers, S.L.C., Sep, S.J.S., Stehouwer, C.D.A., & Biessels, G.J. (2015). Glucose regulation, cognition, and brain MRI in type 2 diabetes: a systematic review. *The Lancet Diabetes & Endocrinology, 3*(1), 75-89. DOI: 10.1016/S2213-8587(14)70148-2.

Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS),* 315-323.

Gonzalez-Moreira, E., Torres-Boza, D., Kairuz, H.A., Ferrer, C., Garcia-Zamora, M., Espinoza-Cuadros, F., & Hernandez-Gómez, L.A. (2015). Automatic prosodic analysis to identify mild dementia. *BioMed Research International*, 916356. DOI: 10.1155/2015/916356.

Gosztolya, G., Tóth, L., Grósz, T., Vincze, V., Hoffmann, I., Szatlóczki, G., Pákáski, M., & Kálmán, J. (2016). Detecting mild cognitive impairment from spontaneous speech by correlation-based phonetic feature selection. *Proceedings of Interspeech, San Francisco, USA,* 107-111. DOI: 10.21437/Interspeech.2016-384.

Gosztolya, G., Vincze, V., Tóth, L., Pákáski, M., Kálmán, J., & Hoffmann, I. (2019). Identifying mild cognitive impairment and mild Alzheimer's disease based on spontaneous

speech using ASR and linguistic features. *Computer Speech & Language, 53,* 181-197. DOI: 10.1016/j.csl.2018.07.007.

Gosztolya, G., Balogh, R., Imre, N., Egas-López, J.V., Hoffmann, I., Vincze, V., Tóth, L., Devanand, D.P., Pákáski, M., & Kálmán, J. (2021). Cross-lingual detection of mild cognitive impairment based on temporal parameters of spontaneous speech. *Computer, Speech & Language, 69*(1), 101215. DOI: 10.1016/j.csl.2021.101215.

Guo, Z., Ling, Z., & Li, Y. (2019). Detecting Alzheimer's disease from continuous speech using language models. *Journal of Alzheimer's Disease, 70*(4), 1163-1174. DOI: 10.3233/JAD-190452.

Hazzard, W.R., & Halter, J.B. (2009). *Hazzard's Geriatric Medicine and Gerontology, 6th edition.* New York, NY: McGraw-Hill.

Hernández-Domínguez, L., Ratté, S., Sierra-Martínez, G., & Roche-Bergua, A. (2018). Computer-based evaluation of Alzheimer's disease and mild cognitive impairment patients during a picture desctiption task. *Alzheimer's and Dementia (Amsterdam), 10,* 260-268. DOI: 10.1016/j.dadm.2018.02.004.

Hinton, G., Deng, L., Yu, D., Dahl, G.E., Mohamed, A.-R., Jaitly, N., Senior, A., Vanhoucke, V., Patrick, N., Sainath, T.N., & Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Processing Magazine, 29*(6), 82-97. DOI: 10.1109/MSP.2012.2205597.

Hlavnička, J., Čmejla, R., Tykalová, T., Šonka, K., Růžička, E., & Rusz, J. (2017). Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Scientific Reports,* 7(1):1-12. DOI: 10.1038/s41598-017-00047-5.

Hoffmann, I., Nemeth, D., Dye, C.D., Pákáski, M., Irinyi, T., & Kálmán, J. (2010). Temporal parameters of spontaneous speech in Alzheimer's disease. *International Journal of Speech-Language Pathology, 12*(1), 29-34. DOI: 10.3109/17549500903137256.

Hungarian Central Statistical Office (2022). *Resident population by age group, county and region* [Data base]. Retrieved from: https://www.ksh.hu/stadat_files/nep/en/nep0035.html. Last accessed: July 25, 2022.

Jarrold, W., Peintner, B., Wilkins, D., Vergryi, D., Richey, C., Gorno-Tempini, M.L., & Ogar, J. (2014). Aided diagnosis of dementia type through computer-based analysis of spontaneous

speech. *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology (CLPsych)*, 27–37. DOI: 10.3115/v1/W14-3204.

Jongsiriyanyong, S., & Limpawattana, P. (2018). Mild cognitive impairment in clinical practice: a review article. *American Journal of Alzheimer's Disease & Other Dementias, 33*(8), 500-507. DOI: 10.1177/1533317518791401.

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*, 1-14. DOI: 10.2174/1567205019666220418155130.

Kandimalla, R., Thirumala, V., & Reddy, P.H. (2017). Is Alzheimer's disease a type 3 diabetes? A critical appraisal. *Biochimica et Biophysica Acta. Molecular Basis of Disease, 1863*(5), 1078-1089. DOI: 10.1016/j.bbadis.2016.08.018.

Kato, S., Endo, H., Homma, A., Sakuma, T., & Watanabe, K. (2013). Early detection of cognitive impairment in the elderly based on Bayesian mining using speech prosody and cerebral blood flow activation. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 5813-5816. DOI: 10.1109/EMBC.2013.6610873.

Khodabakhsh, A., & Demiroglu, C. (2015). Analysis of speech-based measures for detecting and monitoring Alzheimer's disease. *Methods in Molecular Biology, 1246*, 159-173. DOI: 10.1007/978-1-4939-1985-7_11.

Killin, L.O.J., Starr, J.M., Shiue, I.J., & Russ, T.C. (2016). Environmental risk factors for dementia: a systematic review. *BMC Geriatrics, 16*(1), 175. DOI: 10.1186/s12877-016-0342-y.

Kobayashi, M., Kosugi, A., Takagi, H., Nemoto, M., Nemoto, K., Arai, T., & Yamada, Y. (2019). Effects of age-related cognitive decline on elderly user interactions with voice-based dialogue systems. *IFIP Conference on Human-Computer Interaction*, 53-74. DOI: 10.1007/978-3-030-29390-1_4.

König, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., Manera, V., Verhey, F., Aalten, P., Robert, P.H., & David, R. (2015). Automatic speech analysis for the

assessment of patients with predementia and Alzheimer's disease. *Alzheimer's & Dementia (Amsterdam), 1*(1), 112–124. DOI: 10.1016/j.dadm.2014.11.012.

König, A., Satt, A., Sorin, A., Hoory, R., Derreumaux, A., David, R., & Robert, P.H. (2018). Use of speech analyses within a mobile application for the assessment of cognitive impairment in elderly people. *Current Alzheimer Research, 15*(2), 120-129. DOI: 10.2174/1567205014666170829111942.

Kumari, M., & Marmot, M. (2005). Diabetes and cognitive function in a middle-aged cohort: findings from the Whitehall II study. *Neurology, 65*(10), 1597-1603. DOI: 10.1212/01.wnl.0000184521.80820.e4.

Laske, C., Sohrabi, H.R., Frost, S.M., López-de-Ipiña, K., Garrard, P., Buscema, M., Dauwels, J., Soekadar, S.R., Mueller, S., Linnemann, C., Bridenbaugh, S.A., Kanagasingam, Y., Martins, R.N., & O'Bryant, S.E.O. (2015). Innovative diagnostic tools for early detection of Alzheimer's disease. Alzheimer's & Dementia, 11(5), 561-578. DOI: 10.1016/j.jalz.2014.06.004.

Lee, H., Gayraud, F., Hirsh, F., & Barkat-Defradas, M. (2011). Speech dysfluencies in normal and pathological aging: a comparison between Alzheimer patients and healthy elderly subjects. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS),* 1174–1177.

Leszek, J., Trypka, E., Tarasov, V.V., Ashraf, G.M., & Aliev, G. (2017). Type 3 diabetes mellitus: a novel implication of Alzheimers disease. *Current Topics in Medicinal Chemistry, 17*(12), 1331-1335. DOI: 10.1016/j.euroneuro.2014.06.008.

López-de-Ipiña, K., Alonso, J.-B., Travieso, C.M., Solé-Casals, J., Egiraun, H., Faundez-Zanuy, M., Ezeiza, A., Barroso, N., Ecay-Torres, M., Martinez-Lage, P., & de Lizardui, U.M. (2013). On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors (Basel), 13*(5), 6730-6745. DOI: 10.3390/s130506730.

Low, L.-F., & Purwaningrum, F. (2020). Negative stereotypes, fear and social distance: a systematic review of depictions of dementia in popular culture in the context of stigma. *BMC Geriatrics, 20*(1), 477. DOI: 10.1186/s12877-020-01754-x.

Luck, T., Luppa, M., Briel, S., & Riedel-Heller, S.G. (2010). Incidence of mild cognitive impairment: a systematic review. *Dementia and Geriatric Cognitive Disorders, 29*(2), 164-175. DOI: 10.1159/000272424.

Luz, S., de la Fuente, S., & Albert, P. (2018). A method for analysis of patient speech in dialogue for dementia detection. *arXiv,* 1811.09919.

Manos, P., & Wu, R. (1994). The ten-point clock test: A quick screen and grading method for cognitive impairment in medical and surgical patients. *International Journal of Psychiatry in Medicine, 24*(3), 229-244. DOI: 10.2190/5A0F-936P-VG8N-0F5R.

Martínez-Nicolás, I., Llorente, T.E., Martínez-Sánchez, F., & Meilán, J.J.G. (2021). Ten years of research on automatic voice and speech analysis of people with Alzheimer's disease and mild cognitive impairment: a systematic review article. *Frontiers in Psychology,* 620251. DOI: 10.3389/fpsyg.2021.620251.

Martínez-Sánchez, F., Meilán, J.J.G., & García-Sevilla, J. (2013). Oral reading fluency analysis in patients with Alzheimer disease and asymptomatic control subjects. *Neurología,* 28(6), 325-331. DOI: 10.1016/j.nrl.2012.07.012.

Martínez-Sánchez, F., Meilán, J.J.G., Carro, J., & Ivanova, O. (2018). A prototype for the voice analysis diagnosis of Alzheimer's disease. *Journal of Alzheimer's Disease, 64*(2), 473-481. DOI: 10.3233/JAD-180037.

Mayeda, E.R., Haan, M.N., Neuhaus, J., Yaffe, K., Knopman, D.S., Sharrett, A.R., Griswold, M.E., & Mosley, T.H. (2014). Type 2 diabetes and cognitive decline over 14 years in middle-aged African Americans and whites: The ARIC Brain MRI Study. *Neuroepidemiology, 43*(3-4), 220-227. DOI: 10.1159/000366506.

McCrimmon, R.J., Ryan, C.M., & Frier, B.M. (2012). Diabetes and cognitive dysfunction. *Lancet, 379*(9833):2291-2299. DOI: 10.1016/S0140-6736(12)60360-2.

McCullagh, C.D., Craig, D., McIlroy, S.P., & Passmore, A.P. (2018). Risk factors for dementia. *Advances in Psychiatric Treatment, 7*(1), 24-31.

McMenemey, W.H. (1963). Alzheimer's disease: problems concerning its concept and nature. *Acta Neurologica Scandinavica, 39*(4), 369-380. DOI: 10.1111/j.1600-0404.1963.tb05337.x.

Meilán, J.J.G., Martínez-Sánchez, F., Carrol, J., Sánchez, J.A., & Pérez, E. (2012). Acoustic markers associated with impairment in language processing in Alzheimer's disease. *The Spanish Journal of Psychology, 15*(2), 487-494. DOI: 10.5209/rev_sjop.2012.v15.n2.38859.

Meilán, J.J.G., Martínez-Sánchez, F., Carro, J., López, D.E., Millian-Morell, L., & Arana, J.M. (2014). Speech in Alzheimer's disease: can temporal and acoustic parameters discriminate dementia? *Dementia and Geriatric Cognitive Disorders, 37*(5-6), 327-334. DOI: 10.1159/000356726.

Meilán, J.J.G., Martínez-Sánchez, F., Martínez-Nicolás, I., Llorente, T.E., & Carro, J. (2020). Changes in the rhythm of speech difference between people with nondegenerative mild cognitive impairment and with preclinical dementia. *Behavioural Neurology,* 4683573, 1-10. DOI: 10.1155/2020/4683573.

Mirzaei, S., El Yacoubi, M., Garcia-Salicetti, S., Boudy, J., Kahindo, C., Cristancho-Lacroix, V., Kerhervé, H., & Rigaud, A.-S. (2018). Two-stage feature selection of voice parameters for early Alzheimer's disease prediction. *IRBM 39*(6), 430-435. DOI: 10.1016/j.irbm.2018.10.016.

Morelli, N.L., Cachioni, M., Lopes, A., Batistoni, S.S.T., Falcão, D.V.D.S., Neri, A.L., & Yassuda, M.S. (2017). Verbal fluency in elderly with and without hypertension and diabetes from the FIBRA study in Ermelino Matarazzo. *Dementia & Neuropsychologia, 11*(4), 413-418. DOI: 10.1590/1980-57642016dn11-040011.

Mortensen, L., Meyer, A.S., & Humphreys, G.W. (2006). Age-related effects on speech production: A review. *Language and Cognitive Processes, 21*(1-3), 238-290. DOI: 10.1080/01690960444000278.

Mueller, K.D., Hermann, B., Mecollari, J., & Turkstra, L.S. (2018). Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. Journal of Clinical and Experimental Neuropsychology, 40(9), 917-939. DOI: 10.1080/13803395.2018.1446513.

Nasrolahzadeh, M., Mohammadpoory, Z., & Haddadnia, J. (2018). Higher-order spectral analysis of spontaneous speech signals in Alzheimer's disease. *Cognitive Neurodynamics, 12*(6), 583-596. DOI: 10.1007/s11571-018-9499-8.

Neuberger, T., Gyarmathy, D., Gráczi, T.E., Horváth, V., Gósy, M., & Beke, A. (2014). Development of a large spontaneous speech database of agglutinative Hungarian language. *Text, Speech, and Dialogue*, 424-431. DOI: 10.1007/978-3-319-10816-2_51.

Ninomiya, T. (2019). Epidemiological evidence of the relationship between diabetes and dementia. *Advances in Experimental Medicine and Biology, 1128*, 13-25. DOI: 10.1007/978-981-13-3540-2_2.

Pakhomov, S., Chacon, D., Wicklund, M., & Gundel, J. (2011). Computerized assessment of syntactic complexity in Alzheimer's disease: a case study of Iris Murdoch's writing. *Behavior Research Methods, 43*(1), 136-144. DOI: 10.3758/s13428-010-0037-9.

Palta, P., Carlson, M.C., Crum, R.M., Colantuoni, E., Sharrett, A.R., Yasar, S., Nahin, R.L., DeKosky, S.T., Snitz, B., Lopez, O., Williamson, J.D., Furberg, C.D., Rapp, S.R., & Golden, S.H. (2017). Diabetes and cognitive decline in older adults: the Ginkgo Evaluation of Memory Study. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 73*(1), 123-130. DOI: 10.1093/gerona/glx076.

Patterson, C., Feightner, J., Garcia, A., & MacKnight, C. (2007). General risk factors for dementia: a systematic evidence review. *Alzheimer's & Dementia, 3*(4), 341-347. DOI: 10.1016/j.jalz.2007.07.001.

Petersen, R.C., Smith, G.E., Waring, S.C., Ivnik, R.J., Tangalos, E.G., & Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Archives of Neurology, 56*(3), 303–308. DOI: 10.1001/archneur.56.3.303.

Petersen, R.C., Doody, R., Kurz, A., Mohs, R.C., Morris, J.C., Rabins, P.V., Ritchie, K., Rossor, M., Thal, L., & Winblad, B. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology, 58*(12), 1985–1992. DOI: 10.1001/archneur.58.12.1985.

Petersen, R,C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: A concept in evolution. Journal of Internal Medicine, 275(3), 214-228. DOI: 10.1111/joim.12190.

Petersen, R.C., Lopez, O., Armstrong, M.J., Getchius, T.S.D., Ganguli, M., Gloss, D., Gronseth, G.S., Marson, D., Pringsheim, T., Day, G.S., Sager, M., Stevens, J., & Rae-Grant, A. (2018). Practice guideline update summary: Mild cognitive impairment. *Neurology, 90*(3), 126-35. DOI: 10.1212/WNL.0000000000004826.

Petti, U., Baker, S., & Korhonen, A. (2020). A systematic literature review of automatic Alzheimer's disease detection from speech and language. *Journal of the American Medical Informatics Association, 27*(11), 1784-1797. DOI: 10.1093/jamia/ocaa174.

Pistono, A., Jucla, M., Barbeau, E.J., Saint-Aubert, L., Lemesle, B., Calvet, B., Köpke, B., Puel, M., & Pariente, J. (2016). Pauses during autobiographical discourse reflect episodic memory processes in early Alzheimer's disease. *Journal of Alzheimer's Disease, 50*(3), 687-698. DOI: 10.3233/JAD-150408.

Prince, M., Wimo, A., Guerchet, M., Ali, G.-C., Wu, Y.-T., Prina, M., & Alzheimer's Disease International. (2015). *World Alzheimer Report 2015. The global impact of dementia: an analysis of prevalence, incidence, cost and trends.* London, England: Alzheimer's Disease International. https://www.alzint.org/u/WorldAlzheimerReport2015.pdf. Last accessed: July 25, 2022.

Rabinovici, G.D., Carrillo, M.C., Forman, M., DeSanti, S., Miller, D.S., Kozauer, N., Petersen, R.C., Randolph, C., Knopman, D.S., Smith, E.E., Isaac, M., Mattsson, N., Bain, L.J., Hendrix, J.A., & Sims, J.R. (2017). Multiple comorbid neuropathologies in the setting of Alzheimer's disease neuropathology and implications for drug development. *Alzheimer's & Dementia: Translational Research & Clinical Interventions, 3*(1), 83-91. DOI: 10.1016/j.trci.2016.09.002.

