CLINICAL LINGUISTICS & PHONETICS
https://doi.org/10.1080/02699206.2020.1827043

Taylor & Francis
Taylor & Francis Group

Telltale silence: temporal speech parameters discriminate between prodromal dementia and mild Alzheimer's disease

Veronika Vincze[a], Gréta Szatlóczki[b], László Tóth[c], Gábor Gosztolya[a], Magdolna Pákáski[b], Ildikó Hoffmann[d,e], and János Kálmán[b]

[a]MTA-SZTE Research Group on Artifical Intelligence, Szeged, Hungary; [b]Department of Psychiatry, University of Szeged, Szeged, Hungary; [c]Institute of Informatics, University of Szeged, Szeged, Hungary; [d]Department of Linguistics, University of Szeged, Szeged, Hungary; [e]Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, Hungary

ABSTRACT
This study presents a novel approach for the early detection of mild cognitive impairment (MCI) and mild Alzheimer's disease (mAD) in the elderly. Participants were 25 elderly controls (C), 25 clinically diagnosed MCI and 25 mAD patients, included after a clinical diagnosis validated by CT or MRI and cognitive tests. Our linguistic protocol involved three connected speech tasks that stimulate different memory systems, which were recorded, then analyzed linguistically by using the PRAAT software. The temporal speech-related parameters successfully differentiate MCI from mAD and C, such as speech rate, number and length of pauses, the rate of pause and signal. Parameters pauses/duration and silent pauses/duration linearly decreased among the groups, in other words, the percentage of pauses in the total duration of speech continuously grows as dementia progresses. Thus, the proposed approach may be an effective tool for screening MCI and mAD.

ARTICLE HISTORY
Received 4 February 2020
Revised 11 September 2020
Accepted 18 September 2020

KEYWORDS
Mild cognitive impairment; Alzheimer's disease; connected speech; differential diagnosis; temporal speech parameters

Introduction

The diagnostic criteria of Alzheimer's disease (AD) and its prodromal stage Mild Cognitive Impairment (MCI) have been known for a long time (American Psychiatric Association, 2013; Petersen, 2004), nevertheless, they often remain undetected early since patients and relatives tend to ignore the first clinical manifestations. Although memory decline is the main symptom of AD, language impairment can also be an important marker. It is known that AD influences the temporal characteristics of speech (F. F. Martínez-Sánchez et al., 2018; Fraser et al., 2016; Hoffmann et al., 2010; De Ipiña et al., 2015; Meilán et al., 2018, 2012, 2020; Tóth et al., 2018) and these phonetic changes occur even in mild AD (mAD) (Forbes-McKay & Venneri, 2005; Kato et al., 2018; König et al., 2015; Laske et al., 2015; Misiewicz et al., 2018) and in patients with MCI (Kato et al., 2018; Roark et al., 2011; Satt et al., 2014). In these earlier studies, quite different methods were used to obtain the speech output needed for linguistic analysis. For example, some authors examined the connected speech derived from telling pleasant stories or interacting with participants in friendly conversations about topics related to hometown, childhood, etc. (De Ipiña et al., 2015; Kato

CONTACT Veronika Vincze, vinczev@inf.u-szeged.hu MTA-SZTE Research Group on Artifical Intelligence, Szeged, Hungary

et al., 2018). Others preferred the analysis of connected speech by applying the picture description task (Forbes-McKay & Venneri, 2005; Fraser et al., 2016; König et al., 2015). Automatic speech recognition (ASR) can effectively contribute to the analysis of such recordings. Our earlier experiments showed that an acoustic feature set focusing on articulation rate, speech tempo and other descriptors of silent and filled pauses present in the speech is useful for detecting both MCI and mAD (Gosztolya et al., 2016; Tóth et al., 2015, 2018). However, it is important to know whether the circumstances of speech production influence the sensitivity of linguistic analysis. To date, no study has been performed comparing the efficiency of speech tasks involving different memory functions or the immediate and delayed recall. In the case of delayed recall, where a distractor is also applied, the cognitive load on the working memory increases, hence recall requires more cognitive efforts, resulting in longer pauses and less continuous speech, which may be indicative of cognitive impairment.

A further question arises whether the acoustic parameters of speech induced by different tasks show specific characteristics suitable for separating MCI or mAD. Here, we seek to identify speech parameters that can effectively distinguish prodromal from mAD. For this purpose, we will carry out a temporal analysis of different task-restricted connected speech samples produced by MCI and mAD patients, compared with data from elderly controls.

The relationship of language and cognitive functions in aging

There are several psycho- and neurolinguistic approaches to speech planning and production. Psycholinguistic studies on healthy individuals have indicated that differences in cognitive abilities and in the working memory may affect speech production and planning because speakers are able to adapt the scope of speech production planning depending on their cognitive abilities (Ferreira & Swets, 2002; Levelt & Meyer, 2000; Swets et al., 2014).

Another direction of research is the investigation of lexical semantic memory (Balthazar et al., 2007). In cognitive tests, controls, amnestic MCI and mAD patients showed gradually decreasing cognitive ability. This indicates that verbal fluency might have been influenced by short-term memory. As the disease progresses, other cortical areas including the temporal cortex are involved, which could explain the deterioration of semantic knowledge in mAD. It has been shown that amnestic MCI impaired episodic memory, but the lexical semantic system was unaffected, however, the latter might be affected later, in the early phase of AD. In summary, deficits in language and memory functions, especially in semantic memory, are commonly found in patients even with mAD (Szatlóczki et al., 2015).

