Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

Does sleepiness influence reading pauses in hypersomniac patients?

Vincent P. Martin[1], *Brice Arnaud*[1], *Jean-Luc Rouas*[1], *Pierre Philip*[2]

[1]LaBRI - CNRS UMR 5800 - Univ. Bordeaux - Bordeaux INP, Talence ,France
[2]SANPSY - Univ. Bordeaux - CNRS USR 3413, CHU de Bordeaux, France

vincent.martin@labri.fr,
brice.arnaud33@gmail.com,rouas@labri.fr,pierre.philip@u-bordeaux.fr

Abstract

Sleepiness and Excessive Daytime Sleepiness (EDS) are major public health concerns impacting the daily life and performances of subjects experiencing them. Usually measured by electroencephalographic measures or self-reported questionnaires, previous studies have shown that it is possible to measure them through voice analysis. In this article, we propose to investigate potential new vocal biomarkers of sleepiness and EDS on 93 patients affected by hypersomnia, namely reading pauses. We analyze the location and duration of the pauses annotated by a fully automated system and propose a new set of speech features. Based on these 12 descriptors, we have identified seven reading behavior profiles, that are almost fully explained by the physical and medical characteristics of the patients, including their level of EDS. Regarding short-term sleepiness, the observed differences are mainly due to the differences of texts and have a weak correlation with objective and subjective sleepiness measures.

Index Terms: computational paralinguistics, reading pauses, sleepiness

1. Introduction

1.1. Sleepiness and Excessive Sleepiness

Sleepiness is a normal psychophysiological state experienced by most individuals over any given 24h period. However, when it interferes with daily life (e.g. by occurring at inappropriate times), sleepiness is considered as a symptom called Excessive Daytime Sleepiness (EDS). EDS is often associated with a wide range of diseases including metabolic, cardiovascular, neurological and psychiatric diseases, leading to disability and increased risk of mortality [1, 2]. EDS is thus constituting a major public health problem [3, 4] that is potentially dangerous for both personal and public safety, especially regarding sleepiness-related accidents at the wheel [5].

Hence, sleepiness can be studied in two different temporal modalities [6]. On the one hand, short-term sleepiness is a state of the subject that lasts for seconds or minutes. It is usually measured by questionnaires such as the Karolinska Sleepiness Scale – KSS [7] or electroencephalography-based measures (EEG), such as the VIGALL algorithm [8]. On the other hand, EDS is a trait of the subject that lasts for weeks or months. It is usually measured by polysomnographic measurements (PSG) such as the Multiple Sleep Latency Test – MSLT [9, 10] or medical questionnaires, such as the Epworth Sleepiness Scale – ESS [11]. Therefore, sleepiness and EDS are usually measured either with EEG or PSG recordings, or subjectively, using psychometric scales.

1.2. Sleepiness detection through voice

We propose a third modality to detect EDS based on speech analysis. Contrary to the previous ones, this modality is not invasive, costs less than traditional measurements, and benefits from its possibility of implementation in real conditions (e.g. with smartphone microphones) and could complete the information gathered by smartphone-based virtual agents, that are very well accepted by patients on a regular use [12, 13].

Until now, sleepiness and EDS have been studied though two vocal markers families: acoustic features, measuring the quality of the vocalizations [14, 15, 16], and reading-based features, measuring the impact of sleepiness on patients' neurolinguistic abilities [19, 20, 21]. This study proposes to investigate a new set of features aiming at measuring the interaction between reading capacities and sleepiness. To do so, we evaluate the phrasing of the patients when they are reading, and more specifically the pauses between groups of words. Indeed, while some patients have a correct reading flow, pausing where it is 'natural', others artificially lengthen some pauses to plan the reading of the following words, reflecting their difficulties in cognitive planning.

Pause-based features have already been successfully used for the detection of depression in read speech [22] and the detection of mild cognitive impairment in spontaneous speech [23]. Sleepiness however remains an unexplored field for these features, apart from a previous article of the authors describing the automatic extraction of these pauses [24]. Accordingly, the aim of the present article is to unravel the influence of sleepiness and EDS on reading pauses of hypersomniac patients.