Rawlings, A.M., Sharrett, A.R., Schneider, A.L.C., Coresh, J., Albert, M., Couper, D., Griswold, M., Gottesman, R.F., Wagenknecht, L.E., Windham, B.G., & Selvin, E. (2014). Diabetes in midlife and cognitive change over 20 years: the Atherosclerosis Risk in Communities Neurocognitive Study. *Annals of Internal Medicine, 161*(11), 785-793. DOI: 10.7326/M14-0737.

Riley, K.P., Snowdon, D.A., Desrosiers, M.F., & Markesbery, W.R. (2005). Early life linguistic ability, late life cognitive function, and neuropathology: findings from the Nun Study. *Neurobiology of Aging, 26*(3), 341-347. DOI: 10.1016/j.neurobiolaging.2004.06.019.

Roark, B., Mitchell, M., Hosom, J.P., Hollingshead, K., & Kaye, J. (2011). Spoken language derived measures for detecting mild cognitive impairment. *IEEE Transactions on Audio, Speech, and Language Processing, 19*(7), 2081–2090. DOI: 10.1109/TASL.2011.2112351.

Roca, M., Torralva, T., Gleichgerrcht, E., Woolgar, A., Thompson, R., Duncan, J., & Manes, F. (2011). The role of Area 10 (BA10) in human multitasking and in social cognition: a

lesion study. *Neuropsychologia, 49*(13), 3525-3531 DOI: 10.1016/j.neuropsychologia.2011.09.003.

Rosen, W.G., Mohs, R.C., & Davis, K.L. (1984). A new rating scale for Alzheimer's disease. *The American Journal of Psychiatry, 141*(11), 1356–1364. DOI: 10.1176/ajp.141.11.1356.

Rousseau, A., Delglise, P., & Estve, Y. (2012). TED-LIUM: an automatic speech recognition dedicated corpus. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 125–129.

Sadanand, S., Balachandar, R., & Bharath, S. (2016). Memory and executive functions in persons with type 2 diabetes: a meta-analysis. *Diabetes/Metabolism Research and Reviews, 32*(2), 132-142. DOI: 10.1002/dmrr.2664.

Saeedi, P., Petersohn, I., Salpea, P., Malanda, B., Karuranga, S., Unwin, N., Colagiuri, S., Guariguata, L., Motala, A.A., Ogurtsova, K., Shaw, J.E., Bright, D., Williams, R., & IDF Diabetes Atlas Committee (2019). Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Research and Clinical Practice, 157*:107843. DOI: 10.1016/j.diabres.2019.107843.

Sajjadi, S.A., Patterson, K., Tomek, M., & Nestor, P.J. (2012). Abnormalities of connected speech in semantic dementia vs Alzheimer's disease. *Aphasiology, 26*(6), 847-866. DOI: 10.1080/02687038.2012.654933.

Sarazin, M., Chauviré, V., Gerardin, E., Colliot, O., Kinkingnéhun, S., Cruz de Sousa, L., Hugonot-Diener, L., Garnero, L., Lehéricy, S., Chupin, M., & Dubois, B. (2010). The amnestic syndrome of hippocampal type in Alzheimer's disease: an MRI study. *Journal of Alzheimer's Disease, 22*(1), 285-294. DOI: 10.3233/JAD-2010-091150.

Satt, A., Sorin, A., Toledo-Ronen, O., Barkan, O., Kompatsiaris, I., Kokonozi, A., & Tsolaki, M. (2013). Evaluation of speech-based protocol for detection of early-stage dementia. *Proceedings of Interspeech, Lyon, France*, 1692-1696. DOI: 10.21437/Interspeech.2013-32.

Sheikh, J.I., & Yesavage, J.A. (1986). Geriatric Depression Scale (GDS) – Recent evidence and development of a shorter version. *Clinical Gerontologist, 5*(1-2), 165-173. DOI: 10.1300/J018v05n01_09.

Simons, J.S., Gilbert, S.J., Owen, A.M., Fletcher, P.C., & Burgess, P.W. (2005). Distinct roles for lateral and medial anterior prefrontal cortex in contextual recollection. *Journal of Neurophysiology, 94*(1), 813-820. DOI: 10.1152/jn.01200.2004.

Singh, S., Bucks, R.S., & Cuerden, J.M. (2001). Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. *Aphasiology, 15*(6), 571-583. DOI: 10.1080/02687040143000041.

Sluis, R.A., Angus, D., Wiles, J., Back, A., Gibson, T.A., Liddle, J., Worthy, P., Copland, D., & Angwin, A.J. (2020). An automated approach to examining pausing in the speech of people with dementia. *American Journal of Alzheimer's Disease and Other Dementias, 35*, 1533317520939773. DOI: 10.1177/1533317520939773.

Snowdon, D.A., Kemper, S.J., Mortimer, J.A., Greiner, L.H., Wekstein, D.R., & Markesbery, W.R. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life. Findings from the Nun Study. *JAMA, 275*(7), 528-532.

Solomon, A., Mangialasche, F., Richard, E., Andrieu, S., Bennett, D.A., Breteler, M., Fratiglioni, L., Hooshmand, B., Khachaturian, A.S., Schneider, L.S., Skoog, I., & Kivipelto, M. (2014). Advances in the prevention of Alzheimer's disease and dementia. *Journal of Internal Medicine, 275*(3), 229-250. DOI: 10.1111/joim.12178.

Stranahan, A.M., Arumugam, T.V., Cutler, R.G., Lee, K., Egan, J.M., & Mattson, M.P. (2008). Diabetes impairs hippocampal function through glucocorticoid-mediated effects on new and mature neurons. *Nature Neuroscience, 11*(3):309-317. DOI: 10.1038/nn2055.

Szatloczki, G., Hoffmann, I., Vincze, V., Kalman, J., & Pakaski, M. (2015). Speaking in Alzheimer's disease, is that an early sign? Importance of changes in language abilities in Alzheimer's disease. *Frontiers in Aging Neuroscience, 7*(195), 1-7. DOI: 10.3389/fnagi.2015.00195.

Taler, V., & Phillips, N.A. (2008). Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology, 30*(5), 501–556. DOI: 10.1080/13803390701550128.

Tanaka, H., Adachi, H., Ukita, N., Ikeda, M., Kazui, H., Kudo, T., & Nakamura, S. (2017). Detecting dementia through interactive computer avatars. *IEEE Journal of Translational Engineering in Health and Medicine, 5*, 2200111. DOI: 10.1109/JTEHM.2017.2752152.

Themistocleous, C., Eckerström, M., & Kokkinakis, D. (2020). Voice quality and speech fluency distinguish individuals with mild cognitive impairment from healthy controls. *PLoS One, 15*(7), e0236009. DOI: 10.1371/journal.pone.0236009.

Tóth, L. (2015). Phone recognition with hierarchical convolutional deep maxout networks. *EURASIP Journal on Audio, Speech, and Music Processing, 25*, 1-3. DOI: 10.1186/s13636-015-0068-3.

Tóth, L., Gosztolya, G., Vincze, V., Hoffmann, I., Szatlóczki, G., Biró, E., Zsura, F., Pákáski, M., & Kálmán, J. (2015). Automatic detection of mild cognitive impairment from spontaneous speech using ASR. *Proceedings of Interspeech, Dresden, Germany*, 2694-2698. DOI: 10.21437/Interspeech.2015-568.

Tóth, L., Hoffmann, I., Gosztolya, G., Vincze, V., Szatlóczki, G., Bánréti, Z., Pákáski, M., & Kálmán, J. (2018a). A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech. *Current Alzheimer Research, 15*(2), 130-138. DOI: 10.2174/1567205014666171121114930.

Tóth, G., Szabó, D., Sándor, G.L., Nagy, Z.Z., Karadeniz, S., Limburg, H., & Németh, J. (2018b). Diabetes and blindness in people with diabetes in Hungary. *European Journal of Ophthalmology, 29*(2), 141-147. DOI: 10.1177/1120672118811738.

Trejo-Lopez, J.A., Yachnis, A.T., & Prokop, S. (2022). Neuropathology of Alzheimer's disease. *Neurotherapeutics, 19*(1), 173-185. DOI: 10.1007/s13311-021-01146-y.

Tröger, J., Linz, N., König, A., Robert, P., Alexandersson, J., Peter, J., & Kray, J. (2019). Exploitation vs. exploration – computational temporal and semantic analysis explains semantic verbal fluency impairment in Alzheimer's disease. *Neuropsychologia, 131*, 53-61 DOI: 10.1016/j.neuropsychologia.2019.05.007.

Vatanabe, I.P., Manzine, P., & Cominetti, M. R. (2020). Historic concepts of dementia and Alzheimer's disease: from ancient times to the present. *Revue Neurologique, 176*(3), 140-147. DOI: 10.1016/j.neurol.2019.03.004.

Venneri, A. (2005). Impoverishment of spontaneous language and the prediction of Alzheimer's disease. *Brain, 128*(4), E27. DOI: 10.1093/brain/awh419.

Vigo, I., Coelho, L., & Reis, S. (2022). Speech- and language-based classification of Alzheimer's disease: a systematic review. *Bioengineering (Basel), 9*(1), 27. DOI: 10.3390/bioengineering9010027.

Vincze, V., Gosztolya, G., Tóth, L., Hoffmann, I., Szatlóczki, G., Bánréti, Z., Pákáski, M., & Kálmán, J. (2016). Detecting mild cognitive impairment by exploiting linguistic information from transcripts. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, 181-187. DOI: 10.18653/v1/P16-2030.

Vincze, V., Szatlóczki, G., Tóth, L., Gosztolya, G., Pákáski, M., Hoffmann, I., & Kálmán, J. (2020). Telltale silence: temporal speech parameters discriminate between prodromal dementia and mild Alzheimer's disease. *Clinical Linguistics & Phonetics*, 1-16. DOI: 10.1080/02699206.2020.1827043.

Ward, A., Tardiff, S., Dye, C., & Arrighi, H.M. (2013). Rate of conversion from prodromal Alzheimer's disease to Alzheimer's dementia: A systematic review of the literature. *Dementia and Geriatric Cognitive Disorders Extra, 3*(1), 320-332. DOI: 10.1159/000354370.

Wechsler, D. (1981). *The Wechsler Adult Intelligence Scale – Revised.* The Psychological Corporation: New York.

Wennberg, A.M.V., Gottesman, R.F., Kaufmann, C.N., Albert, M.S., Chen-Edinboro, L.P., Rebok, G.W., Kasper, J.D., & Spira, A.P. (2014). Diabetes and cognitive outcomes in a nationally representative sample: the National Health and Aging Trends Study. *International Psychogeriatrics, 26(*10), 1729-1735. DOI: 10.1017/S1041610214001380.

Wimo, A., Jönsson, L., Bond, J., Prince, M., Winblad, B., & Alzheimer Disease International (2013). The worldwide economic impact of dementia 2010. *Alzheimer's & Dementia, 9*(1), 1-11. DOI: 10.1016/j.jalz.2012.11.006.

World Health Organization (2019). *International Statistical Classification of Diseases and Related Health Problems* (11th ed.). Retrieved from: https://icd.who.int/en.

World Health Organization (2021a). *Dementia Fact Sheet.* https://www.who.int/news-room/fact-sheets/detail/dementia. Last accessed: July 25, 2022.

World Health Organization (2021b). *Diabetes Fact Sheet.* https://www.who.int/news-room/fact-sheets/detail/diabetes. Last accessed: July 25, 2022.

Wysokinski, A., Zboralski, K., Orzechowska, A., Galecki, P., Florkowski, A., & Talarowska, M. (2010). Normalization of the Verbal Fluency Test on the basis of results for healthy subjects, patients with schizophrenia, patients with organic lesions of the chronic nervous

system and patients with type 1 and 2 diabetes. *Archives of Medical Science, 6*(3), 438-446. DOI: 10.5114/aoms.2010.14268.

Yamada, Y., Shinkawa, K., Kobayashi, M., Nishimura, M., Nemoto, M., Tsukada, E., Ota, M., Nemoto, K., & Arai, T. (2021). Tablet-based automatic assessment for early detection of Alzheimer's disease using speech responses to daily life questions. *Frontiers in Digital Health, 3*, 1-9. DOI: 10.3389/fdgth.2021.653904.

Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., & Leirer, V.O. (1983). Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatric Research, 17*(1), 37-49. DOI: 10.1016/0022-3956(82)90033-4.

Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., & Woodland, P. (2006). *The HTK Book.* Cambridge University: Cambridge, UK.

I.

Send Orders for Reprints to reprints@benthamscience.net

Current Alzheimer Research, XXXX, *XX*, 1-14

RESEARCH ARTICLE

Temporal Speech Parameters Detect Mild Cognitive Impairment in Different Languages: Validation and Comparison of the Speech-GAP Test® in English and Hungarian

János Kálmán[1,*], Davangere P. Devanand[2,3], Gábor Gosztolya[4], Réka Balogh[1], Nóra Imre[1], László Tóth[5], Ildikó Hoffmann[6,7], Ildikó Kovács[1], Veronika Vincze[4] and Magdolna Pákáski[1]

[1]Department of Psychiatry, Albert Szent-Györgyi Medical School, University of Szeged, Szeged, Hungary; [2]Department of Psychiatry, Columbia University Medical Center, New York, NY, United States; [3]Division of Geriatric Psychiatry, New York State Psychiatric Institute, New York, NY, United States; [4]MTA-SZTE Research Group on Artificial Intelligence, Faculty of Science and Informatics, University of Szeged, Szeged, Hungary; [5]Institute of Informatics, Faculty of Science and Informatics, University of Szeged, Szeged, Hungary; [6]Department of Hungarian Linguistics, Faculty of Humanities and Social Sciences, University of Szeged, Szeged, Hungary; [7]Hungarian Research Centre for Linguistics, Eötvös Loránd Research Network, Budapest, Hungary

ARTICLE HISTORY

Received: December 20, 2021
Revised: February 08, 2022
Accepted: February 17, 2022

DOI:
10.2174/1567205019666220418155130

Abstract: ***Background*:** The development of automatic speech recognition (ASR) technology allows the analysis of temporal (time-based) speech parameters characteristic of mild cognitive impairment (MCI). However, no information has been available on whether the analysis of spontaneous speech can be used with the same efficiency in different language environments.

***Objective*:** The main goal of this international pilot study is to address the question of whether the Speech-Gap Test® (S-GAP Test®), previously tested in the Hungarian language, is appropriate for and applicable to the recognition of MCI in other languages such as English.

***Methods*:** After an initial screening of 88 individuals, English-speaking (n = 33) and Hungarian-speaking (n = 33) participants were classified as having MCI or as healthy controls (HC) based on Petersen's criteria. The speech of each participant was recorded *via* a spontaneous speech task. Fifteen temporal parameters were determined and calculated through ASR.

***Results*:** Seven temporal parameters in the English-speaking sample and 5 in the Hungarian-speaking sample showed significant differences between the MCI and the HC groups. Receiver operating characteristics (ROC) analysis clearly distinguished the English-speaking MCI cases from the HC group based on speech tempo and articulation tempo with 100% sensitivity, and on three more temporal parameters with high sensitivity (85.7%). In the Hungarian-speaking sample, the ROC analysis showed similar sensitivity rates (92.3%).

***Conclusion*:** The results of this study in different native-speaking populations suggest that changes in acoustic parameters detected by the S-GAP Test® might be present across different languages.

Keywords: Mild cognitive impairment, neurocognitive disorder, language, speech analysis, temporal parameters, early recognition.

1. INTRODUCTION

Language changes occur in various types of neurocognitive disorders and are sensitive indicators of cortical dysfunction [1, 2]. The characteristic disruption in the language domain has been identified not only in different stages of dementia [3, 4], but also in its prodromal stage, mild cognitive impairment (MCI) [5]. However, recognition of the first clinical manifestations is still challenging since patients often do not recognize or minimize their deficits. In the early diagnostic procedure, there is an increasing need for non-invasive and cost-effective tools to identify individuals with minor neurocognitive disorders [4]. Since subtle changes in language and communication abilities may be apparent in the early course of such disorders [6], the detection of linguistic impairment could be a viable screening option [7-9]. Recordability of spoken language gives an opportunity to easily collect speech recordings, as biological samples. The purpose of our research group was to develop a new mobile application that would be capable of recording the examined person's telephone conversation and then analyzing the

*Address correspondence to this author at the Department of Psychiatry, Albert Szent-Györgyi Medical School, University of Szeged, Szeged, Hungary; Korányi fasor 8-10., H-6720, Szeged, Hungary; Tel: +36 62341600; E-mail: kalman.janos@med.u-szeged.hu

1567-2050/XX $65.00+.00

acoustic properties of spontaneous speech. Using this information technology (IT) technique, an individual can be examined through everyday activity, namely, a telephone conversation which is an ecologically valid way of assessment, decreases the time spent on neuropsychological tests, and eliminates test-induced anxiety for the user.

The first interest of our research team (beginning the now 10-year long research project on exploring the association between language function and cognition) was to identify speech parameters that might distinguish Hungarian patients with mild Alzheimer's disease (AD) from healthy controls (HC). Significant differences between the mild AD and the HC groups regarding speech tempo and hesitation ratio were first published by our research team [10]. However, in this early study, the transcription and annotation of speech signals were performed manually using the Praat software tool [11]. As the manual calculation of acoustic biomarkers is extremely time-consuming, its applicability in recognizing mild stages of cognitive deficits in clinical routine is rather limited.

However, the deterioration of acoustic language parameters can also be examined by implementing automatic speech recognition (ASR) techniques. ASR is a relatively simple and reliable method that has the potential to analyze large language datasets rapidly using machine-learning methods. Based on this technology, our research team developed the Speech-Gap Test®(S-GAP Test®) which identifies temporal (time-based) speech parameters using the extracted phonetic-level segmentation produced by ASR. In earlier studies applying the S-GAP Test®, we were able to distinguish MCI patients from HC subjects [12-17] based on several temporal parameters which demonstrated that the proposed acoustic characteristics indeed carry clinically relevant information in spontaneous speech [13].