Another study found that language and memory functions are closely related in AD, since linguistic functioning requires proper memory functions. Difficulties in speech production, speech comprehension and memory functions may overlap. Senile changes in language comprehension and expression impose a decline in global speech performance, which hinders active vocabulary (Kempler, 2004). In a retrospective study, the relationship between certain language functions and cognitive impairment in AD was estimated by the Mini-Mental State Examination (MMSE) test. A significant relationship was found between MMSE scores and all language measures except hyperfluency. Impairment in language fluency is common, especially in the case of impaired cognitive and global performance (Weiner et al., 2008). It has also been shown that patients with both AD and MCI have difficulties in performing tasks that require semantic knowledge, such as naming, and verbal fluency (Jarrold et al., 2014; Quaranta et al., 2019; Roark et al., 2011). These symptoms appear early on, and they increase during the course of the disease, suggesting early and progressive impairment of the semantic memory of these patients (Nebes et al., 1989). However, according to a recent systematic review of connected speech changes in AD as assessed by picture description tasks, results are more robust at later stages of AD, but are more fragile and inconclusive at the MCI stage (Slegers et al., 2018).

In a post mortem study, the number of synapses was estimated in the stratum radiatum of the human hippocampal CA1 subfield, – which is a region linked to learning and memory – for MCI and mAD patients and a control group with no cognitive deficits. According to the results, the total number of synapses showed a correlation with several cognitive tests including immediate and delayed recall (Scheff et al., 2007). Another study also indicated significant differences in some delayed recall tasks (phonological fluency, Rey Auditory Verbal Learning tests) for healthy elderly, MCI and mAD groups (Barban et al., 2016). All in all, the number of synapses seems to decrease with the progress of cognitive impairment, which affects the performance on cognitive tasks such as immediate or delayed recall.

Based on the above, memory functions are strongly influenced by language production (Balthazar et al., 2007; Kempler, 2004; Swets et al., 2014; Weiner et al., 2008), and speakers' cognitive abilities affect their speech production (Ferreira & Swets, 2002; Levelt & Meyer, 2000). Cognitive abilities may show a linear decline in pathological aging (Balthazar et al., 2007; Barban et al., 2016; Scheff et al., 2007), hence it is worth investigating whether this decline has any effect on the change in linguistic factors.

Speech-based aspects of MCI and mAD

Speech analysis has already been used in patients with dementia (Canning et al., 2004; Forbes-McKay & Venneri, 2005) to find potential vocal markers of the disease. The most characteristic speech-based features of MCI are the lengthening of hesitations with regard to the total length of speech and a slower speech rate (Hoffmann et al., 2010; Jarrold et al., 2014; Roark et al., 2011; Satt et al., 2014). Moreover, patients diagnosed with MCI often have problems with finding the correct word in connected speech (Fraser et al., 2014; Garrard et al., 2014). Moreover, it has been shown that connected speech production is consistently reported to be affected in early AD, and therefore, hesitations and pauses are useful indicators to detect dementia (Vincze et al., 2016). Further affected speech characteristics in patients with AD seem to be related to articulation (König et al., 2015), prosody in terms of temporal and acoustic measures, which includes alterations in rhythm (Horley et al., 2010; Martínez-Sánchez et al., 2012), and eventually, in later stages, phonological fluency (Henry et al., 2004).

There is no general consensus among researchers concerning which parameters of spontaneous or connected speech can effectively distinguish the group of mAD patients from controls. Some studies highlight the role of articulation rate, speech rate and the rate of hesitations in the total amount of speech (Hoffmann et al., 2010), while others emphasise the importance of articulation rate, speech rate and variations in syllable timing (Martínez-Sánchez et al., 2013). Some other studies focus on the analysis of pauses in speech (López-de-Ipiña et al., 2013). Overall, there is a consensus across languages that the temporal analysis of speech can be used effectively to detect mAD; however, it is still unclear what speech task and what temporal parameters of speech are the most effective predictors. In our study, we seek to identify these parameters.

Aims and hypothesis

The goal of this study is to compare the speech abilities of three groups, namely the cognitively healthy elderly, patients with MCI and patients with mAD. More specifically, we seek to find out whether speech features indicating decline in the elderly could be used to diagnose the MCI group **(H1)**. It is also an important goal to identify the optimal speech task that provides the best recognition of MCI or mAD. We hypothesise that a speech task involving episodic memory function (describing events from the previous day) results in a more reliable detection of MCI patients during the analysis of connected speech than film description tasks (Hoffmann et al., 2010; Jarrold et al., 2014; Roark et al., 2011; Satt et al., 2014). Here we predict that speech rate can effectively differentiate the groups examined **(H2)**. Moreover, pauses can effectively separate MCI patients from healthy controls and mAD patients (Hoffmann et al., 2010; López-de-Ipiña et al., 2013; Tóth et al., 2015). We would like to stress that filled pauses and silent pauses are treated independently in our analysis. For the duration of both types (filled and silent pauses) significant differences are supposed to exist between groups **(H3)**. Moreover, we aim to prove that some novel indicators (e.g., the ratio of pauses/duration, silent pauses/duration and filled pauses/duration) can also show statistically significant differences between the groups **(H4)**.

Materials and methods

Participants

A total of 75 subjects participated in the study: 25 MCI patients, 25 mAD patients and 25 cognitively healthy controls. All the subjects with MCI and with mAD were right-handed and native speakers of Hungarian. The exclusion criteria for controls were drug or alcohol consumption; being under pharmacological treatment affecting cognitive functions; a medical history of head injuries, depression or psychosis, and visual or auditory deficits. MCI and mAD patients were selected after a medical diagnosis supported by neurological and psychiatric examination, CT or MRI and cognitive ability tests. Types of MCI were not considered here so as to have a sufficient number of patients in each group. Patients indicating any signs of dementia of some other origin were not enrolled in this study.