This article is organized as follows. In Section 2, we introduce the Multiple Sleep Latency Test corpus (MSLTc). Then, in Section 3, we introduce the pause-based features automatically extracted from the audio files. In Section 4, we draw profiles of the readers based on their pausing behavior and interpret them in light of a Principal Components Analysis. We present in Section 5 three statistical analyses based on deviation of mean, ANOVA and regression. We discuss and interpret the link between the obtained reading profiles and the other speakers characteristics in light of the statistical analyses in Section 6, and conclude in Section 7.

2. Multiple Sleep Latency Test Corpus

This study is based on a subset of the Multiple Sleep Latency Test Corpus – MSLTc [25, 6], containing the recordings of 93 native french-speaking hypersomniac patients of the Sleep Clinic of Bordeaux, France. These patients undertake a Multiple Sleep Latency Test, which consists of asking them to take five naps every two hours from 9 am to 5 pm. During these naps, their sleep latency (noted MSLT it.) is measured by polysomnography. The average sleep latency (noted MSLT)

10.21437/SpeechProsody.2022-13

is the reference diagnostic criteria for numerous sleep disorders [10]. Before each nap, the patients are recorded reading out loud texts extracted from *Le Petit Prince* (A. de Saint-Exupéry) and fill a Karolinska Sleepiness Scale (KSS).

During their stay, the patients also fill numerous regarding sleep-related health questionnaires: in this study, we will only consider the Epworth Sleepiness Scale (ESS), a subjective measure of EDS [11] and the Hospital Anxiety and Depression questionnaire (HAD), made up of two sub-scales measuring anxiety and depression levels [26]. In addition to these questionnaires, the corpus also contains the sex, Body Mass Index (BMI), neck circumference, and socio-demographic level of patients (measured by the years of study after the French Certificate of general education).

3. Pause-based features

3.1. Automatic estimation of pauses position and duration

Annotating manually all the pauses made by the readers in all the recordings of the corpus is a time-consuming and tedious task. Hence, we designed a system, based on Automatic Speech Recognition, to automatically align the audio files and the read texts, in order to annotate the location and duration of the pauses made by the readers. The full system is presented in [24]. Briefly, the system proceeds in three steps. First, based on the denoised audio files, an energy-based Voice Activity Detection system is trained on a manually annotated subset of 25 files picked from the MSLTc. Second, the output of a chain-based ASR model is realigned using the VAD system output. Finally, the corrected ASR hypothesis is aligned with the reference text based on the Smith-Waterman distance between the ASR chunks and their equivalent in the reference text, allowing to process ASR errors and repetition of words. This process is then used to extract both the duration ad the location of the pauses made by the patients during their reading.

3.2. Corpus annotation

In order to quantify the reading behavior of the patients, each of the five read texts has been independently annotated by three speech therapist students, with the instruction to 'annotate the inter-words silences with how much it is natural to pause, from -10: very unnatural to 10: very natural, 0 being neutral'. Hence, knowing the location and duration of a pause, we can quantify how much pausing at a given location a given duration is part of a 'normal' reading behavior or if it is deviant from the average reader. We consider here the mean of the three annotations. The annotated texts are available online[1].

3.3. Features

From the location and duration of the pauses, we derive four sets of speech features. Each of them is computed on all the pauses (noted 'all'), on the pauses annotated with a positive score (noted '+') and on their negative equivalent ('-'). These features are the number of pauses on each audio recording N, the mean duration D, the mean score S of a pause across the recording, and the average of the score weighted by the duration of the pauses WS. For example, noting d_i and s_i the duration and score of the i^{th} pause, we have:

$$\text{WS}^+ = \frac{1}{\text{N}^+} \frac{1}{d_{tot}^+} \sum_{i \in +} d_i s_i \quad \text{with: } d_{tot}^+ = \sum_{i \in +} d_i$$

[1] https://zenodo.org/record/4644021

An example of negative and positive pauses duration, score, and weighted score is given in Figure 1. When studying patients characteristics, these features are averaged across the five naps of the MSLT.

