



BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Principle and protocol for using a smartphone application to study speech biomarkers of Parkinson's disease and other synucleinopathies: SMARTSPEECH

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-059871
Article Type:	Protocol
Date Submitted by the Author:	03-Dec-2021
Complete List of Authors:	Kouba, Tomáš; Czech Technical University in Prague, Department of Circuit Theory Illner, Vojtěch; Czech Technical University in Prague, Department of Circuit Theory Rusz, Jan; Czech Technical University in Prague, Department of Circuit Theory; Charles University, Department of Neurology
Keywords:	Parkinson-s disease < NEUROLOGY, Speech pathology < OTOLARYNGOLOGY, Audiology < OTOLARYNGOLOGY

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Principle and protocol for using a smartphone application to study speech biomarkers of Parkinson's disease and other synucleinopathies: SMARTSPEECH

Tomáš Kouba,^{1§} Vojtěch Illner,^{1§} Jan Rusz^{1*}

¹Department of Circuit Theory, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 160 00, Prague 6, Czech Republic

***Corresponding author address:**

Jan Rusz, MSc, Ph.D., Department of Circuit Theory, Czech Technical University in Prague, Technická 2, 160 00, Praha 6, Czech Republic.

E-mail: rusz.mz@gmail.com

Phone: (+420) 224-352-287

Fax: (+420) 224-311-081

§ both authors contributed equally to this work

Word count: 4282

Keywords: REM sleep behavior disorder, Parkinson's disease, parkinsonism, speech, voice, dysarthria, acoustic analysis.

Declaration of interest: None

Abstract

Introduction: Early identification of Parkinson's disease (PD) in its prodromal stage has fundamental implications for the future development of neuroprotective therapies. However, no sufficiently accurate biomarkers of prodromal PD are currently available to facilitate early identification. The vocal assessment of patients with isolated rapid eye movement sleep behavior disorder (iRBD) and PD appears to have intriguing potential as a diagnostic and progressive biomarker of PD and related synucleinopathies.

Methods and analysis: Speech patterns in the spontaneous speech of iRBD, early PD, and control participants' voice calls will be collected from data acquired via a developed smartphone application over a period of two years. A significant increase in several aspects of PD-related speech disorders is expected, and is anticipated to reflect the underlying neurodegeneration processes.

Ethics and dissemination: The study has been approved by the Ethics Committee of the General University Hospital in Prague, Czech Republic, and all the participants will provide written, informed consent prior to their inclusion in the research. The application satisfies the General Data Protection Regulation law requirements of the European Union. The study findings will be published in peer-reviewed journals and presented at international scientific conferences.

Strengths and limitations of this study

- It is the first study to employ the smartphone-based, cross-sectional passive speech monitoring of iRBD patients who represent a special case of prodromal PD.
- The study will involve two years of longitudinal, home-based monitoring of iRBD and early stage PD patients.
- Automatic passive speech assessment may provide a natural, progressive biomarker reflecting the detrimental processes in the brain's neuronal structures while being directly physiologically interpretable, non-invasive, inexpensive, and scalable to a large population.
- The results of this proof-of-concept study will be based solely on the Czech language, although multilingual scalability of the proposed speech assessment approach is warranted.
- As only one type of smartphone will be used in the study, the impact of different devices on speech outcomes will not be examined.

INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder characterized by the loss of dopaminergic neurons in the substantia nigra.[1] The incidence of PD is approximately 1.8% in persons older than 65 years.[2] There is no treatment that halts or slows the progression of PD, and the pharmacotherapy and neurosurgical interventions that are available only mitigate specific symptoms. A PD diagnosis is typically made when cardinal motor manifestations appear; by this point, up to 50% of the neurons in the substantia nigra may already be irrecoverably damaged.[3] Unfortunately, no sufficiently accurate biomarkers of PD are currently available, although the existence of such biomarkers would allow for the measurement of the effectiveness of experimental treatments in slowing the progression of the disease. Furthermore, no reliable method for identifying people who are at high risk of developing PD exists at present. Thus, establishing a suitable biomarker would be a crucial breakthrough that would impact on diagnoses and future PD treatments, and is one of the most important topics in PD-related research.[4] In particular, motor-related dysfunctions are strong predictors of conversion to synucleinopathy, and have one of the most significant hazard ratios of 3.16 among the available predictive markers to date.[5]

Isolated rapid eye movement sleep behavior disorder (iRBD) is a parasomnia characterized by dream-enactment behaviors associated with REM sleep without muscle atonia[6], and represents a prodromal marker of neurodegenerative synucleinopathies, mainly PD and dementia with Lewy bodies.[7] The risk of developing a neurodegenerative disease is exceptionally high (>80%) in subjects with iRBD.[8, 9] Since symptoms of iRBD precede Parkinsonism, research focused on iRBD is essential for the development of a neuroprotective therapy to counter synucleinopathy,[10] as no other preclinical marker has a predictive value that is comparable to iRBD.[11]

As the most complex acquired human motor skill, and involving over 100 muscles, speech is a sensitive marker of damage to neural structures engaged in the brain's motor system control.[12] Up to 90% of PD patients develop distinctive speech and voice abnormalities, collectively termed hypokinetic dysarthria, which are mainly characterized by a decrease in the quality of the voice, hypokinetic articulation, hypophonia, monopitch, monoloudness, and deficits in timing and phrasing.[13] Based on the recent findings of a multi-language, multi-centric study using an objective acoustical analysis of 150 iRBD patients, it is clear that speech disorders are one of the earliest motor signs of PD.[14] Thus, vocal assessment appears to have intriguing potential as a preclinical diagnostic and progressive biomarker of PD and related neurodegenerations; it is also inexpensive, non-invasive, easy to administer, and has the potential to being conducted remotely from the patient's home (for example, by using a smartphone[15]).

In current practice, measures of PD are primarily subjective, rater-dependent, and require in-clinic assessments.[16] The trials using these measures are lengthy and expensive and, as they are usually based on a once-off assessment, might produce false results. Given the technical and accessibility advancements in smartphone technologies, evaluating motor symptoms of PD via mobile devices continues to be an increasing focus among the research community.[17-19] The potential of smartphones has already been demonstrated in quantifying the severity of PD based on the remote assessment of five tasks (voice, finger

tapping, gait, balance, and reaction time) using an application that was developed for this purpose.[20-22] However, such an assessment requires the patients (or research subjects) to repeatedly perform a set of predefined activities. It is thus reasonable to expect that most patients will not be willing to perform such artificially constructed activities on a daily basis for several years.

By contrast, the extraction of speech patterns from smartphone calls would provide a natural, passive biomarker that does not require additional effort on the part of the subjects. Moreover, such a speech-based application could easily be scaled to a larger population, thus allowing for high-throughput screening, followed by a more detailed analysis if the screen is abnormal. However, the reliability of smartphones in detecting prodromal PD (that is, iRBD) via smartphone calls in realistic scenarios with ambient noise level environments has not yet been investigated.

Therefore, the aim of this study is to develop a fully automated and noise-resistant smartphone-based system that is able to monitor the distinctive speech patterns of neurodegeneration on a daily basis using acoustic data obtained in various environments (SMARTSPEECH). Such a system would have tremendous potential to revolutionize the diagnostic process of PD and could provide a robust biomarker of the progression of the disease. To fully exploit the possibilities of smartphones, a set of active speech and motor tasks is included for sensitivity analysis and for comparison.

METHODS

Collection of speech data

To avoid potential conflicts due to different microphone characteristics across various manufacturers, the speech data will be collected using the same smartphone device, HONOR 9X Lite (Shenzhen Zhixin New Information Technology Co., Ltd.), operating on the Android 9 system. The application will record the subject's telephone speech with a high degree of quality using a sampling frequency of 44.1 kHz and 16-bit quantification.

The distant-speaker filtering algorithm

Due to ethical reasons and the General Data Protection Regulation law, the subjects' speech partners will need to be removed entirely from the recordings, which will be accomplished by employing real-time adaptive filtering of the input audio. The audio will be collected via two different microphones, representing two channels in the stereo mix (see Figure 1). The primary microphone (MIC 1) is the closest to the speaker's mouth and thus captures the subject's speech with absolute power dominance. However, the distant speaker's talk might still be present, such as in segments in which the subject is silent or speaks quietly, or when the distant speaker is speaking very loudly. The secondary microphone (MIC 2) mainly captures the speech of the subject, but also captures the distant speaker (coming from a call speaker nearby, hence at a greater power than via MIC1) and surrounding noise. However, these settings are valid only if the loudspeaker is not activated. In this case, the power of both speakers is more or less equal in the channels. Since the algorithm is not

guaranteed of functioning properly in such scenario, the recording process will be canceled immediately when the user activates the loudspeaker.

In order to remove the distant speaker from the final mix completely, a real-time adaptive algorithm was designed based on the hardware and software settings of the smartphone. Specifically, the computational complexity needed to be optimized because intensive background processes tend to eventually be suppressed by the phone's operating system. The principle, which is cross-channel thresholding using smoothed energy estimates from a specific spectral band, is based on the Neumann-Pearson Criterion. The algorithm's illustrative schema is displayed in Figure 2, and the procedure is as follows:

The input is a stereo call recording $s[n]$ consisting of channel 1 (from MIC 1), $s_1[n]$, and channel 2 (from MIC 2), $s_2[n]$. Both are processed in subsequent windows of a given length of L seconds. In the current frame, the signals are decimated by two, thus reducing the computational burden of the entire procedure while maintaining adequate resolution. The signals are then processed through a Butterworth band-pass filter with a range of 300 to 8000 Hz, as this is the range in which the fundamental speech information is expected to be found. A power estimate is computed from the filtered channels using the L-1 norm, which is simply the sum of the absolute values of the samples in a given segment with a length of w . The estimated power trajectory is then smoothed using a normalized integrator with a forgetting factor λ , producing $S_1[n]$ and $S_2[n]$. The segments containing the subject's speech are then detected when the following condition is met:

$$S_1[n] > k \cdot \text{mean}(S_2[n]), \quad (1)$$

where k represents the extent of the channels' power difference as a ratio of the $s_1[n]$ and $s_2[n]$ standard deviation, calculated as $k = \sqrt{\text{std}(s_1[n])/\text{std}(s_2[n])}$. The standard deviation and the mean in equation (1) are both computed per window. At the end of the window processing section, the timestamps are shifted to compensate for the delay introduced by the normalized integrator. The segments that are not considered to consist of the subject's speech are masked by zeros, and only the output from the first channel is subject to further processing.

A final check is made to ensure that the end of the window does not conflict with the subject's speech segment, which might produce disruptive artifacts. If the result is positive, the output from the current window is discarded, and the process is rerun with an adjusted, shorter L to avoid the conflict. The final output is a mono audio recording containing only the subject's speech.

After the trial testing was conducted, the default value for L was set to 30 seconds as the optimal value for reduced computational power, more precise results, and lower possibilities of conflicts at the end of the window. The segment length for the L-1 norm was set to $w = 128$ samples with an overlap of 50%, which resulted in fast processing with sufficient precision. The forgetting factor was set to $\lambda = 0.99$, producing the optimal smoothing effect for the given scenario.

SMARTSPEECH application for monitoring passive speech via regular smartphone calls

After either an incoming or outgoing call is concluded, the user will be presented with a screen containing an option to delete or save the call (Figure 3a). If the second option is selected, the recording will be filtered immediately using a distant-speaker filtering algorithm, and will be saved to the private application folder on the device. It will be possible to replay or delete each recording from the recent calls list (Figure 3b) in the application within the following 24 hours. After 24 hours, the recording will be compressed and sent to a secured server using SSH and REST API. The original call will then be deleted from the device. The aim is to record at least four five-minute recordings per month from each participant. A setting section on which the user can configure the application to suit their needs will be included (Figure 3c).

Active tasks

The application has a module that invites participants to record three active functional vocal tasks once every 14 days; the participants are instructed to capture the data in everyday environments with a low ambient noise level. In addition, the participants are asked to perform five motor tasks that are commonly used in PD research.[20-23] These tasks will be used to compare the sensitivity of the active and the passive data. Upon launching the active tasks protocol the users are presented with the guidelines and elapsed time since last run (Figure 3d). Before each task, the user will be presented with short, written instructions, including an option to replay an audio or video example of the given task. In the voice section, the recordings can be replayed after being recorded. If the user is dissatisfied with the result, an option to repeat the task is available.

The entire set of active tasks consists of prolonged phonation (Figure 4a), /pa/-/ta/-/ka/ syllable repetition (Figure 4b), reading a passage (Figure 4c, Figure 4d), a tapping game (Figure 5a, Figure 5b), alternated tapping (Figure 5c, Figure 5d), a writing task (Figure 5e, Figure 5f), resting hand tremor (Figure 6a, Figure 6b), and a gait with turnaround (Figure 6c, Figure 6d). Each task will be executed twice, and the motor tasks are to be executed first using the right hand and then the left hand.