The following clinical tests were applied to assess the cognitive state of the subjects: Mini-Mental State Examination (MMSE; Folstein et al., 1975; standardised in Hungarian by Janka et al., 1988), Clock Drawing Test (CDT; Manos & Wu, 1994; standardised in Hungarian by Kálmán et al., 1995) and Alzheimer's Disease Assessment Scale (ADAS-Cog; Rosen et al., 1984; standardised in Hungarian by Pákáski et al., 2012). Performance scores in these tests revealed significant differences for all three groups based on ANOVA. Another exclusion criterion was a score above 10 in the Geriatric Depression Scale (Millian-Morell et al., 2018). The three groups ($F(2;74) = 2.202$; $p = .118$) were aligned with regard to gender ($\chi^2(2) = 1.389$; $p = .499$), age ($F(2;74) = 2.321$; $p = .105$) and years of education ($F(2;74) = 2.202$; $p = .118$). The clinical characteristics of the control, the MCI and the mAD group are summarised in Table 1.

Table 1. The characteristics of the three groups of the study participants.

	Control (n = 25)	MCI (n = 25)	mAD (n = 25)	Statistics (F and *p*-values)
Age (mean ± SD)	70.72 ± 5.004	72.4 ± 3.594	73.96 ± 6.846	F(2;74) = 2.321 $p = .105$
Years of education (mean ± SD)	12.08 ± 2.326	10.84 ± 2.304	10.76 ± 2.818	F(2;74) = 2.202 $p = .118$
MMSE (mean ± SD)	29.24 ± 0.523	27.16 ± 0.898	23.92 ± 2.488	F(2;74) = 76.213 $p < .001$
CDT (mean ± SD)	8.88 ± 2.007	6.44 ± 3.429	5.88 ± 3.244	F(2;74) = 7.254 $p = .001$
ADAS-Cog (mean ± SD)	8.575 ± 2.374	12.044 ± 3.205	18.675 ± 5.818	F(2;74) = 38.35 $p < .001$

Groups: MCI = mild cognitive impairment; mAD = mild Alzheimer's disease. Tests: MMSE = Mini-Mental State Examination; CDT = Clock Drawing Test; ADAS-Cog = Alzheimer's Disease Assessment Scale

All the tests were carried out at the Memory Clinic at the Department of Psychiatry of the University of Szeged, Hungary. All procedures were performed according to the Declaration of Helsinki, with the approval of the University of Szeged Ethical Committee and Regional Human Investigation Review Board and written informed consent was obtained from all participants.

Experimental protocol

In our methodology, it was an important goal to create different speech tasks eliciting connected speech that stimulate different memory systems; to this end, our linguistic protocol involved analysing three different speech tasks. After the presentation of a specially designed one-minute long silhouette animation designed and edited by our group, the subjects were asked to talk about the events seen during the film ("*immediate recall of film description task*"). Afterwards, the subjects were asked to talk about their previous day ("*previous day task*"). In the last task, the subjects were shown another animation, which was followed by a one-minute-long silent pause, and then they were asked to talk about the second film ("*delayed recall of film description task*"). The duration of the silent break was determined in such a way that it could be successfully used in the case of the mobile application in the future, as well. In our experimental setting, a 1 minute long break was chosen instead of a distractor task, so that the cognitive load could be increased, without the need to focus on another task, which might also influence the results. In this way, we could examine whether the interruption itself has any effect on performance. Also, having a break in the conversation can also be easily executed in a mobile application.

The first task, involving a direct recall, targets the working memory of the speaker. To handle the second task, besides working memory, episodic memory is also needed. In the third task, we included a one-minute distractor of the delayed recall to increase the demand for working memory capacity, and to make it more burdensome for the central executive system than direct recall.

This way, we obtained three audio recordings from each subject. The recording was performed with a digital voice recorder and a lapel microphone. After a careful listening, the recordings were transcribed both orthographically and phonetically; then the manual transcriptions were verified by another researcher. The manual segmentation of the signals was carried out using the PRAAT tool (Boersma, 2001), by one of the authors (G. Sz.), while raw data on the length of pauses and speech segments were extracted separately by two researchers (I. H. and G. Sz.).

The acoustic parameters we examined were the temporal variables listed in Table 2. Hesitation was defined as the absence of speech for more than 30 ms that occurs in between words (Gósy, 1998). It should be noted, however, that the definition of hesitation varies in the literature. For instance, hesitations and fillers are often viewed as part of the same category (cf. Corley & Stewart, 2008; Glücksmannová, 2008); for others, hesitation includes filled pauses only (Berger & Jordan, 1992; Gosztolya et al., 2016; Hoffmann et al., 2010; Tóth et al., 2015; Vincze et al., 2016), while other studies identify hesitation with silent pauses and other vocalizations (Gósy, 2008).

Results

Results of temporal analysis

First, we will present the average values obtained from analysis. For this purpose, we used the SPSS Statistics 24 (IBM Corp, 2016) software package. In this study, we were interested in finding the significant differences for each group, hence we compared the mean ± SD values of the different speaker groups and different temporal speech parameters via one-way ANOVA.

Results for articulation and speech rate

Figure 1 shows the average values of the three speaker groups for the articulation rate and speech rate parameters for all three tasks. It can be seen that in the case of the two recall tasks, articulation rate and speech rate both tend to decrease along with the progress of dementia; however, for the *previous day* task, we do not see such a change.

Table 2. The examined temporal speech parameters.

Name	Description
Articulation rate	The number of phonemes per second during speech (excluding hesitations)
Speech rate	The number of phonemes per second during speech (including hesitations); the number of total phonemes uttered, divided by the total duration of the utterance
Number of total pauses	The number of total (silent and filled) pause occurrences
Number of silent pauses	The number of silent pause occurrences
Number of filled pauses	The number of filled pause occurrences
Length of total pauses	The total duration of total pauses (sec)
Length of silent pauses	The total duration of silent pauses (sec)
Length of filled pauses	The total duration of filled pauses (sec)
Pauses/duration	The ratio of total pause duration and the length of the utterance (%)
Silent pauses/duration	The ratio of total silent pause duration and the length of the utterance (%)
Filled pauses/duration	The ratio of total filled pause duration and the length of the utterance (%)