Among the most informative temporal parameters, articulation and speech tempo, number and length of silent/filled pauses, and length of utterance were measured. Articulation/speech tempo is the number of phonemes per second during speech excluding/including hesitations, respectively. Hesitation is defined as the absence of speech and has two categories: silent pauses (silences that are not attributable to articulation constraints) and filled pauses (vocalizations like 'uhm', 'er'). A novelty of our studies was the focus on both silent and filled pauses along with the measurement of separated articulation and speech tempo. As our database of MCI patients was continuously growing and machine learning techniques were also exploited, the differentiation between MCI subjects and control probands gradually became more accurate (sensitivity: 81.3%; specificity: 66.7%) [16].

It is a basic requirement for diagnostic procedures used for the detection of MCI to be internationally applicable [18]. Particularly, in the case of procedures testing linguistic functions, the question arises of whether they have similar sensitivities in different languages. A recent systematic review emphasized that the methodology of speech-based studies in different native languages is quite heterogeneous [19]. Until now, phonetical-phonological analyses of speech for the assessment of cognitive impairment have been independently performed on native speakers of languages such as Chinese [20], English [21-27], French [28-31], Greek [32], Hungarian [13, 16, 17, 33, 34], Italian [35], Japanese [36-39], Persian [40], Spanish [2, 41-44], Swedish [45, 46], or Turkish [47]. However, until our present investigation, no information has been available on how the temporal characteristics of spontaneous speech compared between MCI *vs.* HC subjects in different language environments.

The main goal of this international pilot study was to explore the S-GAP-related temporal parameters of spontaneous speech in the English language with the purpose of MCI detection, and to address the methodological question of whether the S-GAP Test®, previously tested for Hungarian speakers, is appropriate for the recognition of speech parameters indicating MCI in the English language. Comparison of speech data obtained from native English- and Hungarian-speaking populations and assessing the effectiveness of the S-GAP Test® in these two different language environments would be the first step in the international application of this MCI-screening method. An IT application based on the S-GAP Test® could be a low-cost, non-invasive, and non-stressful method that could be applied in a rapid and easy way, without personal contact, and in a large population. The need for noncontact, remote assessment has also gained special urgency in light of the current COVID-19 pandemic.

2. MATERIALS AND METHODS

2.1. Participants and Study Design

Elderly individuals were recruited in parallel at two institutions: 1) Memory Disorders Center of the Department of Psychiatry, New York State Psychiatric Institute and Columbia University (New York, NY, USA) and 2) Memory Clinic, Department of Psychiatry, University of Szeged (Szeged, Hungary).

The ethnicity of the participants was not an inclusion or exclusion criterion and differed across the two study sites. The Hungarian participants were all Caucasian, while in the English-speaking group at Columbia University the individuals were Caucasian (69.7%), African-American (24.2%), and Hispanic (6.1%).

From the two outpatient clinics, a total of 88 individuals were recruited, 66 of whom were eligible for final inclusion (Fig. **1**). Both the English-speaking (n = 33) and Hungarian-speaking (n = 33) participants were classified as either MCI or as HC. The classification was based on Petersen's criteria [48] in both languages, with the Mini-Mental State Examination (MMSE) [49] serving as a measure for objective cognitive impairment (30-28 points: HC; 27-24 points: MCI).

To get an overview of participant characteristics and eligibility data, an interview focused on demographic features and medical history, as well as a brief neuropsychological test battery was administered (including the MMSE, the Clock Drawing Test (CDT) [50] and the Geriatric Depression Scale (GDS). All individuals were screened for possible dementia using the MMSE and those with a score under 24 were not involved in further participation. Corresponding to institutional protocols, the possibility of depression was also evaluated based on the 30-items [51] or the 15-items [52] version of the GDS (GDS-30/GDS-15; for the English-speaking/Hungarian-speaking sample, respectively): patients

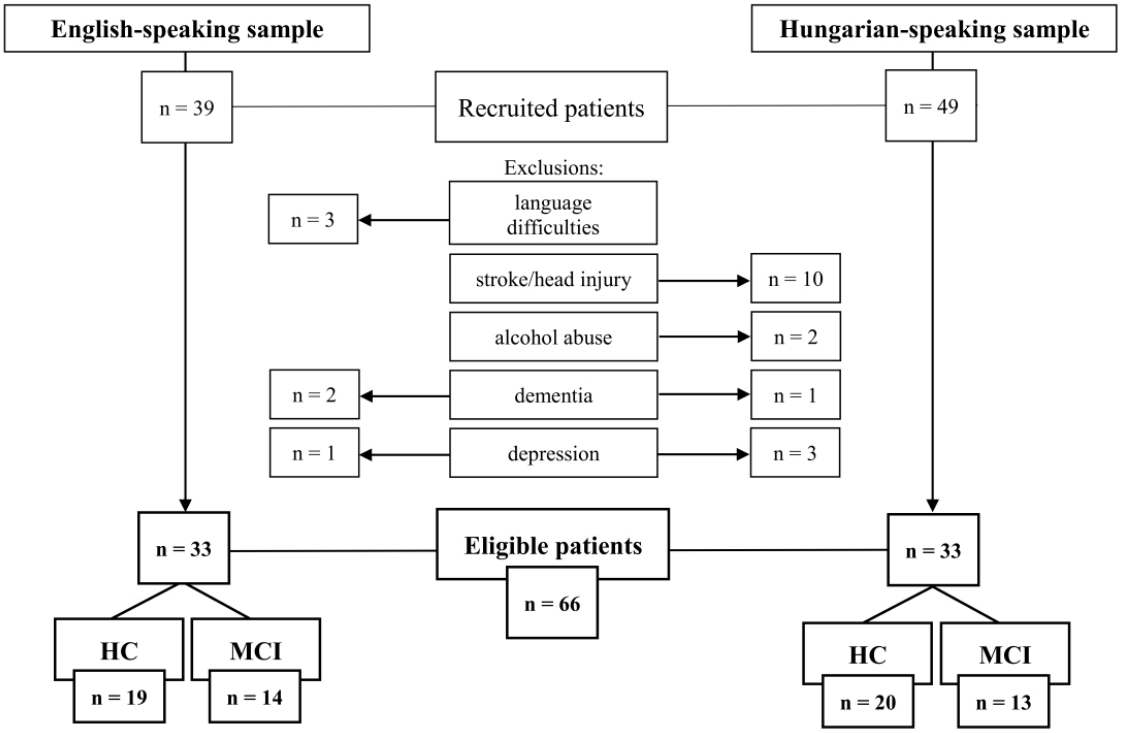

Fig. (1). Flowchart of participants' inclusion and exclusion process. **Abbreviations:** HC: healthy control; MCI: mild cognitive impairment.

scoring above 10 on GDS-30 or above 5 on GDS-15 were excluded.

Inclusion/exclusion criteria were the same at both sites. Inclusion criteria were a minimum age of 60 years, a minimum of 8 years of formal education, and English/Hungarian as native language (corresponding to the country of recruitment; bilingualism was not taken into account). Exclusion criteria included major hearing problems (*e.g.* uncorrected hearing loss), manifest speech problems (any form of aphasia), significant articulation problems (*e.g.* stutter), history of a substance use disorder, previous CT/MRI showing evidence of significant abnormality suggesting another potential etiology for MCI or dementia (*e.g.* micro- or macrohemorrhages, lacunar infarcts or single large infarct), evidence of cerebral contusion, encephalomalacia, aneurysm, vascular malformations, or clinically significant space-occupying lesions.

2.2. S-GAP Protocol and Preparation of Speech Samples

Following the clinical evaluation, speech samples were obtained from all participants. Spontaneous speech was elicited in the following way: Investigator 1, pointing to a mobile phone, informed the participant that a colleague (Investigator 2) would call from another room and provide instructions for a new task. Investigator 1 also told the participant that the conversation would be recorded and the task would only take a few minutes. Investigator 2 called the mobile phone, and after introduction, asked the participant to talk about his/her previous day. The standardized instruction was: *'Please tell me about your previous day in as much detail as you can.'* After the instruction, the investigators could not give verbal prompts, nor could they repeat the instruction; they remained silent throughout the call until the participant finished the task.

Each participants monologue was recorded by a call recorder application installed on the mobile phone device. The obtained recordings were then converted into an uncompressed PCM mono, 16-bit wav format with a sampling rate of 8,000 Hz. A professional expert linguist (I.H.) checked the quality of the recordings.

2.3. Analysis of Speech Samples

Pauses were defined as the disruption of speech for more than 30 ms (either silent segments in the case of silent-, or vocalizations in the case of filled pauses). Both silent and filled pauses were identified in each recording using ASR technology. Our ASR system was built on a modified version of the HTK tool [53], where we used the Hidden Markov model, but replaced the acoustic model with a Deep Neural Network (DNN) based one. This way, we utilized a standard HMM/DNN hybrid model, which is known to outperform traditional HMM models [54]. To realize the DNN acoustic model, we employed a custom DNN implementation [55] written in Visual C++ and utilized the CUDA library to speed up both model training and evaluation.

As acoustic features, we were using 40 raw Mel-frequency filter bank energy values along with the global log-energy, which was extended with the first and second-order derivatives ("FBANK + Δ + ΔΔ"), resulting in 123 acoustic features overall. Training and evaluation were done on a 150 ms (15 frames) wide sliding window, leading to 1,845 input neurons in the actual acoustic models. Then the acoustic model DNNs contained 5 fully connected hidden layers, each consisting of 1,024 neurons employing the

ReLU activation function [56], while they have a softmax final layer with a number of neurons equal to the phonetic units in the given language. The DNN acoustic models were trained for phoneme identification on two audio datasets consisting of spontaneous speech (as this type is expected to contain filled pauses), to match the language used by the subjects. For the speech samples in English, a subset of the TEDlium speech corpus [57] was used (100 speakers, approximately 15 hours of recording). For the Hungarian speech samples, a subset of the BEA corpus was employed (116 speakers, approximately 44 hours of recording) [58]. Before training, both corpora were down sampled to 8,000 Hz to match the sampling rate of the recordings in the study.

This ASR model was used to perform phoneme-level recognition, in which we also treated filled pauses as a special "phoneme". As language models, we employed simple phone bigrams both for English and Hungarian. This procedure produces a time-aligned phoneme sequence for each recording; that is, it supplies a hypothesis of the sequence of phones uttered, along with the starting and ending time indices. From this output, the 15 S-GAP parameters can be obtained *via* simple calculations (*e.g.* by counting the number of pauses and the total number of phones, and dividing the two values by each other; or by doing the same with the total duration of the pauses and all phones). We measured the accuracy of this workflow on a holdout set of the BEA corpus, consisting of 3 hours and 23 minutes, containing the speech of 10 subjects. Based on this, Pearson's correlation values of the speech tempo attributes calculated by our workflow and those derived from the transcripts were 0.857, while for articulation tempo this value was 0.866, indicating a precise (although not perfect) estimation. Most of the mismatching values were present in short speech segments: evaluating these values only for the segments with at least 2 seconds of duration led to Pearson's correlation values of 0.914 and 0.920, for articulation tempo and speech tempo, respectively. Furthermore, silent pauses were almost perfectly detected (precision: 96.1%, recall: 94.9%, F-measure: 95.5), while filled pauses were also identified with a high performance (precision: 83.2%, recall: 69.6%, F-measure: 75.8). In most cases, filled pauses were confused with prolongations of certain phonemes (*e.g.* m / n / a), which are acoustically similar and are often used by the speakers for similar purposes as filled pauses [59, 60].

The output of the ASR system was the phonetic segmentation and labeling of the input signal, which included filled pauses. Based on this output, we extracted 15 S-GAP-related temporal speech parameters using simple calculations (Table **1**).

2.4. Statistical Analysis

Descriptive statistics were used to examine the demographic, neuropsychological, and speech characteristics of participants. In both the English- and Hungarian-speaking samples, comparisons between the MCI *vs.* HC groups were executed using either the independent samples *t*-test/Welch's *t*-test (based on equality of variances), the Mann-Whitney U test (for cases when the normality assumption was not fulfilled according to the Shapiro-Wilk test of normality) or the Chi-square test (for categorical variables). For the examination of inter-language differences (English-speaking HC *vs.* Hungarian-speaking HC; English-speaking MCI *vs.* Hungarian-speaking MCI), independent samples *t*-test/Welch's *t*-test or the Mann-Whitney U test was carried out.

Table 1. List and definitions of the 15 S-GAP-related temporal parameters of spontaneous speech.

S-GAP-related Parameters	Description
Utterance length (s)	Total length of the utterance (s)
Articulation tempo (1/s)	Total number of phonemes (without hesitations) (count) / total length of the utterance (s)
Speech tempo (1/s)	Total number of phonemes (including hesitations) (count) / total length of the utterance (s)
Silent pause occurrence rate (%)	Total number of silent pauses (count) x 100 / total number of phonemes (count)
Filled pause occurrence rate (%)	Total number of filled pauses (count) x 100 / total number of phonemes (count)
Total pause occurrence rate (%)	Total number of silent and filled pauses (count) x 100 / total number of phonemes (count)
Silent pause duration rate (%)	Total length of silent pauses (s) x 100 / total length of the utterance (s)
Filled pause duration rate (%)	Total length of filled pauses (s) x 100 / total length of the utterance (s)
Total pause duration rate (%)	Total length of silent and filled pauses (s) x 100 / total length of the utterance (s)
Silent pause frequency (1/s)	Total number of silent pauses (count) / total length of the utterance (s)
Filled pause frequency (1/s)	Total number of filled pauses (count) / total length of the utterance (s)
Total pause frequency (1/s)	Total number of silent and filled pauses (count) / total length of the utterance (s)
Silent pause average duration (s)	Total length of silent pauses (s) / total number of silent pauses (count)
Filled pause average duration (s)	Total length of filled pauses (s) / total number of filled pauses (count)
Total pause average duration (s)	Total length of silent and filled pauses (s) / total number of silent and filled pauses (count)

Abbreviations: s: second.

Receiver operating characteristics (ROC) analysis was applied to assess which S-GAP-related parameters have the most promising classification abilities based on their area under the curve (AUC) in the two languages. Sensitivity and specificity (true positive rate and true negative rate) were

calculated using threshold values that yielded the highest possible sensitivity (while keeping specificity above 50%). For comparison of the S-GAP parameters' classification ability between the two languages, the comparison of the independent ROC curves module of the MedCalc software was used.

All statistical analyses were performed using SPSS v.24 (SPSS Inc., Chicago, IL, USA), except for the inter-language comparison of AUCs for which MedCalc v.19.4 was applied (MedCalc Software Ltd., Ostend, Belgium). For all statistical comparisons, the level of significance was set at the 0.05 level.

3. RESULTS

3.1. Demographics and Neuropsychological Test Performances

Detailed demographic characteristics and neuropsychological test scores of all groups (means and standard deviations) are presented in Table **2**. Regarding demographics (gender, age, and years of education) and the CDT test, there were no statistically significant differences between the MCI and the HC group in either languages. However, regarding the other neuropsychological tests, MCI patients showed significantly poorer performance in the MMSE than HCs (English-speaking sample: $U = 62.500$; $Z = -2.703$; $p = 0.009$; Hungarian-speaking sample: $U = 0.000$; $Z = -4.879$; $p < 0.001$), and they also had higher scores in the GDS in both languages (English-speaking sample: $U = 71.000$; $Z = -2.277$; $p = 0.024$; Hungarian-speaking sample: $U = 59.000$; $Z = -2.736$; $p = 0.008$).

3.2. S-GAP-related Temporal Parameters and Sensitivity Measures in the English-Speaking Sample

Regarding the English-speaking sample, 7 of the total 15 S-GAP-related temporal parameters displayed significant differences between the MCI and the HC groups. Patients with MCI showed significantly lower articulation tempo and speech tempo as well, while they produced a significantly higher occurrence rate of total pauses, duration rate of silent pauses and total pauses, as well as the average duration of silent pauses and total pauses (Table **3**).

To determine which S-GAP-related temporal speech parameters would be the most precise in classifying patients, ROC analysis was executed. The ROC analysis revealed that the following 8 parameters had statistically significant classification abilities (starting with the highest AUC): speech tempo, articulation tempo, total pause duration rate, silent pause duration rate, silent pause average duration, total pause average duration, total pause occurrence rate, and filled pause occurrence rate. Sensitivity was above 90% both for speech tempo (sensitivity: 100%; specificity: 63.2%) and for articulation tempo (sensitivity: 100%; specificity: 57.9%).

Sensitivity and specificity measures of the statistically significant S-GAP-related temporal parameters (calculated using threshold values optimal for early screening) are detailed in Table **4**; ROC curves are plotted in Fig. (**2**).

3.3. S-GAP-related Temporal Parameters and Sensitivity Measures in the Hungarian-Speaking Sample

Regarding the Hungarian-speaking sample, 5 of the total 15 S-GAP-related temporal parameters turned out to be statistically different between the MCI and the HC group. MCI patients' utterance length was significantly shorter, while a higher duration rate of silent pauses and total pauses, as well as a higher average duration of silent pauses and total pauses characterized their speech (Table **5**).

With regard to the ROC analysis, the following 5 parameters turned out to be statistically significant (from highest to lowest AUCs): silent pause duration rate, utterance length, total pause duration rate, silent pause average duration, and total pause average duration. Sensitivity was above 90% in

Table 2. Means (standard deviations) of participants' demographic characteristics and neuropsychological test scores in the English-speaking and Hungarian-speaking samples.

