

Deriving acoustic voice quality measures from subglottal neck-surface acceleration

Daryush D. Mehta^{1,2,3}; Jarrad H. Van Stan^{1,3}; Robert E. Hillman^{1,2,3}

¹Ctr. for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital; ²Dept. of Surgery, Harvard Medical School; ³Institute of Health Professions, Massachusetts General Hospital

Abstract

This study explores the relationship between **voice quality measures derived from subglottal neck-surface acceleration** and the corresponding acoustic voice signal. Results are reported for adult participants with normal (32 subjects) and disordered (60 subjects) voices. All participants produced sustained vowels at comfortable pitch and loudness during the simultaneous recording of radiated acoustic pressure and subglottal neck-surface acceleration above the sternal notch. Statistically significant **correlations were found between acoustic-based and accelerometer-based jitter, shimmer, harmonics-to-noise ratio, spectral tilt, and cepstral peak magnitude**. As expected, timing-based measures of jitter correlated highly between acoustic and neck-surface acceleration waveforms ($r \leq 0.99$), whereas amplitude-based measures of shimmer correlated less strongly ($r \leq 0.67$). Expected weaker relationships were exhibited by the spectral-based measures of harmonics-to-noise ratio ($r \leq 0.51$) and tilt ($r = 0.33$), whereas cepstral peak magnitude exhibited a stronger correlation ($r = 0.89$).

Motivation

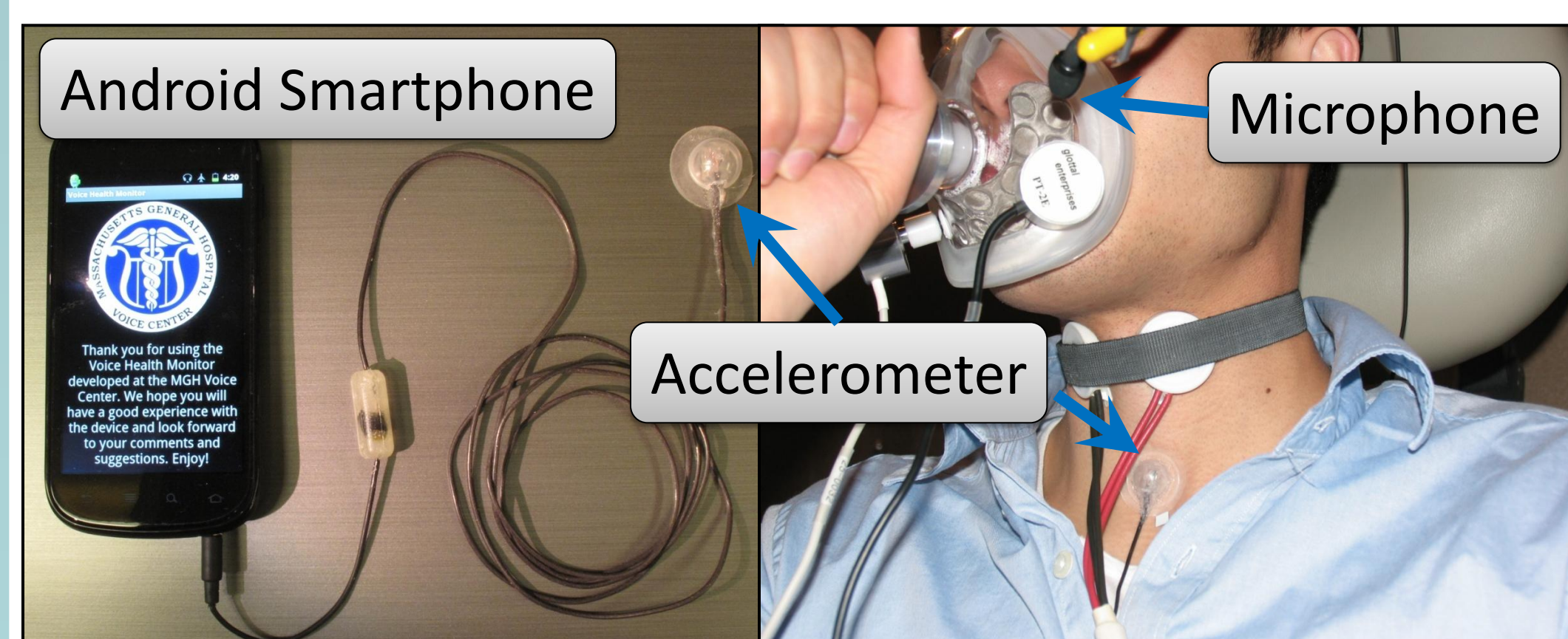
With accelerometer sensors becoming a popular way of tracking voice use and subglottal acoustic characteristics, this study was motivated by the need to extract as much information as possible from the neck-surface acceleration signal recorded subglottally. In particular, the purpose is to **determine whether common acoustic voice quality measures in clinical voice assessment can be derived from the subglottal neck-surface acceleration signal** in participants with and without voice disorders.