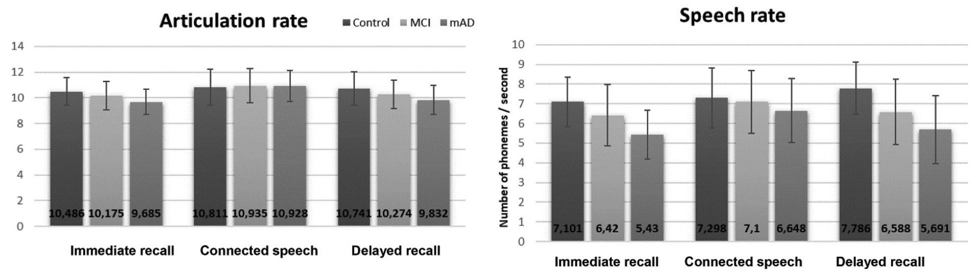

Figure 1. Average values of articulation rate and speech rate markers (number of phonemes/second) of the three speech tasks for the three speaker groups. Groups: MCI = mild cognitive impairment; mAD = mild Alzheimer's disease. Error bars indicate standard deviation. Columns indicate mean values for different groups.

Results for the number of different types of pauses

The average number of the different types of pauses for the three speaker groups and for all three speech tasks can be seen in Figure 2. It should be observed that speakers having MCI tend to have more pauses (total, silent and filled) than either control subjects or those having mAD; this is perhaps the most apparent in the previous day task, although we can observe this phenomenon in the two film description tasks as well.

Results for the length of different types of pauses

The average length of the different types of pauses for the three speaker groups and for all the three speech tasks can be seen in Figure 3. We can see that, just as in the case of number of pauses, MCI speakers tend to produce the highest (total) length of pause values, especially in the *previous day task*. This observation, combined with the same phenomenon regarding the number of pauses parameters, might indicate that MCI speakers tend to use more pauses.

Results for the percentage of different types of pauses in the whole speech unit

The average rate of utterances, which consisted of the different types of pauses for the three speaker groups and for all three speech tasks can be seen in Figure 4. We see that, while Figures 2 and 3 displayed similar trends, in Figure 4 we observe a different pattern. Recall

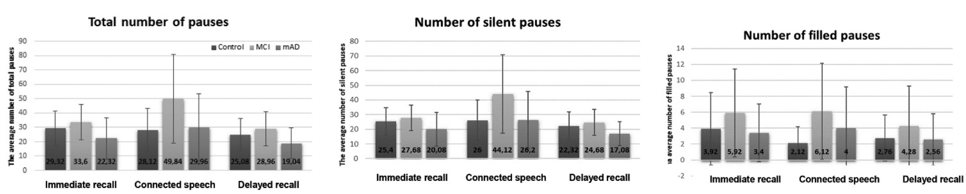

Figure 2. Average values of the total number of pauses, number of silent pauses and number of filled pauses markers of the three speech tasks for the three speaker groups. Groups: MCI = mild cognitive impairment; mAD = mild Alzheimer's disease. Error bars indicate standard deviation. Columns indicate mean values for different groups.

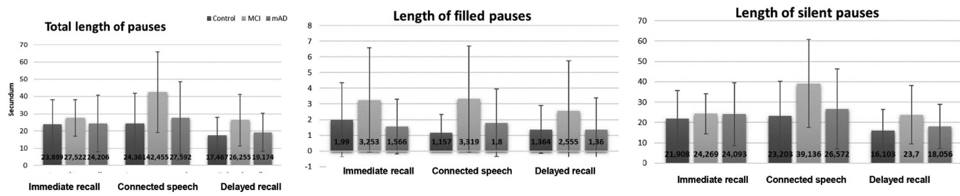

Figure 3. Average values of the total length of pauses, length of silent pauses and length of filled pauses markers of the three speech tasks for the three speaker groups. Groups: MCI = mild cognitive impairment; mAD = mild Alzheimer's disease. Error bars indicate standard deviation. Columns indicate mean values for different groups.

that the only difference between the length of pauses and pauses/duration temporal parameters is that the first category expresses the total duration of some pauses in seconds, while the latter expresses the same phenomenon as the ratio of the duration of the whole recording. This reflects that MCI patients tend to produce slightly more pauses than mAD speakers in general, but the utterances produced by the mAD speakers are much shorter. This is most apparent in the case of the *previous day task*, and it was justified by our further analysis. For this task, the average length of recordings were 75.6 ± 39.6 SD seconds,

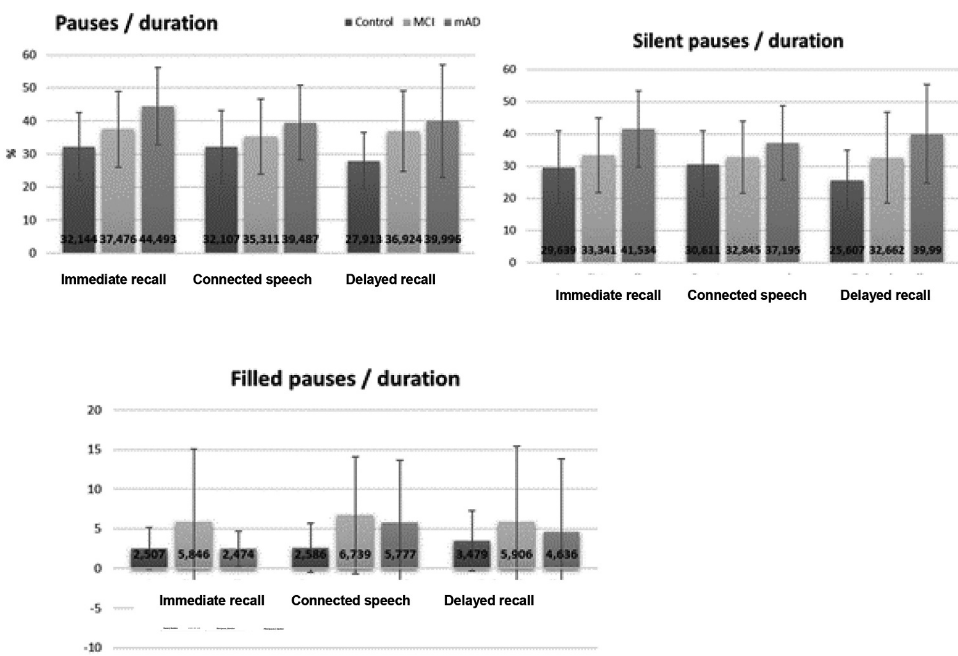

Figure 4. Average values of the pauses/duration, silent pauses/duration and filled pauses/duration markers of the three speech tasks for the three speaker groups. Groups: MCI = mild cognitive impairment; mAD = mild Alzheimer's disease. Error bars indicate standard deviation. Columns indicate mean values for different groups.