English-Speaking Sample		-	**Hungarian-Speaking Sample**	
HC (n = 19)	**MCI (n = 14)**	-	**HC (n = 20)**	**MCI (n = 13)**
-		**Demographic characteristics**	-	
5/14	6/8	**Gender (male/female)**	3/17	4/9
74.47 (7.321)	72.36 (6.857)	**Age (years)**	69.90 (5.609)	73.77 (4.969)
17.84 (3.532)	16.79 (3.118)	**Education (years)**	13.15 (2.455)	11.77 (2.743)
-		**Neuropsychological test scores**	-	
29.16 (1.015)	27.71 (1.773)	**MMSE**	28.85 (0.813)	26.31 (0.751)
8.89 (1.197)	9.21 (1.188)	**CDT**	7.60 (3.152)	7.92 (2.178)
3.16 (2.853)	5.50 (2.822)	**GDS-30 / GDS-15**	1.65 (1.387)	2.77 (1.013)

Abbreviations: HC: healthy control; MCI: mild cognitive impairment; MMSE: Mini-Mental State Examination; CDT: Clock Drawing Test; GDS-30: Geriatric Depression Scale (30-item); GDS-15: Geriatric Depression Scale (15-item).

Table 3. Descriptive statistics (means and standard deviations) and group comparisons in the English-speaking sample using the independent samples *t*-test / Mann-Whitney U test.

English-Speaking Sample	M (SD)		Test STATISTICS	
S-GAP-Related Parameters	**HC (*n* = 19)**	**MCI (*n* = 14)**	***t*-test / Mann-Whitney U TEST**	***p***
Utterance length (s)	275.33 (120.02)	201.94 (135.07)	$U = 82.000$; $Z = -1.858$	0.065
Articulation tempo (1/s)	8.88 (1.21)	6.78 (1.32)	$t(31) = 4.732$	0.000*
Speech tempo (1/s)	10.07 (1.10)	8.02 (1.34)	$t(31) = 4.810$	0.000*
Silent pause occurrence rate (%)	9.43 (3.17)	12.11 (4.35)	$U = 85.000$; $Z = -1.748$	0.084
Filled pause occurrence rate (%)	2.55 (1.08)	3.63 (1.73)	$U = 79.000$; $Z = -1.967$	0.050
Total pause occurrence rate (%)	11.98 (3.55)	15.75 (4.34)	$t(31) = -2.736$	0.010*
Silent pause duration rate (%)	31.43 (8.72)	45.61 (12.05)	$t(31) = -3.927$	0.000*
Filled pause duration rate (%)	5.64 (3.23)	6.56 (5.22)	$U = 126.000$; $Z = -0.255$	0.815
Total pause duration rate (%)	37.07 (9.27)	52.17 (11.23)	$t(31) = -4.228$	0.000*
Silent pause frequency (1/s)	0.93 (0.30)	0.95 (0.28)	$t(31) = -0.139$	0.890
Filled pause frequency (1/s)	0.25 (0.09)	0.28 (0.14)	$U = 122.000$; $Z = -0.401$	0.706
Total pause frequency (1/s)	1.18 (0.33)	1.24 (0.30)	$t(31) = -0.453$	0.653
Silent pause average duration (s)	0.34 (0.07)	0.51 (0.18)	$t(15.802) = -3.108$	0.007*
Filled pause average duration (s)	0.21 (0.05)	0.21 (0.09)	$U = 105.000$;$Z = -1.020$	0.321
Total pause average duration (s)	0.31 (0.05)	0.44 (0.14)	$t(15.968) = -3.007$	0.008*

Abbreviations: M: mean; SD: standard deviation; HC: healthy control; MCI: mild cognitive impairment; *p-values indicating statistically significant differences (level of significance was set at $p < 0.05$).

Table 4. Accuracy measures of S-GAP-related temporal parameters with statistically significant classification ability in the English-speaking sample using ROC analysis.

English-Speaking Sample	Accuracy Measures						
S-GAP-Related Parameters	***p***	**AUC**	**95% CI-**	**95% CI+**	**Threshold Value**	**Sensitivity (%)**	**Specificity (%)**
Speech tempo (1/s)	0.000	0.891	0.784	0.998	9.843	100	63.2
Articulation tempo (1/s)	0.000	0.891	0.779	1.000	8.772	100	57.9
Total pause duration rate (%)	0.001	0.846	0.711	0.980	36.689	85.7	52.6
Silent pause duration rate (%)	0.001	0.835	0.695	0.974	32.398	85.7	63.2
Silent pause average duration (s)	0.003	0.808	0.654	0.963	0.346	85.7	52.6
Total pause average duration (s)	0.006	0.782	0.614	0.950	0.329	78.6	57.9
Total pause occurrence rate (%)	0.016	0.748	0.578	0.918	12.078	78.6	52.6
Filled pause occurrence rate (%)	0.049	0.703	0.524	0.882	2.567	78.6	52.6

Abbreviations: ROC: receiver operating characteristics; AUC: area under the curve; CI: confidence interval; (level of significance was set at $p < 0.05$).

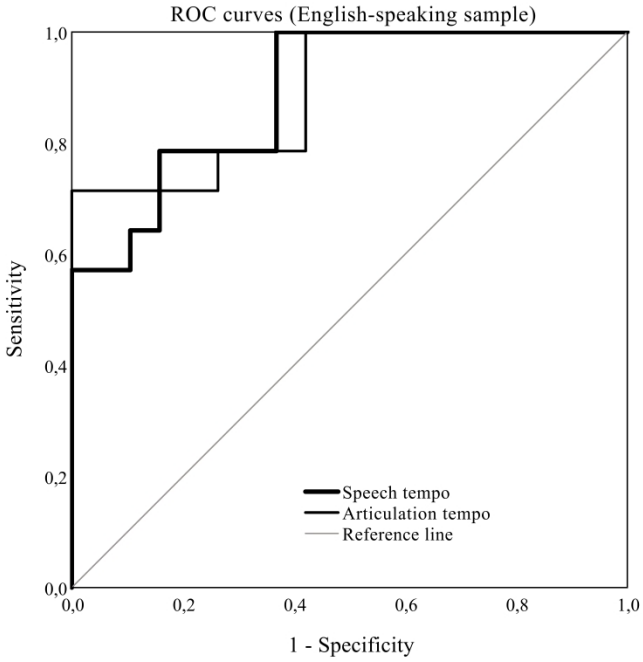

Fig. (2). ROC curves for S-GAP-related temporal parameters with the highest (above 90%) sensitivity for discriminating between MCI and HC participants in the English-speaking sample (speech tempo and articulation tempo). **Abbreviations:** ROC: receiver operating characteristics; HC: healthy control; MCI: mild cognitive impairment.

Table 5. Descriptive statistics (means and standard deviations) and group comparisons in the Hungarian-speaking sample using the independent samples *t*-test / Mann-Whitney U test.

Hungarian-Speaking Sample	**M (SD)**		**Test Statistics**	
S-GAP-Related Parameters	**HC (*n* = 20)**	**MCI (*n* = 13)**	***t*-test / Mann-Whitney U Test**	***p***
Utterance length (s)	155.06 (70.21)	107.82 (87.65)	$U = 66.000$; $Z = -2.358$	0.018*
Articulation tempo (1/s)	9.90 (1.97)	8.63 (1.75)	$t(31) = 1.878$	0.070
Speech tempo (1/s)	10.67 (1.87)	9.47 (1.62)	$t(31) = 1.894$	0.068
Silent pause occurrence rate (%)	4.88 (1.64)	5.91 (1.83)	$t(31) = -1.678$	0.103
Filled pause occurrence rate (%)	2.69 (1.83)	3.28 (2.10)	$U = 112.000$; $Z = -0.663$	0.524
Total pause occurrence rate (%)	7.58 (3.13)	9.20 (3.37)	$U = 94.500$; $Z = -1.308$	0.194
Silent pause duration rate (%)	23.49 (9.72)	32.46 (8.16)	$t(31) = -2.750$	0.010*
Filled pause duration rate (%)	6.26 (4.10)	7.03 (4.68)	$t(31) = -0.494$	0.625
Total pause duration rate (%)	29.76 (11.81)	39.49 (11.07)	$t(31) = -2.367$	0.024*
Silent pause frequency (1/s)	0.49 (0.11)	0.54 (0.13)	$t(31) = -1.008$	0.321
Filled pause frequency (1/s)	0.26 (0.14)	0.28 (0.15)	$t(31) = -0.336$	0.739
Total pause frequency (1/s)	0.76 (0.21)	0.83 (0.22)	$U = 108.000$; $Z = -0.811$	0.434
Silent pause average duration (s)	0.47 (0.18)	0.62 (0.17)	$U = 70.000$; $Z = -2.211$	0.027*
Filled pause average duration (s)	0.21 (0.06)	0.24 (0.10)	$U = 123.000$; $Z = -0.258$	0.813
Total pause average duration (s)	0.39 (0.14)	0.48 (0.10)	$U = 73.000$; $Z = -2.100$	0.036*

Abbreviations: M: mean; SD: standard deviation; HC: healthy control; MCI: mild cognitive impairment; *p-values indicating statistically significant differences (level of significance was set at $p < 0.05$).

Table 6. Accuracy measures of S-GAP-related temporal parameters with statistically significant classification ability in the Hungarian-speaking sample using ROC analysis.

Hungarian-Speaking Sample	Accuracy Measures						
S-GAP-Related Parameters	***p***	**AUC**	**95% CI-**	**95% CI+**	**Threshold Value**	**Sensitivity (%)**	**Specificity (%)**
Silent pause duration rate (%)	0.018	0.746	0.579	0.914	24.191	92.3	60.0
Utterance length (s)	0.018	0.746	0.558	0.934	132.345	76.9	60.0
Total pause duration rate (%)	0.020	0.742	0.573	0.912	27.280	92.3	55.0
Silent pause average duration (s)	0.027	0.731	0.551	0.910	0.438	84.6	55.0
Total pause average duration (s)	0.036	0.719	0.537	0.902	0.349	92.3	55.0

Abbreviations: ROC: receiver operating characteristics; AUC: area under the curve; CI: confidence interval; (level of significance was set at $p < 0.05$).

the case of three parameters, with the highest specificity for silent pause duration rate (sensitivity: 92.3%; specificity: 60.0%) while lower for total pause duration rate (sensitivity: 92.3%; specificity: 55.0%), and total pause average duration (sensitivity: 92.3%; specificity: 55.0%).

Sensitivity and specificity measures of the statistically significant temporal parameters (calculated at optimal threshold values) are detailed in Table **6**; ROC curves are plotted in Fig. (**3**).

To examine whether the S-GAP-related parameters have different classification abilities in the two languages, pairwise comparisons of AUCs were executed between the English- and Hungarian-speaking samples. The analysis showed that the AUCs did not differ significantly regarding any of the 15 S-GAP-related parameters between the two language groups (Table **7**).

3.4. Inter-Language Group Comparisons of S-GAP-related Temporal Parameters

Besides our main goal of exploring the S-GAP-related temporal parameters separately in the two language samples, inter-language comparisons were also carried out as additional analyses between the English-speaking *vs.* Hungarian-speaking HC group and the English-speaking *vs.* Hungarian-speaking MCI group (Table **8**). Regarding the HC group, 8 S-GAP-related parameters showed statistically significant differences between the English- and Hungarian-speaking samples, which were the following: utterance length (E-HC > H-HC), silent pause occurrence rate (E-HC > H-HC), total pause occurrence rate (E-HC > H-HC), silent pause duration rate (E-HC > H-HC), total pause duration rate (E-HC > H-HC), silent pause frequency (E-HC > H-HC), total pause frequency (E-HC > H-HC), and silent pause average duration (H-HC > E-HC). Regarding the MCI group, 9 significantly different parameters were revealed again: utterance length (E-MCI > H-MCI), articulation tempo (H-MCI > E-MCI), speech tempo (H-MCI > E-MCI), silent pause occurrence rate (E-MCI > H-MCI), total pause occurrence rate (E-MCI > H-MCI), silent pause duration rate (E-MCI > H-MCI), total pause duration rate (E-MCI > H-MCI), silent pause frequency (E-MCI > H-MCI), and total pause frequency (E-MCI > H-MCI).

4. DISCUSSION

The aim of this international study was to validate the S-GAP Test®, a novel spontaneous speech analyzer (originally developed for the Hungarian language), in an English-speaking sample for the purpose of MCI-recognition. The major objective was to develop a neuropsychological screening method, which is sensitive to in multiple languages and provides clinicians with a simple and quick way for the screening of MCI. For this purpose, automatic analysis of spontaneous speech was carried out by applying ASR. This is the first study conducted with both English- and Hungarian native speakers in which the same method was applied to explore the acoustic parameters of spontaneous speech in MCI and HC subjects.

To summarize the 10-year development process of the S-GAP Test®, the first main finding was the discovery of significant differences between the mild stage of AD and HC regarding speech tempo and hesitation ratio [10]; subsequently, its usefulness was also demonstrated in the prodromal stage of AD since the proposed acoustic biomarkers carried significant information on the separation of MCI from HC [13]. In parallel with the introduction of MCI as a target group, another important step in the development process was the implementation of automatic analysis instead of manual counting. Through the efforts toward the automatic extraction of acoustic features, a machine learning model was constructed [13, 15]. The automatically selected feature sets were found to be superior to the manually constructed ones used for MCI detection [14]. Extending the previous studies, the applicability of the S-GAP Test® was demonstrated in differentiation not only between MCI and HC but also between MCI and mild AD patients by relying on automatically extracted acoustic markers of spontaneous speech [17, 33]. Before the present study, the S-GAP Test® was applied to a total of 95 HC, 105 MCI, and 35 mild AD individuals.

4.1. Main Findings

Present results indicated that analysis of spontaneous speech using the S-GAP Test® is sensitive to detect MCI cases not only in native Hungarian-speaking but in native English-speaking populations as well. Four temporal parameters that differed significantly between the HC and MCI groups both in the English-speaking and in the Hungarian-

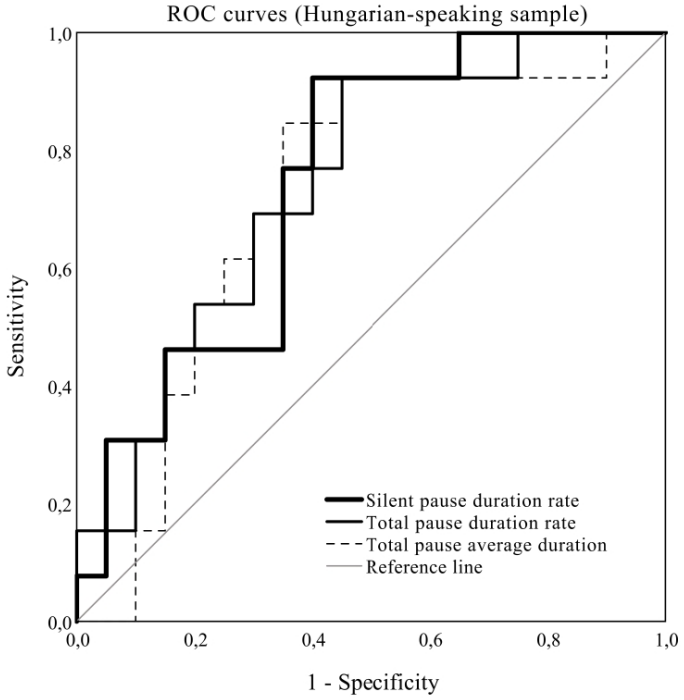

Fig. (3). ROC curves for temporal parameters with the highest (above 90%) sensitivity for discriminating between MCI and HC participants in the Hungarian-speaking sample (silent pause duration rate, total pause duration rate, and total pause average duration). **Abbreviations:** ROC: receiver operating characteristics; HC: healthy control; MCI: mild cognitive impairment.

Table 7. Pairwise comparison of the English- and Hungarian-speaking samples' AUCs regarding the 15 S-GAP-related temporal parameters of speech.

-	AUC		Pairwise Comparisons	
S-GAP-Related Parameters	**English-Speaking Sample**	**Hungarian-Speaking Sample**	***z- statistic***	***p***
Utterance length (s)	0.692	0.746	0.384	0.701
Articulation tempo (1/s)	0.891	0.692	1.741	0.082
Speech tempo (1/s)	0.891	0.685	1.828	0.068
Silent pause occurrence rate (%)	0.680	0.658	0.163	0.871
Filled pause occurrence rate (%)	0.703	0.569	0.931	0.352
Total pause occurrence rate (%)	0.748	0.637	0.827	0.408
Silent pause duration rate (%)	0.835	0.746	0.784	0.433
Filled pause duration rate (%)	0.528	0.508	0.120	0.904
Total pause duration rate (%)	0.846	0.743	0.927	0.354
Silent pause frequency (1/s)	0.541	0.631	0.600	0.548
Filled pause frequency (1/s)	0.541	0.523	0.119	0.905
Total pause frequency (1/s)	0.560	0.585	0.169	0.866
Silent pause average duration (s)	0.808	0.731	0.630	0.529
Filled pause average duration (s)	0.605	0.527	0.492	0.623
Total pause average duration (s)	0.782	0.719	0.486	0.627

Abbreviations: AUC: area under the curve; (level of significance was set at $p < 0.05$).

Table 8. Inter-language comparisons of the S-GAP-related temporal parameters of speech using the independent samples *t*-test / Mann-Whitney U test.

