A Protocol for Acoustic Voice Analysis in Young Adults

Authors: Paris Binos, CIRCLE Lab, Cyprus University of Technology
Elina Nirgianaki, Keele University
Loukia Taxitari, Neapolis University Pafos
Abstract ID: 402
Event: Linguistics 2026 Paris
Subject: 4. Phonetics
Presenter Name: Paris Binos

Objectives

Acoustic voice analysis is widely used to quantify phonatory function in phonetic research and clinical voice assessment. However, the validity and interpretability of acoustic measures depend on methodological decisions involving participant selection, recording conditions, elicitation, signal quality, segment definition, and analysis procedures. Variation at these stages may introduce measurement error and reduce cross-study comparability. This study aimed to develop a rigorously specified protocol for acoustic voice analysis in vocally healthy young adults aged 18–30 years, emphasizing methodological transparency, repeatability, and applicability to research and clinical contexts.

Methods

Voice recordings were obtained in a sound-controlled audiology facility, with background noise maintained below 35 dB. High-quality digital recordings were acquired using a Sony PCM-D50 recorder at a sampling rate of 44.1 kHz. Participants produced a sustained /a/ at habitual pitch and comfortable loudness in a standardized recording position. Acoustic analysis was performed in Praat.

A central component of the protocol was the systematic identification of a continuous 3-second period of stable phonation. Each recording underwent combined auditory and visual inspection using the waveform, spectrogram, pitch contour, intensity contour, detected vocal pulses, and formant trajectories. The selected segment was required to demonstrate stable fundamental frequency, regular vocal-cycle periodicity, stable waveform amplitude and intensity, and stable F1 and F2 trajectories. Transitional onset and offset regions were excluded, and segment selection began only after pitch, intensity, periodicity, and formant structure had stabilised. Pitch tracking and vocal pulses were additionally inspected for potential tracking errors or irregular periodicity. The selected 3-second segment was used to extract fundamental frequency (f0), local jitter, local shimmer, harmonics-to-noise ratio (HNR), and CPPS, ensuring that all measures were derived from the same inspected period of stable phonation.

Results

The development process produced an operational protocol covering participant eligibility, controlled voice acquisition, signal-quality assessment, stable-segment identification, acoustic measurement, and quality control. Predefined acoustic criteria guided segment selection before parameter extraction, reducing dependence on arbitrary temporal selection and making the analytical decision process explicit and reproducible.

Discussion

The proposed protocol conceptualizes acoustic voice analysis as a controlled measurement framework rather than a software-driven extraction procedure. Its principal contribution is the systematic specification of methodological decisions that can affect acoustic outcomes. By integrating controlled recording conditions, a defined 3-second phonatory sample, multimodal signal inspection, criterion-based segment selection, and complementary perturbation, harmonicity, and cepstral measures, the protocol provides a rigorous framework for acoustic characterization of young-adult voice. It may also support future validation in clinical populations and improve methodological comparability across phonetic and voice research.