Data validation and monitoring

Once the call and task data arrive on our server, they are saved on a database and the basic information becomes visible on the web interface to allow the data parameters to be reviewed for the quality, frequency, and recording duration. In the event that the participant does not make any calls or perform passive tasks within the required period, the participant will be notified via a push notification sent to their email address.

A gateway validation of the incoming data needs to be performed before the data are stored securely on the database to minimize the possibility of corrupted, irrelevant data entering the system, which would produce misleading results.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

First, recordings that are shorter than two seconds will be discarded because they do not contain much information that is relevant and often contain errors. Subsequently, the speaker-detection algorithm will be employed to prevent data contamination when the user hands the smartphone to someone else for the call. The Kaldi toolkit[24] was used to construct the procedure. We employed a high-order Gaussian mixture model trained on the Czech section of the GLOBALPHONE database[25] using a set of mel-frequency cepstral coefficients. New subjects in the study will be enrolled in the model once a sufficient amount of data has been collected. I-vectors are extracted upon detection, and are used to compute the probabilistic linear discriminant analysis score. Using this score, a classifier decides whether the recording belongs to the correct subject. The data that will be assessed as not originating from the given subject will be moved to a specific folder and are not included in the data analysis.

Note that the purpose of the algorithm is also a form of quality control, as it excludes recordings that are of poor quality, such as utterances that are significantly affected by disruptive noise at a low signal-to-noise ratio (SNR). In these cases, the subject's speech is so corrupted that the computed features have little similarity to the subject's model. In addition, an SNR estimator based on minimum statistics[26] is employed in the last validation phase. Segments with an instantaneous SNR lower than the preselected threshold will be excluded from the analysis because the methods for extracting speech features cease to be reliable at certain very low SNR levels (such as pitch detection algorithms at SNR less than 6 dB[27]).

A voice activity detector is applied to the samples that have passed the validation procedure. The output indicates long calls during which the subject only speaks for a tiny fraction of the overall duration. This allows to have a better grasp of the amount of speech data gathered via the system and to notify the subject if the amount of data is insufficient.

Extraction of acoustic speech and motor features

Hypokinetic dysarthria is due to dysfunction in the basal ganglia motor circuit, leading to impairments in the regulation of the initiation, the amplitude, and the velocity of movements.[12,28] Therefore, this project aims to collect, develop, and extract different acoustic features describing motor aspects of speech that have well-defined PD pathophysiology[17] and correspond to the perceptual description of hypokinetic dysarthria defined by Darley et al.[29]

- The following features represent the most anticipated candidates for the final analysis:
- (i) Disruptions in phonation caused by dysfunctions in the vocal folds can be captured using the acoustic measure of Cepstral Peak Prominence, which correlates with the auditory perception of decreased voice quality/breathiness.[30]
 - (ii) Articulation deficits, which are perceived as a decrease in intelligibility, can be described using metrics such as the Vowel Space Area.[31]
 - (iii) Dysprosody is captured by the reduced amplitude of vocal cord movements, which correlates with the impression of monopitch, and can be assessed using the acoustic measure of pitch Variability.[14]
 - (iv) Decreased ability to maintain the speech motor sequence or to alternate quickly between responses can be reflected by the acoustic measures of the Net Speech

Rate and Duration of Pause Intervals, which reflect the perceived auditory timing of speech and may describe deficits such as a slow articulation rate and a reduced ability to intermit and initiate speech.[32]

- (v) Linguistic deficits, including limited vocabulary and a decrease in its range, which might indicate potential underlying mild cognitive impairment, can be assessed using a set of lexical features such as Content Density.[33]
- (vi) To assess the overall degree of speech impairment, powerful deep neural networks could be employed, including spectrogram analysis to evaluate specific features when traditional methods might be insufficient. Insight into the detailed method behavior would be supportive, and would reveal critical physiological details.[34]

Several methods for the automatic analysis of key dimensions of speech in PD patients have already been developed.[32,35,36] However, all the methods need to undergo experimental and theoretical testing for noise robustness and reliability to validate their usefulness.[27] A set of features commonly used in the existing literature will be extracted from the active motor tasks;[20-22] the set includes tapping velocity, intratap duration variability, reaction time, tremor acceleration skewness, and velocity.

Study design and participants

During this project, each participant will be given a full explanation of the project's purpose and aims, will be informed about the procedure, and will be given the opportunity to ask questions before deciding whether to sign the informed consent form. In total, we plan to recruit 25 iRBD subjects, 25 early stage PD patients, and 25 healthy controls as part of this longitudinal study. As we expect a drop-out rate of about 20% per year[17], up to 50 iRBD patients and the same number of early stage PD patients will be recruited at the baseline. These data might be used for a better-sampled cross-sectional study. All the iRBD subjects will fulfill the criteria listed in the International Classification of Sleep Disorders, third edition diagnostic criteria,[37] including confirmation of REM sleep without atonia via polysomnography. The inclusion criteria for iRBD will be

- (i) onset of iRBD after 50 years of age,
- (ii) no history of major neurological disease (such as epilepsy or strokes) or other significant diseases that could affect study participation or voice analysis (for example, active cancer, drug abuse, or diseased vocal cords),
- (iii) no significant cognitive decline or severe depression, and
- (iv) no history of therapy with antiparkinsonian medication.

All the PD patients will meet the Movement Disorders Society's clinical diagnostic criteria for PD,[39] and will be investigated during the on-medication state. The inclusion criteria for PD will be:

- (i) onset of PD after 50 years of age,
- (ii) Hoehn and Yahr stage 1–2 in the on-medication state,
- (iii) disease duration more than five years after diagnosis,

- (iv) no history of a major neurological disease other than PD (such as epilepsy or strokes) or other significant diseases that could affect study participation or voice analysis (for example, active cancer, drug abuse, or diseased vocal cords),
- (v) no significant cognitive decline or severe depression, and
- (vi) no involvement in any speech therapy during the duration of the project.

The inclusion criterion for controls will be that the participants have no history of neurological or communication disorders, and no history of parasomnias or other sleep disorders. PD and healthy control subjects will be age- and gender-matched to the iRBD group.

Each participant will be required to perform passive speech recordings and active tasks using the provided smartphone for two years. In addition, each of the subjects will undergo three examinations at the clinic: at the baseline, after one year, and after two years (at the end of the project). The clinical examinations will consist of taking a structured personal history, quantitative testing of motor and non-motor symptoms of PD based on the Movement Disorder Society-Unified Parkinson's Disease Rating Scale, [38] cognitive testing using the Montreal Cognitive Assessment,[39] autonomic testing using the Scales for Outcomes in Parkinson's Disease-Autonomic Dysfunction,[40] the evaluation of depressive symptoms using the Beck Depression Inventory II,[41] and speech examinations according to the dysarthria research guidelines for acoustic analyses.[23] Recruitment of the participants will begin in October 2021.

Patient and public involvement statement

Most of the patients have been involved in previous studies and are familiar with the team, the research topic, and the methods. The SMARTSPEECH protocol was designed with the aid of a short questionnaire that was completed by a selected representative group of 33 patients with iRBD or PD, which provided insight into their requirements (for example, a dual-SIM phone and monetary compensation for mobile phone tariffs), doubts and motivations (such as understanding how to operate the phone and data security), and the expected frequency of phone usage and calls.

Sample size estimation and statistical analysis

For the primary endpoint of the project, an ad-hoc power analysis for a given large effect size (Cohen's d of 0.8), with the error probability (α) set at 0.05 and a false negative rate (β) set at 0.2 (that is, the power of 0.8), based on a three-group analysis of variance with one covariate (group), determined a minimum sample size of 66 subjects with at least 22 subjects in each subgroup (that is, 22 iRBD patients, 22 PD patients, and 22 controls).

A one-way analysis of variance with a group (PD versus iRBD versus controls) as a between-subject factor will be used to calculate the differences for each parameter of interest. The Pearson correlation coefficient will be applied to search for correlations among the variables. The minimum level of significance will be set at $p < 0.05$ with an appropriate Bonferroni adjustment, and a binary logistic regression followed by a leave-one-subject-out

cross-validation will be used to assess the sensitivity/specificity of the proposed features to differentiate the groups.

ETHICS

The study has been approved by the Ethics Committee of the General University Hospital in Prague, Czech Republic (no. 30/19 Grant AZV VES 2020 VFN), and will be performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All the participants will provide written, informed consent prior to their inclusion.

The following steps will be taken to mitigate potential ethical concerns: The researchers performing the analyses will have no known relationship with the research subjects, and the recordings will be de-identified to the extent possible. The analyses of audio recordings will be automated to avoid the recordings being listened to by a human; furthermore, the audio recordings will be encrypted, will only be available to the authorized researchers, and will be deleted at the request of any participant without the need for justification. All the steps will be conducted according to the directive on personal data protection legislation in the Czech Republic and the approval of the Ethics Committee.

DISCUSSION

It is widely known that speech performance varies depending on the actual mood, time of day, fatigue, cold, attitude toward a specific person, and other factors. However, we assume there will be a considerable decline in the speech quality of the iRBD and PD subjects over the two-year duration of the continuous measurement, whereas no trend toward a decline will be manifested by the healthy controls (Figure 7). A significant decrease in motor performance has been observed in iRBD patients several years before they developed Parkinsonism.[42] In addition, a period of just one year was found to be sufficiently long to capture a significant decline in the speech of iRBD subjects.[14] In summary, a two-year duration for recording represents a clinically relevant period for the detection of longitudinal speech impairment, as well as being a comfortable study duration for the research subjects; for example, a three-year period would be likely to be associated with a considerably higher drop-out rate.

The SMARTSPEECH project may provide a natural digital biomarker of disease progression based on longitudinal data acquired without any cost or time burden on the patients and investigators. To the best of our knowledge, there is currently no biomarker that would provide similar opportunities to speech assessment via a smartphone. The advantages of the proposed research approach are that (i) it is highly sensitive, as most PD patients already manifest specific speech disorders at the time of diagnosis,[43] (ii) it has up to 100% specificity, as it is highly likely that the decline of speech observed in subjects with iRBD over several months or years represents the underlying neurodegenerative process, (iii) no cost is involved, as it could be applied using any device that allows remote speech communication in the future, (iv) it is non-invasive, entirely passive, and requires no additional effort on the part of investigated subjects, (v) it has a long lead time, as

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

longitudinal speech changes in iRBD were estimated as being the first motor signs to develop.[42]

In the future, SMARTSPEECH algorithms can be implemented directly on smartphones and analyzed on the device offline; thus, only statistical trends related to speech performance will be stored, which means that no personal data will need to be shared or transferred. However, the robustness of the available features for the analysis of speech data obtained via low-level microphones in various environments has yet to be tested, as in the studies investigating noise resilience and effect of microphone type on pitch variability.[15,27]

The SMARTSPEECH methods could also be applied in the case of already diagnosed PD patients to enhance the current treatment strategies. Quick, inexpensive, and non-invasive vocal assessments using smartphones may allow for a personalized implementation of therapeutic strategies by providing rapid feedback after exercise, monitoring the effects of pharmacological therapies (including advanced drug delivery systems and modifying medication doses according to immediate needs), and the modification of speech-related side-effects of deep brain stimulation via the re-programming and optimization of stimulation parameters.

CONCLUSIONS

The SMARTSPEECH project, driven by the advances in smartphones and adapted speech-processing algorithms, has the potential to introduce a novel, game-changing methodology into the diagnosis, monitoring, and treatment of PD and other synucleinopathies. Longitudinally tracking specific speech markers and other motor features can sensitively capture the underlying neurodegeneration processes at early stages and assess the severity and effects of neuroprotective treatment while being inexpensive, non-invasive, not demanding for the patients, and scalable to a large population.

AUTHORS' CONTRIBUTION

JR conceived of the study and the study design. TK and VI drafted and revised the manuscript. TK and VI designed the application. TK, VI and JR approved the final manuscript.

FUNDING

This study was supported by the Czech Ministry of Health (grant no. NU20-08-00445).