Methods

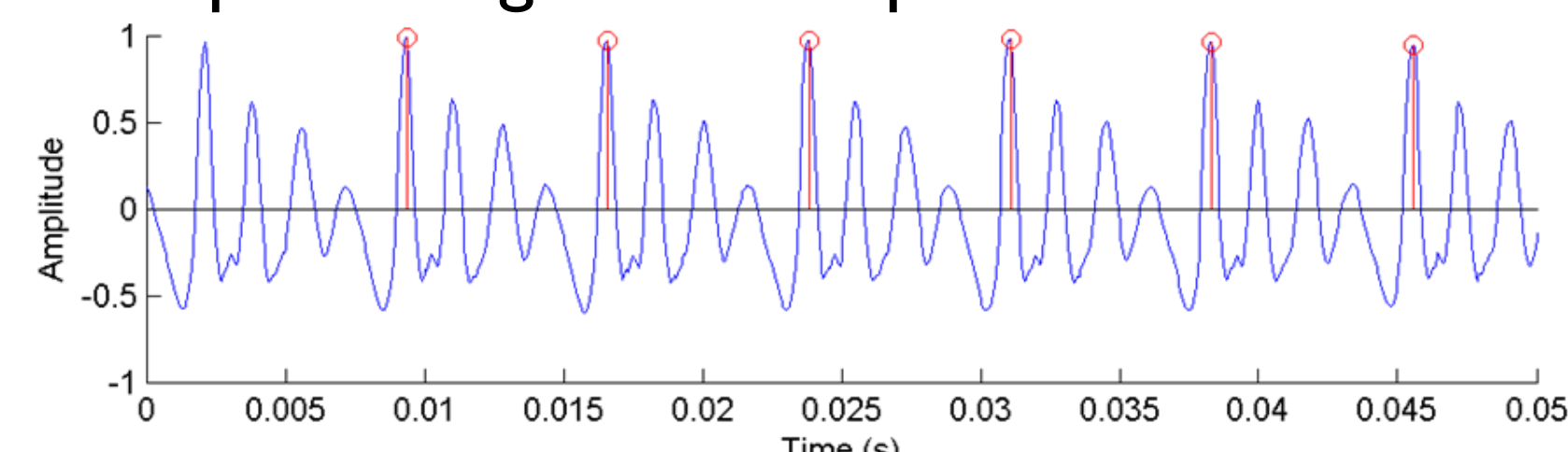
Subject Enrollment