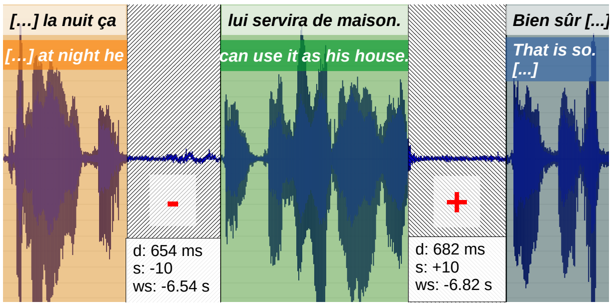

Figure 1: *Example of misplaced and correctly-placed pauses and their duration, score and weighted score.*

4. Reader profiles

4.1. Method

To draw reader profiles, we apply a K-Means clustering algorithm on the normalized speakers' features. The number of clusters $k < 10$ is chosen such as it maximizes the silhouette score, which is $k = 7$ for our database (silhouette score: 0.253).

4.2. Principal Components Analysis

As a first step to analyze the obtained profiles, we apply a Principal Component Analysis with n=2 components on the z-scaled pause-based features. The coefficients given by the PCA to each feature are reported in Table 1. Keeping only two dimensions replicates 75.5% of the original variance, and the seven clusters previously identified, plotted in Figure 2, have good contrast (silhouette score: 0.246). Hence, the two major dimensions of this set of markers are directed on the one hand by the total number of pauses and their length, independently of their correctness (Dim. 1), and on the other hand by the difference in occurrence and duration of positive and negative pauses (Dim. 2). These two dimensions are anti-correlated with S^- and WS^-, and they both represent misreading behaviors: Dim. 1 corresponds to a very high stopping frequency and excessive stopping times, while Dim. 2 captures the propensity to pause at incorrect locations.

Table 1: *PCA coefficients for the different pause-based features*

	+				-				all			
	N	D	S	WS	N	D	S	WS	N	D	S	WS
Dim. 1	.28	.36	.20	.34	.25	.33	-.27	-.27	.34	.36	-.20	.16
Dim. 2	-.23	-.20	-.06	-.23	.36	.21	-.25	-.25	-.04	-.20	-.44	-.49

5. Statistical Analysis

The aim of this paper is to unravel the relation between the speakers' physical and medical characteristics and their reading behavior. To do so, we process in three steps.

5.1. Deviation from the average

First, for each of the seven reader profiles and for each reading feature, we report in Table 2 the distance between the average of

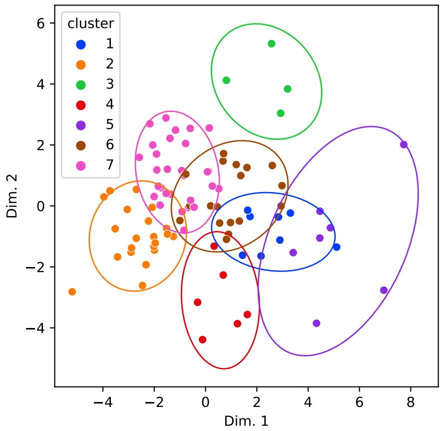

Figure 2: *Visual representation of the KMeans clusters on the plane directed by the two dimensions of the PCA. The ellipses have been drawn using radiuses equals to* 2σ*.*

the cluster and the global average value on all the database. We also report at the bottom of this table the same metrics computed on the medical and physical information of the patients. These results are discussed in Section 6.

5.2. ANOVA

5.2.1. *Method*

Then, to identify the factors influencing the most the pauses in reading, we process a multivariate ANOVA with repeated measures, independently from any profile: in this manner, we attempt to explain the inter- and intra-speaker variations of the features with their characteristics described in Section 2. In order to bring out interesting results without losing too much statistical power, we calculate only the effect of simple variables, and first- and second-order interactions between ESS, MSLT, and speakers' cofactors. The result of this statistical test is presented in Table 3.

5.2.2. *Results*

The intra-speaker variations of all the features except the duration are strongly influenced by the session ($p < 0.001$): the variations of the features across a session are either influenced by the time of day or the read text, that changes at each iteration of the MSLT. Subjective sleepiness is partly responsible for the variation of D^+ and D^-, N^+, and of WS^+, WS^- and WS^{all}. As the individual sleep latency does not influence the evolution of any feature, it has not been reported in the table. Finally, the joint interaction of KSS and MSLT sleep latency has a small effect on the score of the positive pauses.