REFERENCES

1 Poewe W, Seppi K, Tanner C, et al. Parkinson disease. *Nat Rev Dis Primers* 2017;3. doi:10.1038/nrdp.2017.13

- 2 de Rijk M, Launer L, Berger K, et al. Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. *Neurol* 2000;54:S21-3.http://intl.neurology.org/cgi/content/abstract/54/11_suppl_5/S21
- 3 Rodriguez-Oroz M, Jahanshahi M, Krack P, et al. Initial clinical manifestations of Parkinson's disease: features and pathophysiological mechanisms. *Lancet Neurol* 2009;8:1128-1139. doi:10.1016/S1474-4422(09)70293-5
- 4 Postuma R, Berg D. Advances in markers of prodromal Parkinson disease. *Nat Rev Neurol* 2016;12:622-634. doi:10.1038/nrneurol.2016.152
- 5 Postuma R, Iranzo A, Hu M, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. *Brain* 2019;142:744-759. doi:10.1093/brain/awz030
- 6 Boeve B. REM sleep behavior disorder. *Ann N Y Acad Sci* 2010;1184:15-54. doi:10.1111/j.1749-6632.2009.05115.x
- 7 Miglis M, Adler C, Antelmi E, et al. Biomarkers of conversion to α -synucleinopathy in isolated rapid-eye-movement sleep behaviour disorder. *Lancet Neurol* 2021;20:671-684. doi:10.1016/S1474-4422(21)00176-9
- 8 Iranzo A, Tolosa E, Gelpi E, et al. Neurodegenerative disease status and post-mortem pathology in idiopathic rapid-eye-movement sleep behaviour disorder: an observational cohort study. *Lancet Neurol* 2013;12:443-453. doi:10.1016/S1474-4422(13)70056-5
- 9 Schenck C, Boeve B, Mahowald M. Delayed emergence of a parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder: a 16-year update on a previously reported series. *Sleep Med* 2013;14:744-748. doi:10.1016/j.sleep.2012.10.009
- 10 Schenck C, Montplaisir J, Frauscher B, et al. Rapid eye movement sleep behavior disorder: devising controlled active treatment studies for symptomatic and neuroprotective therapy—a consensus statement from the International Rapid Eye Movement Sleep Behavior Disorder Study Group. *Sleep Med* 2013;14:795-806. doi:10.1016/j.sleep.2013.02.016
- 11 Postuma R, Aarsland D, Barone P, et al. Identifying prodromal Parkinson's disease: Pre-Motor disorders in Parkinson's disease. *Mov Disord* 2012;27:617-626. doi:10.1002/mds.24996
- 12 Duffy J. *Motor speech disorders: substrates, differential diagnosis, and management*. Fourth. Maryland Heights: : Elsevier 2019.
- 13 Ho A, Iansek R, Marigliani C, et al. Speech Impairment in a Large Sample of Patients with Parkinson's Disease. *Behav Neurol* 1999;11:131-137. doi:10.1155/1999/327643
- 14 Rusz J, Hlavnička J, Novotný M, et al. Speech Biomarkers in Rapid Eye Movement Sleep Behavior Disorder and Parkinson Disease. *Ann Neurol* 2021;90:62-75. doi:10.1002/ana.26085
- 15 Rusz J, Hlavnicka J, Tykalova T, et al. Smartphone Allows Capture of Speech Abnormalities Associated With High Risk of Developing Parkinson's Disease. *IEEE Trans Neural Syst Rehabil Eng* 2018;26:1495-1507. doi:10.1109/TNSRE.2018.2851787

- 16 Goetz C, Tilley B, Shaftman S, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord* 2008;23:2129-2170. doi:10.1002/mds.22340
- 17 Arora S, Venkataraman V, Zhan A, et al. Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study. *Parkinsonism Relat Disord* 2015;21:650-653. doi:10.1016/j.parkreldis.2015.02.026
- 18 Espay A, Bonato P, Nahab F, et al. Technology in Parkinson's disease: Challenges and opportunities. *Mov Disord* 2016;31:1272-1282. doi:10.1002/mds.26642
- 19 Artusi C, Mishra M, Latimer P, et al. Integration of technology-based outcome measures in clinical trials of Parkinson and other neurodegenerative diseases. *Parkinsonism Relat Disord* 2018;46:S53-S56. doi:10.1016/j.parkreldis.2017.07.022
- 20 Zhan A, Mohan S, Tarolli C, et al. Using Smartphones and Machine Learning to Quantify Parkinson Disease Severity. *JAMA Neurol* 2018;75. doi:10.1001/jamaneurol.2018.0809
- 21 Lipsmeier F, Taylor K, Kilchenmann T, et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. *Mov Disord* 2018;33:1287-1297. doi:10.1002/mds.27376
- 22 Arora S, Baig F, Lo C, et al. Smartphone motor testing to distinguish idiopathic REM sleep behavior disorder, controls, and PD. *Neurol* 2018;91:e1528-e1538. doi:10.1212/WNL.0000000000006366
- 23 Rusz J, Tykalova T, Ramig L, et al. Guidelines for Speech Recording and Acoustic Analyses in Dysarthrias of Movement Disorders. *Mov Disord* 2021;36:803-814. doi:10.1002/mds.28465
- 24 Povey D, Ghoshal A, Boulianne G, et al. The Kaldi Speech Recognition Toolkit. *Proc IEEE Workshop Autom Speech Recognit Underst* 2011.
- 25 Schultz T, Vu N, Schlippe T. GlobalPhone: A multilingual text & speech database in 20 languages. In: *2013 Proc IEEE Int Conf Acoust Speech Signal Process* 2013. 8126-8130. doi:10.1109/ICASSP.2013.6639248
- 26 Martin R. Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Trans Speech Audio Process*;9:504-512. doi:10.1109/89.928915
- 27 Illner V, Sovka P, Rusz J Validation of freely-available pitch detection algorithms across various noise levels in assessing speech captured by smartphone in Parkinson's disease. *Biomed Signal Process Control* 2020;58. doi:10.1016/j.bspc.2019.101831
- 28 Ho A, Iansek R, Marigliani C, et al. Speech Impairment in a Large Sample of Patients with Parkinson's Disease. *Behav Neurol* 1999;11:131-137. doi:10.1155/1999/327643
- 29 Darley F, Aronson A, Brown J. Differential Diagnostic Patterns of Dysarthria. *J Speech Lang Hear Res* 1969;12:246-269. doi:10.1044/jshr.1202.246
- 30 Jannetts S, Lowit A. Cepstral Analysis of Hypokinetic and Ataxic Voices: Correlations With Perceptual and Other Acoustic Measures. *J Voice* 2014;28:673-680. doi:10.1016/j.jvoice.2014.01.013
- 31 Rusz J, Cmejla R, Tykalova T, et al. Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task. *J Acoust Soc Am* 2013;134:2171-2181. doi:10.1121/1.4816541
- 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. doi:10.1038/s41598-017-00047-5

- 33 Beltrami D, Gagliardi G, Rossini Favretti R, et al. Speech Analysis by Natural Language Processing Techniques: A Possible Tool for Very Early Detection of Cognitive Decline?. *Front Aging Neurosci* 2018;10. doi:10.3389/fnagi.2018.00369
- 34 Laganas C, Iakovakis D, Hadjimitsiou S, et al. Parkinson's Disease Detection Based on Running Speech Data From Phone Calls. *IEEE Trans Biomed Eng* 2021;:1-1. doi:10.1109/TBME.2021.3116935
- 35 Rusz J, Cmejla R, Ruzickova H, et al. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *J Acoust Soc Am* 2011;129:350-367. doi:10.1121/1.3514381
- 36 Novotny M, Rusz J, Cmejla R, et al. Automatic Evaluation of Articulatory Disorders in Parkinson's Disease. *IEEE/ACM Trans Audio Speech Lang Process* 2014;22:1366-1378. doi:10.1109/TASLP.2014.2329734
- 37 American Academy of Sleep Medicine *International Classification of Sleep Disorders, Third Edition: Diagnostic and Coding Manual*. Chicago, Illinois: : American Academy of Sleep Medicine 2014.
- 38 Postuma R, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2015;30:1591-1601. doi:10.1002/mds.26424
- 39 Nasreddine Z, Phillips N, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc* 2005;53:695-699. doi:10.1111/j.1532-5415.2005.53221.x
- 40 Visser M, Marinus J, Stiggelbout A, et al. Assessment of autonomic dysfunction in Parkinson's disease: The SCOPA-AUT. *Mov Disord* 2004;19:1306-1312. doi:10.1002/mds.20153
- 41 Beck A, Steer R, Brown G. Manual for the Beck Depression Inventory-II. *San Antonio, TX: Psychological Corporation* 1996. doi:10.1007/978-1-4419-1005-9_441
- 42 Postuma R, Lang A, Gagnon J, et al. How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behaviour disorder. *Brain* 2012;135:1860-1870. doi:10.1093/brain/aws093
- 43 Rusz J, Tykalová T, Novotný M, et al. Automated speech analysis in early untreated Parkinson's disease: Relation to gender and dopaminergic transporter imaging. *Eur J Neurol* 2021. doi:10.1111/ene.15099

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FIGURE CAPTIONS

Figure 1. Schematics of the HONOR 9X smartphone audio inputs and outputs.

Figure 2. An illustrative schema of the distant-speaker filtering algorithm.

Figure 3. (a) Option screen to delete or save the recently completed call for speech analysis. (b) List of calls from the previous 24 hours with details and the options of replaying and deletion. (c) The setting screen of the app. (d) The active task introduction screen, which shows a guide to performing the active tasks and the time elapsed since the last active task run.

Figure 4. (a) Prolonged phonation task instruction screen (task duration ~15 seconds). (b) Sequential motion rates instruction screen (task duration 7 seconds). (c) Reading passage instruction screen. (d) Reading passage sample text that is chosen randomly from six samples and contains approximately 80 words (task duration ~40 seconds).

Figure 5. (a) A tapping game introduction screen. (b) A tapping game task. The delay between spawning the circle and a successful tap is measured. The action takes place 20 times (task duration ~20 seconds). (c) An alternated tapping introduction screen. (d) An alternated tapping task. The timestamps of the individual taps are measured (task duration 15 seconds). (e) Writing task instruction screen. (f) A writing task. The time spent rewriting each word is measured. The patient is shown five words in each run. Eight different datasets that are assembled in the same manner are chosen randomly in each run. Examples include the Czech words "SESTRICKA", "NEJLEPSI", "LECENI", "ZABAVA", and "PREDSTAVA". The task can be skipped after 80 seconds have elapsed (task duration ~40 seconds).

Figure 6. (a) Rest tremor instruction screen. (b) A rest tremor task (task duration 15 seconds). The patient's hand tremor is measured using an accelerometer in all three axes. The z-axis is live plotted for user feedback. (c) A gait task instruction screen. (d) A gait task (task duration ~30 seconds). The z-axis rotation is measured and plotted on the screen for user feedback via an internal gyroscope.

Figure 7. Fictive graphical demonstration of the expected result of the measurements using SMARTSPEECH. The solid line represents the overall trend in the decline of speech performance over two years. The points represent the average performance of pitch variability for each one-minute duration of the telephone call.

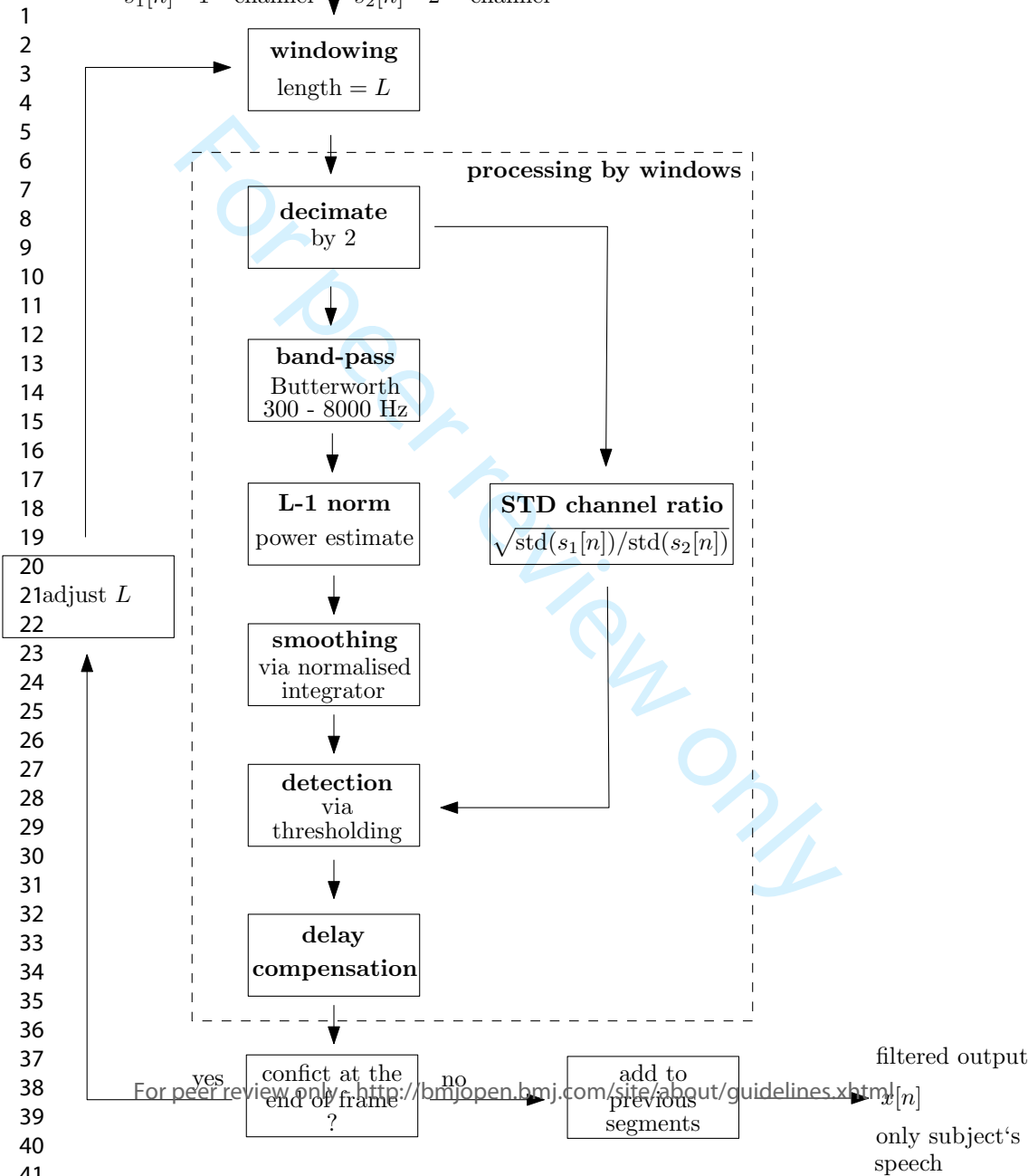


Figure 1

66x80mm (300 x 300 DPI)

