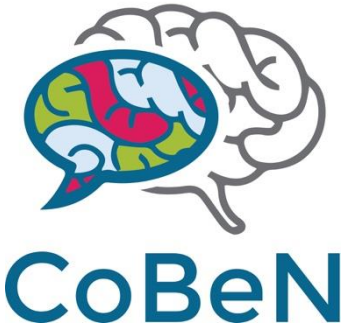

Novel Network-Based Approaches for Studying Cognitive Dysfunction in Behavioural Neurology

H2020-MSCA-RISE-2016-734718

D2.4 Novel parametrization software for acoustic analysis

Work Package:	WP2
Task:	-
Deliverable due date:	31.10.2019
Responsible partner:	MU
Deliverable number:	D13
Deliverable type:	R/O
Dissemination level:	PU

1 Introduction

Since hundreds of parameters may be analysed by using acoustic analysis of speech data and it is very time-consuming we will develop a novel parametrization software to speed the whole process and we will identify and choose only those speech parameters that reflect both dysarthria and apraxia of speech in the most efficient way.

In the frame of this deliverable, we aimed to design and implement a novel speech/voice parameterization system that will enable to perform a complex quantification of several speech disorders manifested mainly in the field of phonation, articulation and prosody. The software is freely available upon request (http://coben.ceitec.cz/software) and will help the community to perform objective acoustic analysis of disorders such as hypokinetic dysarthria (HD) or apraxia of speech (AOS).

2 Methods

Based on a comprehensive literature review we firstly identified all possible speech/voice disorders that could be present in HD or AOS. In the consequent step, we identified acoustic features (measures) that could quantify these disorders. These two steps were performed in the frame of deliverable D2.1. Finally, all the features were coded in Matlab (also depending on some third-party libraries) and the functions were grouped into one toolbox (deliverable D2.10).

3 Results

CoBeN speech and voice parameterization toolbox is a set of Matlab functions that enables to perform complex acoustic analysis of speech disorders that are mainly manifested in different types of dysarthria or in apraxia of speech (AOS). A list of these features and related Matlab functions can be found in Table 1. This toolbox will help neurologists or clinical pathologists to get a quick and comprehensive insight about a pathology present in a speech/voice recording (the toolbox is tuned for 16 kHz sampling frequency). It also contains 2 demo files demonstrating parameterization of connected speech and sustained phonation.

The toolbox is available upon request (irena.rektorova@ceitec.muni.cz) under **MIT license**. It has been developed for Windows and is partially dependent on Praat libraries (praatcon.exe, http://www.fon.hum.uva.nl/praat/download_win.html).

Table 1: HD/AOS dimensions, specific disorders, and Matlab functions that are used to quantify them

HD/AOS dimension and specific disorder	**Speech tasks**	**Acoustic feature**	**Feature definition** *Name of the related Matlab function*
Phonation			
Airflow insufficiency	Sustained phonation	MPT	Maximum phonation time, aerodynamic efficiency of the vocal tract measured as the maximum duration of the prolonged vowel. *Function (file)*: derive_speech_rate.m *Third-party dependency*: Praat