S-GAP-Related Parameters	E-HC *vs.* H-HC		E-MCI *vs.* H-MCI	
	***t*-test / Mann-Whitney U Test**	***p***	***t*-test / Mann-Whitney U Test**	***p***
Utterance length (s)	$t(28.729) = 3.794$	0.001*	$U = 47.000$; $Z = -2.135$	0.033*
Articulation tempo (1/s)	$t(31.801) = -1.949$	0.060	$t(25) = -3.120$	0.005*
Speech tempo (1/s)	$t(31.081) = -1.219$	0.232	$t(25) = -2.529$	0.018*
Silent pause occurrence rate (%)	$t(26.715) = 5.570$	0.000*	$U = 8.000$; $Z = -4.028$	0.000*
Filled pause occurrence rate (%)	$U = 179.000$; $Z = -0.309$	0.771	$U = 74.000$; $Z = -0.825$	0.430
Total pause occurrence rate (%)	$U = 65.000$; $Z = -3.512$	0.000*	$t(25) = 4.347$	0.000*
Silent pause duration rate (%)	$t(37) = 2.678$	0.011*	$t(25) = 3.293$	0.003*
Filled pause duration rate (%)	$U = 174.000$; $Z = -0.450$	0.667	$U = 85.000$; $Z = -0.291$	0.793
Total pause duration rate (%)	$t(37) = 2.142$	0.039*	$t(25) = 2.951$	0.007*
Silent pause frequency (1/s)	$t(23.309) = 5.898$	0.000*	$t(25) = 4.652$	0.000*
Filled pause frequency (1/s)	$t(37) = -0.400$	0.691	$U = 89.000$; $Z = -0.097$	0.943
Total pause frequency (1/s)	$U = 55.000$; $Z = -3.793$	0.000*	$U = 17.000$; $Z = -3.591$	0.000*
Silent pause average duration (s)	$U = 110.000$; $Z = -2.248$	0.024*	$t(25) = -1.537$	0.137
Filled pause average duration (s)	$U = 148.000$; $Z = -1.180$	0.247	$U = 73.000$; $Z = -0.873$	0.402
Total pause average duration (s)	$U = 141.000$; $Z = -1.377$	0.175	$t(25) = -0.789$	0.437

Abbreviations: E-HC: English-speaking sample - healthy control; E-MCI: English-speaking sample - mild cognitive impairment; H-HC: Hungarian-speaking sample - healthy control; H-MCI: Hungarian-speaking sample - mild cognitive impairment; **p*-values indicating statistically significant differences (level of significance was set at $p < 0.05$).

speaking samples are: MCI patients showed higher silent pause duration rate, total pause duration rate, silent pause average duration, and total pause average duration. Based on this finding, these parameters might be sensitive biomarkers of MCI in both languages.

Additional to the above-mentioned four temporal parameters, the English-speaking MCI group also showed lower articulation tempo and speech tempo compared to HC. The importance of these linguistic features in mild AD or MCI has been previously demonstrated in the Hungarian language, using both manual calculation and automatic analysis [10, 13, 16, 17, 33]. Interestingly, in our previous studies, Hungarian-speaking MCI/AD patients also showed a reduction in articulation and speech tempo, while in the present sample this difference was only tendentious. A possible explanation of this might be the difference in the task that was implemented for speech elicitation: namely, in our previous studies, a film description task was used in which the participants had to retell the events of a specially designed, one-minute long silhouette animation, instead of the previous day task applied in the present study.

ROC analysis clearly distinguished the English-speaking MCI cases from HCs based on speech tempo and articulation tempo with 100% sensitivity and further three parameters with very high sensitivity (85.7%) at moderate specificity. In the Hungarian-speaking groups, ROC analysis showed high sensitivity values for silent and total pause duration rate and also for total pause average duration (92.3%). These results suggest that the S-GAP Test® might indicate MCI more sensitively in the English-speaking than in the Hungarian-speaking sample.

Higher number and/or length of pauses, and the decrease of articulation/speech tempo have been described in a number of studies examining varying degrees of cognitive impairment, however, with different methodologies and using various types of tasks such as spontaneous speech [61-63], narrative recall [21], picture description [32], or reading aloud [2, 24, 43].

Pause-related features indicate retrieval difficulties [12] related to degeneration in hippocampal brain regions [64], and they are also associated with atrophy of grey matter in the frontopolar (or Brodmann) area [65] which has a role in higher-order cognitive functions like memory retrieval [66] or multitasking [67]. It is hypothesized that an increase in the number or duration of pauses demonstrates the increase in the cognitive load required for maintaining one's train of thought during speech [28]. Although these changes might not always be perceptible to the ear, speech analysis indicates that silence might be a significant marker of planning, word-retrieval, and executive difficulties due to cognitive deterioration [2, 61]. Language functions in general (*e.g.* measured by naming or verbal fluency tasks) also show a correlation with grey matter volume of the left temporal lobe in MCI and AD [68].

It is important to note that metrics regarding the diagnostic accuracy of language functions have been reported in variants of primary progressive aphasia [69] but earlier investigations of MCI/AD did not focus on this [21, 22]. However, in recent years more data related to this field have been reported [16, 17, 28, 29, 33, 70]. For example, the classification sensitivity of linguistic and phonetic features of connected speech by automated assessment of the Cookie Theft picture description task was 85% between HC and AD/MCI cases in a Canadian-Mexican co-operation study [70]. The diagnostic utility of automatic speech analysis for recorded vocal tasks has also been previously demonstrated in a French-speaking population with 79% or 86% classification accuracy between HC and MCI [28, 29]. As for the previous investigations of our research group, the irregularities in MCI speech and language were demonstrated by 68% sensitivity in the differentiating MCI from HC [17]. This result, however, was based on a different language elicitation task, *i.e.* a video description (as mentioned earlier). Comparing results of the present study with previous ones, the S-GAP Test® applied for English- and Hungarian-speaking MCI populations has shown a relatively high sensitivity.

4.2. Limitations and Considerations

A limitation of this pilot study was the small sample size and particularly the low number of MCI participants, which represent the main drawback regarding statistical power. However, this disadvantage was compensated by careful examination of the patients included in the study with the aim of excluding other confounding factors. Taking into consideration that this research was intended as a pilot to find those temporal parameters (from the full set of 15) with the highest differentiating potential for embedding in a future mobile application, multiple correction testing was not applied for the statistical comparisons. This needs to be taken into account when interpreting the results.

Regarding the sensitivity and specificity of temporal parameters, the optimal threshold values were defined to maximize sensitivity, which, as a result of a trade-off between the two measures, decreased specificity (although it exceeded 50% in every case). Given that the goal was to create an early MCI screening tool specifically targeting high-risk individuals (*e.g.* people above the age of 60) and considering the serious consequences of undiagnosed MCI (mainly the possibility of converting to dementia [71], reaching a high true positive rate was prioritized.

Before applying the S-GAP Test® internationally in clinical settings, the observed inter-language differences (E-HC *vs.* H-HC; E-MCI *vs.* H-MCI) emphasize the need for gathering normative data for international adaptations. In our present sample, English-speaking individuals on average produced longer monologues regarding their previous day, while they talked slower and their speech contained more pauses compared to the Hungarian-speaking participants. These language differences will have to be taken into account during the setting of screening thresholds in different countries as temporal features indicative of HC/MCI speech can have substantially different mean values in each language.

CONCLUSION

In summary, the results of the S-GAP Test® in the English- and Hungarian native speaker populations suggest that similar changes in temporal parameters of spontaneous speech detected by ASR can be observed across different languages. Based on these findings, it could be suggested 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. An early and accurate diagnosis of cognitive deficits would be of much help for patients and their families in order to plan for the future and to start early treatment. However, it is important to state that this method can only be the first step in the diagnostic process of MCI, as it is not intended to be a complete substitute for a detailed clinical examination.

In the future, an S-GAP Test®-based speech analysis might permit the screening and research evaluation of prodromal stages of different types of dementia through a computerized, interactive smart phone application (which is currently under development in co-operation with the Institute of Informatics at the University of Szeged, Hungary). This could be a low-cost, noninvasive, non-stressful method that allows quick, easy, and remote assessment. A further advantage of this method is that the recording of spontaneous speech (in a phone call-like setting) is less stressful for the patient than a neuropsychiatric test. Additionally, this approach might also serve as an objective measurement for the efficacy of pharmacotherapy and drug candidate molecules in cognitive impairment.

LIST OF ABBREVIATIONS

AD	=	Alzheimer's Disease
ASR	=	Automatic Speech Recognition
AUC	=	Area Under Curve
CDT	=	Clock Drawing Test
CT	=	Computed Tomography
DNN	=	Deep Neural Network
GDS	=	Geriatric Depression Scale
HC	=	Healthy Control
HMM	=	Hidden Markov Model
IT	=	Information Technology
M	=	Mean
MCI	=	Mild Cognitive Impairment
MRI	=	Magnetic Resonance Imaging
MMSE	=	Mini-Mental State Examination
ROC	=	Receiver Operating Characteristic
s	=	Second
SD	=	Standard Deviation
S-GAP Test®	=	Speech-Gap Test®

AUTHORS' CONTRIBUTION

J.K. and I.H. conceived and designed the study. D.P.D. and M.P. evaluated the subjects, while I.K., R.B. and N.I. collected the data. R.B., N.I., G.G., L.T., V.V. analyzed the data and performed the figures and tables. J.K., M.P., G.G., R.B. and N.I. wrote the manuscript. All authors contributed to the article and approved the final manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The American part of the study was approved by the Institutional Review Board of the New York State Psychiatric Institute – Columbia University Department of Psychiatry (protocol number: 7611). The Hungarian part of the study was approved by the Regional Human Biomedical Research Ethics Committee of the University of Szeged, Hungary (reference number: 231/2017-SZTE).

HUMAN AND ANIMAL RIGHTS

No animals were used in this study. All human research procedures followed were in accordance with the guidelines of the Declaration of Helsinki of 1975, as revised in 2013 (http://ethics.iit.edu/ecodes/node/3931).

CONSENT FOR PUBLICATION

Written informed consent was obtained from all participants at both research sites.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

This work was supported by Merck & Co., Inc., the National Research, Development, and Innovation Office of Hungary (NKFIH-FK-124413), and the Hungarian Ministry for Innovation and Technology (TUDFO/47138-1/2019-ITM). The authors were also supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (G. Gosztolya: BO/00653/19), and the Hungarian Ministry for Innovation and Technology (G. Gosztolya: NKFIH-1279-2/2020, NKFIH-FK-124413, ÚNKP-21-5; N. Imre: ÚNKP-21-4-I).

CONFLICT OF INTEREST

Dr. D. P. Devanand is a consultant to Acadia, BXCel, Genentech, Corium, and Grifols. The authors have no competing interests to declare.

ACKNOWLEDGEMENTS

The authors wish to thank Nóra Domján for her help in performing the neuropsychiatric tests and recording the audio files of the participants' speeches.

REFERENCES

[1] Braaten AJ, Parsons TD, McCue R, Sellers A, Burns WJ. Neurocognitive differential diagnosis of dementing diseases: Alzheimer's dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. Int J Neurosci 2006; 116(11): 1271-93. http://dx.doi.org/10.1080/00207450600920928 PMID: 17000529

[2] Meilán JJG, Martínez-Sánchez F, Carro J, Sánchez JA, Pérez E. Acoustic markers associated with impairment in language processing in Alzheimer's Disease. Span J Psychol 2012; 15(2): 487-94. http://dx.doi.org/10.5209/rev_SJOP.2012.v15.n2.38859 PMID: 22774422

[3] Forbes KE, Venneri A, Shanks MF. Distinct patterns of spontaneous speech deterioration: An early predictor of Alzheimer's disease. Brain Cogn 2002; 48(2-3): 356-61. http://dx.doi.org/10.1006/brcg.2001.1377 PMID: 12030467

[4] Laske C, Sohrabi HR, Frost SM, *et al.* Innovative diagnostic tools for early detection of Alzheimer's disease. Alzheimers Dement 2015; 11(5): 561-78. http://dx.doi.org/10.1016/j.jalz.2014.06.004 PMID: 25443858

[5] Taler V, Phillips NA. Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. J Clin Exp Neuropsychol 2008; 30(5): 501-56. http://dx.doi.org/10.1080/13803390701550128 PMID: 18569251

[6] Ahmed S, Haigh AM, de Jager CA, Garrard P. Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. Brain 2013; 136(Pt 12): 3727-37. http://dx.doi.org/10.1093/brain/awt269 PMID: 24142144

[7] Petti U, Baker S, Korhonen A. A systematic literature review of automatic Alzheimer's disease detection from speech and language. J Am Med Inform Assoc 2020; 27(11): 1784-97. http://dx.doi.org/10.1093/jamia/ocaa174 PMID: 32929494

[8] Martínez-Nicolás I, Llorente TE, Martínez-Sánchez F, Meilán JJG. Ten years of research on automatic voice and speech analysis of people with Alzheimer's disease and mild cognitive impairment: A systematic review article. Front Psychol 2021; 12: 620251. http://dx.doi.org/10.3389/fpsyg.2021.620251 PMID: 33833713

[9] Vigo I, Coelho L, Reis S. Speech- and language-based classification of Alzheimer's disease: A systematic review. Bioengineering (Basel) 2022; 9(1): 27. http://dx.doi.org/10.3390/bioengineering9010027 PMID: 35049736

[10] Hoffmann I, Nemeth D, Dye CD, Pákáski M, Irinyi T, Kálmán J. Temporal parameters of spontaneous speech in Alzheimer's disease. Int J Speech Lang Pathol 2010; 12(1): 29-34. http://dx.doi.org/10.3109/17549500903137256 PMID: 20380247

[11] Boersma P. Praat, a system for doing phonetics by computer. Glot Int 2002; 5: 341-5.

[12] Szatloczki G, Hoffmann I, Vincze V, Kalman J, Pakaski M. Speaking in Alzheimer's disease, is that an early sign? Importance of changes in language abilities in Alzheimer's disease. Front Aging Neurosci 2015; 7: 195-202. http://dx.doi.org/10.3389/fnagi.2015.00195 PMID: 26539107

[13] Tóth L, Gosztolya G, Vincze V, Hoffmann I, Szatlóczki G, Biró E, *et al.* Automatic detection of mild cognitive impairment from spontaneous speech using ASR. Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech). 2694-8. http://dx.doi.org/10.21437/Interspeech.2015-568

[14] Gosztolya G, Tóth L, Grósz T, *et al.* Detecting mild cognitive impairment from spontaneous speech by correlation-based phonetic feature selection. Proceedings of Interspeech 2016; 107-111: 107-11. http://dx.doi.org/10.21437/Interspeech.2016-384

[15] Vincze V, Gosztolya G, Tóth L, Hoffmann I, Szatlóczki G, Bánréti Z, *et al.* Detecting mild cognitive impairment by exploiting linguistic information from transcripts. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL). 181-7. http://dx.doi.org/10.18653/v1/P16-2030

[16] Tóth L, Hoffmann I, Gosztolya G, *et al.* A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech. Curr Alzheimer Res 2018; 15(2): 130-8. http://dx.doi.org/10.2174/1567205014666171121114930 PMID: 29165085

[17] Gosztolya G, Vincze V, Tóth L, Pákáski M, Kálmán J, Hoffmann I. Identifying mild cognitive impairment and mild Alzheimer's

disease based on spontaneous speech using ASR and linguistic features. Comput Speech Lang 2019; 53: 181-97. http://dx.doi.org/10.1016/j.csl.2018.07.007

[18] Solomon A, Mangialasche F, Richard E, *et al.* Advances in the prevention of Alzheimer's disease and dementia. J Intern Med 2014; 275(3): 229-50. http://dx.doi.org/10.1111/joim.12178 PMID: 24605807

[19] de la Fuente Garcia S, Ritchie CW, Luz S. Artificial intelligence, speech, and language processing approaches to monitoring Alzheimer's disease: A systematic review. J Alzheimers Dis 2020; 78(4): 1547-74. http://dx.doi.org/10.3233/JAD-200888 PMID: 33185605

[20] Chien Y, Hong S, Cheah W, Fu L, Chang Y. An assessment system for Alzheimer's disease based on speech using a novel feature sequence design and recurrent neural network. IEEE International Conference on Systems, Man, and Cybernetics (SMC). 3289-94. http://dx.doi.org/10.1109/SMC.2018.00557

[21] Roark B, Mitchell M, Hosom JP, Hollingshead K, Kaye J. Spoken language derived measures for detecting mild cognitive impairment. IEEE Trans Audio Speech Lang Process 2011; 19(7): 2081-90. http://dx.doi.org/10.1109/TASL.2011.2112351 PMID: 22199464

[22] Sajjadi SA, Patterson K, Tomek M, Nestor PJ. Abnormalities of connected speech in semantic dementia vs Alzheimer's disease. Aphasiology 2012; 26(6): 847-66. http://dx.doi.org/10.1080/02687038.2012.654933

[23] Jarrold W, Peintner B, Wilkins D, Vergryi D, Richey C, Gorno-Tempini ML, *et al.* Aided diagnosis of dementia type through computer-based analysis of spontaneous speech. Proceedings of the Workshop on Computational Linguistics and Clinical Psychology (CLPsych). 27-37. http://dx.doi.org/10.3115/v1/W14-3204

[24] De Looze C, Kelly F, Crosby L, *et al.* Changes in speech chunking in reading aloud is a marker of mild cognitive impairment and mild-to-moderate Alzheimer's disease. Curr Alzheimer Res 2018; 15(9): 828-47. http://dx.doi.org/10.2174/1567205015666180404165017 PMID: 29623841

[25] Luz S, de la Fuente S, Albert P. A method for analysis of patient speech in dialogue for dementia detection. arXiv 2018; 1811.09919.