Subject Type	Male	Female
Normal voice	5	27
Vocal fold nodules or polyps	4	31
Muscle tension dysphonia	8	17

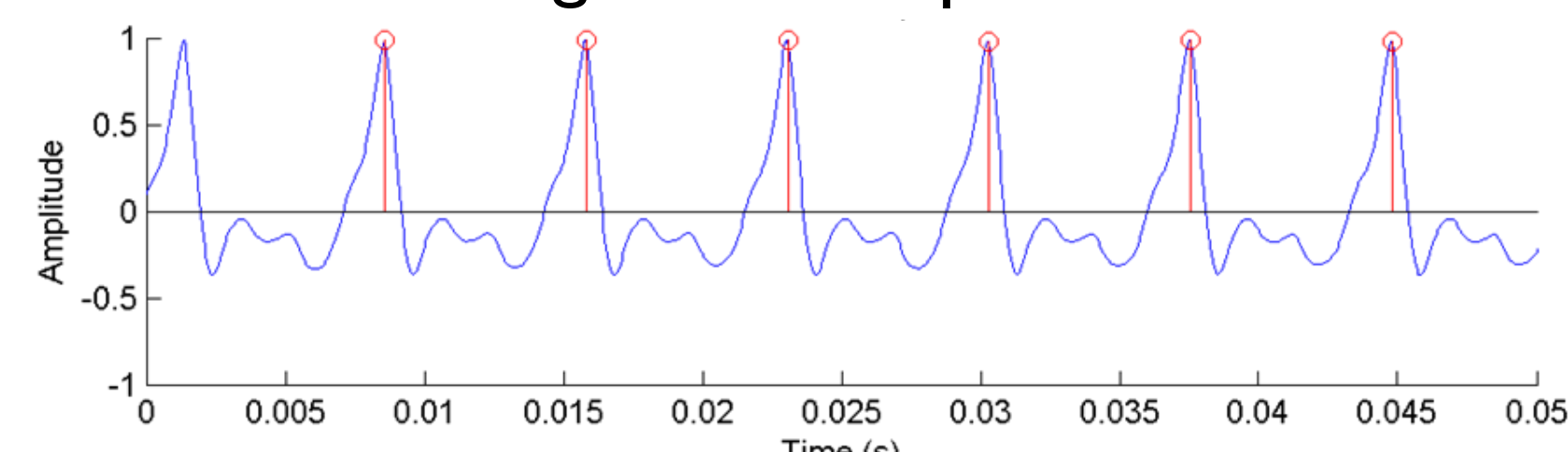
Data Collection (N = 128 /a/ vowels)



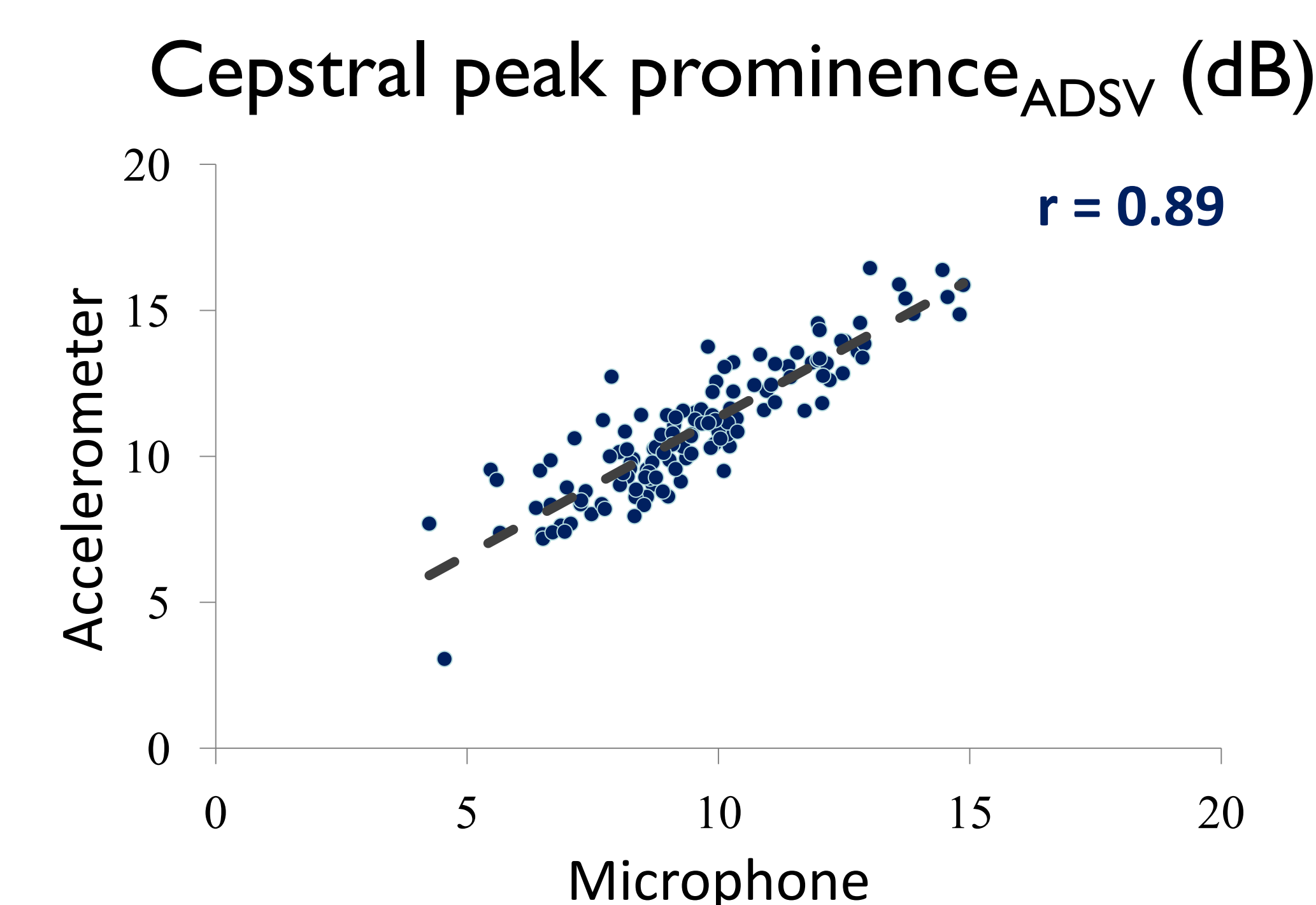