$s[n]$

$s_1[n]$ - 1st channel $s_2[n]$ - 2nd channel



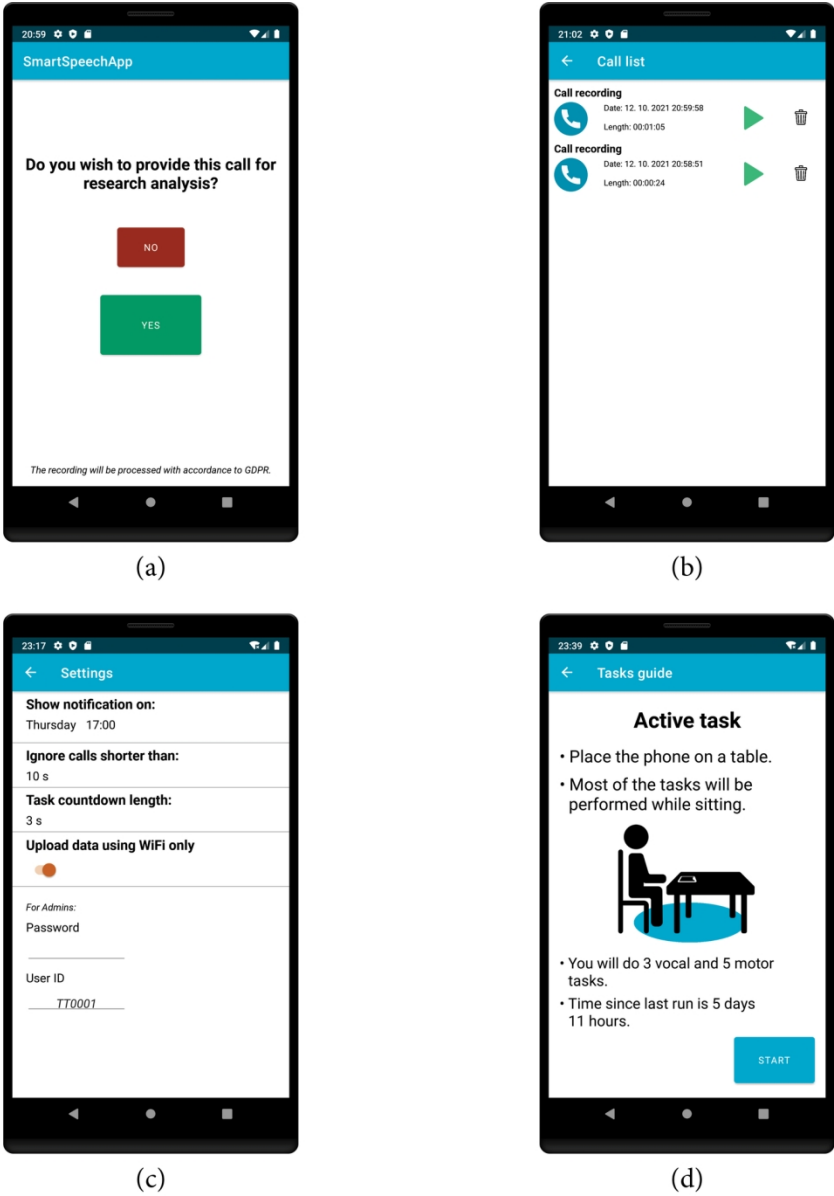


Figure 3

143x183mm (300 x 300 DPI)

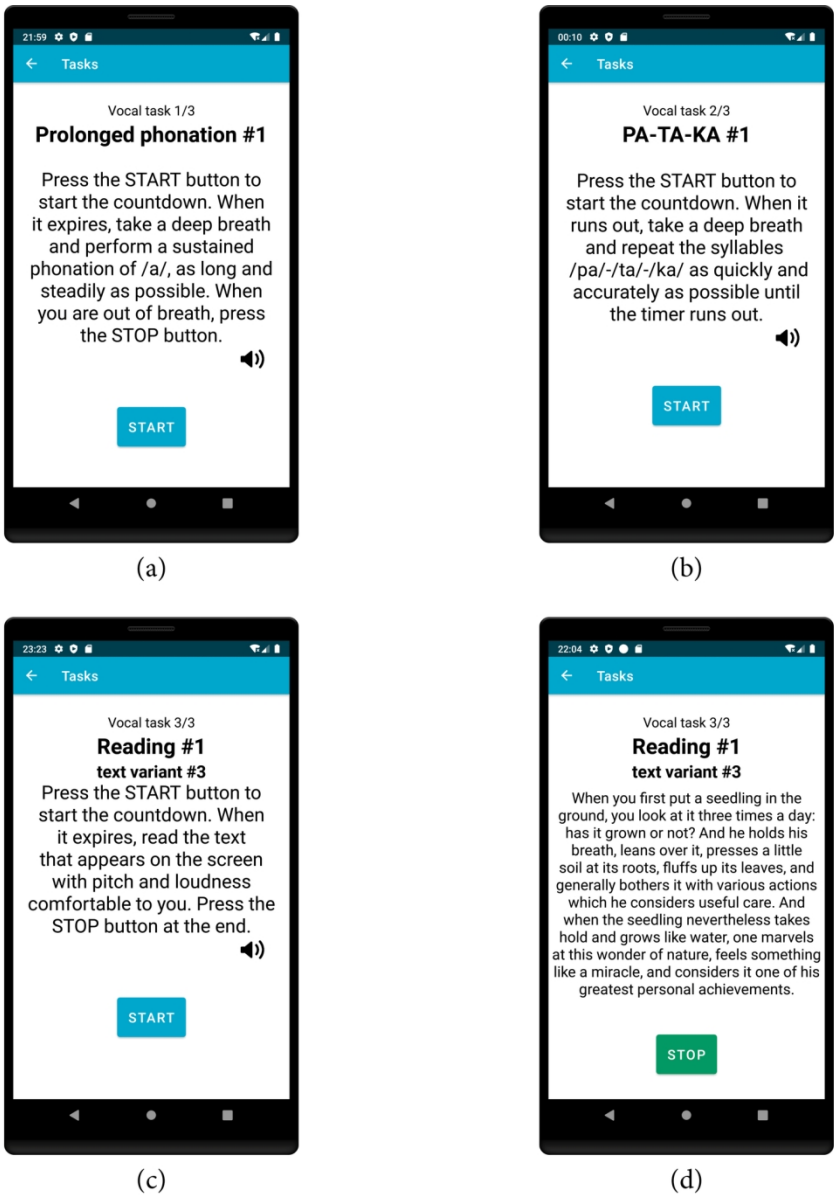


Figure 4
143x183mm (300 x 300 DPI)

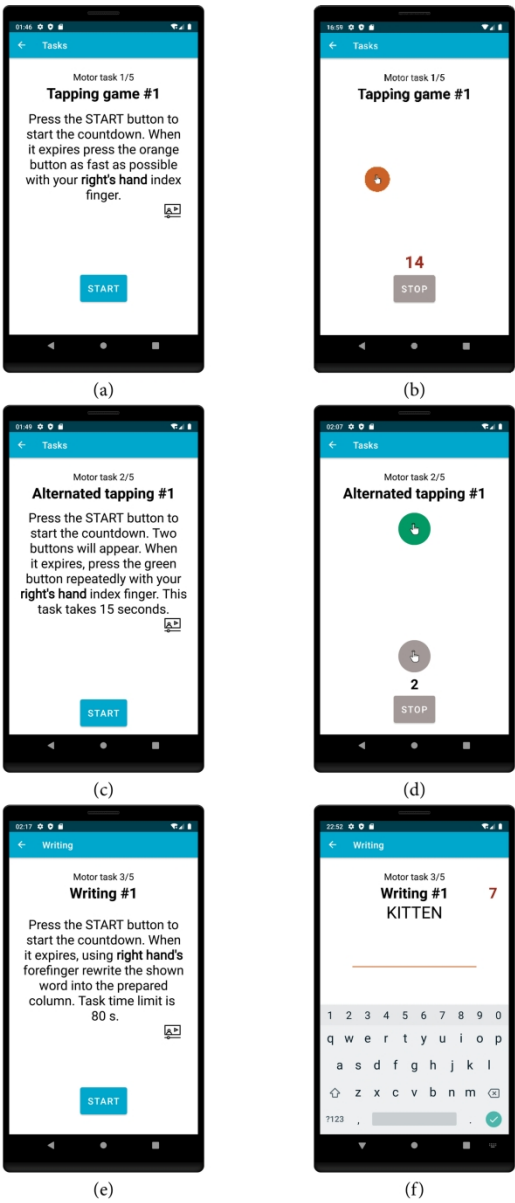
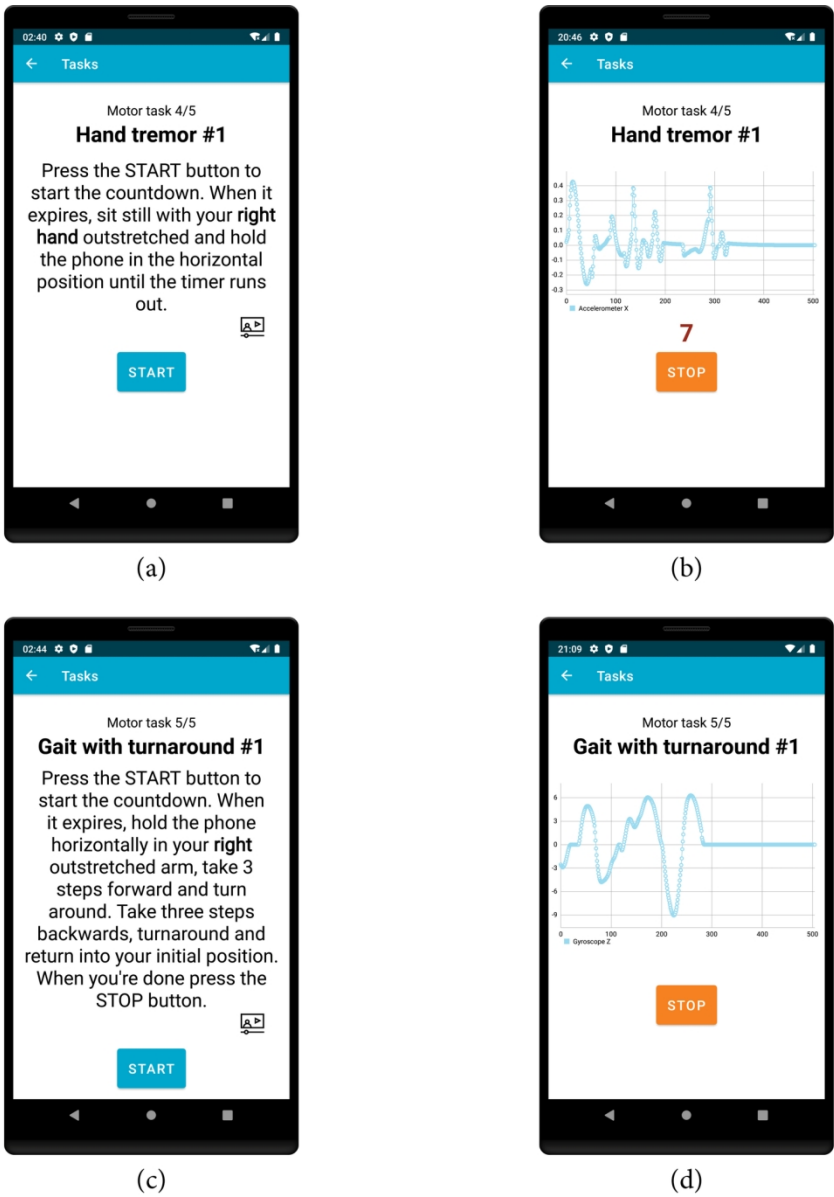