122.9 ± 63.2 SD seconds and 73.8 ± 57.3 SD seconds, control, MCI and mAD speakers, respectively.

Statistical differences of speech parameters

The average values of the three groups were compared with each other with regard to the different temporal speech parameters via a statistical analysis (one way ANOVA) using the SPSS Statistics 24 software package (IBM Corp, 2016). Multiple comparisons were performed by Tukey's HDS test (which was developed specifically to account for multiple comparison and maintains alpha at the specified level). This examination provides an answer to our key research question, namely which temporal speech parameters are suitable for distinguishing the three groups of patients. Table 3 contains the significance levels for the three tasks and the temporal speech parameter values. *p* values below 0.05 are shown as **bold**.

Upon examining the significance levels in Table 3, perhaps the most apparent point is that the results for *the immediate recall and delayed recall of film description tasks* are quite similar. For instance, the temporal speech parameters of articulation rate and speech rate were both found to differ significantly (i.e. $p < .05$), while this was not so for the *previous day task*. (Notice that this is in accordance with the results in Figure 1.) With regard to the pause-related parameters, the total number of pauses and pauses/duration parameters were also found to be significant for the two *film description tasks* after inspecting the silent pauses or both types of pauses; nevertheless, we found no significant result with any filled pauses-related temporal parameters for these two tasks. For the *task of previous day*, however, two temporal parameters (i.e. total number of pauses and total length of pauses) displayed significant differences for both silent pauses and filled pauses.

Statistical differences for speaker groups

The temporal speech parameter values that revealed significant differences among the three groups were further investigated. We felt this further step was essential as our previous statistical tests supplied significance information only about the temporal parameters, and

Table 3. Significance (F and *p*-values) of manual analysis for the three tasks.

	Immediate recall		*Connected speech*		*Delayed recall*	
	F	*p*	F	*p*	F	*p*
Articulation rate	3.666	*0.03*	0.07	*0.932*	3.671	*0.03*
Speech rate	9.498	*p < .001*	1.111	*0.335*	11.07	*p < .001*
Number of pauses	4.735	*0.012*	6.273	*0.003*	4.976	*0.009*
Length of pauses	0.52	*0.597*	5.417	*0.006*	3.586	*0.033*
Pauses/duration	7.668	*0.001*	2.728	*0.072*	5.758	*0.005*
Number of silent pauses	3.839	*0.026*	6.286	*0.003*	4.764	*0.011*
Length of silent pauses	0.247	*0.782*	4.619	*0.013*	2.719	*0.073*
Silent pauses/duration	6.904	*0.002*	2.314	*0.106*	7.446	*0.001*
Number of filled pauses	2.064	*0.134*	4.419	*0.015*	1.511	*0.228*
Length of filled pauses	2.912	*0.061*	5.289	*0.007*	2.141	*0.125*
Filled pauses/duration	2.88	*0.63*	2.823	*0.66*	0.585	*0.56*

Significant results are bold.

we were also interested in identifying the speaker groups among which the given temporal parameters differed to a significant extent.

Our analysis revealed that – for instance, – articulation rate showed a significant difference among the speaker groups. More specifically, the post-hoc analysis (i.e. ANOVA for each pair of groups) proved that controls and mAD patients show a significant difference but there is no significant difference among either MCI and control groups or MCI and mAD groups. We carried out this analysis for the other temporal speech parameters also, which revealed what parameters show significant differences among groups. This provides an answer to our research question, namely, which parameters are distinctive in identifying the three groups of patients. Table 4 lists the statistically significant results for the analysis of features in the three different speech tasks, with special regard to differences between MCI and C groups and MCI and mAD groups, respectively. For the sake of simplicity, differences between C and mAD groups are not listed here, moreover, they are irrelevant to our research questions.

An interesting observation is that in the *immediate recall of film description task*, there are significant differences for four parameters in the case of MCI and mAD groups, but there are no significant differences between the MCI and C groups. In the *previous day task*, six temporal parameters (of the total eleven) can be used to distinguish the C and MCI groups, and three temporal parameters showed significant differences among the MCI and mAD patients as well. In the *delayed recall of film description task*, there are significant differences for the speech rate, length of total pauses, and pauses/duration parameters between the C and MCI groups, and for two parameters (number of total pauses and silent pauses) between the MCI and mAD groups.

Discussion

Our aim was to find speech parameters to distinguish the three speaker groups: we have attempted to find out whether features indicating language decline triggered by other

Table 4. *P*-values of statistically significant differences of the investigated temporal speech parameters for the three different tasks, for the MCI and C, and the MCI and mAD groups, respectively.