While sex influences all the characteristics of the positive pauses, age affects only D^+, and WS^+. On the contrary, the BMI seems to affect only D^- and S^-. The neck size interferes with N^+ and S^+, and D^{all}. The socio-demographic level has the same effect on the positive pauses, and the level of anxiety affects N^+ and S^{all}. The level of depression has no influence on pauses features, thus the HAD-D has not been reported in this table. Regarding sleepiness, D^- is influenced by the ESS, N^- and S^{all} are affected by the joint interaction of the MSLT, ESS, and BMI. Finally, the joint interaction of the MSLT, ESS, and Socio-demographic level interacts with S^-.

Table 2: *Profiles of the readers based on their pauses.* m*: mean value on the total population,* σ *standard deviations on the total population. Notation depending on the mean of the cluster* m_c*. '=':* $m_c \in [m-0.25\sigma; m+0.25\sigma]$*; '-':* $m_c \in [m-0.5\sigma; m-0.25\sigma]$*; '- -':* $m_c \in [m-\sigma; m-0.5\sigma]$*; '- - -':* $m_c < m-\sigma$*. '+', '++' and '+++' are the equivalent notations for* $m_c > m$

Feature		m (σ)	**n°1** *n=8*	**n°2** *n=21*	**n°3** *n=4*	**n°4** *n=6*	**n°5** *n=7*	**n°6** *n=18*	**n°7** *n=29*
+	N	23.3 (5.4)	=	- -	=	+++	+++	++	- -
	D	615.5 ms (143.0)	+++	- -	=	++	+++	=	- -
	S	8.0 (0.3)	=	+++	- - -	-	- - -	- -	- -
	WS	5.0 (1.1)	+++	- -	-	++	+++	=	- -
-	N	3.5 (2.7)	=	- -	+++	- -	+++	=	=
	D	480.3 ms (140.3)	+++	- - -	+++	- -	+++	+	=
	S	-7.3 (1.7)	- -	+++	- -	+++	=	-	- -
	WS	-4.0 (1.3)	- - -	+++	- -	+++	- -	=	=
All	N	26.8 (6.4)	=	- -	+++	++	+++	++	- -
	D	605.3 ms (132.7)	+++	- -	=	++	+++	=	- -
	S	6.0 (1.3)	=	++	- - -	++	- -	=	=
	WS	3.8 (1.1)	+++	=	- - -	+++	+++	=	- -
Sex %F		62.4	25.0	85.7	75.0	50.0	14.3	44.4	79.3
Age		36.6 (14.5)	++	-	- -	++	++	+	=
BMI		23.9 (5.1)	=	=	-	=	=	=	=
Neck		37.8 (4.4)	+	-	- -	=	++	+	=
Socio-D.		5.5 (2.6)	- -	+	- -	+	=	-	=
HAD-A		8.5 (4.2)	=	=	+++	=	-	=	=
HAD-D		6.7 (3.8)	- - -	-	+	+	=	+	=
ESS		14.6 (4.7)	-	=	-	++	- -	=	=
MSLT		11.6 (4.6)	- -	=	=	=	=	=	=

5.3. Regression

We also reported in Table 3 the significant univariate correlation between the cofactors and the pauses features. At the sample level, the sleep latency does not correlate with any feature and has not been reported in the table. In contrast, subjective sleepiness correlates weakly but significantly with S^-, D^+, D^- and D^{all}, and WS^+ and anti-correlates with S^+ and S^{all}.

At the speaker level, only age, neck size, and socio-demographic level show a significant correlation with pauses characteristics. Thus, age correlates with N^+, D^+, WS^+, D^{all} and WS^{all}. The socio-demographic level and the neck size have antagonist effects on D^-, D^{all}, N^+ and N^-, S^+, and S^{all}. The neck size is also correlated to N^+, WS^+ and anti-correlated to WS^-.

6. Discussion

6.1. Patients with impaired pauses behavior

The 8 speakers gathered in profile n°1 have two impaired reading behaviors: not only they lengthen their pauses both at natural and unnatural locations in the text (Dim. 1 of the PCA), but the location of their negative pauses are worst than the average (Dim. 2 of the PCA). Regarding their positive pauses, these patients are older and have a bigger neck than the average, which is linked to the higher values for D^+ and WS^+. Regarding their negative reading behavior, the higher D^- could be due to the correlation of this factor with neck size, whereas the lower S^- and WS^- could be due to the interaction of the MSLT, ESS, and socio-demographic level, that are lower than the average.