Figure 5

131x269mm (300 x 300 DPI)



(a)

(b)

(c)

(d)

Figure 6

143x183mm (300 x 300 DPI)

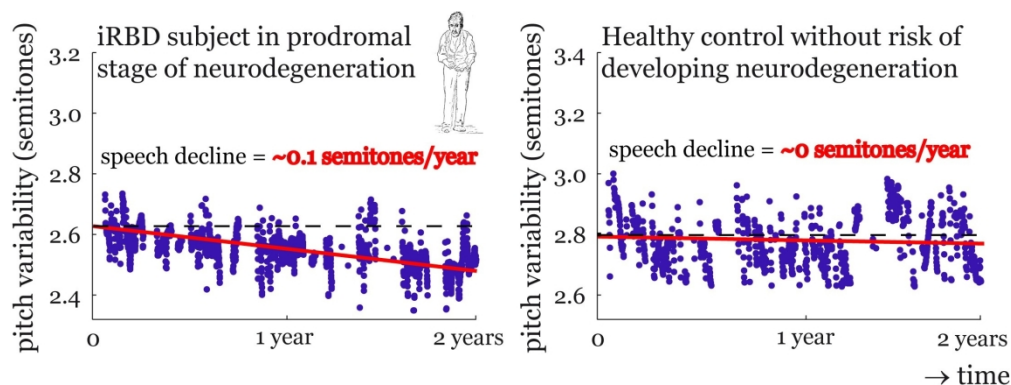


Figure 7

186x70mm (300 x 300 DPI)

BMJ Open

Study protocol for using a smartphone application to investigate speech biomarkers of Parkinson's disease and other synucleinopathies: SMARTSPEECH

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-059871.R1
Article Type:	Protocol
Date Submitted by the Author:	22-Feb-2022
Complete List of Authors:	Kouba, Tomáš; Czech Technical University in Prague, Department of Circuit Theory Illner, Vojtěch; Czech Technical University in Prague, Department of Circuit Theory Rusz, Jan; Czech Technical University in Prague, Department of Circuit Theory; Charles University, Department of Neurology
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Ear, nose and throat/otolaryngology
Keywords:	Parkinson-s disease < NEUROLOGY, Speech pathology < OTOLARYNGOLOGY, Audiology < OTOLARYNGOLOGY

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Study protocol for using a smartphone application to investigate speech biomarkers of Parkinson's disease and other synucleinopathies: SMARTSPEECH

Tomáš Kouba,^{1§} Vojtěch Illner,^{1§} Jan Ruz^{1*}

¹Department of Circuit Theory, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 160 00, Prague 6, Czech Republic

***Corresponding author address:**

Jan Ruz, MSc, Ph.D., Department of Circuit Theory, Czech Technical University in Prague, Technická 2, 160 00, Praha 6, Czech Republic.

E-mail: ruz.mz@gmail.com

Phone: (+420) 224-352-287

Fax: (+420) 224-311-081

§ both authors contributed equally to this work

Word count: 4542

Keywords: REM sleep behavior disorder, Parkinson's disease, parkinsonism, speech, voice, dysarthria, acoustic analysis.

Declaration of interest: None

Abstract

Introduction: Early identification of Parkinson's disease (PD) in its prodromal stage has fundamental implications for the future development of neuroprotective therapies. However, no sufficiently accurate biomarkers of prodromal PD are currently available to facilitate early identification. The vocal assessment of patients with isolated rapid eye movement sleep behavior disorder (iRBD) and PD appears to have intriguing potential as a diagnostic and progressive biomarker of PD and related synucleinopathies.

Methods and analysis: Speech patterns in the spontaneous speech of iRBD, early PD, and control participants' voice calls will be collected from data acquired via a developed smartphone application over a period of two years. A significant increase in several aspects of PD-related speech disorders is expected, and is anticipated to reflect the underlying neurodegeneration processes.

Ethics and dissemination: The study has been approved by the Ethics Committee of the General University Hospital in Prague, Czech Republic, and all the participants will provide written, informed consent prior to their inclusion in the research. The application satisfies the General Data Protection Regulation law requirements of the European Union. The study findings will be published in peer-reviewed journals and presented at international scientific conferences.

Strengths and limitations of this study

- This study aims to employ the remote longitudinal speech monitoring of prodromal and early-stage patients with Parkinson's disease via smartphone calls.
- An additional set of active speech and motor tasks will be captured.
- The results of this study will be based solely on the Czech language.
- This study will be implemented in one type of smartphone, and the impact of different devices on speech outcomes is yet to be determined.

INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder characterized by the loss of dopaminergic neurons in the substantia nigra.[1] The incidence of PD is approximately 1.8% in persons older than 65 years.[2] There is no treatment that halts or slows the progression of PD, and the pharmacotherapy and neurosurgical interventions that are available only mitigate specific symptoms. A PD diagnosis is typically made when cardinal motor manifestations appear; by this point, up to 50% of the neurons in the substantia nigra may already be irrecoverably damaged.[3] Unfortunately, no sufficiently accurate biomarkers of PD are currently available, although the existence of such biomarkers would allow for the measurement of the effectiveness of experimental treatments in slowing the progression of the disease. Furthermore, no reliable method for identifying people who are at high risk of developing PD exists at present. Thus, establishing a suitable biomarker would be a crucial breakthrough that would impact on diagnoses and future PD treatments, and is one of the most important topics in PD-related research.[4] In particular, motor-related dysfunctions are strong predictors of conversion to synucleinopathy, and have one of the most significant hazard ratios of 3.16 among the available predictive markers to date.[5]

Isolated rapid eye movement sleep behavior disorder (iRBD) is a parasomnia characterized by dream-enactment behaviors associated with REM sleep without muscle atonia[6], and represents a prodromal marker of neurodegenerative synucleinopathies, mainly PD and dementia with Lewy bodies.[7] The risk of developing a neurodegenerative disease is exceptionally high (>80%) in subjects with iRBD.[8, 9] Since symptoms of iRBD precede Parkinsonism, research focused on iRBD is essential for the development of a neuroprotective therapy to counter synucleinopathy,[10] as no other preclinical marker has a predictive value that is comparable to iRBD.[11]

As the most complex acquired human motor skill, and involving over 100 muscles, speech is a sensitive marker of damage to neural structures engaged in the brain's motor system control.[12] Up to 90% of PD patients develop distinctive speech and voice abnormalities, collectively termed hypokinetic dysarthria, which are mainly characterized by a decrease in the quality of the voice, hypokinetic articulation, hypophonia, monopitch, monoloudness, and deficits in timing and phrasing.[13] Based on the recent findings of a multi-language, multi-centric study using an objective acoustical analysis of 150 iRBD patients, it is clear that speech disorders are one of the earliest motor signs of PD.[14] Thus, vocal assessment appears to have intriguing potential as a preclinical diagnostic and progressive biomarker of PD and related neurodegenerations; it is also inexpensive, non-invasive, easy to administer, and has the potential to being conducted remotely from the patient's home (for example, by using a smartphone[15]).

In current practice, measures of PD are primarily subjective, rater-dependent, and require in-clinic assessments.[16] The trials using these measures are lengthy and expensive and, as they are usually based on a once-off assessment, might produce false results. Given the technical and accessibility advancements in smartphone technologies, evaluating motor symptoms of PD via mobile devices continues to be an increasing focus among the research community.[17-19] The potential of smartphones has already been demonstrated in quantifying the severity of PD based on the remote assessment of five tasks (voice, finger

tapping, gait, balance, and reaction time) using an application that was developed for this purpose.[20-22] However, such an assessment requires the patients (or research subjects) to repeatedly perform a set of predefined activities. It is thus reasonable to expect that most patients will not be willing to perform such artificially constructed activities on a daily basis for several years.

By contrast, the extraction of speech patterns from smartphone calls would provide a natural, passive biomarker that does not require additional effort on the part of the subjects. Moreover, such a speech-based application could easily be scaled to a larger population, thus allowing for high-throughput screening, followed by a more detailed analysis if the screen is abnormal. However, the reliability of smartphones in detecting prodromal PD (that is, iRBD) via smartphone calls in realistic scenarios with ambient noise level environments has not yet been investigated.

Therefore, the aim of this study is to develop a fully automated and noise-resistant smartphone-based system that is able to monitor the distinctive speech patterns of neurodegeneration on a daily basis using acoustic data obtained in various environments (SMARTSPEECH). Such a system would have tremendous potential to revolutionize the diagnostic process of PD and could provide a robust biomarker of the progression of the disease. To fully exploit the possibilities of smartphones, a set of active speech and motor tasks is included for sensitivity analysis and for comparison.

METHODS

Objectives

The main objective of the study is to demonstrate that, using the SMARTSPEECH system, speech performance elicited during regular phone calls through a smartphone can provide principal biomarkers for diagnosis and monitoring the progression of prodromal PD.

The specific objectives are to:

- (i) develop a smartphone application that will be able to capture the subjects' data,
- (ii) collect up to two-year longitudinal speech data from subjects with iRBD, patients with early-stage PD, and healthy control subjects of comparable age,
- (iii) collate existing approaches and develop novel methods allowing assessment speech markers of neurodegeneration in PD, including tests of their robustness against noise and recording conditions and selecting the most appropriate parameters for smartphone-based monitoring,
- (iv) build the concept of a complex system (SMARTSPEECH) for detection of speech abnormalities in PD and other synucleinopathies, including its statistical power evaluation in differentiation between PD, iRBD, and control groups and identify its relationship to essential clinical markers reflecting disease progression such as the Movement Disorder Society-Unified Parkinson's Disease Rating Scale,
- (v) analyze the sensitivity of a set of active speech and motor tasks and compare the outcomes to data acquired through SMARTSPEECH passive calls monitoring.

Collection of speech data

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

To avoid potential conflicts due to different microphone characteristics across various manufacturers, the speech data will be collected using the same smartphone device, HONOR 9X Lite (Shenzhen Zhixin New Information Technology Co., Ltd.), operating on the Android 9 system. The application will record the subject's telephone speech with a high degree of quality using a sampling frequency of 44.1 kHz and 16-bit quantification.

The distant-speaker filtering algorithm

Due to ethical reasons and the General Data Protection Regulation law, the subjects' speech partners will need to be removed entirely from the recordings, which will be accomplished by employing real-time adaptive filtering of the input audio. The audio will be collected via two different microphones, representing two channels in the stereo mix (see Figure 1). The primary microphone (MIC 1) is the closest to the speaker's mouth and thus captures the subject's speech with absolute power dominance. However, the distant speaker's talk might still be present, such as in segments in which the subject is silent or speaks quietly, or when the distant speaker is speaking very loudly. The secondary microphone (MIC 2) mainly captures the speech of the subject, but also captures the distant speaker (coming from a call speaker nearby, hence at a greater power than via MIC1) and surrounding noise. However, these settings are valid only if the loudspeaker is not activated. In this case, the power of both speakers is more or less equal in the channels. Since the algorithm is not guaranteed of functioning properly in such scenario, the recording process will be canceled immediately when the user activates the loudspeaker.

In order to remove the distant speaker from the final mix completely, a real-time adaptive algorithm was designed based on the hardware and software settings of the smartphone. Specifically, the computational complexity needed to be optimized because intensive background processes tend to eventually be suppressed by the phone's operating system. The principle, which is cross-channel thresholding using smoothed energy estimates from a specific spectral band, is based on the Neumann-Pearson Criterion. The algorithm's illustrative schema is displayed in Figure 2, and the procedure is as follows:

The input is a stereo call recording $s[n]$ consisting of channel 1 (from MIC 1), $s_1[n]$, and channel 2 (from MIC 2), $s_2[n]$. Both are processed in subsequent windows of a given length of L seconds. In the current frame, the signals are decimated by two, thus reducing the computational burden of the entire procedure while maintaining adequate resolution. The signals are then processed through a Butterworth band-pass filter with a range of 300 to 8000 Hz, as this is the range in which the fundamental speech information is expected to be found. A power estimate is computed from the filtered channels using the L-1 norm, which is simply the sum of the absolute values of the samples in a given segment with a length of w . The estimated power trajectory is then smoothed using a normalized integrator with a forgetting factor λ , producing $S_1[n]$ and $S_2[n]$. The segments containing the subject's speech are then detected when the following condition is met:

$$S_1[n] > k \cdot \text{mean}(S_2[n]), \quad (1)$$

where k represents the extent of the channels' power difference as a ratio of the $s_1[n]$ and $s_2[n]$ standard deviation, calculated as $k = \sqrt{\text{std}(s_1[n])/\text{std}(s_2[n])}$. The standard deviation and the mean in equation (1) are both computed per window. At the end of the window processing section, the timestamps are shifted to compensate for the delay introduced by the normalized integrator. The segments that are not considered to consist of the subject's speech are masked by zeros, and only the output from the first channel is subject to further processing.