Irregular pitch fluctuations	Sustained phonation	relF0SD	Standard deviation of fundamental frequency relative to its mean, variation in frequency of vocal fold vibration. *Function (file)*: praat_F0.m *Third-party dependency*: Praat
Microperturbations in frequency	Sustained phonation	Jitter (PPQ)	Frequency perturbation, extent of variation of the voice range. Jitter is defined as the variability of the F0 of speech from one cycle to the next. *Function (file)*: praat_info.m *Third-party dependency*: Praat
Microperturbations in amplitude	Sustained phonation	Shimmer (APQ)	Amplitude perturbation, representing rough speech. Shimmer is defined as the sequence of maximum extent of the signal amplitude within each vocal cycle. *Function (file)*: praat_info.m *Third-party dependency*: Praat
Increased noise	Sustained phonation	HNR	Harmonics-to-noise ratio, the amount of noise in the speech signal, mainly due to incomplete vocal fold closure. HNR is defined as the amplitude of noise relative to tonal components in speech. *Function (file)*: praat_harmonicity.m *Third-party dependency*: Praat
Aperiodicity	Sustained phonation	DUV	Degree of unvoiced segments, the fraction of pitch frames marked as unvoiced. *Function (file)*: praat_info.m *Third-party dependency*: Praat
Tremor of jaw	Sustained phonation	relF1SD, relF2SD	Standard deviation of first (F1) and second (F2) formant relative to its mean. Formants are related to resonances of the oro-naso-pharyngeal tract and are modified by position of tongue and jaw. *Function (file)*: praat_formants.m *Third-party dependency*: Praat
Articulation			
Decreased tongue movement (imprecise vowels)	Monologue, reading	VAI	Vowel articulation index, based on formant centralization, defined as VAI = (F1a + F2i)/(F1i + F1u + F2a + F2u). *Function (file)*: vai.m *Third-party dependency*: no
Rigidity of tongue and jaw	Monologue, reading	relF1SD, relF2SD	Standard deviation of first (F1) and second (F2) formant relative to its mean. *Function (file)*: praat_formants.m *Third-party dependency*: Praat
Slow alternating motion rate	Diadochokinetic task	PR	Pace rate, representing the number of syllable vocalizations per second. Considering first 30 syllables. *Function (file)*: ddk_features.m *Third-party dependency*: no
Instability of diadochokinetic pace	Diadochokinetic task	$COV_{4–10}$	Coefficient of variation, defined as $100 \times SD_{C4–10}/mean_{C1–3}$, where $SD_{C4–10}$ stands for standard deviation of duration of 4^{th}–10^{th} DDK cycles (pa-ta-ka), and $mean_{C1–3}$ means average duration of the first three DDK cycle. *Function (file)*: ddk_features.m *Third-party dependency*: no

Instability of diadochokinetic pace	Diadochokinetic task	RI	Rhythm instability, defined as sum of absolute deviations from a regression line modelling each DDK cycle duration, weighted to the total DDK performance time. *Function (file)*: ddk_features.m *Third-party dependency*: no
Acceleration of diadochokinetic pace	Diadochokinetic task	PA	Pace acceleration, defined as PA = 100 × (avCycDur$_{4–6}$ – avCycDur$_{7–9}$)/avCycDur$_{1–3}$, where avCycDur$_{X–Y}$ is average duration of cycles *X*–*Y*. *Function (file)*: ddk_features.m *Third-party dependency*: no
Acceleration of diadochokinetic pace	Diadochokinetic task	RA	Rhythm acceleration, defined as gradient of regression line modelling DDK cycle durations (positive values mean acceleration). *Function (file)*: ddk_features.m *Third-party dependency*: no
Prosody			
Monoloudness	Monologue, reading	relSEOSD	Speech loudness variation, defined as a standard deviation of intensity contour relative to its mean after removing silences exceeding 50 ms. *Function (file)*: energy.m *Third-party dependency*: Praat
Monopitch	Monologue, reading	relF0SD	Pitch variation, defined as a standard deviation of F0 contour relative to its mean. *Function (file)*: praat_F0.m *Third-party dependency*: Praat
Inappropriate silences	Reading	SPIR	Number of pauses relative to total speech time after removing periods of silence lasting less than 50 ms. *Function (file)*: derive_speech_rate.m *Third-party dependency*: Praat
Higher proportion of silence time	Reading	PPR	Percentual pause ratio, defined as total duration of silences (longer than 50 ms)/total duration of speech. *Function (file)*: derive_speech_rate.m *Third-party dependency*: Praat
Longer duration of silences	Reading	DurMED	Median duration of silences longer than 50 ms. *Function (file)*: derive_speech_rate.m *Third-party dependency*: Praat
Higher variability of silence duration	Reading	DurMAD	Median absolute deviation of silence duration (silences longer than 50 ms). *Function (file)*: derive_speech_rate.m *Third-party dependency*: Praat
Unnatural speech rate	Reading	AR	Number of speech sounds produced per second after pauses longer than 50 ms were removed. *Function (file)*: derive_speech_rate.m *Third-party dependency*: Praat

4 Conclusion

In the frame of this deliverable, we introduced a new software of speech/voice disorders parameterization. The software is novel especially in a way that it is focused on a narrow field of dysarthrias and on apraxia of speech, enabling to perform complex quantification in the field of phonation, articulation and prosody. The features were selected based on a

comprehensive review, are clinically interpretable and can support neurologists or clinical pathologists in e.g. differential analysis, early diagnosis, treatment effect monitoring, etc.