		Tasks					
		Immediate recall		*Connected speech*		*Delayed recall*	
		MCI – C	MCI – mAD	MCI – C	MCI – mAD	MCI – C	MCI – mAD
Speech parameters	Articulation rate	-		-		-	
	Speech rate		0.033	-		0.024	
	Number of total pauses		0.009	0.006	p = .013		0.007
	Number of silent pauses		p = .023	0.008	0.009		0.01
	Number of filled pauses	-		p = .011		-	
	Length of total pauses	-		0.008	MCI – mAD (p = .035)	0.036	
	Length of silent pauses	-		0.014		-	
	Length of filled pauses	-		0.006		-	
	Pauses/duration	-		-		0.045	
	Silent pauses/duration		0.039	-		-	
	Filled pauses/duration	-		-		-	

Groups: C = control; MCI = mild cognitive impairment; mAD = mild Alzheimer's disease.

cognitive functions in the elderly could be used to identify the MCI group. Our hypothesis that speech parameters derived from the analysis of connected speech can help distinguish the groups of MCI patients, mAD patients and controls was proven to be well-founded. Earlier results on the influence of the speaker's cognitive abilities on speech production (Braaten et al., 2006; Ferreira & Swets, 2002; Levelt & Meyer, 2000) were confirmed here as each parameter yielded significant differences between the C and the mAD groups and most of the parameters did the same between the C and MCI groups. Moreover, we could confirm that the phonetic changes exist in mAD (Forbes-McKay & Venneri, 2005; Laske et al., 2015) as well as inAD and MCI (Ahmed et al., 2013; Roark et al., 2011; Satt et al., 2014). Based on our results, we can conclude that speech based methods can effectively distinguish controls, patients with MCI and patients with mAD (Drummond et al., 2015; König et al., 2015) **(H1)**.

Articulation and speech rates (see Figure 1) tend to decrease in *film description tasks* (Task 1 and 3), as dementia progresses. It seems that these tasks (immediate and delayed recall) slow down thinking, which requires more cognitive efforts, hence different speech parameters can also be detected for the groups. Our hypothesis that speech rate will effectively differentiate the MCI group from the others was partially proved: in Task 1, there are significant differences between MCI and mAD groups while in Task 3, there are significant differences for MCI and C groups (see Table 4) **(H2)**.

As for the number and length of pauses, the C and mAD groups were similar to each other while the MCI group exhibited differences (see Figures 2 and 3), which proved to be significant in the *previous day task*. This strongly suggests that the duration of pauses is a strong characteristic of MCI. Our results confirm the earlier hypothesis that pauses had key roles in separating these groups (Hoffmann et al., 2010; López-de-Ipiña et al., 2013) **(H3)**. However, it must be noted that when differentiating filled and silent pauses, we expected somewhat different results as in a previous study, we obtained statistically significant differences for the number and length of silent and filled pauses (among other parameters) between the MCI and C groups (Tóth et al., 2015). In our current study, most parameters show significant differences for the MCI and mAD groups (see Table 4), however, filled pauses have distinctive power only in the *previous day task*, while silent pauses do the same in all tasks.

The parameters pauses/duration and silent pauses/duration linearly decreased among the groups (see Figure 4), in other words, the percentage of pauses in the total duration of speech continuously grows as dementia progresses. According to our results, the more the cognitive deficit is, the greater percent of speech contains pauses. This difference is significant for Tasks 1 and 3, which in part proves that these novel parameters can effectively distinguish the groups examined **(H4)**. Such differences could not be found for the filled pauses/duration parameter, which might be explained by the inaccurate manual segmentation of filled pauses. Nevertheless, we think that the distinctive power of this parameter should be further investigated in future studies.

Our results prove that the number and length of pauses can distinguish the groups effectively in the previous day task. In film description tasks, the parameters speech rate, number of silent pauses and pauses/duration showed significant differences among the groups (see Table 4). Thus, it seems that it is the silent pause that yields statistically significant differences among the three groups of speakers for all tasks. These findings imply that the two recall tasks can be used mainly to identify speakers with mAD, while connected speech is mostly effective to distinguish MCI speakers from the other two groups, with special regard to the C group.

Based on our results, we may conclude that speech-based differences are present in MCI, and they can be observed in mAD as well. Hence, changes in language functions could play a diagnostic role in these types of dementia. Despite this, there have been relatively few published linguistic methods that were able to diagnose AD at an early stage. We found that temporal parameters of connected speech are one of the earliest detectable language signs in MCI and mAD, and we hypothesise that a computer-based analysis of connected speech might be a promising method for an early and quick diagnosis of dementia (see e.g., Tóth et al., 2015, 2018), substituting a costly and slow manual data analysis, which we would like to implement in the future. An Automatic Speech Recognition (ASR) system working at the phonetic- or at the word-level, where the phonetic set or the dictionary (i.e. the list of words) includes filled pauses as well, might allow for the estimation of our temporal speech parameters. On the basis of the temporal speech parameters, machine learning methods (e.g., Support Vector Machines, random forests, Deep Neural Networks, etc.). could be utilized to predict the category of the actual speaker. Based on these previous results, it can be assumed that a new mobile application can be developed to screen for a potential risk of mAD or MCI. By detecting acoustic features of mAD or MCI, this mobile application could indicate the need for further medical examination.

With the joint application of telemedicine-based tools and linguistic tests, an interactive test or mobile application could permit software-based filtering for MCI and AD. This might lay the groundwork for extensive neurolinguistic research and the development of a widespread screening application as well. Furthermore, automatic speech analysis is important because using real-life situations and applying fewer intrusive methods that do not require specialised personnel would also be advantageous.

Besides the above findings, this study also has some limitations. For instance, we experimented with Hungarian only, hence a more comprehensive study of other languages would also be required to investigate whether our findings hold for other languages as well. Second, we analysed speech data from 75 participants. A wider pool of subjects would definitely contribute to the deeper understanding of the speech performance of the different speaker groups examined here. Third, we just focused on three speech eliciting tasks here, whereas the output of other similar tasks (e.g., picture description) could also highlight interesting tendencies that are characteristic of the speech of different groups.

Our study provides evidence that the analysis of connected speech is a useful tool for distinguishing MCI patients, mAD patients and controls. Although our results are promising, future studies involving more subjects and other languages are needed to gain a better insight into the linguistic dysfunctions of dementia.

Disclosure of interest

The authors report no conflict of interest.