A final check is made to ensure that the end of the window does not conflict with the subject's speech segment, which might produce disruptive artifacts. If the result is positive, the output from the current window is discarded, and the process is rerun with an adjusted, shorter L to avoid the conflict. The final output is a mono audio recording containing only the subject's speech.

After the trial testing was conducted, the default value for L was set to 30 seconds as the optimal value for reduced computational power, more precise results, and lower possibilities of conflicts at the end of the window. The segment length for the L-1 norm was set to $w = 128$ samples with an overlap of 50%, which resulted in fast processing with sufficient precision. The forgetting factor was set to $\lambda = 0.99$, producing the optimal smoothing effect for the given scenario.

SMARTSPEECH application for monitoring passive speech via regular smartphone calls

After either an incoming or outgoing call is concluded, the user will be presented with a screen containing an option to delete or save the call (Figure 3a). If the second option is selected, the recording will be filtered immediately using a distant-speaker filtering algorithm, and will be saved to the private application folder on the device. It will be possible to replay or delete each recording from the recent calls list (Figure 3b) in the application within the following 24 hours. After 24 hours, the recording will be compressed and sent to a secured server using SSH and REST API. The original call will then be deleted from the device. The aim is to record at least four five-minute recordings per month from each participant. A setting section on which the user can configure the application to suit their needs will be included (Figure 3c).

Active tasks

The application has a module that invites participants to record three active functional vocal tasks once every 14 days; the participants are instructed to capture the data in everyday environments with a low ambient noise level. In addition, the participants are asked to perform five motor tasks that are commonly used in PD research.[20-23] These tasks will be used to compare the sensitivity of the active and the passive data. Upon launching the active tasks protocol the users are presented with the guidelines and elapsed time since last run

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(Figure 3d). Before each task, the user will be presented with short, written instructions, including an option to replay an audio or video example of the given task. In the voice section, the recordings can be replayed after being recorded. If the user is dissatisfied with the result, an option to repeat the task is available.

The entire set of active tasks consists of prolonged phonation (Figure 4a), /pa/-/ta/-/ka/ syllable repetition (Figure 4b), reading a passage (Figure 4c, Figure 4d), a tapping game (Figure 5a, Figure 5b), alternated tapping (Figure 5c, Figure 5d), a writing task (Figure 5e, Figure 5f), resting hand tremor (Figure 6a, Figure 6b), and a gait with turnaround (Figure 6c, Figure 6d). Each task will be executed twice, and the motor tasks are to be executed first using the right hand and then the left hand.

Data validation and monitoring

Once the call and task data arrive on our server, they are saved on a database and the basic information becomes visible on the web interface to allow the data parameters to be reviewed for the quality, frequency, and recording duration. In the event that the participant does not make any calls or perform passive tasks within the required period, the participant will be notified via a push notification sent to their email address.

A gateway validation of the incoming data needs to be performed before the data are stored securely on the database to minimize the possibility of corrupted, irrelevant data entering the system, which would produce misleading results.

First, recordings that are shorter than two seconds will be discarded because they do not contain much information that is relevant and often contain errors. Subsequently, the speaker-detection algorithm will be employed to prevent data contamination when the user hands the smartphone to someone else for the call. The Kaldi toolkit[24] was used to construct the procedure. We employed a high-order Gaussian mixture model trained on the Czech section of the GLOBALPHONE database[25] using a set of mel-frequency cepstral coefficients. New subjects in the study will be enrolled in the model once a sufficient amount of data has been collected. I-vectors are extracted upon detection, and are used to compute the probabilistic linear discriminant analysis score. Using this score, a classifier decides whether the recording belongs to the correct subject. The data that will be assessed as not originating from the given subject will be moved to a specific folder and are not included in the data analysis.

Note that the purpose of the algorithm is also a form of quality control, as it excludes recordings that are of poor quality, such as utterances that are significantly affected by disruptive noise at a low signal-to-noise ratio (SNR). In these cases, the subject's speech is so corrupted that the computed features have little similarity to the subject's model. In addition, an SNR estimator based on minimum statistics[26] is employed in the last validation phase. Segments with an instantaneous SNR lower than the preselected threshold will be excluded from the analysis because the methods for extracting speech features cease to be reliable at certain very low SNR levels (such as pitch detection algorithms at SNR less than 6 dB[27]).

A voice activity detector is applied to the samples that have passed the validation procedure. The output indicates long calls during which the subject only speaks for a tiny

fraction of the overall duration. This allows to have a better grasp of the amount of speech data gathered via the system and to notify the subject if the amount of data is insufficient.

Extraction of acoustic speech and motor features

Hypokinetic dysarthria is due to dysfunction in the basal ganglia motor circuit, leading to impairments in the regulation of the initiation, the amplitude, and the velocity of movements.[12,28] Therefore, this project aims to collect, develop, and extract different acoustic features describing motor aspects of speech that have well-defined PD pathophysiology[17] and correspond to the perceptual description of hypokinetic dysarthria defined by Darley et al.[29]

The following features represent the most anticipated candidates for the final analysis:

- (i) Disruptions in phonation caused by dysfunctions in the vocal folds can be captured using the acoustic measure of Cepstral Peak Prominence, which correlates with the auditory perception of decreased voice quality/breathiness.[30]
- (ii) Articulation deficits, which are perceived as a decrease in intelligibility, can be described using metrics such as the Vowel Space Area.[31]
- (iii) Dysprosody is captured by the reduced amplitude of vocal cord movements, which correlates with the impression of monopitch, and can be assessed using the acoustic measure of pitch variability.[14]
- (iv) Decreased ability to maintain the speech motor sequence or to alternate quickly between responses can be reflected by the acoustic measures of the Net Speech Rate and Duration of Pause Intervals, which reflect the perceived auditory timing of speech and may describe deficits such as a slow articulation rate and a reduced ability to intermit and initiate speech.[32]
- (v) Linguistic deficits, including limited vocabulary and a decrease in its range, which might indicate potential underlying mild cognitive impairment, can be assessed using a set of lexical features such as Content Density.[33]
- (vi) To assess the overall degree of speech impairment, powerful deep neural networks could be employed, including spectrogram analysis to evaluate specific features when traditional methods might be insufficient. Insight into the detailed method behavior would be supportive, and would reveal critical physiological details.[34]

Several methods for the automatic analysis of key dimensions of speech in PD patients have already been developed.[32,35,36] However, all the methods need to undergo experimental and theoretical testing for noise robustness and reliability to validate their usefulness.[27] A set of features commonly used in the existing literature will be extracted from the active motor tasks;[20-22] the set includes tapping velocity, intratap duration variability, reaction time, tremor acceleration skewness, and velocity.

Endpoints

The primary endpoint will be represented by composite dysarthria index, reflecting the severity of speech impairment, which will be based on a combination of several distinct acoustic speech features associated with hypokinetic dysarthria in PD. Secondary endpoints

will be represented by individual acoustic speech features associated with hypokinetic dysarthria in PD and linguistic features associated with potential cognitive decline.

Study design and participants

During this project, each participant will be given a full explanation of the project’s purpose and aims, will be informed about the procedure, and will be given the opportunity to ask questions before deciding whether to sign the informed consent form. In total, we plan to recruit 25 iRBD subjects, 25 early stage PD patients, and 25 healthy controls as part of this longitudinal study. As we expect a drop-out rate of about 20% per year[17], up to 50 iRBD patients and the same number of early stage PD patients will be recruited at the baseline. These data might be used for a better-sampled cross-sectional study. All the iRBD subjects will fulfill the criteria listed in the International Classification of Sleep Disorders, third edition diagnostic criteria,[37] including confirmation of REM sleep without atonia via polysomnography. The inclusion criteria for iRBD will be

- (i) onset of iRBD after 50 years of age,
- (ii) no history of major neurological disease (such as epilepsy or strokes) or other significant diseases that could affect study participation or voice analysis (for example, active cancer, drug abuse, or diseased vocal cords),
- (iii) no significant cognitive decline or severe depression, and
- (iv) no history of therapy with antiparkinsonian medication.

All the PD patients will meet the Movement Disorders Society’s clinical diagnostic criteria for PD,[38] and will be investigated during the on-medication state. The inclusion criteria for PD will be:

- (i) onset of PD after 50 years of age,
- (ii) Hoehn and Yahr stage 1–2 in the on-medication state,
- (iii) disease duration more than five years after diagnosis,
- (iv) no history of a major neurological disease other than PD (such as epilepsy or strokes) or other significant diseases that could affect study participation or voice analysis (for example, active cancer, drug abuse, or diseased vocal cords),
- (v) no significant cognitive decline or severe depression, and
- (vi) no involvement in any speech therapy during the duration of the project.

The inclusion criterion for controls will be that the participants have no history of neurological or communication disorders, and no history of parasomnias or other sleep disorders. PD and healthy control subjects will be age- and gender-matched to the iRBD group.

Each participant will be required to perform passive speech recordings and active tasks using the provided smartphone for two years. In addition, each of the subjects will undergo three examinations at the clinic: at the baseline, after one year, and after two years (at the end of the project). The clinical examinations will consist of taking a structured personal history, quantitative testing of motor and non-motor symptoms of PD based on the Movement Disorder Society-Unified Parkinson's Disease Rating Scale, [16] cognitive testing using the Montreal Cognitive Assessment,[39] autonomic testing using the Scales for Outcomes in Parkinson's Disease-Autonomic Dysfunction,[40] the evaluation of depressive

symptoms using the Beck Depression Inventory II,[41] and speech examinations according to the dysarthria research guidelines for acoustic analyses.[23] Recruitment of the participants will begin in October 2021.

Patient and public involvement statement

Most of the patients have been involved in previous studies and are familiar with the team, the research topic, and the methods. The SMARTSPEECH protocol was designed with the aid of a short questionnaire that was completed by a selected representative group of 33 patients with iRBD or PD, which provided insight into their requirements (for example, a dual-SIM phone and monetary compensation for mobile phone tariffs), doubts and motivations (such as understanding how to operate the phone and data security), and the expected frequency of phone usage and calls.

Sample size estimation

For the primary endpoint of the project, an ad-hoc power analysis for a given large effect size (Cohen's d of 0.8), with the Type I error probability (α) set at 0.05 and power of 80%, based on a three-group analysis of variance with one covariate (group), determined a minimum sample size of 66 subjects with at least 22 subjects in each subgroup (that is, 22 iRBD patients, 22 PD patients, and 22 controls).

Statistical analysis

A one-way analysis of variance with a group (PD versus iRBD versus controls) as a between-subject factor will be used to calculate the differences for each parameter of interest. The Pearson correlation coefficient will be applied to search for correlations among the variables. The minimum level of significance will be set at $p < 0.05$ with an appropriate Bonferroni adjustment. In addition, a binary logistic regression followed by a leave-one-subject-out cross-validation will be used to assess the sensitivity/specificity of the proposed features to differentiate the iRBD subjects from controls, PD patients from controls, and PD patients from iRBD subjects.

ETHICS AND DISSEMINATION

The study has been approved by the Ethics Committee of the General University Hospital in Prague, Czech Republic (no. 30/19 Grant AZV VES 2020 VFN), and will be performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All the participants will provide written, informed consent prior to their inclusion.

The following steps will be taken to mitigate potential ethical concerns: The researchers performing the analyses will have no known relationship with the research subjects, and the recordings will be de-identified to the extent possible. The analyses of audio recordings will be automated to avoid the recordings being listened to by a human;

furthermore, the audio recordings will be encrypted, will only be available to the authorized researchers, and will be deleted at the request of any participant without the need for justification. All the steps will be conducted according to the directive on personal data protection legislation in the Czech Republic and the approval of the Ethics Committee. Any amendments will be agreed on by the research steering committee and submitted for ethics committee approval prior to implementation.

This project may provide a natural digital biomarker of disease progression based on longitudinal data acquired without any cost or time burden on the patients and investigators. Observing disease progression over a short period using well-defined and disease-specific speech biomarkers may significantly aid in recruiting appropriate cases into large clinical trials for disease-modifying drugs and allows monitoring possible disease-modifying effects of treatment in prodromal PD [42,43]. In the future, speech biomarkers may also bolster early presymptomatic diagnosis and enable rapid access to neuroprotective therapy once available. Results will be presented at national and international conferences, published in peer-reviewed journals, and disseminated to the researchers and Parkinson’s community.

AUTHORS’ CONTRIBUTION

JR conceived of the study and the study design. TK and VI drafted and revised the manuscript. TK and VI designed the application. TK, VI and JR approved the final manuscript.

FUNDING

This study was supported by the Czech Ministry of Health (grant no. NU20-08-00445).