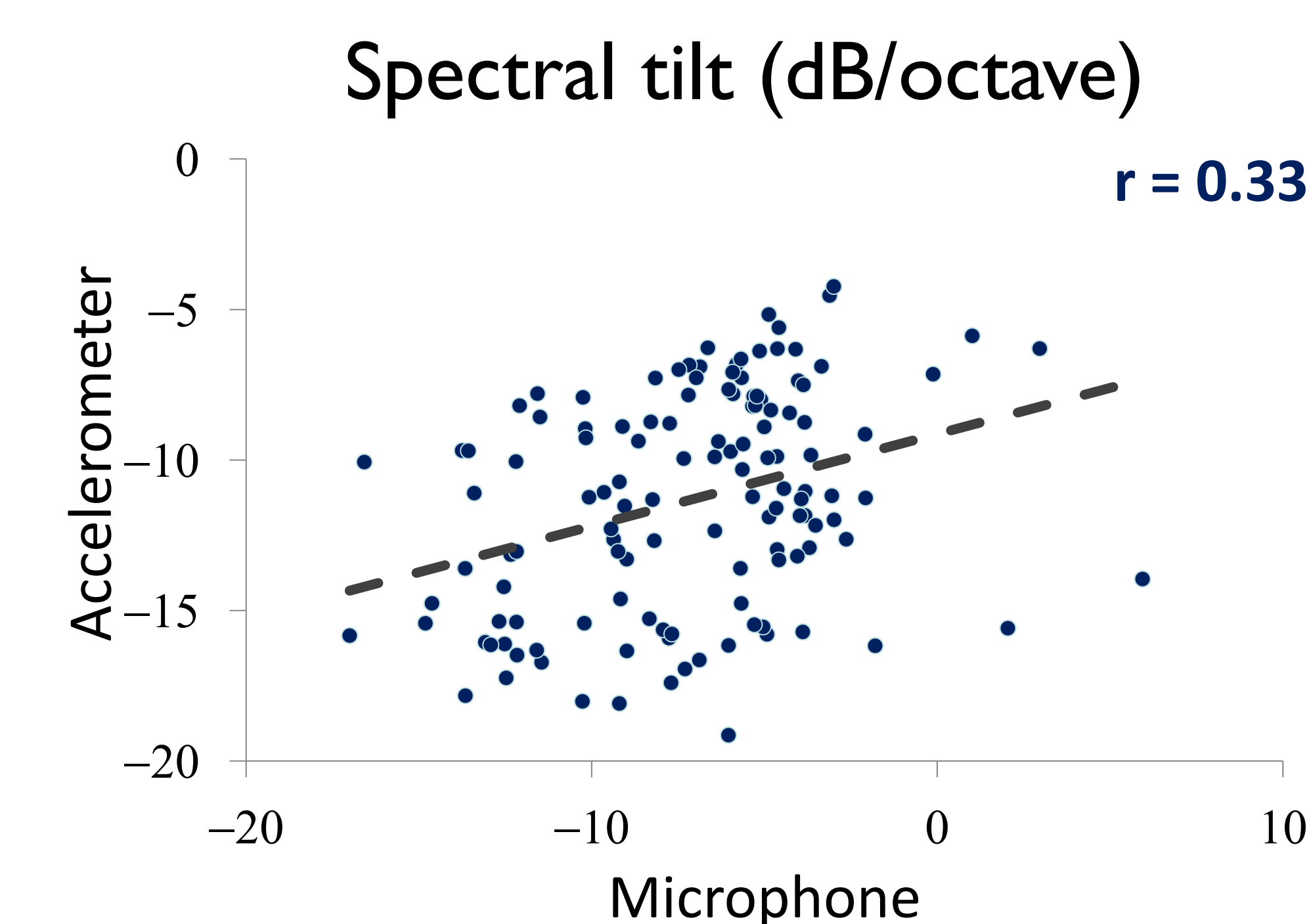
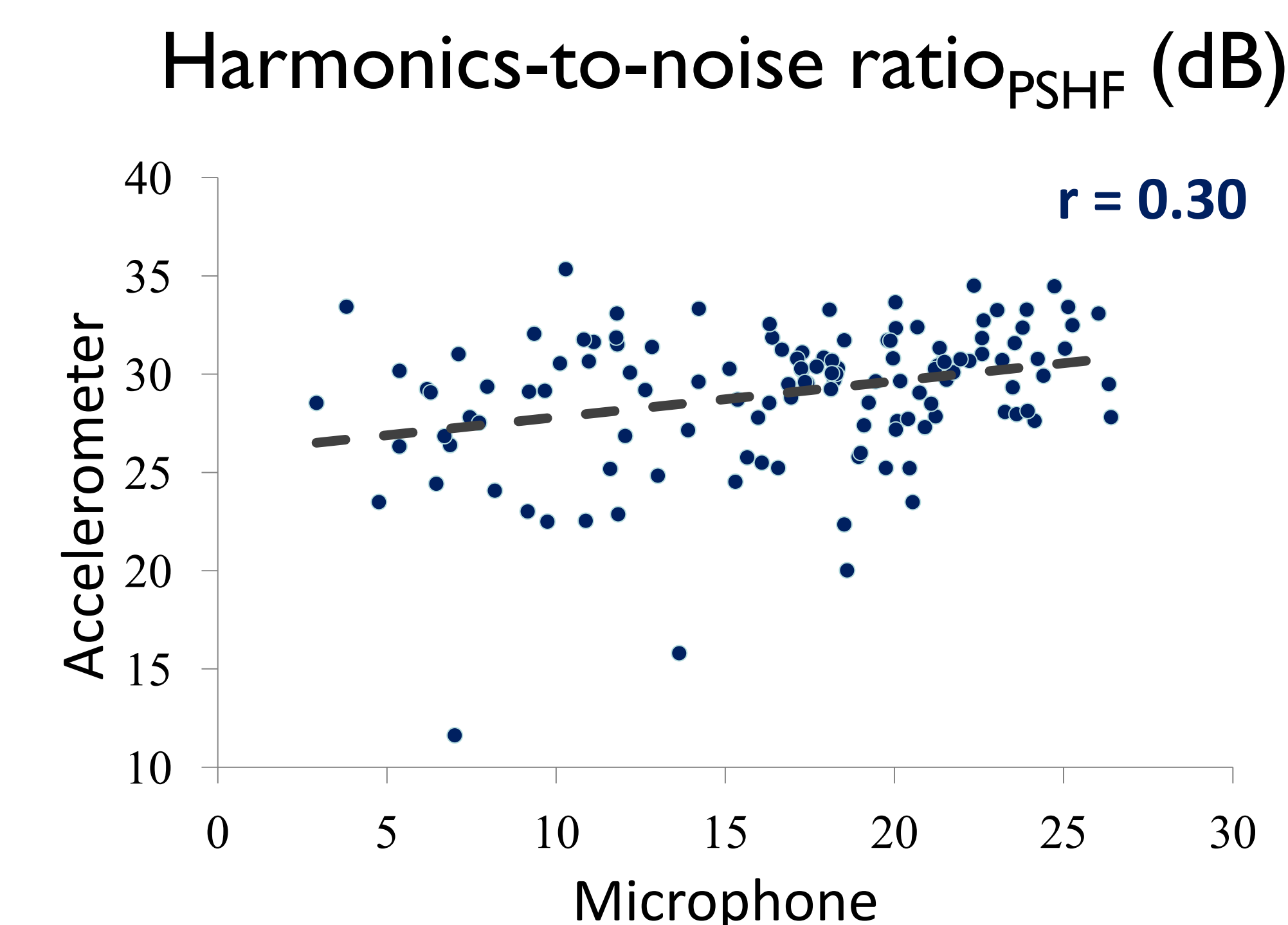