[26] Guo Z, Ling Z, Li Y. Detecting Alzheimer's disease from continuous speech using language models. J Alzheimers Dis 2019; 70(4): 1163-74. http://dx.doi.org/10.3233/JAD-190452 PMID: 31322577

[27] Sluis RA, Angus D, Wiles J, *et al.* An automated approach to examining pausing in the speech of people with dementia. Am J Alzheimers Dis Other Demen 2020; 35: 1533317520939773. http://dx.doi.org/10.1177/1533317520939773 PMID: 32648470

[28] König A, Satt A, Sorin A, *et al.* Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. Alzheimers Dement (Amst) 2015; 1(1): 112-24. http://dx.doi.org/10.1016/j.dadm.2014.11.012 PMID: 27239498

[29] König A, Satt A, Sorin A, *et al.* Use of speech analyses within a mobile application for the assessment of cognitive impairment in elderly people. Curr Alzheimer Res 2018; 15(2): 120-9. http://dx.doi.org/10.2174/1567205014666170829111942 PMID: 28847279

[30] Mirzaei S, El Yacoubi M, Garcia-Salicetti S, *et al.* Two-stage feature selection of voice parameters for early Alzheimer's disease prediction. IRBM 2018; 39(6): 430-5. http://dx.doi.org/10.1016/j.irbm.2018.10.016

[31] Tröger J, Linz N, König A, *et al.* Exploitation *vs.* exploration-computational temporal and semantic analysis explains semantic verbal fluency impairment in Alzheimer's disease. Neuropsychologia 2019; 131: 53-61. http://dx.doi.org/10.1016/j.neuropsychologia.2019.05.007 PMID: 31121184

[32] Satt A, Sorin A, Toledo-Ronen O, Barkan O, Kompatsiaris I, Kokonozi A, *et al.* Evaluation of speech-based protocol for detection of early-stage dementia. Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech). 1692-6. http://dx.doi.org/10.21437/Interspeech.2013-32

[33] Gosztolya G, Balogh R, Imre N, *et al.* Cross-lingual detection of mild cognitive impairment based on temporal parameters of spontaneous speech. Comput Speech Lang 2021; 69: 101215. http://dx.doi.org/10.1016/j.csl.2021.101215

[34] Balogh R, Imre N, Gosztolya G, Hoffmann L, Pákáski M, Kálmán J. The role of silence in verbal fluency tasks – a new approach for the detection of mild cognitive impairment. J Int Neuropsychol Soc 2022; 1-13. http://dx.doi.org/10.1017/S1355617721001454 PMID: 35067261

[35] Beltrami D, Gagliardi G, Rossini Favretti R, Ghidoni E, Tamburini F, Calzà L. Speech analysis by natural language processing techniques: A possible tool for very early detection of cognitive decline? Front Aging Neurosci 2018; 10: 369. http://dx.doi.org/10.3389/fnagi.2018.00369

[36] Kato S, Endo H, Homma A, Sakuma T, Watanabe K. Early detection of cognitive impairment in the elderly based on Bayesian mining using speech prosody and cerebral blood flow activation. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 5813-6. http://dx.doi.org/10.1109/EMBC.2013.6610873

[37] Tanaka H, Adachi H, Ukita N, *et al.* Detecting dementia through interactive computer avatars. IEEE J Transl Eng Health Med 2017; 5: 2200111. http://dx.doi.org/10.1109/JTEHM.2017.2752152 PMID: 29018636

[38] Kobayashi M, Kosugi A, Takagi H, Nemoto M, Nemoto K, Arai T, *et al.* Effects of age-related cognitive decline on elderly user interactions with voice-based dialogue systems. IFIP Conference on Human-Computer Interaction. 53-74. http://dx.doi.org/10.1007/978-3-030-29390-1_4

[39] Yamada Y, Shinkawa K, Kobayashi M, *et al.* Tablet-based automatic assessment for early detection of Alzheimer's disease using speech responses to daily life questions. Front Digit Health 2021; 3: 653904. http://dx.doi.org/10.3389/fdgth.2021.653904 PMID: 34713127

[40] Nasrolahzadeh M, Mohammadpoory Z, Haddadnia J. Higher-order spectral analysis of spontaneous speech signals in Alzheimer's disease. Cogn Neurodynamics 2018; 12(6): 583-96. http://dx.doi.org/10.1007/s11571-018-9499-8 PMID: 30483366

[41] Meilán JJG, Martínez-Sánchez F, Carro J, López DE, Millian-Morell L, Arana JM. Speech in Alzheimer's disease: Can temporal and acoustic parameters discriminate dementia? Dement Geriatr Cogn Disord 2014; 37(5-6): 327-34. http://dx.doi.org/10.1159/000356726 PMID: 24481220

[42] López-de-Ipiña K, Alonso JB, Travieso CM, *et al.* On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. Sensors (Basel) 2013; 13(5): 6730-45. http://dx.doi.org/10.3390/s130506730 PMID: 23698268

[43] Espinoza-Cuadros F, Garcia-Zamora MA, Torres-Boza D, Ferrer-Riesgo CA, Montero-Benavides A, Gonzales-Moreira E, *et al.* A spoken language database for research on moderate cognitive impairment: Design and preliminary analysis. Lect Notes Comput Sci 2014; 8854: 219-28. [IberSpeech]. http://dx.doi.org/10.1007/978-3-319-13623-3_23

[44] Gonzalez-Moreira E, Torres-Boza D, Kairuz HA, *et al.* Automatic prosodic analysis to identify mild dementia. BioMed Res Int 2015; 2015: 916356. http://dx.doi.org/10.1155/2015/916356 PMID: 26558287

[45] Fraser KC, Lundholm Fors K, Eckerström M, Öhman F, Kokkinakis D. Predicting MCI status from multimodal language data using cascaded classifiers. Front Aging Neurosci 2019; 11: 205. http://dx.doi.org/10.3389/fnagi.2019.00205 PMID: 31427959

[46] Themistocleous C, Eckerström M, Kokkinakis D. Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls. PLoS One 2020; 15(7): e0236009. http://dx.doi.org/10.1371/journal.pone.0236009 PMID: 32658934

[47] Khodabakhsh A, Demiroglu C. Analysis of speech-based measures for detecting and monitoring Alzheimer's disease. Methods Mol Biol 2015; 1246: 159-73. http://dx.doi.org/10.1007/978-1-4939-1985-7_11 PMID: 25417086

[48] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: Clinical characterization and outcome. Arch Neurol 1999; 56(3): 303-8. http://dx.doi.org/10.1001/archneur.56.3.303 PMID: 10190820

[49] 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(3): 189-98. http://dx.doi.org/10.1016/0022-3956(75)90026-6 PMID: 1202204

[50] Manos PJ, Wu R. The ten point clock test: A quick screen and grading method for cognitive impairment in medical and surgical patients. Int J Psychiatry Med 1994; 24(3): 229-44. http://dx.doi.org/10.2190/5A0F-936P-VG8N-0F5R PMID: 7890481

[51] Yesavage JA, Brink TL, Rose TL, *et al.* Development and validation of a geriatric depression screening scale: A preliminary report. J Psychiatr Res 1982-1983; 17(1): 37-49. http://dx.doi.org/10.1016/0022-3956(82)90033-4 PMID: 7183759

[52] Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS) – Recent evidence and development of a shorter version. Clin Gerontol 1986; 5: 165-73. http://dx.doi.org/10.1300/J018v05n01_09

[53] Young S, Evermann G, Gales MJF, Hain T, Kershaw D, Liu X, *et al.* The HTK Book. Cambridge, UK: Cambridge University 2006.

[54] Hinton G, Deng L, Yu D, *et al.* Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Process Mag 2012; 29(6): 82-97. http://dx.doi.org/10.1109/MSP.2012.2205597

[55] Tóth L. Phone recognition with hierarchical convolutional deep maxout networks. EURASIP J Audio Speech Music Process 2015; 25(1): 1-3. http://dx.doi.org/10.1186/s13636-015-0068-3

[56] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. Proceedings of International Conference on Artificial Intelligence and Statistics (AISTATS). 315-23.

[57] Rousseau A, Delglise P, Estve Y. TED-LIUM: An automatic speech recognition dedicated corpus. Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12). 125-9.

[58] Neuberger T, Gyarmathy D, Gráczi TE, Horváth V, Gósy M, Beke A. Development of a large spontaneous speech database of agglutinative Hungarian language. Text Speech and Dialogue 2014; pp. 424-31. http://dx.doi.org/10.1007/978-3-319-10816-2_51

[59] Eklund R. Prolongations: A dark horse in the disfluency stable. Proceedings of Disfluency in Spontaneous Speech (DiSS) 2001; 5-8.

[60] Deme A, Markó A. Lengthenings and filled pauses in Hungarian adults' and children's speech. Proceedings of Disfluency in Spontaneous Speech (DiSS) 2013; 21-4.

[61] Gayraud F, Lee HR, Barkat-Defradas M. Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. Clin Linguist Phon 2011; 25(3): 198-209. http://dx.doi.org/10.3109/02699206.2010.521612 PMID: 21080826

[62] Lee H, Gayraud F, Hirsh F, Barkat-Defradas M. Speech dysfluencies in normal and pathological aging: A comparison between Alzheimer patients and healthy elderly subjects. Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS). 1174-7.

[63] Singh S, Bucks RS, Cuerden JM. Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. Aphasiology 2001; 15(6): 571-83. http://dx.doi.org/10.1080/02687040143000041

[64] Sarazin M, Chauviré V, Gerardin E, *et al.* The amnestic syndrome of hippocampal type in Alzheimer's disease: An MRI study. J Alzheimers Dis 2010; 22(1): 285-94. http://dx.doi.org/10.3233/JAD-2010-091150 PMID: 20847406

[65] Pistono A, Jucla M, Barbeau EJ, *et al.* Pauses during autobiographical discourse reflect episodic memory processes in early Alzheimer's disease. J Alzheimers Dis 2016; 50(3): 687-98. http://dx.doi.org/10.3233/JAD-150408 PMID: 26757034

[66] Simons JS, Gilbert SJ, Owen AM, Fletcher PC, Burgess PW. Distinct roles for lateral and medial anterior prefrontal cortex in contextual recollection. J Neurophysiol 2005; 94(1): 813-20. http://dx.doi.org/10.1152/jn.01200.2004 PMID: 15728761

[67] Roca M, Torralva T, Gleichgerrcht E, *et al.* The role of Area 10 (BA10) in human multitasking and in social cognition: A lesion study. Neuropsychologia 2011; 49(13): 3525-31. http://dx.doi.org/10.1016/j.neuropsychologia.2011.09.003 PMID: 21930138

[68] Arlt S, Buchert R, Spies L, Eichenlaub M, Lehmbeck JT, Jahn H. Association between fully automated MRI-based volumetry of different brain regions and neuropsychological test performance in patients with amnestic mild cognitive impairment and Alzheimer's disease. Eur Arch Psychiatry Clin Neurosci 2013; 263(4): 335-44. http://dx.doi.org/10.1007/s00406-012-0350-7 PMID: 22940716

[69] Wilson SM, Henry ML, Besbris M, *et al.* Connected speech production in three variants of primary progressive aphasia. Brain 2010; 133(Pt 7): 2069-88. http://dx.doi.org/10.1093/brain/awq129 PMID: 20542982

[70] Hernández-Domínguez L, Ratté S, Sierra-Martínez G, Roche-Bergua A. Computer-based evaluation of Alzheimer's disease and mild cognitive impairment patients during a picture description task. Alzheimers Dement (Amst) 2018; 10(1): 260-8. http://dx.doi.org/10.1016/j.dadm.2018.02.004 PMID: 29780871

[71] Gauthier S, Reisberg B, Zaudig M, *et al.* Mild cognitive impairment. Lancet 2006; 367(9518): 1262-70. http://dx.doi.org/10.1016/S0140-6736(06)68542-5 PMID: 16631882

II.

OPEN

Temporal Speech Parameters Indicate Early Cognitive Decline in Elderly Patients With Type 2 Diabetes Mellitus

Nóra Imre, MA, Réka Balogh, MA,* Gábor Gosztolya, PhD,†*
László Tóth, PhD,† Ildikó Hoffmann, PhD,‡§ Tamás Várkonyi, MD, PhD,‖
*Csaba Lengyel, MD, PhD,‖ Magdolna Pákáski, MD, PhD,**
*and János Kálmán, MD, PhD, DSc**

Introduction: The earliest signs of cognitive decline include deficits in temporal (time-based) speech characteristics. Type 2 diabetes mellitus (T2DM) patients are more prone to mild cognitive impairment (MCI). The aim of this study was to compare the temporal speech characteristics of elderly (above 50 y) T2DM patients with age-matched nondiabetic subjects.

Materials and Methods: A total of 160 individuals were screened, 100 of whom were eligible (T2DM: n = 51; nondiabetic: n = 49). Participants were classified either as having healthy cognition (HC) or showing signs of MCI. Speech recordings were collected through a phone call. Based on automatic speech recognition, 15 temporal parameters were calculated.

Results: The HC with T2DM group showed significantly shorter utterance length, higher duration rate of silent pause and total pause, and higher average duration of silent pause and total pause compared with the HC without T2DM group. Regarding the MCI participants, parameters were similar between the T2DM and the nondiabetic subgroups.

Conclusions: Temporal speech characteristics of T2DM patients showed early signs of altered cognitive functioning, whereas neuropsychological tests did not detect deterioration. This method is useful for identifying the T2DM patients most at risk for manifest MCI, and could serve as a remote cognitive screening tool.

Key Words: mild cognitive impairment, type 2 diabetes mellitus, cognitive screening, neuropsychology, early detection, cognitive dysfunction, language functions, speech analysis, temporal speech characteristics, automatic speech recognition

(*Alzheimer Dis Assoc Disord* 2022;36:148–155)

Increasing evidence confirms the heightened risk of cognitive disorders in elderly patients living with type 2 diabetes mellitus (T2DM), compared with nondiabetic individuals.[1,2] T2DM not only doubles the odds of Alzheimers disease (AD) and vascular dementia (VD),[3] but also increases the incidence of mild cognitive impairment (MCI), the clinical condition between healthy aging and dementia.[4] MCI patients experience subtle cognitive symptoms (eg, deficits in language and executive functions, attention, or memory), which can cause problems with more complex activities of daily living but do not interfere with basic everyday functioning.[5] This association with cognitive decline poses a significant risk worldwide, as the global prevalence of T2DM is more than 9.3% of all adults today.[6] Although the exact pathophysiological pathways are under investigation, diabetes has been reported to accelerate the aging process of the brain through alterations in the metabolism of glucose, insulin, and amyloid, which can act as serious biological risk factors for dementia.[7] Cognition in T2DM was found to be impaired in several domains, like learning, verbal memory, attention, executive functions, processing and psychomotor speed, and language.[8]

Decline in language functions have been found to be one of the earliest signs of cognitive deterioration.[9] Especially, the temporal (time-based) organization of speech reflects the functioning of several underlying cognitive processes, including the planning of speech production, the access to vocabulary, working memory, and, depending on the specific task, even episodic memory.[10] Studies using temporal analysis of speech found increased signs of disfluency (eg, word finding delays), or decreased speech rate in cognitively impaired individuals (eg, patients with AD or MCI).[11–13] Increased number/duration of pauses in speech is hypothesized to reflect the increased cognitive load required for maintaining one's train of thought[14] and the general slowing down of word-retrieval.[9]

Since temporal analyses of speech provide highly valuable information regarding cognitive processes, and there is a strong association between cognitive deficits and T2DM, it is of great significance to explore temporal speech characteristics among a high risk group, the elderly with T2DM. In the present study, an automated speech analysis method, the Speech-Gap Test (S-GAP Test) was applied on speech recordings of T2DM participants. This method, built on automatic speech recognition (ASR) techniques, was sensitive to distinguish between MCI patients and elderly

Received for publication May 7, 2021; accepted December 28, 2021.
From the Departments of *Psychiatry; ‡Hungarian Linguistics; ‖Internal Medicine; †MTA-SZTE Research Group on Artificial Intelligence, University of Szeged, Szeged; and §Hungarian Research Centre for Linguistics, Eötvös Loránd Research Network, Budapest, Hungary.
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 (N.I.: ÚNKP-21-4-I; G.G.: ÚNKP-21-5); the University of Szeged, Hungary (N.I., R.B.: EFOP-3.6.3-VEKOP-16-2017-00009); the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (G.G.: BO/00653/19); and the Hungarian Ministry for Innovation and Technology (G.G.: NKFIH-1279-2/2020, NKFIH-FK-124413).
The authors declare no conflicts of interest.
Reprints: Nóra Imre, MA, Department of Psychiatry, Albert Szent-Györgyi Medical School, University of Szeged, Korányi fasor 8-10, Szeged H-6720, Hungary (e-mail: imre.nora@med.u-szeged.hu).
Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, www.alzheimerjournal.com.

individuals with healthy cognition (HC), both for Hungarian[15–19] and for English native speakers.[20]

The objective of the present study was (1) to explore whether elderly HC individuals with and without T2DM differ in temporal speech characteristics, which may reflect subtle differences in cognition as well; and (2) to also understand how the same temporal speech characteristics compare between MCI patients with and without T2DM.

MATERIALS AND METHODS

Participants

Based on the initial inclusion criteria, a total of 160 individuals were enrolled. After the exclusion process (Fig. 1), 100 of them were eligible for participation. Data collection took place at 2 departments of the Albert Szent-Györgyi Health Center, University of Szeged, Hungary: (1) T2DM patients were recruited at the Division of Diabetology of the Department of Internal Medicine, while (2) nondiabetic subjects were studied at the Memory Clinic of the Department of Psychiatry. The investigation took place within a 25-month time frame between 2018 and 2020.