Speakers belonging to the profile n°3 (n=4) show an excessive misplacement of their pauses, with an emphasis on the un-

Table 3: *Statistical analysis of the proposed features with regards to the patients' characteristics. Top: significance of the factors included in a multivariate ANOVA with repeated measures designed to explain the inter-speaker and intra-speaker variations of each pause metric. Bottom: Correlation between the features and the patient's characteristics. Correlations on the speaker level are computed with the average of the features across the MSLT sessions. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$*

Factor	**Number**			**Duration (ms.)**			**Score**			**Weighted Score**		
	+	-	all	+	-	all	+	-	all	+	-	all
Mean	23.3	3.5	26.8	615.5	480.3	605	8.0	-7.3	6.0	5.0	-4.0	3.8
Std	6.3	3.4	7.5	156.5	241.9	144.7	.53	3.2	1.7	1.3	2.33	1.4
ANOVA												
Sex	**		**	**		**	*			**		*
Age				***		***				***		***
BMI					**			*			**	
Neck	*		*	*	**	*	*					
Socio-D.	*		*				*					
HAD-A	*								*			
ESS					*						*	
MSLT:ESS:BMI		*							*			
MSLT:ESS:Socio-Dem								**			*	
Session	***	***	***	*			***	***	***	***	***	***
KSS				td	**			*		*	**	*
KSS:MSLT it							td					
Correlations: Spearman's ρ (sig.)												
Age	.21 (*)			.44 (***)		.42 (***)				.45 (***)		.40 (***)
Neck	.3 (**)	.26 (*)	.29 (**)	.32 (**)	.33 (**)	.33 (**)	-.32 (**)		-.23 (*)	.28 (**)	-.21 (*)	
Socio-Dem		-.29 (**)	-.28 (**)	-.21 (*)	-.21 (*)		.30 (**)		.27 (**)			
KSS		.13 (**)		.12 (**)	.09 (*)	.13 (**)	-.09 (*)		-.15 (**)	.09 (*)		

natural ones, which are more numerous and last longer than the average patient (Dim. 2 of the PCA). The augmentation of N^- and D^- and the diminution of S^+ can be explained by their lower socio-demographic level, which is not enough compensated by their narrower neck. Moreover, these patients are characterized by a higher level of anxiety, which can explain their lower S^{all}. Similarly to profile n°1, the lower S^- and WS^- are explained by the joint influence of the MSLT, ESS, and socio-demographic level.

Speakers classified in profile n°5 (n=7) offer a third category of misreading behavior. On the one hand, they make more well-placed pauses, but with a lower average score (Dim. 1 of the PCA): this seems to indicate that the reader makes an effort of nuance, stopping at the implicit pauses of the text in addition to those indicated by the punctuation. This behavior seems to be linked with the age of the speakers included in this cluster. On the other hand, their pauses (both positive and negative) are longer than the average, and they stop where they should not more often but without particularly misplacing their negative pauses, contrary to profiles n°1 and n°3. This behavior seems to find its origin in the bigger neck of the speakers for S^+, N^- and D^-, and in a lower ESS for the latter.

6.2. Patients with above-average reading capacities

The identified profiles do not only contain readers with impaired pausing behaviors. Indeed, the patients of the profile n°2 (n=21) have a lower number of correctly placed pauses than the average, but they are on average better placed (Dim 1. of the PCA). They particularly have a characteristic that they share with the readers of profile n°4 (n=4), namely a low number of negative pauses, which are shorter and have on average higher scores and higher weighted scores (Dim 2. of the PCA). This reflects a lower number of pauses errors, which are more quickly corrected.

While for profile n°2 the diminution of N^+ can be explained by a narrower neck, the lower D^+ and WS^+ are linked to a lower mean age.For this profile, the augmentation of S^+ and diminution of N^- and D^- can be explained by the narrower neck and higher socio-demographic level of these readers. The higher S^- and WS^- can not be explained by our analyses.

Regarding profile n°4, the lower age explains the higher values of N^+ and D^+, and the higher socio-demographic level could be responsible for the lower N^- and D^- values. The latter is also explained by the ESS, which has higher values in this profile. Finally, the higher values for S^- and WS^- could be explained by the interaction between the MSLT, the ESS, and the BMI.