REFERENCES

- 1 Poewe W, Seppi K, Tanner C, et al. Parkinson disease. *Nat Rev Dis Primers* 2017;3. doi:10.1038/nrdp.2017.13
- 2 de Rijk M, Launer L, Berger K, et al. Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group: A collaborative study of population-based cohorts. *Neurologic Diseases in the Elderly Research Group. Neurol* 2000;54:S21-3.http://intl.neurology.org/cgi/content/abstract/54/11_suppl_5/S21
- 3 Rodriguez-Oroz M, Jahanshahi M, Krack P, et al. Initial clinical manifestations of Parkinson's disease: features and pathophysiological mechanisms. *Lancet Neurol* 2009;8:1128-1139. doi:10.1016/S1474-4422(09)70293-5
- 4 Postuma R, Berg D. Advances in markers of prodromal Parkinson disease. *Nat Rev Neurol* 2016;12:622-634. doi:10.1038/nrneurol.2016.152
- 5 Postuma R, Iranzo A, Hu M, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. *Brain* 2019;142:744-759. doi:10.1093/brain/awz030

- 6 Boeve B. REM sleep behavior disorder. *Ann N Y Acad Sci* 2010;1184:15-54. doi:10.1111/j.1749-6632.2009.05115.x
- 7 Miglis M, Adler C, Antelmi E, et al. Biomarkers of conversion to α -synucleinopathy in isolated rapid-eye-movement sleep behaviour disorder. *Lancet Neurol* 2021;20:671-684. doi:10.1016/S1474-4422(21)00176-9
- 8 Iranzo A, Tolosa E, Gelpi E, et al. Neurodegenerative disease status and post-mortem pathology in idiopathic rapid-eye-movement sleep behaviour disorder: an observational cohort study. *Lancet Neurol* 2013;12:443-453. doi:10.1016/S1474-4422(13)70056-5
- 9 Schenck C, Boeve B, Mahowald M. Delayed emergence of a parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder: a 16-year update on a previously reported series. *Sleep Med* 2013;14:744-748. doi:10.1016/j.sleep.2012.10.009
- 10 Schenck C, Montplaisir J, Frauscher B, et al. Rapid eye movement sleep behavior disorder: devising controlled active treatment studies for symptomatic and neuroprotective therapy—a consensus statement from the International Rapid Eye Movement Sleep Behavior Disorder Study Group. *Sleep Med* 2013;14:795-806. doi:10.1016/j.sleep.2013.02.016
- 11 Postuma R, Aarsland D, Barone P, et al. Identifying prodromal Parkinson's disease: Pre-Motor disorders in Parkinson's disease. *Mov Disord* 2012;27:617-626. doi:10.1002/mds.24996
- 12 Duffy J. *Motor speech disorders: substrates, differential diagnosis, and management*. Fourth. Maryland Heights: : Elsevier 2019.
- 13 Ho A, Iansek R, Marigliani C, et al. Speech Impairment in a Large Sample of Patients with Parkinson's Disease. *Behav Neurol* 1999;11:131-137. doi:10.1155/1999/327643
- 14 Rusz J, Hlavnička J, Novotný M, et al. Speech Biomarkers in Rapid Eye Movement Sleep Behavior Disorder and Parkinson Disease. *Ann Neurol* 2021;90:62-75. doi:10.1002/ana.26085
- 15 Rusz J, Hlavnicka J, Tykalova T, et al. Smartphone Allows Capture of Speech Abnormalities Associated With High Risk of Developing Parkinson's Disease. *IEEE Trans Neural Syst Rehabil Eng* 2018;26:1495-1507. doi:10.1109/TNSRE.2018.2851787
- 16 Goetz C, Tilley B, Shaftman S, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord* 2008;23:2129-2170. doi:10.1002/mds.22340
- 17 Arora S, Venkataraman V, Zhan A, et al. Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study. *Parkinsonism Relat Disord* 2015;21:650-653. doi:10.1016/j.parkreldis.2015.02.026
- 18 Espay A, Bonato P, Nahab F, et al. Technology in Parkinson's disease: Challenges and opportunities. *Mov Disord* 2016;31:1272-1282. doi:10.1002/mds.26642
- 19 Artusi C, Mishra M, Latimer P, et al. Integration of technology-based outcome measures in clinical trials of Parkinson and other neurodegenerative diseases. *Parkinsonism Relat Disord* 2018;46:S53-S56. doi:10.1016/j.parkreldis.2017.07.022
- 20 Zhan A, Mohan S, Tarolli C, et al. Using Smartphones and Machine Learning to Quantify Parkinson Disease Severity. *JAMA Neurol* 2018;75. doi:10.1001/jamaneurol.2018.0809

- 21 Lipsmeier F, Taylor K, Kilchenmann T, et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. *Mov Disord* 2018;33:1287-1297. doi:10.1002/mds.27376
- 22 Arora S, Baig F, Lo C, et al. Smartphone motor testing to distinguish idiopathic REM sleep behavior disorder, controls, and PD. *Neurol* 2018;91:e1528-e1538. doi:10.1212/WNL.00000000000006366
- 23 Rusz J, Tykalova T, Ramig L, et al. Guidelines for Speech Recording and Acoustic Analyses in Dysarthrias of Movement Disorders. *Mov Disord* 2021;36:803-814. doi:10.1002/mds.28465
- 24 Povey D, Ghoshal A, Boulianne G, et al. The Kaldi Speech Recognition Toolkit. *Proc IEEE Workshop Autom Speech Recognit Underst* 2011.
- 25 Schultz T, Vu N, Schlippe T. GlobalPhone: A multilingual text & speech database in 20 languages. In: *2013 Proc IEEE Int Conf Acoust Speech Signal Process* 2013. 8126-8130. doi:10.1109/ICASSP.2013.6639248
- 26 Martin R. Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Trans Speech Audio Process*;9:504-512. doi:10.1109/89.928915
- 27 Illner V, Sovka P, Rusz J. Validation of freely-available pitch detection algorithms across various noise levels in assessing speech captured by smartphone in Parkinson's disease. *Biomed Signal Process Control* 2020;58. doi:10.1016/j.bspc.2019.101831
- 28 Ho A, Iansek R, Marigliani C, et al. Speech Impairment in a Large Sample of Patients with Parkinson's Disease. *Behav Neurol* 1999;11:131-137. doi:10.1155/1999/327643
- 29 Darley F, Aronson A, Brown J. Differential Diagnostic Patterns of Dysarthria. *J Speech Lang Hear Res* 1969;12:246-269. doi:10.1044/jshr.1202.246
- 30 Jannetts S, Lowit A. Cepstral Analysis of Hypokinetic and Ataxic Voices: Correlations With Perceptual and Other Acoustic Measures. *J Voice* 2014;28:673-680. doi:10.1016/j.jvoice.2014.01.013
- 31 Rusz J, Cmejla R, Tykalova T, et al. Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task. *J Acoust Soc Am* 2013;134:2171-2181. doi:10.1121/1.4816541
- 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. doi:10.1038/s41598-017-00047-5
- 33 Beltrami D, Gagliardi G, Rossini Favretti R, et al. Speech Analysis by Natural Language Processing Techniques: A Possible Tool for Very Early Detection of Cognitive Decline?. *Front Aging Neurosci* 2018;10. doi:10.3389/fnagi.2018.00369
- 34 Laganas C, Iakovakis D, Hadjidimitriou S, et al. Parkinson's Disease Detection Based on Running Speech Data From Phone Calls. *IEEE Trans Biomed Eng* 2021;:1-1. doi:10.1109/TBME.2021.3116935
- 35 Rusz J, Cmejla R, Ruzickova H, et al. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *J Acoust Soc Am* 2011;129:350-367. doi:10.1121/1.3514381
- 36 Novotny M, Rusz J, Cmejla R, et al. Automatic Evaluation of Articulatory Disorders in Parkinson's Disease. *IEEE/ACM Trans Audio Speech Lang Process* 2014;22:1366-1378. doi:10.1109/TASLP.2014.2329734

- 37 American Academy of Sleep Medicine *International Classification of Sleep Disorders, Third Edition: Diagnostic and Coding Manual*. Chicago, Illinois: : American Academy of Sleep Medicine 2014.
- 38 Postuma R, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2015;30:1591-1601. doi:10.1002/mds.26424
- 39 Nasreddine Z, Phillips N, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc* 2005;53:695-699. doi:10.1111/j.1532-5415.2005.53221.x
- 40 Visser M, Marinus J, Stiggelbout A, et al. Assessment of autonomic dysfunction in Parkinson's disease: The SCOPA-AUT. *Mov Disord* 2004;19:1306-1312. doi:10.1002/mds.20153
- 41 Beck A, Steer R, Brown G. Manual for the Beck Depression Inventory-II. *San Antonio, TX: Psychological Corporation* 1996. doi:10.1007/978-1-4419-1005-9_441
- 42 Paolini Paoletti F, Gaetani L, Parnetti L. The Challenge of Disease-Modifying Therapies in Parkinson's Disease: Role of CSF Biomarkers. *Biomolecules*. 2020 Feb 19;10(2):335. doi: 10.3390/biom10020335
- 43 Espay AJ, Schwarzschild MA, Tanner CM, Fernandez HH, Simon DK, Leverenz JB, Merola A, Chen-Plotkin A, Brundin P, Kauffman MA, Erro R, Kiebertz K, Woo D, Macklin EA, Standaert DG, Lang AE. Biomarker-driven phenotyping in Parkinson's disease: A translational missing link in disease-modifying clinical trials. *Mov Disord*. 2017 Mar;32(3):319-324. doi: 10.1002/mds.26913

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FIGURE CAPTIONS

Figure 1. Schematics of the HONOR 9X smartphone audio inputs and outputs.

Figure 2. An illustrative schema of the distant-speaker filtering algorithm.

Figure 3. (a) Option screen to delete or save the recently completed call for speech analysis. (b) List of calls from the previous 24 hours with details and the options of replaying and deletion. (c) The setting screen of the app. (d) The active task introduction screen, which shows a guide to performing the active tasks and the time elapsed since the last active task run.

Figure 4. (a) Prolonged phonation task instruction screen (task duration ~15 seconds). (b) Sequential motion rates instruction screen (task duration 7 seconds). (c) Reading passage instruction screen. (d) Reading passage sample text that is chosen randomly from six samples and contains approximately 80 words (task duration ~40 seconds).

Figure 5. (a) A tapping game introduction screen. (b) A tapping game task. The delay between spawning the circle and a successful tap is measured. The action takes place 20 times (task duration ~20 seconds). (c) An alternated tapping introduction screen. (d) An alternated tapping task. The timestamps of the individual taps are measured (task duration 15 seconds). (e) Writing task instruction screen. (f) A writing task. The time spent rewriting each word is measured. The patient is shown five words in each run. Eight different datasets that are assembled in the same manner are chosen randomly in each run. Examples include the Czech words "SESTRICKA", "NEJLEPSI", "LECENI", "ZABAVA", and "PREDSTAVA". The task can be skipped after 80 seconds have elapsed (task duration ~40 seconds).

Figure 6. (a) Rest tremor instruction screen. (b) A rest tremor task (task duration 15 seconds). The patient's hand tremor is measured using an accelerometer in all three axes. The z-axis is live plotted for user feedback. (c) A gait task instruction screen. (d) A gait task (task duration ~30 seconds). The z-axis rotation is measured and plotted on the screen for user feedback via an internal gyroscope.