Participation was voluntary after giving written informed consent. Participants did not receive financial compensation. The study was approved by the Regional Human Biomedical Research Ethics Committee of the University of Szeged, Hungary (231/2017-SZTE). The study was conducted in compliance with the principles of the Declaration of Helsinki.

All participants were evaluated by means of a neuropsychological battery (under *Study protocol* in detail). The battery included the Mini-Mental State Examination (MMSE),[21] which served as the measure of objective cognitive status. Based on the MMSE, participants were classified as either HC (30 to 28 points) or as having MCI (27 to 25 points). Finally, 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 and Exclusion Process

Diabetes-related Criteria

In the T2DM sample, medical diagnosis of T2DM was the initial inclusion criterion. Diagnosis was based on current international guidelines of the American Diabetes Association.[22] Patients with type 1 diabetes mellitus, prediabetes, or chronic hyperglycemia of any other etiology were not enrolled. Average duration of diabetes was 11.4 years (SD = 8.08); treatment was either oral medication (50.9%; n = 26), insulin (25.5%; n = 13), combined oral medication and insulin (17.6%; n = 9), or only diet (5.9%; n = 3).

Other Criteria

For all participants, initial inclusion criteria were a minimum age of 50 years, a minimum of 8 years of formal education, and Hungarian as native language. Exclusion criteria included the following: major hearing problems/deafness, acute depression, dementia, history of substance use disorder, head injuries, major neuropsychiatric disorders, previous computed tomography/magnetic resonance imaging showing evidence of significant abnormality suggesting another

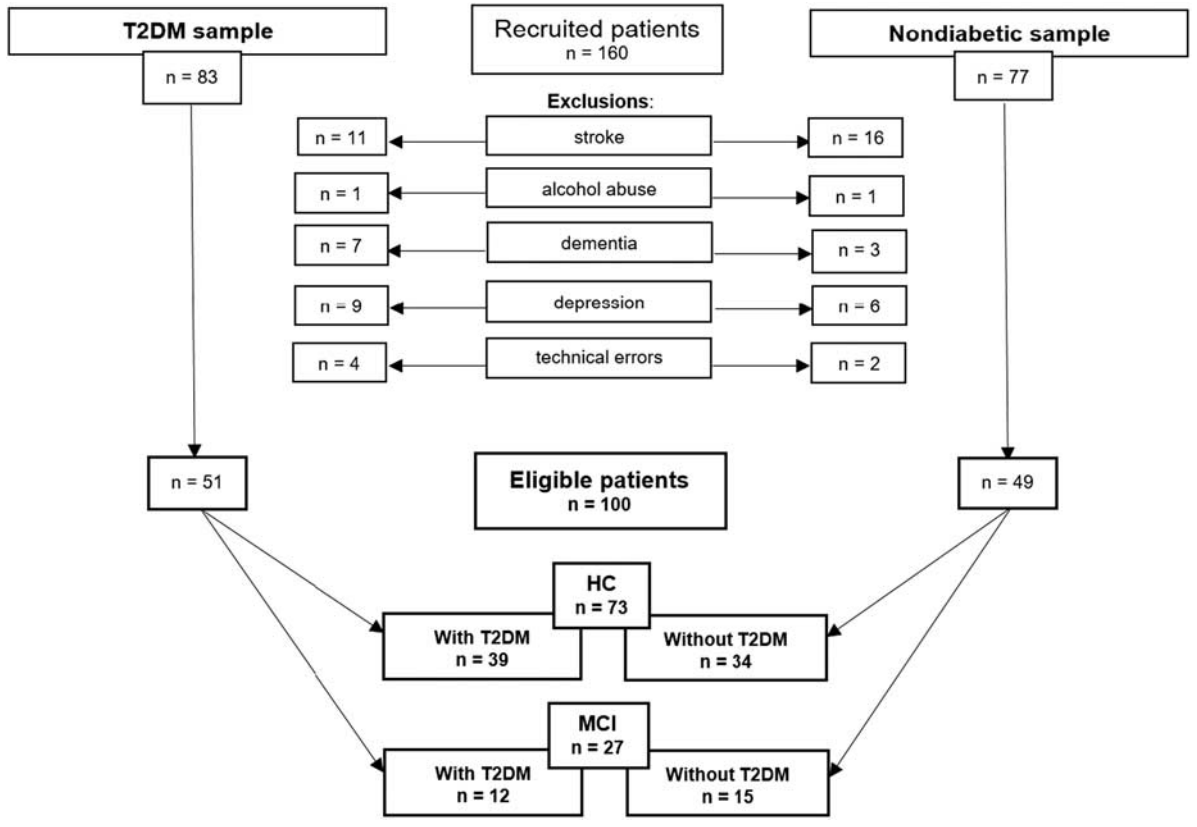

FIGURE 1. Demonstration of the inclusion/exclusion process, and the final sample sizes of the four study groups: HC with and without T2DM; MCI with and without T2DM. HC indicates healthy cognition; MCI, mild cognitive impairment; T2DM, type 2 diabetes mellitus.

potential etiology for MCI (eg, prior macrohemorrhage/microhemorrhages, lacunar infarcts or single large infarct), evidence of cerebral contusion, encephalomalacia, aneurysm, vascular malformations or clinically significant space-occupying lesions. Finally, individuals whose speech could not be properly recorded due to technical errors were also excluded from further analysis (Fig. 1).

To check all inclusion and exclusion criteria, patient history was gathered from an initial interview and from available medical records. Furthermore, dementia and depression were screened on-site at the beginning of the protocol. The MMSE was used for dementia screening, and patients with a score under 25 were excluded. The presence/absence of acute depressive symptoms was evaluated by applying the 15-item Geriatric Depression Scale (GDS-15),[23] with a cut-off score of 6 above which individuals were not considered eligible.

Study Protocol

Neuropsychological Tests

Following a brief demographic and eligibility interview, a neuropsychological test sequence was administered, comprised of 8 instruments. These included 3 test batteries measuring current cognitive state: MMSE, Clock Drawing Test (CDT),[24] and Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog)[25]; 4 tests measuring working memory and executive functions: digit span test forward and backward,[26] nonword repetition test,[27] and listening span test[28]; and one scale for measuring current depressive symptoms: GDS-15. The test order was fixed for all participants and had been assembled to ensure that tasks requiring the same cognitive function were separated (eg, working memory tasks did not directly follow each other).

Speech Task

A speech task was also administered to collect spontaneous (unplanned) speech samples for the temporal speech analysis. This task was chosen because it requires both working and episodic memory, allows remote and repeated testing, and was found to be sensitive in discriminating between MCI and controls.[19] In order to prevent fatigue, this speech task was administered approximately at the 15-minute mark of the 1-hour protocol. Speech was elicited in the following manner: the lead investigator (Investigator 1) told the participant that another researcher (Investigator 2), who was in a different room was to call them on a mobile phone and provide instructions for a new task. Following this cue, Investigator 2 called the participant and after a brief 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." Following the instruction, both Investigator 1 (in the room) and Investigator 2 (on the phone) remained silent until the participant finished the task. The elicited monologue was recorded by a call recorder application installed on the mobile phone.

Speech Sample Preparation and Analysis

The obtained speech recordings were independently screened before analysis by 2 investigators: a linguist specialized in language pathologies (I.H.) screened the overall quality of the recording, while a researcher of computational speech analysis (G.G.) provided technical control. Those recordings that were not of suitable quality (n = 4 in the T2DM, and n = 2 in the nondiabetic groups) were excluded (Fig. 1). The remaining 100 recordings were converted into an uncompressed PCM mono, 16-bit wave format with a sampling rate of 8000 Hz, and were edited in the beginning and at the end so that only the participants' speech remained; the opening/closing formulas and the instructions were removed.

After these preparations, ASR techniques were employed to identify pauses, both silent and filled, in each recording. Pauses were defined as the interruption of speech by either complete silence (silent pause) or by filler words like "um" or "er" (filled pause) lasting longer than 30 ms. The acoustic model was trained on a subset of the BEA audio corpus[29] that consisted of spontaneous speech, as this type of speech is expected to contain filled pauses (for the training of the ASR system, see Gosztolya et al[20]). For training, the speech of 116 speakers was utilized, which amounted to ~44 hours of recordings. This ASR model performed phone-level recognition, with labeling of the input signal (including filled pauses, treated as a special "phoneme") and the output of a phonetic segmentation. Based on the raw parameters from the ASR output, 15 temporal speech parameters were extracted using simple calculations established in previous works of our research group.[16,20] The calculations and definitions of the parameters are available as supplements (Supplemental Digital Content 1, http://links.lww.com/WAD/A379).

Statistical Analysis

Descriptive statistical data are expressed as means, medians, and SD for each group. The Shapiro-Wilk test demonstrated non-normality of data in most scale variables, thus the Mann-Whitney U test was employed to assess between-group differences on demographic data, neuropsychological test scores and temporal speech parameters. For categorical variables, Fisher exact test was applied. To further examine the abilities of each speech parameter in identifying T2DM patients, receiver operating characteristic (ROC) analysis was applied. Sensitivity and specificity (true positive and true negative rate) were calculated using threshold values that yielded the highest possible sensitivity (while keeping specificity above 50%). The level of significance was set at $P < 0.05$ for all statistical tests. Analyses were performed using IBM SPSS 24.0 (SPSS Inc., Chicago, IL).

RESULTS

Demographic and Neuropsychological Characteristics

Demographic and neuropsychological test scores in the HC and MCI groups are presented in Table 1, respectively. Within the HC sample, participants with T2DM and without T2DM did not differ statistically significantly in either of the demographic factors, or any of the neuropsychological tests. However, within the MCI sample, digit span (backwards) performance turned out to be significantly lower among the T2DM patients, compared with the nondiabetic participants.

Temporal Speech Parameters in the HC and MCI Groups According to Diabetic Status

Comparison between the T2DM and the nondiabetic groups was applied both within the HC and within the MCI samples. In the HC sample (Table 2), 5 of the 15 parameters differed significantly, as follows: the HC with T2DM group had shorter utterance length, higher duration rate of silent pause and total pause, and also higher average duration of silent pause and total pause, compared with the HC without T2DM group.

TABLE 1. Descriptive and Comparative Statistics of the Demographic Characteristics and Neuropsychological Test Scores in the HC With and Without T2DM, and the MCI With and Without T2DM Groups, Using the Mann-Whitney *U* Test or Fisher Exact Test (in Italics)

	HC With T2DM (n = 39)			HC Without T2DM (n = 34)			Mann-Whitney *U* Test/Fisher Exact Test		
	M	**Mdn**	**SD**	**M**	**Mdn**	**SD**	**U**	**Z**	***P***
Sex (male/female)		13/26			9/25		—	—	*0.613*
Age (y)	65.31	66.00	8.059	67.74	68.00	6.934	548.000	−1.273	0.203
Education (y)	13.03	12.00	2.748	13.29	12.00	2.505	609.500	−0.608	0.543
MMSE	28.72	29.00	0.647	29.00	29.00	0.778	531.000	−1.582	0.114
CDT	7.62	9.00	3.159	7.50	9.00	3.077	612.000	−0.584	0.559
ADAS-Cog	7.08	6.15	2.989	6.61	6.95	2.608	607.500	−0.435	0.664
Digit span: forward	5.56	5.00	0.995	5.85	5.50	1.158	579.500	−0.975	0.330
Digit span: backward	4.13	4.00	0.894	4.18	4.00	0.999	642.000	−0.243	0.808
Nonword repetition	5.18	5.00	1.715	4.74	5.00	1.620	552.000	−1.275	0.202
Listening span	2.53	2.60	0.583	2.75	2.85	0.602	504.500	−1.782	0.075
GDS-15	2.00	1.00	1.717	2.00	2.00	1.595	645.000	−0.205	0.838
	MCI with T2DM (n = 12)			MCI without T2DM (n = 15)			Mann-Whitney *U* Test/Fisher Exact Test		
	M	Mdn	SD	M	Mdn	SD	U	Z	*P*
Sex (male/female)		2/10			5/10		—	—	*0.408*
Age (y)	70.42	73.50	9.120	72.60	74.00	6.311	83.500	−0.318	0.755
Education (y)	11.17	11.50	2.855	11.73	12.00	2.865	76.000	−0.712	0.516
MMSE	26.17	26.00	0.835	26.27	26.00	0.799	84.000	−0.315	0.792
CDT	5.50	4.50	3.529	7.33	8.00	2.870	64.000	−1.281	0.217
ADAS-Cog	9.38	9.00	2.070	10.61	10.60	3.104	64.000	−1.271	0.217
Digit span: forward	5.00	5.00	1.128	5.33	5.00	0.617	60.500	−1.668	0.152
Digit span: backward	3.25	3.00	0.754	3.93	4.00	0.799	49.000	−2.161	**0.047**
Nonword repetition	3.58	5.00	2.575	3.67	4.00	1.718	81.000	−0.450	0.683
Listening span	2.32	2.15	0.476	2.23	2.30	0.434	87.000	−0.151	0.905
GDS-15	1.92	2.00	1.505	2.53	2.00	1.187	62.000	−1.436	0.183

The *P*-values indicating statistically significant differences (at the $P < 0.05$ level) are in bold.

ADAS-Cog indicates Alzheimer's Disease Assessment Scale-Cognitive Subscale; CDT, Clock Drawing Test; GDS-15, 15-item Geriatric Depression Scale; HC, healthy cognition; *M*, mean; MCI, mild cognitive impairment; Mdn, median; MMSE, Mini-Mental State Examination; T2DM, type 2 diabetes mellitus.

A subsequent ROC analysis was executed in order to explore if HC with T2DM patients could be discriminated from HC without T2DM participants, based on their temporal speech parameters. The results showed that the same 5 parameters demonstrated significant classification potential, with utterance length having the highest area

TABLE 2. Descriptive and Comparative Statistics of the HC With and Without T2DM Groups Using the Mann-Whitney *U* Test

	HC With T2DM (n = 39)			HC Without T2DM (n = 34)			Mann-Whitney *U* Test		
Temporal Speech Parameters	**M**	**Mdn**	**SD**	**M**	**Mdn**	**SD**	**U**	**Z**	***P***
Utterance length (s)	114.00	93.36	68.274	205.68	151.88	235.281	407.000	−2.831	**0.005**
Articulation tempo (1/s)	9.27	9.49	1.907	9.65	9.68	2.001	602.000	−0.675	0.500
Speech tempo (1/s)	10.05	10.30	1.872	10.46	10.48	1.850	597.000	−0.730	0.465
Occurrence rates of pauses									
Silent pause (%)	5.55	5.35	1.562	5.29	4.83	2.458	536.000	−1.404	0.160
Filled pause (%)	2.57	2.15	1.613	3.09	2.56	2.123	573.000	−0.995	0.320
Total pause (%)	8.11	7.32	2.642	8.38	7.41	4.268	639.000	−0.265	0.791
Duration rates of pauses									
Silent pause (%)	32.16	29.40	10.991	25.79	24.13	10.850	429.000	−2.588	**0.010**
Filled pause (%)	5.81	5.04	4.054	6.92	6.03	3.940	556.000	−1.183	0.237
Total pause (%)	37.97	37.90	11.495	32.71	30.79	12.700	474.000	−2.090	**0.037**
Frequency of pauses									
Silent pause (1/s)	0.53	0.53	0.101	0.52	0.48	0.142	580.000	−0.918	0.359
Filled pause (1/s)	0.24	0.23	0.140	0.30	0.27	0.150	516.000	−1.626	0.104
Total pause (1/s)	0.78	0.74	0.174	0.82	0.78	0.241	620.000	−0.476	0.634
Average durations of pauses									
Silent pause (s)	0.62	0.55	0.248	0.50	0.46	0.169	453.000	−2.322	**0.020**
Filled pause (s)	0.22	0.20	0.072	0.22	0.22	0.056	590.500	−0.802	0.423
Total pause (s)	0.50	0.45	0.164	0.41	0.37	0.128	419.000	−2.698	**0.007**

The *P*-values indicating statistically significant differences (at the $P < 0.05$ level) are in bold.

HC indicates healthy cognition; M, mean; Mdn, median; T2DM, type 2 diabetes mellitus.

TABLE 3. Descriptive and Comparative Statistics of the MCI With and Without T2DM Groups Using the Mann-Whitney *U* Test

	MCI With T2DM (n = 12)			MCI Without T2DM (n = 15)			Mann-Whitney *U* test		
Temporal Speech Parameters	**M**	**Mdn**	**SD**	**M**	**Mdn**	**SD**	**U**	**Z**	***P***
Utterance length (s)	119.50	80.10	93.150	131.70	79.40	139.058	83.000	−0.342	0.755
Articulation tempo (1/s)	9.26	9.64	2.644	8.76	8.20	1.703	76.000	−0.683	0.516
Speech tempo (1/s)	9.99	10.37	2.555	9.57	9.09	1.582	77.000	−0.634	0.548
Occurrence rates of pauses									
Silent pause (%)	5.77	5.74	2.504	5.73	5.47	1.841	88.000	−0.098	0.943
Filled pause (%)	2.19	2.74	1.344	3.13	2.72	2.009	67.000	−1.122	0.277
Total pause (%)	7.97	7.98	3.445	8.85	8.63	3.272	77.000	−0.634	0.548
Duration rates of pauses									
Silent pause (%)	33.94	32.68	16.602	31.93	28.69	7.933	89.000	−0.049	0.981
Filled pause (%)	4.41	5.03	2.883	6.84	7.65	4.474	62.000	−1.366	0.183
Total pause (%)	38.35	36.75	17.231	38.77	36.57	10.476	86.000	−0.195	0.867
Frequency of pauses									
Silent pause (1/s)	0.52	0.53	0.128	0.53	0.54	0.135	88.000	−0.098	0.943
Filled pause (1/s)	0.20	0.19	0.129	0.28	0.27	0.152	62.000	−1.366	0.183
Total pause (1/s)	0.73	0.77	0.204	0.81	0.78	0.214	73.000	−0.830	0.427
Average durations of pauses									
Silent pause (s)	0.64	0.56	0.255	0.62	0.62	0.164	82.000	−0.390	0.719
Filled pause (s)	0.21	0.21	0.041	0.24	0.23	0.097	76.000	−0.683	0.516
Total pause (s)	0.53	0.47	0.210	0.49	0.49	0.099	87.000	−0.146	0.905

The *P*-values indicating statistically significant differences (at the $P < 0.05$ level) are in bold.
M indicates mean; MCI, mild cognitive impairment; Mdn, median; T2DM, type 2 diabetes mellitus.

under the curve (AUC) (0.693) and the average duration of total pause yielding the highest sensitivity (79.5%). Sensitivity and specificity measures of temporal parameters were derived from ROC analysis; parameters with an AUC above 0.600 are shown in Table 4.