ORCID

Veronika Vincze http://orcid.org/0000-0002-9844-2194

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. https://doi.org/10.1093/brain/awt269

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing

Balthazar, M. L. F., Martinelli, J. E., Cendes, F., & Damasceno, B. P. (2007). Category verbal fluency performance may be impaired in amnestic mild cognitive impairment. *Dementia & Neuropsychologia*, *1*(2), 161–165. https://doi.org/10.1590/s1980-57642008dn10200008

Barban, F., Annicchiarico, R., Pantelopoulos, S., Federici, A., Perri, R., Fadda, L., Carlesimo, G. A., Ricci, C., Giuli, S., Scalici, F., Turchetta, C. S., Adriano, F., Lombardi, M. G., Zaccarelli, C., Cirillo, G., Passuti, S., Mattarelli, P., Lymperopoulou, O., Sakka, P., Ntanasi, E., … Caltagirone, C. (2016). Protecting cognition from aging and Alzheimer's disease: A computerized cognitive training combined with reminiscence therapy. *International Journal of Geriatric Psychiatry*, *31*(4), 340–348. https://doi.org/10.1002/gps.4328

Berger, C. R., & Jordan, J. M. (1992). Planning sources, planning difficulty and verbal fluency. *Communication Monographs*, *59*(2), 130–149. https://doi.org/10.1080/03637759209376257

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345.

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. *International Journal of Neuroscience*, *116*(11), 1271–1293. https://doi.org/10.1080/00207450600920928

Canning, S. J., Leach, L., Stuss, D., Ngo, L., & Black, S. E. (2004). Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. *Neurology*, *62*(4), 556–562. https://doi.org/10.1212/wnl.62.4.556

Corley, M., & Stewart, O. W. (2008). Hesitation Disfluencies in Spontaneous Speech: The Meaning of um. *Language and Linguistics Compass*, *2*(4), 589–602. https://doi.org/10.1111/j.1749-818X.2008.00068.x

Drummond, C., Coutinho, G., Fonseca, R. P., Assunção, N., Teldeschi, A., de Oliveira-souza, R., Moll, J., Tovar-Moll, F., & Mattos, P. (2015). Deficits in narrative discourse elicited by visual stimuli are already present in patients with mild cognitive impairment. *Frontiers in Aging Neuroscience*, *7*, 1–11. https://doi.org/10.3389/fnagi.2015.00096.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57–84. https://doi.org/10.1006/jmla.2001.2797

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. https://doi.org/10.1016/0022-3956(75)90026-6

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. https://doi.org/10.1007/s10072-005-0467-9

Fraser, K. C., Meltzer, J. A., Graham, N. L., Leonard, C., Hirst, G., Black, S. E., & Rochon, E. (2014). Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex*, *55*, 43–60. https://doi.org/10.1016/j.cortex.2012.12.006

Fraser, K. C., Meltzer, J. A., & Rudzicz, F. (2016). Linguistic features identify Alzheimer's disease in narrative speech. *Journal of Alzheimer's Disease*, *49*(2), 407–422. https://doi.org/10.3233/JAD-150520

Garrard, P., Rentoumi, V., Gesierich, B., Miller, B., & Gorno-Tempini, M. L. (2014). Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse. *Cortex*, *55*, 122–129. https://doi.org/10.1016/j.cortex.2013.05.008

Glücksmannová, H. (2008). Spontaneous speech reconstruction. In J. Safránkoná & Pavlú (Eds.), *WDS '08 processings of contributed pepers, Part I* (pp. 134–138). Metfyzpress.

Gósy, M. (1998). The paradox of speech planning and production. *Magyar Nyelvor (Hungarian Language Guard)*, *122*(1), 3–15. http://real-j.mtak.hu/6072/1/MagyarNyelvor_1998.pdf
Gósy, M. (2008). Önellenőrzési folyamatok a beszédben. *Magyar Nyelv*, *104*(4), 402–426. https://epa.oszk.hu/00000/00032/00039/pdf/gosy.pdf
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, CA, USA, 107–111. https://doi.org/10.21437/Interspeech.2016-384
Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia*, *42*(9), 1212–1222. https://doi.org/10.1016/j.neuropsychologia.2004.02.001
Hoffmann, I., Nemeth, D., Dye, C., Pákáski, M., Irinyi, T., & Kálmán, J. (2010). Temporal features of spontaneous speech in Alzheimer's Disease. *International Journal of Speech-language Pathology*, *12*(1), 29–34. https://doi.org/10.3109/17549500903137256
Horley, K., Reid, A., & Burnham, D. (2010). Emotional prosody perception and production in dementia of the Alzheimer's type. *Journal of Speech, Language, and Hearing Research*, *53*(5), 1132–1146. https://doi.org/10.1044/1092-4388(2010/09-0030)
IBM Corp. (2016). *Released 2016. IBM SPSS Statistics for Windows, Version 24.0.*
Janka, Z., Somogyi, A., Maglóczky, E., Pákáski, M., & Kálmán, J. (1988). Dementia screening by a short cognitive test. *Orvosi Hetilap*, *129*(52), 2797–2800. http://real-j.mtak.hu/11251/10/650.1988.12.04.pdf
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," in *Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality*, Baltimore, Maryland, USA, 27–37. https://doi.org/10.3115/v1/W14-3204
Kálmán, J., Maglóczky, E., & Janka, Z. (1995). Disturbed visuo-spatial orientation in the early stage of Alzheimer's dementia. *Archives of Gerontology and Geriatrics*, *21*(1), 27–34. https://doi.org/10.1016/0167-4943(95)00639-3
Kato, S., Homma, A., & Sakuma, T. (2018). Easy screening for mild Alzheimer's disease and mild cognitive impairment from elderly speech. *Current Alzheimer Research*, *15*(2), 104–110. https://doi.org/10.2174/1567205014666171120144343
Kempler, D. (2004). *Neurocognitive disorders in aging*. Sage Publications.
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: Diagnosis, Assessment & Disease Monitoring*, *1*(1), 112–124. https://doi.org/10.2174/1567205014666171120144343
Laske, C., Sohrabi, H. R., Frost, S. M., López-de-Ipiña, K. L., 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. (2015). Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer & Dementia.*, *11*(5), 561–578. https://doi.org/10.1016/j.jalz.2014.06.004
Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology*, *12*(4), 433–452. https://doi.org/10.1080/095414400750050178
López-de-Ipiña, K., Alonso, J. B., Travieso, C. M., Solé-Casals, J., Eguiraun, H., Faundez-Zanuy, M., Ezeiza, A., Barroso, N., Ecay-Torres, M., Martinez-Lage, P., & Martinez de Lizardui, U. (2013). On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensor*, *13*(5), 6730–6745. https://doi.org/10.3390/s130506730
López-de-Ipiña, K. , J. B. Alonso, J. Solé-Casals, N. Barroso, P. Henriquez, M. Faundez-Zanuy, C. M. Travieso, M. Ecay-Torres, P. Martínez-Lage & H. Eguiraun (2015). On automatic diagnosis of Alzheimer's disease based on spontaneous speech analysis and emotional temperature. *Cognitive Computation*, *7*, 44–55. https://doi.org/10.1007/s12559-013-9229-9.