Readers labeled under profile n°6 (n=18) tend to pause a little more on positive locations, but this reading behavior does not deviate significantly from the average reader of the MSLTc. These small deviations can be explained by the corresponding small variations of age, neck size, and socio-demographic level in the same manner as in the other profiles.

Finally, readers of profile n°7 (n=29) stop a little less on positive locations and tend to misplace their pauses, but in the same manner, as the previous paragraph, this does not deviate significantly from the average population.

7. Conclusions

To conclude, we have drawn seven profiles from the reading pauses of 93 patients affected by hypersomnia. These profiles are directed on the one hand by the total number and total duration of pauses, and on the other hand by the number and duration of negative pauses. These reading behaviors are related to patient characteristics, with a predominant influence of age, neck size and socio-demographic level, but also an influence of excessive daytime sleepiness through the joint influence of objective and subjective sleepiness and socio-demographic level.

At the session level, these features do not correlate with the MSLT sleep latency and weakly with the KSS, the session effect masking all the other differences: these features are better fitted for speaker trait estimation than speaker state.

8. References

[1] M. Jike, O. Itani, N. Watanabe, D. J. Buysse, and Y. Kaneita, "Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression," *Sleep Medicine Reviews*, vol. 39, pp. 25–36, Jun. 2018.

[2] A. J. Scott, T. L. Webb, M. Martyn-St James, G. Rowse, and S. Weich, "Improving sleep quality leads to better mental health: A meta-analysis of randomised controlled trials," *Sleep Medicine Reviews*, vol. 60, p. 101556, Dec. 2021.

[3] C. M. Barnes and N. F. Watson, "Why healthy sleep is good for business," *Sleep Medicine Reviews*, vol. 47, pp. 112–118, Oct. 2019.

[4] D. Léger and C. Stepnowsky, "The economic and societal burden of excessive daytime sleepiness in patients with obstructive sleep apnea," *Sleep Medicine Reviews*, vol. 51, p. 101275, Jun. 2020.

[5] S. Bioulac, J.-A. Micoulaud-Franchi, M. Arnaud, P. Sagaspe, N. Moore, F. Salvo, and P. Philip, "Risk of Motor Vehicle Accidents Related to Sleepiness at the Wheel: A Systematic Review and Meta-Analysis," *Sleep*, vol. 40, no. 10, Oct. 2017. [Online]. Available: https://doi.org/10.1093/sleep/zsx134

[6] V. P. Martin, J.-L. Rouas, J.-A. Micoulaud-Franchi, P. Philip, and J. Krajewski, "How to Design a Relevant Corpus for Sleepiness Detection Through Voice?" *Frontiers in Digital Health*, vol. 3, p. 124, 2021. [Online]. Available: https://www.frontiersin.org/article/10.3389/fdgth.2021.686068

[7] T. Åkerstedt and M. Gillberg, "Subjective and objective sleepiness in the active individual." *Int J Neurosci*, vol. 52, pp. 29–37, 1990.

[8] J. Huang, C. Sander, P. Jawinski, C. Ulke, J. Spada, U. Hegerl, and T. Hensch, "Test-retest reliability of brain arousal regulation as assessed with VIGALL 2.0," *Neuropsychiatric Electrophysiology*, vol. 1, no. 1, p. 13, Dec. 2015. [Online]. Available: http://npepjournal.biomedcentral.com/articles/10.1186/s40810-015-0013-9

[9] M. R. Littner, C. Kushida, M. Wise, D. G. Davila, T. Morgenthaler, T. Lee-Chiong, M. Hirshkowitz, D. L. Loube, D. Bailey, R. B. Berry, S. Kapen, and M. Kramer, "Practice Parameters for Clinical Use of the Multiple Sleep Latency Test and the Maintenance of Wakefulness Test," *Sleep*, vol. 28, no. 1, pp. 113–121, 2005.

[10] D. Arand, M. Bonnet, T. Hurwitz, M. Mitler, R. Rosa, and R. B. Sangal, "The Clinical Use of the MSLT and MWT," *SLEEP*, vol. 28, no. 1, pp. 123–144, 2005.

[11] M. W. Johns, "A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale," *Sleep*, vol. 14, no. 6, pp. 540–545, 1991.