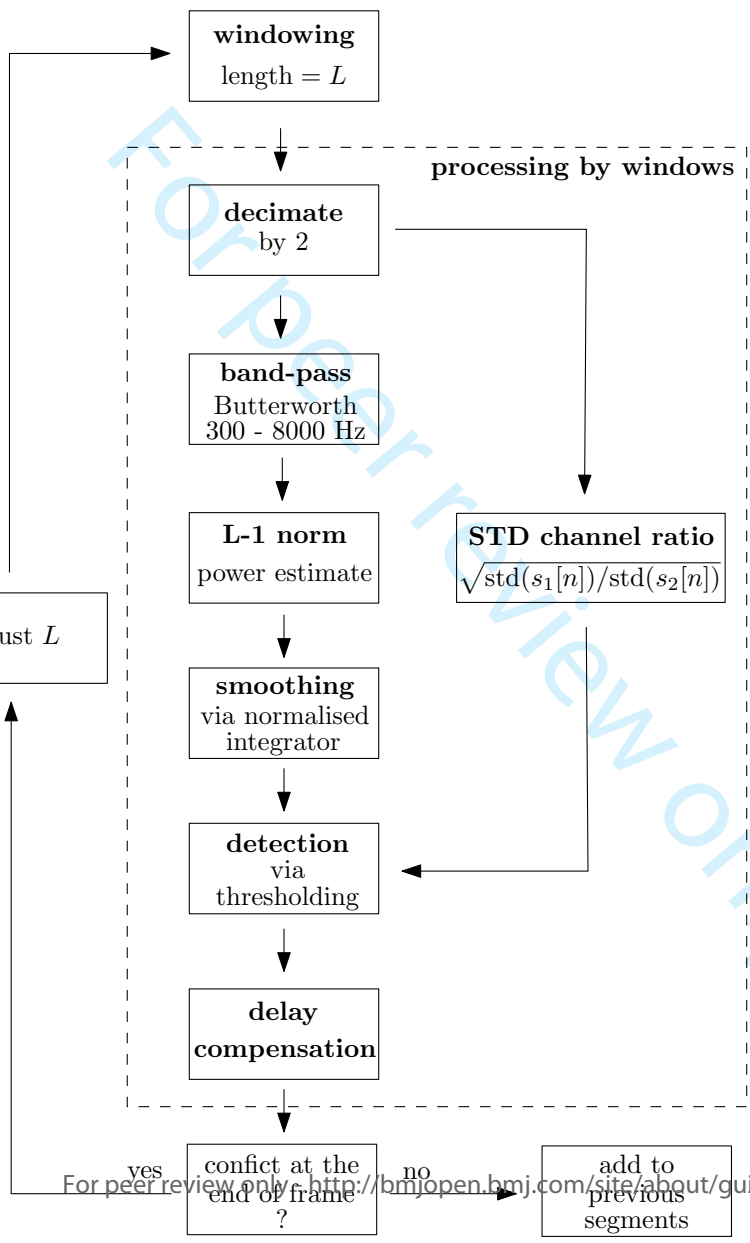


Figure 1

66x80mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

$s_1[n]$ - 1st channel $s_2[n]$ - 2nd channel



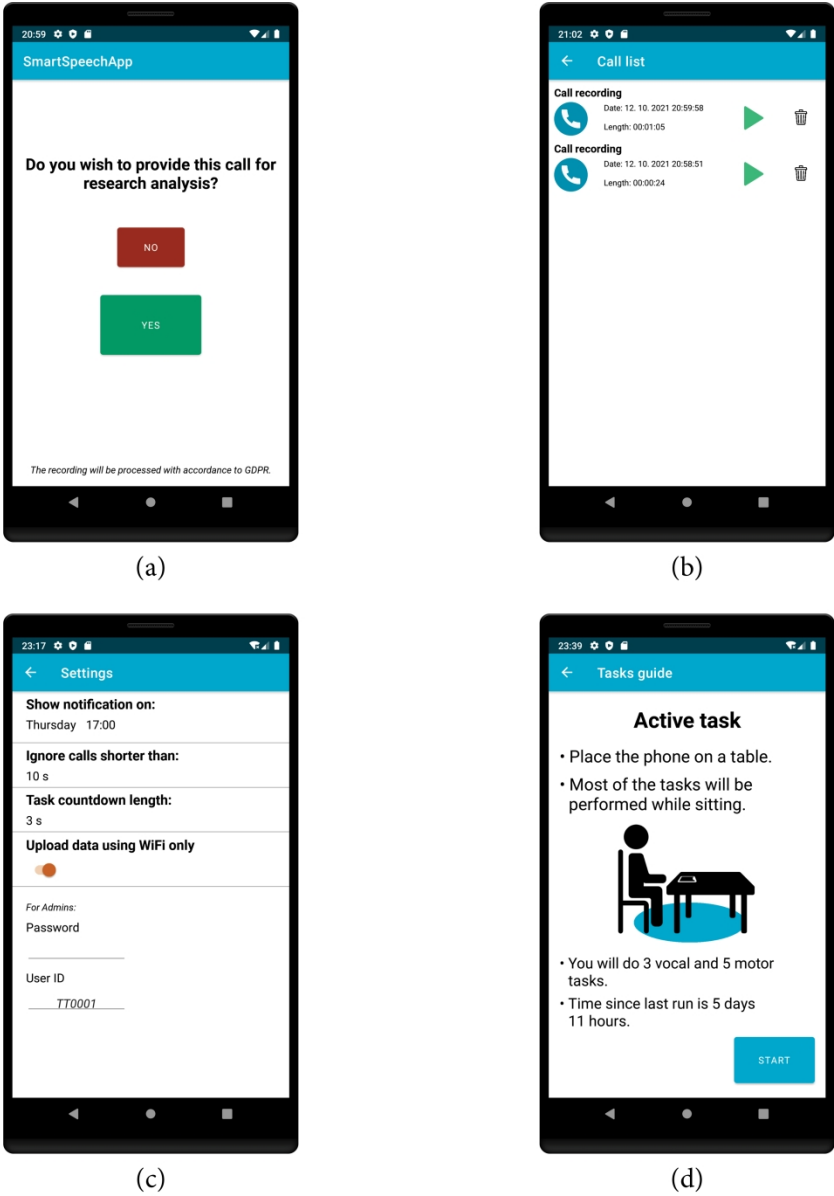


Figure 3
143x183mm (600 x 600 DPI)

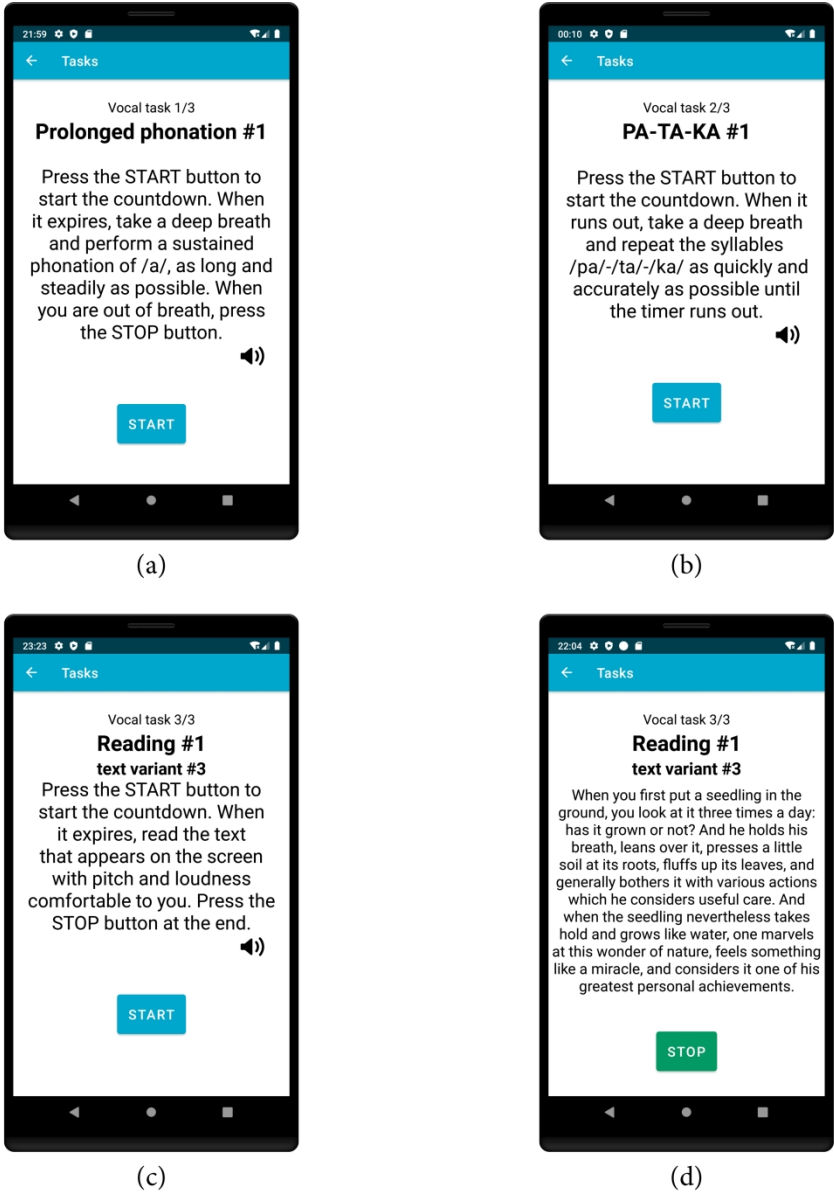


Figure 4

143x183mm (600 x 600 DPI)

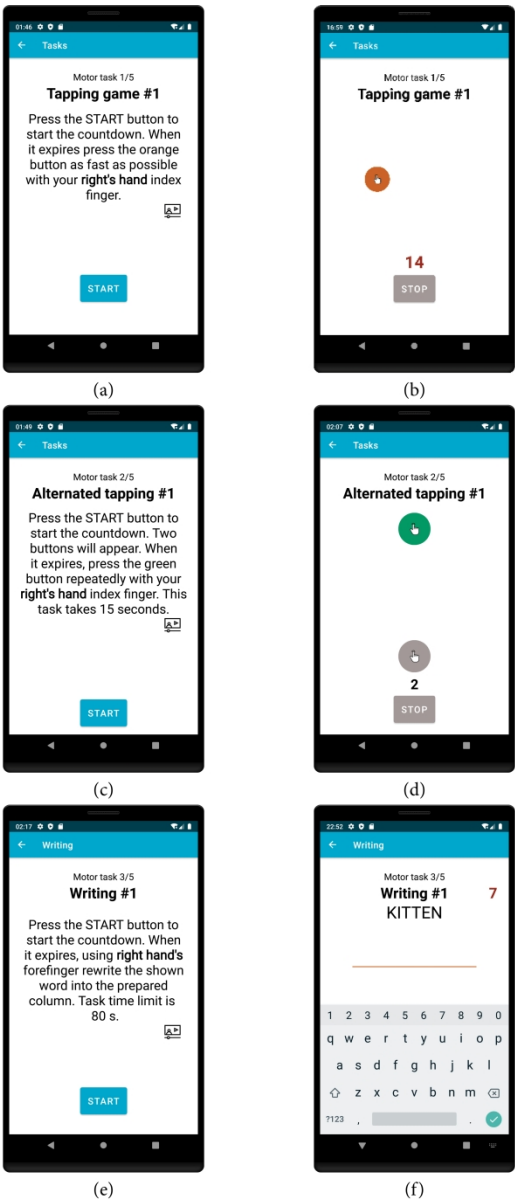


Figure 5

131x269mm (600 x 600 DPI)

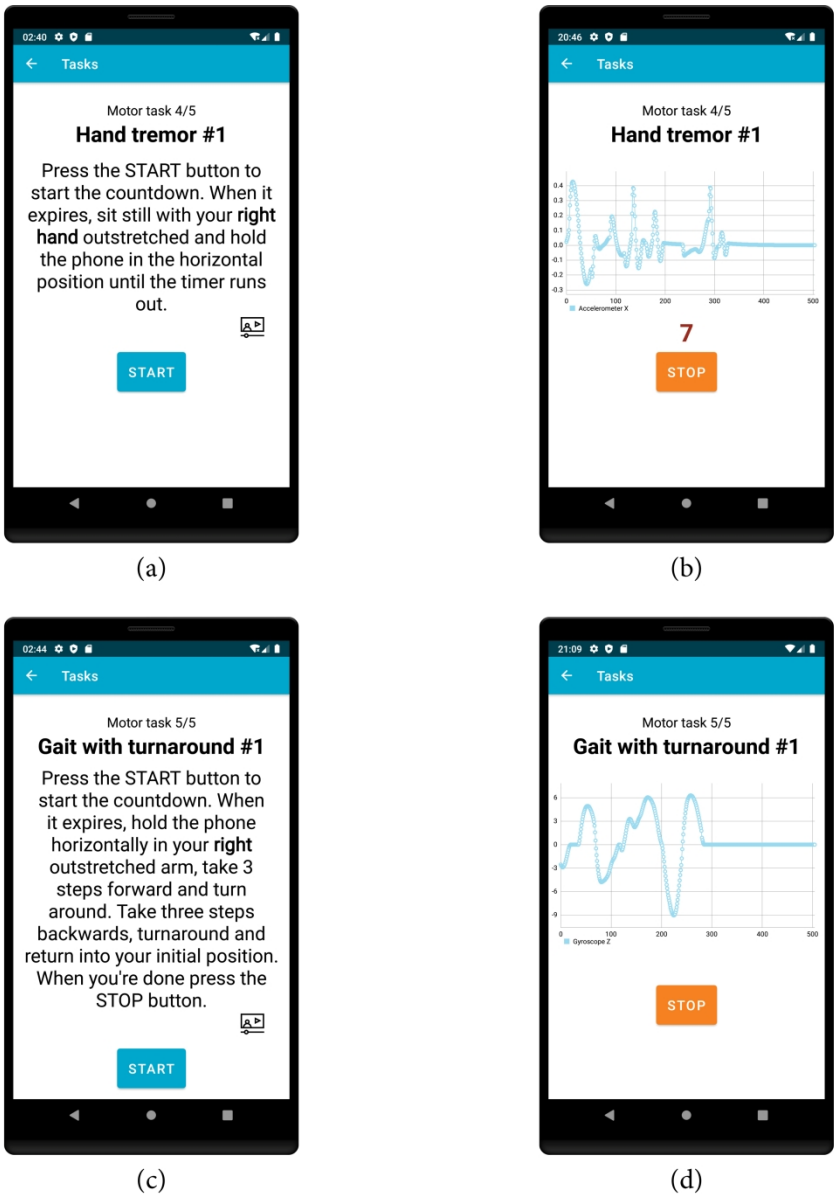


Figure 6
143x183mm (600 x 600 DPI)