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

Martínez-Sánchez, F. F., Meilán, J. J. G., Carro, J., & Ivanona, O. (2018). A prototype for the voice analysis diagnosis of Alzheimer's disease. *Journal of Alzheimer's Disease*, *64*(2), 1–9. https://doi.org/10.3233/JAD-180037

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

Martínez-Sánchez, F. F., Meilán, J. J. G., Pérez, E., Carro, J., & Arana, J. M. (2012). Expressive prosodic patterns in individuals with Alzheimer's disease. *Psicothema*, *24*(1), 16–21. https://pubmed.ncbi.nlm.nih.gov/22269358/

Meilán, J. J. G., Martínez-Sánchez, F., Carrol, J., Carcavilla, N., & Ivanova, O. (2018). Voice markers of lexical access in mild cognitive impairment and Alzheimer's disease. *Current Alzheimer Research*, *15*(2), 111–119. https://doi.org/10.2174/1567205014666170829112439

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), 2081–2090. https://doi.org/10.5209/rev_sjop.2012.v15.n2.38859

Meilán, J. J. G., Martínez-Sánchez, F., Martínez-Nicolás, F., 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. https://doi.org/10.1155/2020/4683573

Millian-Morell, L., López-Alburquerque, T., Rodríguez-Rodríguez, A., Gómez-Nieto, R., Carro, J., Meilán, J. J. G., Martínez-Sánchez, F., Sancho, C., & López, D. E. (2018). Relations between sensorimotor integration and speech disorders in Parkinson's disease. *Current Alzheimer Research*, *15*(2), 149–156. https://doi.org/10.2174/1567205014666170829103019

Misiewicz, S., Brickman, A. M., & Tosto, G. (2018). Prosodic impairment in dementia: Review of the literature. *Current Alzheimer Research*, *15*(2), 157–163. https://doi.org/10.2174/1567205014666171030115624

Nebes, R. D., Brady, C. B., & Huff, F. J. (1989). Automatic and attentional mechanisms of semantic priming in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *11*(2), 219–230. https://doi.org/10.1080/01688638908400884

Pákáski, M., Drótos, G., Janka, Z., & Kálmán, J. (2012). Validation of the Hungarian version of Alzheimer's Disease Assessment Scale–cognitive subscale. *Orvosi Hetilap*, *153*(12), 461–466. https://doi.org/10.1556/OH.2012.29332

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183–194. https://doi.org/10.1111/j.1365-2796.2004.01388.x

Quaranta, D., Piccininni, C., Caprara, A., Malandrino, A., Gainotti, G., & Marra, C. (2019). Semantic relations in a categorical verbal fluency test: An exploratory investigation in mild cognitive impairment. *Frontiers in Psychology*, *10*, 2797. https://doi.org/10.3389/fpsyg.2019.02797

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. https://doi.org/10.1109/TASL.2011.2112351

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. https://doi.org/10.1176/ajp.141.11.1356

Satt, A., Hoory, R., König, A., Aalten, P., & Robert, P. (2014). *Speech-based automatic and robust detection of very early dementia, interspeech.*

Scheff, S. W., Price, D. A., Schmitt, F. A., DeKosky, S. T., & Mufson, E. J. (2007). Synaptic alterations in CA1 in mild Alzheimer disease and mild cognitive impairment. *Neurology*, *68*(18), 1501–1508. https://doi.org/10.1212/01.wnl.0000260698.46517.8f

Slegers, A., Filiou, R. P., Montembeault, M., & Brambati, S. M. (2018). Connected Speech Features from Picture Description in Alzheimer's Disease: A Systematic Review. *Journal of Alzheimers Disease*, *65*(2), 519–542. https://doi.org/10.3233/JAD-170881

Swets, B., Jacovina, M. E., & Gerrig, R. J. (2014). Individual differences in the scope of speech planning: Evidence from eye movements. *Language and Cognition*, *6*(1), 12–44. https://doi.org/10.1017/langcog.2013.5

Szatlóczki, G., Hoffmann, I., Vincze, V., Kálmán, J., & Pákáski, 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. https://doi.org/10.3389/fnagi.2015.00195

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. *Interspeech tutorials & main conference*. Dresden, Germany.

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

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 ACL*, Berlin, Germany, 181–187. https://doi.org/10.18653/v1/P16-2030

Weiner, M. F., Neubecker, K. E., Bret, M. E., & Hynan, L. S. (2008). Language in Alzheimer's disease. *Journal of Clinical Psychiatry*, *69*(8), 1223–1227. https://doi.org/10.4088/jcp.v69n0804