[12] L. Dupuy, C. M. Morin, E. Sevin, J. Taillard, N. Salles, S. Bioulac, M. Auriacombe, J. Micoulaud-Franchi, and P. Philip, "Smartphone-based virtual agents and insomnia management: A proof-of-concept study for new methods of autonomous screening and management of insomnia symptoms in the general population," *Journal of Sleep Research*, Sep. 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1111/jsr.13489

[13] M. Auriacombe, L. Fournet, L. Dupuy, J.-A. Micoulaud-Franchi, E. de Sevin, S. Moriceau, E. Baillet, J.-M. Alexandre, F. Serre, and P. Philip, "Effectiveness and Acceptance of a Smartphone-Based Virtual Agent Screening for Alcohol and Tobacco Problems and Associated Risk Factors During COVID-19 Pandemic in the General Population," *Frontiers in Psychiatry*, vol. 12, p. 693687, Jul. 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyt.2021.693687/full

[14] V. P. Martin, J.-L. Rouas, P. Thivel, and J. Krajewski, "Sleepiness detection on read speech using simple features," in *10th Conference on Speech Technology and Human-Computer Dialogue*, Timisoara, Romania, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8906577

[15] V. P. Martin, J.-L. Rouas, and P. Philip, "Détection de la somnolence dans la voix : nouveaux marqueurs et nouvelles stratégies," *Traitement Automatique des Langues*, vol. 61, no. 2, pp. 67–90, 2020.

[16] J. Krajewski, A. Batliner, and M. Golz, "Acoustic sleepiness detection: Framework and validation of a speech-adapted pattern recognition approach," *Behavior Research Methods*, vol. 41, pp. 785–804, 2009.

[17] B. Schuller, S. Steidl, A. Batliner, F. Schiel, and J. Krajewski, "The INTERSPEECH 2011 Speaker State Challenge," in *Interspeech 2011*, 2011, pp. 3201–3204.

[18] F. Eyben and B. Schuller, "Opensmile," *ACM SIGMultimedia Records*, vol. 6, pp. 4–13, 2015.

[19] V. P. Martin, G. Chapouthier, M. Rieant, J.-L. Rouas, and P. Philip, "Using reading mistakes as features for sleepiness detection in speech," in *10th International Conference on Speech Prosody 2020*, Tokyo, Japan, 2020, pp. 985–989. [Online]. Available: http://dx.doi.org/10.21437/SpeechProsody.2020-201

[20] V. P. Martin, J.-L. Rouas, F. Boyer, and P. Philip, "Automatic Speech Recognition system errors for accident-prone sleepiness detection through voice," in *EUSIPCO 2022*, Dublin (virtual), 2021, pp. 541–545.

[21] ——, "Automatic Speech Recognition systems errors for objective sleepiness detection through voice," in *Interspeech 2021*, Brno, 2021, pp. 2476–2480.

[22] F. Tao, A. Esposito, and A. Vinciarelli, "Spotting the Traces of Depression in Read Speech: An Approach Based on Computational Paralinguistics and Social Signal Processing," in *Interspeech 2020*. ISCA, Oct. 2020, pp. 1828–1832.

[23] L. Tóth, I. Hoffmann, G. Gosztolya, V. Vincze, G. Szatlóczki, Z. Bánréti, M. Pákáski, and J. Kálmán, "A speech recognition-based solution for the automatic detection of mild cognitive impairement from spontaneaous speech," *Current Alzheimer Research*, vol. 5, no. 130-138, 2018.

[24] V. P. Martin, B. Arnaud, J.-L. Rouas, and P. Philip, "Automatic Estimation of Pauses Duration and Location in Read Speech as Biomarkers of Sleepiness," *Submitted to ICASSP 2022*, 2021, publisher: Unpublished. [Online]. Available: http://rgdoi.net/10.13140/RG.2.2.14453.91360

[25] V. P. Martin, J.-L. Rouas, J.-A. Micoulaud-Franchi, and P. Philip, "The Objective and Subjective Sleepiness Voice Corpora," in *12th Language Resources and Evaluation Conference*. Marseille, France: European Language Resources Association, 2020, p. 6525-6533. [Online]. Available: https://www.aclweb.org/anthology/2020.lrec-1.803.pdf

[26] A. S. Zigmond and R. P. Snaith, "The hospital anxiety and depression scale," *Acta Psychiatrica Scandinavica*, vol. 67, no. 6, pp. 361–370, 1983.