Acoustic feature	Description
Source features	*Features reflecting airflow from the lungs through the glottis (i.e. glottal features) or vocal fold vibrations (i.e., voice quality features)[27], which is the sound source later filtered by the vocal tract following the source-filter theory of speech production[57].*
Jitter [%]	Deviations in individual consecutive f0 period lengths, which indicates irregular closure and asymmetric vocal-fold vibrations (for various ways of measuring, see [58]).
Shimmer [%]	Difference of the peak amplitudes of consecutive f0 periods, which indicates irregularities in voice intensity (for various ways of measuring, see [58]).
Tremor [Hz]	Frequency of the most intense low-frequency fundamental frequency-modulating component in a specified analysis range[59].
Harmonics-to-noise ratio (HNR) [dB]	Ratio between f0 and noise components, which indirectly correlates with perceived aspiration. This may be due to reducing laryngeal muscle tension resulting in a more open, turbulent glottis[47].
Frequency disturbance ratio (FDR) [%]	Relative mean value of the frequency disturbance from 5 to 5 periods (five points average)[60].
Amplitude Disturbance ratio (ADR) [%]	Relative mean amplitude value over a set of windows[60].
Quasi-open quotient (QOQ)	Ratio of the vocal folds' opening time. Functional dysphonias often reduce QOQ range. Speaking loudly requires more effort with a low QOQ and sounds more stalled[61].
Normalized amplitude quotient (NAQ)	Ratio between peak-to-peak pulse amplitude and the negative peak of the differentiated flow glottogram and normalized with respect to the period time. It can be an estimate of glottal adduction[62].
Peak slope	Slope of the regression line that is fit to log10 of the maxima of each frame[63].
Filter features	*The resonant properties of the vocal and nasal tracts filter the sound source from the vocal folds: the filter attenuates certain frequencies and strengthens others by the shape of the vocal and nasal tracts.*
F1 mean [Hz]	First peak in the spectrum (especially of voiced utterances such as vowels) that results from a resonance of the human vocal tract.
F2 mean [Hz]	Second peak in the spectrum (especially of voiced utterances such as vowels) that results from a resonance of the human vocal tract.
F1 variability [Hz]	Measures of dispersion of F1 (variance, standard deviation).
F2 variability [Hz]	Measures of dispersion of F2 (variance, standard deviation).
F1 range [Hz]	Difference between the lowest and highest F1 values .
Vowel space	F1 and F2 2D space for the vowels /a/, /i/, /u/[64].
Linear predictive coding (LPC) coefficients	Coefficients that best predict the values of the next time point of the audio signal using the values from the previous n time points, which is used to reconstruct filter properties.

Spectral features	*Features characterizing the spectrum of speech, which is the frequency distribution of the speech signal at a specific time*[27].
Mel-frequency cepstral coefficients (MFCCs)	The coefficients derived by computing a spectrum of the log-magnitude Mel-spectrum of the audio segment. The lower coefficients represent the vocal tract filter and the higher coefficients represent periodic vocal fold sources.
Prosodic features	*Changes over longer segments of time, which is perceived in the rhythm, stress, and intonation of speech*[27].
f0 mean [Hz]	Fundamental frequency: lowest frequency of the speech signal, perceived as pitch (mean, median).
f0 variability [Hz]	Measures of dispersion of f0 (variance, standard deviation).
f0 range [Hz]	Difference between the lowest and highest f0 values.
Intensity [dB]	Defined as the acoustic intensity (i.e., power carried by sound per unit area in a direction perpendicular to that area) in decibels relative to a reference value, perceived as loudness.
Intensity variability [dB]	Measures of dispersion of intensity (variance, standard deviation).
Energy velocity	Measured as the mean-squared central difference across frames and may correlate with motor coordination[65].
Maximum phonation time [s]	The mean of three attempts of the following measure is taken: the maximum time during which phonation of a vowel (usually /a/) is sustained as long as possible with an upright position, deep breath, and a comfortable pitch and loudness[66].
Speech rate	Number of speech utterances per second over the duration of the speech sample (including pauses).
Articulation rate	Number of speech units per second over the duration of the speech sample (excluding pauses).
Time talking [s]	Sum of the duration of all speech segments.
Utterance duration mean [s]	Mean duration of utterance length.
Pause duration mean [s]	Mean duration of pause length.
Pause variability [s]	Measures of dispersion of pause duration (variance, standard deviation).
Pause rate [s]	Total length of pauses divided by the total length of speech (including pauses).
Pauses total [s]	Total duration of pauses.