However, regarding the MCI sample (Table 3), no statistically significant differences could be detected between the with and the without T2DM subgroups. This was further consolidated by the subsequent ROC analysis, which revealed that none of the 15 temporal parameters had statistically significant abilities to discriminate MCI with T2DM from MCI without T2DM participants. Nevertheless, parameters concerning filled pauses produced the highest AUCs. Sensitivity and specificity measures of temporal parameters were derived from ROC analysis; parameters with an AUC above 0.600 are shown in Table 4.

Correlations of Temporal Speech Parameters With Age and Education

Regarding the relationship between age and the 15 temporal speech parameters across the 4 groups, correlation was statistically significant for articulation tempo (HC with T2DM: $\tau_b = -0.221$, $P = 0.050$), for speech tempo (HC with T2DM: $\tau_b = -0.229$, $P = 0.042$), and for silent pause frequency (MCI without T2DM: $\tau_b = 0.390$, $P = 0.046$). With regards to education, weak to moderate but statistically significant correlations were found with utterance length (HC without T2DM: $\tau_b = 0.269$, $P = 0.035$; MCI with T2DM:

TABLE 4. Accuracy Measures of Temporal Parameters With AUC Above 0.600 in the HC and the MCI Samples, Respectively (Containing Both the "With T2DM" and "Without T2DM" Subgroups), Using Receiver Operating Characteristic (ROC) Analysis

HC Groups (With vs. Without T2DM)			**Accuracy Measures**				
Temporal Speech Parameters	***P***	**AUC**	**95% CI−**	**95% CI+**	**Threshold Value**	**Sensitivity (%)**	**Specificity (%)**
Utterance length (s)	**0.005**	0.693	0.572	0.815	131.845	74.4	61.8
Total pause average duration (s)	**0.007**	0.684	0.560	0.808	0.374	79.5	55.9
Silent pause duration rate (%)	**0.010**	0.676	0.553	0.800	24.192	74.4	52.9
Silent pause average duration (s)	**0.020**	0.658	0.532	0.785	0.471	74.4	55.9
Total pause duration rate (%)	**0.037**	0.643	0.514	0.771	31.705	66.7	55.9
Filled pause frequency (1/s)	0.104	0.611	0.481	0.740	0.246	61.5	58.8
MCI groups (with vs. without T2DM)			Accuracy measures				
Temporal speech parameters	*P*	AUC	95% CI−	95% CI+	Threshold value	Sensitivity (%)	Specificity (%)
Filled pause duration rate (%)	0.172	0.656	0.446	0.865	6.754	83.3	53.3
Filled pause frequency (1/s)	0.172	0.656	0.443	0.868	0.229	66.7	60.0
Filled pause occurrence rate (%)	0.262	0.628	0.408	0.848	2.715	50.0	53.3

The *P*-values indicating statistically significant classification abilities (at the $P < 0.05$ level) are in bold.
AUC indicates area under the curve; CI, confidence interval; HC, healthy cognition; MCI, mild cognitive impairment; ROC, receiver operating characteristic; T2DM, type 2 diabetes mellitus.

$\tau_b = 0.478$, $P = 0.044$), articulation tempo (MCI with T2DM: $\tau_b = 0.478$, $P = 0.044$), speech tempo (MCI with T2DM: $\tau_b = 0.546$, $P = 0.021$), filled pause occurrence rate (HC with T2DM: $\tau_b = 0.274$, $P = 0.022$), filled pause duration rate (HC with T2DM: $\tau_b = 0.268$, $P = 0.025$; MCI without T2DM: $\tau_b = 0.596$, $P = 0.004$), silent pause average duration (MCI with T2DM: $\tau_b = -0.580$, $P = 0.014$), filled pause average duration (MCI without T2DM: $\tau_b = -0.618$, $P = 0.003$), and total pause average duration (MCI with T2DM: $\tau_b = -0.615$, $P = 0.010$). The comprehensive table containing all correlations is available as supplement (Supplemental Digital Content 1, http://links.lww.com/WAD/A379).

DISCUSSION

To the best of our knowledge, this was the first study that investigated the speech of T2DM patients with the purpose of detecting signs of subtle cognitive deficits that can manifest as changes in the temporal characteristics of speech. The major finding was that the speech of elderly HC individuals with T2DM compared significantly worse on several temporal characteristics to that of age-matched and education-matched HC individuals without T2DM.

Firstly, we intended to study the temporal speech characteristics of elderly T2DM patients who have been classified as HC based on traditional neuropsychological screening. Our results showed that their speech contains more signs of subtle, underlying cognitive deficits than that of the HC subjects without T2DM. Namely, 5 of 15 temporal speech parameters showed statistically significant differences between the diabetic and nondiabetic groups: HC with T2DM patients had shorter utterance length, higher duration rate of silent pause and total pause, and also higher average duration of silent pause and total pause compared to HC without T2DM participants. [Although it was not the focus of the present study, it is interesting to note that the temporal speech parameters that differentiated between the HC with/without T2DM groups also showed different mean/median values within the nondiabetic sample, between HC and MCI (Table 2 vs. Table 3). This further highlights that from the full set of 15 parameters these would have the most discriminative potential in future clinical applications.]

These differences are in agreement with the results of previous studies using the S-GAP Test and other speech analysis methods: in earlier works, more or longer pauses (signs of disfluency, word-finding difficulties and decreased lexical access) had been reported in the speech of patients with varying levels of cognitive impairment, for example, due to AD,[11,30,31] MCI,[12,13] or Parkinson disease.[32,33] These results, now complemented by the findings of the present study, confirm that pauses in speech provide a highly valuable source of information regarding language functions and thus cognitive state, especially in the introductory stages of neurocognitive disorders when other cognitive domains measured by traditional test batteries have not yet deteriorated in such a magnitude to be detected. In the case of T2DM patients, these subtle cognitive changes may be explained by diabetes-associated changes in the brain, such as impaired insulin signaling, neuronal insulin resistance, inflammation, mitochondrial dysfunction, vascular damage, or disturbances in synaptic plasticity, all of which can lead to an onset of cognitive decline.[7,34,35]

Furthermore, we also compared the temporal speech characteristics of MCI patients with and without T2DM. No significant differences could be detected in any of the 15 analyzed temporal speech parameters, suggesting that these two groups performed similarly. A possible explanation for this could be that the pathophysiological processes in the brain are facilitated by T2DM and, as a consequence, cognitive abilities gradually deteriorate. According to current medical protocol, MCI diagnosis is only given when, besides fulfilling other criteria, cognitive symptoms reach a measurable level and can be confirmed by an objective evaluation tool.[36,37] However, it has been reported that the underlying cognitive deterioration is usually present for a longer period, more or less without clinical symptoms.[38] It could be argued that in the case of T2DM patients, the *onset* of the latent phase of transitioning from HC to MCI might take place earlier, and speech disfluencies might precede the more robust symptoms by a longer period of time than in the case of nondiabetic subjects. Our results also indicate that the temporal speech characteristics of T2DM and nondiabetic subjects tend to be similar when the cognitive deterioration reaches the level of MCI, which would suggest that once the transition to MCI has manifested, the presence of T2DM may not necessarily exacerbate 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, age showed a statistically significant, weak correlation with 3 parameters: a negative correlation with articulation tempo and speech tempo, and a positive correlation with silent pause frequency. Education weakly to moderately correlated with 8 parameters: positively with utterance length, articulation tempo, speech tempo, filled pause occurrence rate, filled pause duration rate, and filled pause average duration; and negatively with silent pause average duration and total pause average duration. Careful examination of the positive and negative directions of the statistically significant correlations reveals that the increased presence of silent pauses (higher frequency or average length) was aligned with the demographic risk factors of cognitive decline (lower education, higher age[39,40]). In contrast, the ability to produce more and faster speech (longer utterance length, higher articulation and speech tempo) was more associated with lower dementia-risk (such as higher education and lower age[39]).

Limitations of the present study include the small number of MCI individuals which might reduce the statistical power of the comparisons, and therefore could contribute to the lack of between-group differences within the MCI sample. As this research was a pilot study for identifying speech parameters with the highest differentiating potential for future telemedicine-based assessments, multiple correction testing was not applied for the statistical comparisons. This needs to be taken into account when interpreting the results. On another note, even though the sampling rate used for speech recording (8000 Hz) might seem relatively low, the S-GAP Test was specifically intended to be applied in real-life settings, potentially in the form of a mobile application. A minimum sampling rate of 8000 Hz is available on most mobile phone devices, enabling wider adoption of the technology. Also, future works could also involve more diabetes-related medical characteristics, which could enable the creation of subgroups based on, for example, diabetes severity, glycemic control, or insulin levels.

The utilization of telemedicine in the management of diabetes is a dynamically emerging area, however, to this date no

such technique has been used for the cognitive examination of diabetic patients. A subtle speech deficiency detected by the S-GAP Test could be an indication for a thorough medical and neuropsychological examination to search for possible underlying causes or for monitoring the patient more closely, for example, with frequent check-ups. Remote assessment is gaining increasing significance in light of the current COVID-19 pandemic, with every medical field facing restrictions of face-to-face appointments. The S-GAP Test is currently being developed in a mobile application format which could serve as a rapid, cost-effective, noninvasive, and no-contact form of cognitive screening for the elderly and, according to the present results, could be implemented for monitoring T2DM patients as well.

CONCLUSIONS

We explored the speech of T2DM patients, building on the shared pathophysiology of T2DM and neurocognitive disorders, as well as the strong association between cognitive deterioration and speech deficits. Even though T2DM patients classified as HC and matched nondiabetic subjects performed similarly on global cognitive and traditional neuropsychological tests, we demonstrated that the speech of T2DM patients contained an increased number and length of silent pauses compared to the nondiabetic group. Therefore, we would suggest that temporal analysis of speech offers a sensitive and ecologically valid tool for monitoring cognitive state in the early, introductory stages of cognitive impairments, and it could be useful for identifying the T2DM individuals who are more at risk of developing manifest MCI or later, dementia.

REFERENCES

1. Cukierman T, Gerstein HC, Williamson JD. Cognitive decline and dementia in diabetes—systematic overview of prospective observational studies. *Diabetologia*. 2005;48:2460–2469.
2. Sadanand S, Balachandar R, Bharath S. Memory and executive functions in persons with type 2 diabetes: a meta-analysis. *Diabetes Metab Res Rev*. 2016;32:132–142.
3. Ahtiluoto S, Polvikoski T, Peltonen M, et al. Diabetes, Alzheimer-disease, and vascular dementia: a population-based neuropathologic study. *Neurology*. 2010;75:1195–1202.
4. Degen C, Toro P, Schönknecht P, et al. Diabetes mellitus type II and cognitive capacity in healthy aging, mild cognitive impairment and Alzheimer's disease. *Psychiatry Res*. 2016;240:42–46.
5. Jekel K, Damian M, Wattmo C, et al. Mild cognitive impairment and deficits in instrumental activities of daily living: a systematic review. *Alzheimers Res Ther*. 2015;7:1–20.
6. Saeedi P, Petersohn I, Paraskevi S, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract*. 2019;157:107843.
7. Biessels GJ, Staekenborg S, Brunner E, et al. Risk of dementia in diabetes mellitus: a systematic review. *Lancet Neurol*. 2006;5:64–74.
8. McCrimmon RJ, Ryan CM, Frier BM. Diabetes and cognitive dysfunction. *Lancet*. 2012;379:2291–2299.
9. Szatloczki G, Hoffmann I, Vincze V, et al. Speaking in Alzheimer's disease, is that an early sign? Importance of changes in language abilities in Alzheimer's disease. *Front Aging Neurosci*. 2015;7:1–7.
10. Mortensen L, Meyer AS, Humphreys GW. Age-related effects on speech production: a review. *Lang Cogn Process*. 2006;21:238–290.
11. Hoffmann I, Nemeth D, Dye CD, et al. Temporal features of spontaneous speech in Alzheimer's disease. *Int J Speech Lang Pathol*. 2010;12:29–34.
12. Roark B, Mitchell M, Hosom JP, et al. Spoken language derived measures for detecting mild cognitive impairment. *IEEE Trans Audio Speech Lang Processing*. 2011;19:2081–2090.
13. Meilán JJG, Martínez-Sánchez F, Martínez-Nicolás I, et al. Changes in the rhythm of speech difference between people with nondegenerative mild cognitive impairment and with preclinical dementia. *Behav Neurol*. 2020;4683573:1–10.
14. König A, Satt A, Sorin A, et al. Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimers Dement*. 2015;1:112–124.
15. Tóth L, Gosztolya G, Vincze V, et al. Automatic detection of mild cognitive impairment from spontaneous speech using ASR. Proceedings of Interspeech, Dresden, Germany. 2015: 2694–2698.
16. Tóth L, Hoffmann I, Gosztolya G, et al. A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech. *Curr Alzheimer Res*. 2018; 15:130–138.
17. Gosztolya G, Tóth L, Grósz T, et al. Detecting mild cognitive impairment from spontaneous speech by correlation-based phonetic feature selection. Proceedings of Interspeech, San Francisco, USA. 2016:107–111.
18. Gosztolya G, Vincze V, Tóth L, et al. Identifying mild cognitive impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Comput Speech Lang*. 2019;53:181–197.
19. Vincze V, Szatlóczki G, Tóth L, et al. Telltale silence: temporal speech parameters discriminate between prodromal dementia and mild Alzheimer's disease. *Clin Linguist Phon*. 2021;35:727–742.
20. Gosztolya G, Balogh R, Imre N, et al. Cross-lingual detection of mild cognitive impairment based on temporal parameters of spontaneous speech. *Comput Speech Lang*. 2021;69:101215.
21. 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;12:189–198.
22. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diab Care*. 2014;37:S81–S90.
23. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS)—recent evidence and development of a shorter version. *Clin Gerontol*. 1986;5:165–173.
24. Shulman KI, Shedletsky R, Silver IL. The challenge of time: clock-drawing and cognitive function in the elderly. *Int J Geriatric Psychiatry*. 1986;1:135–140.
25. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984;141:1356–1364.
26. Wechsler D. *The Wechsler Adult Intelligence Scale—Revised*. New York, NY: The Psychological Corporation; 1981.
27. Gathercole SE, Willis CS, Baddeley AD, et al. The children's test of nonword repetition: a test of phonological working memory. In: Gathercole SE, McCarthy RA, eds. *Memory Tests and Techniques*. Hove: Lawrence Erlbaum Associates; 1994: 103–127.
28. Daneman M, Carpenter PA. Individual differences in working memory and reading. *J Verbal Learning Verbal Behav*. 1980; 19:450–466.
29. Neuberger T, Gyarmathy D, Gráczi TE, et al. Development of a large spontaneous speech database of agglutinative Hungarian language. In: Sojka P, Horák A, Kopeček I, eds. *Text, Speech and Dialogue*. Brno, Czech Republic: Springer; 2014: 424–431.
30. López-de-Ipiña K, Alonso J-B, Travieso CM, et al. On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors (Basel)*. 2013;13:6730–6745.
31. Martínez-Sánchez F, Meilán JJG, García-Sevilla J. Oral reading fluency analysis in patients with Alzheimer disease and asymptomatic control subjects. *Neurología*. 2013;28:325–331.
32. Hlavnička J, Čmejla R, Tykalová T, et al. Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Sci Rep*. 2017;7:1–12.
33. Alvar AM, Lee J, Huber JE. Filled pauses as a special case of automatic speech behaviors and the effect of Parkinson's disease. *Am J Speech Lang Pathol*. 2019;28(2S):835–843.

34. Bello-Chavolla OY, Antonio-Villa NE, Vargas-Vázquez A, et al. Pathophysiological mechanisms linking type 2 diabetes and dementia: review of evidence from clinical, translational and epidemiological research. *Curr Diabetes Rev*. 2019;15: 456–470.
35. Stranahan AM, Arumugam TV, Cutler RG, et al. Diabetes impairs hippocampal function through glucocorticoid-mediated effects on new and mature neurons. *Nat Neurosci*. 2008;11:309–317.
36. Petersen RC, Smith GE, Warning SC, et al. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol*. 1999;56:303–308.
37. Petersen RC, Doody R, Kurz A, et al. Current concepts in mild cognitive impairment. *Arch Neurol*. 2001;58:1985–1992.
38. Albert M, Yuxin Z, Moghekar A, et al. Predicting progression from normal cognition to mild cognitive impairment for individuals at 5 years. *Brain*. 2018;141:877–887.
39. Luck T, Luppa M, Briel S, et al. Incidence of mild cognitive impairment: a systematic review. *Dement Geriatr Cogn Disord*. 2010;29:164–175.
40. Patterson C, Feightner J, Garcia A, et al. General risk factors for dementia: a systematic evidence review. *Alzheimers Dement*. 2007;3:341–347.