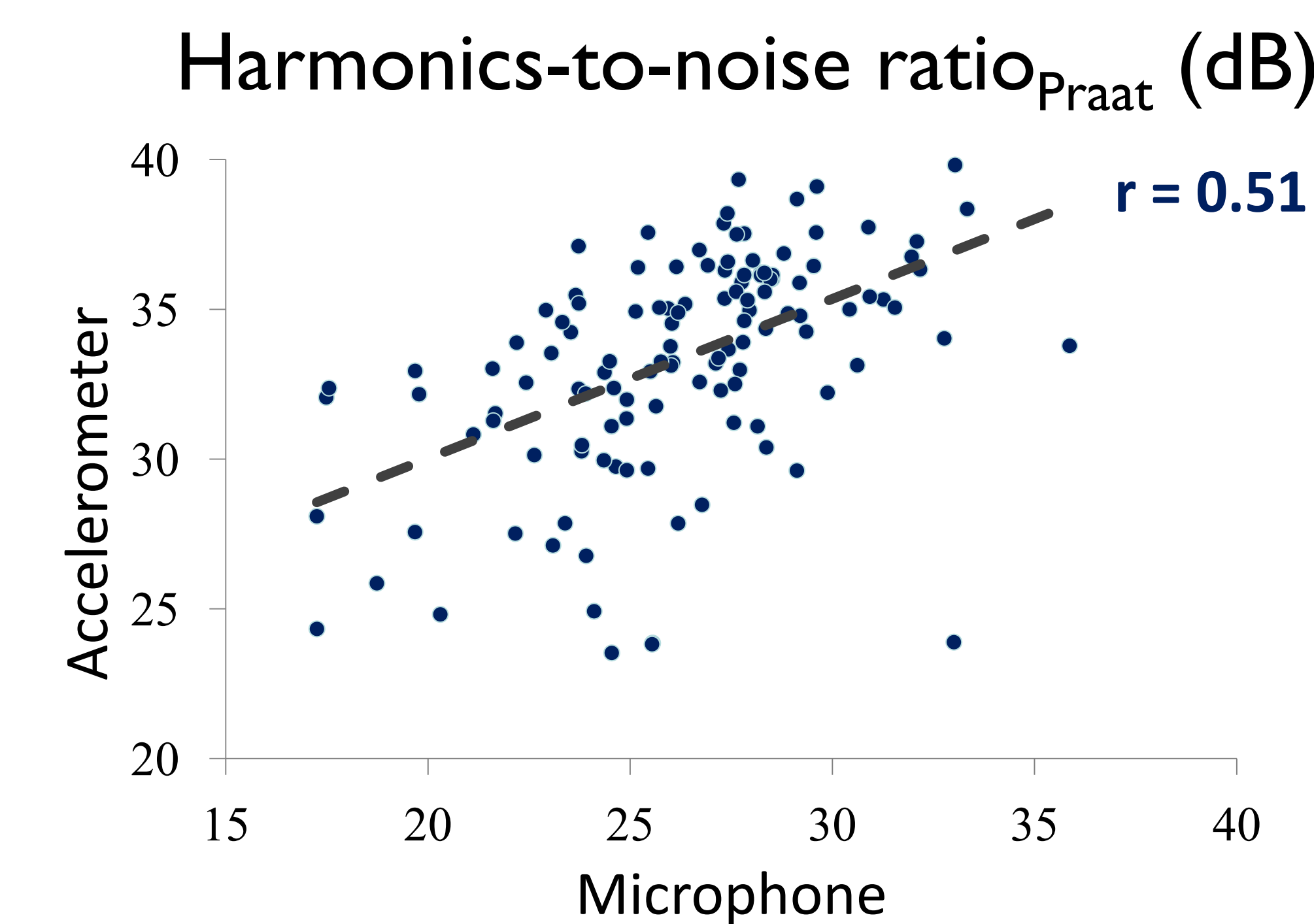
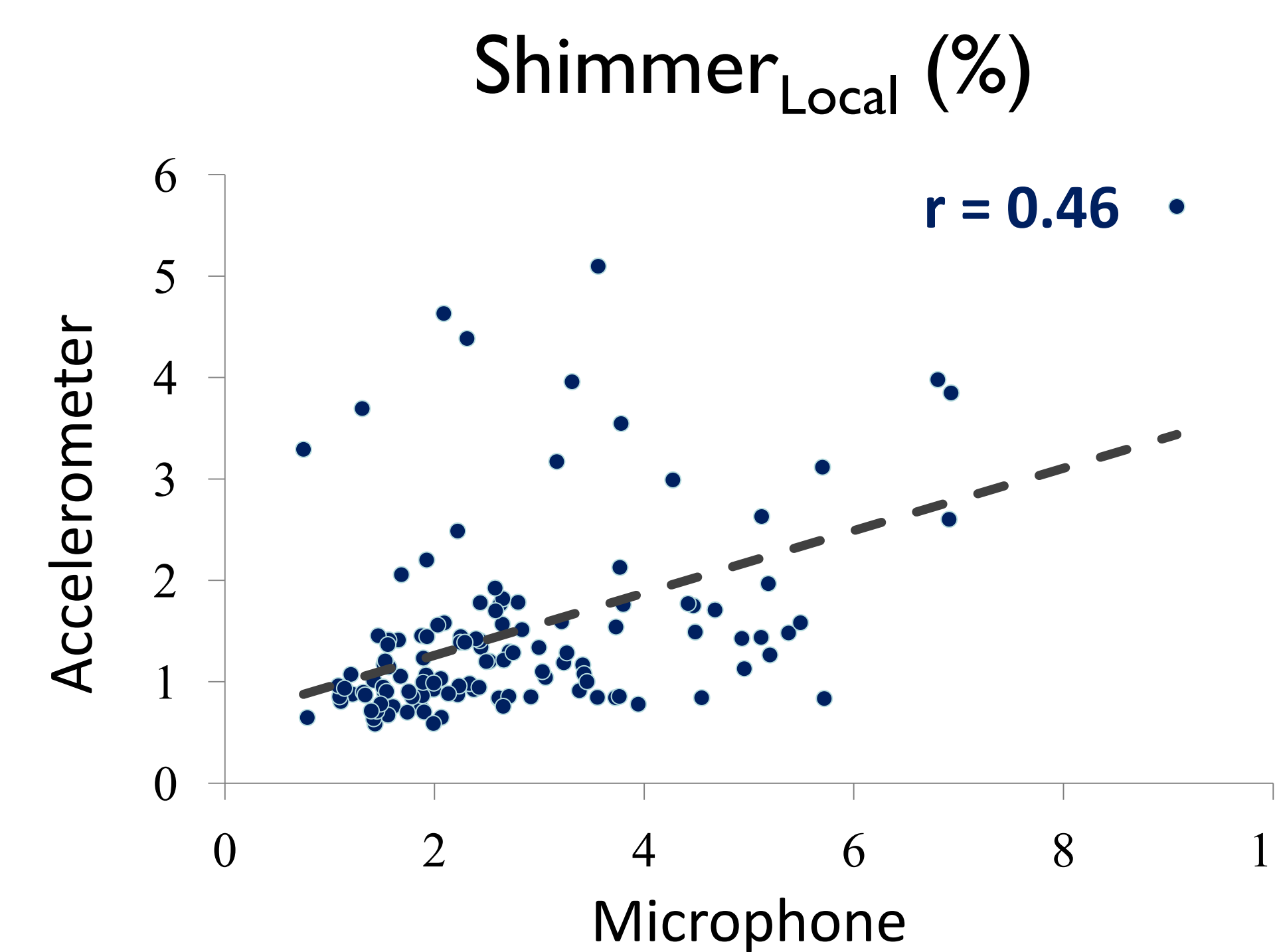
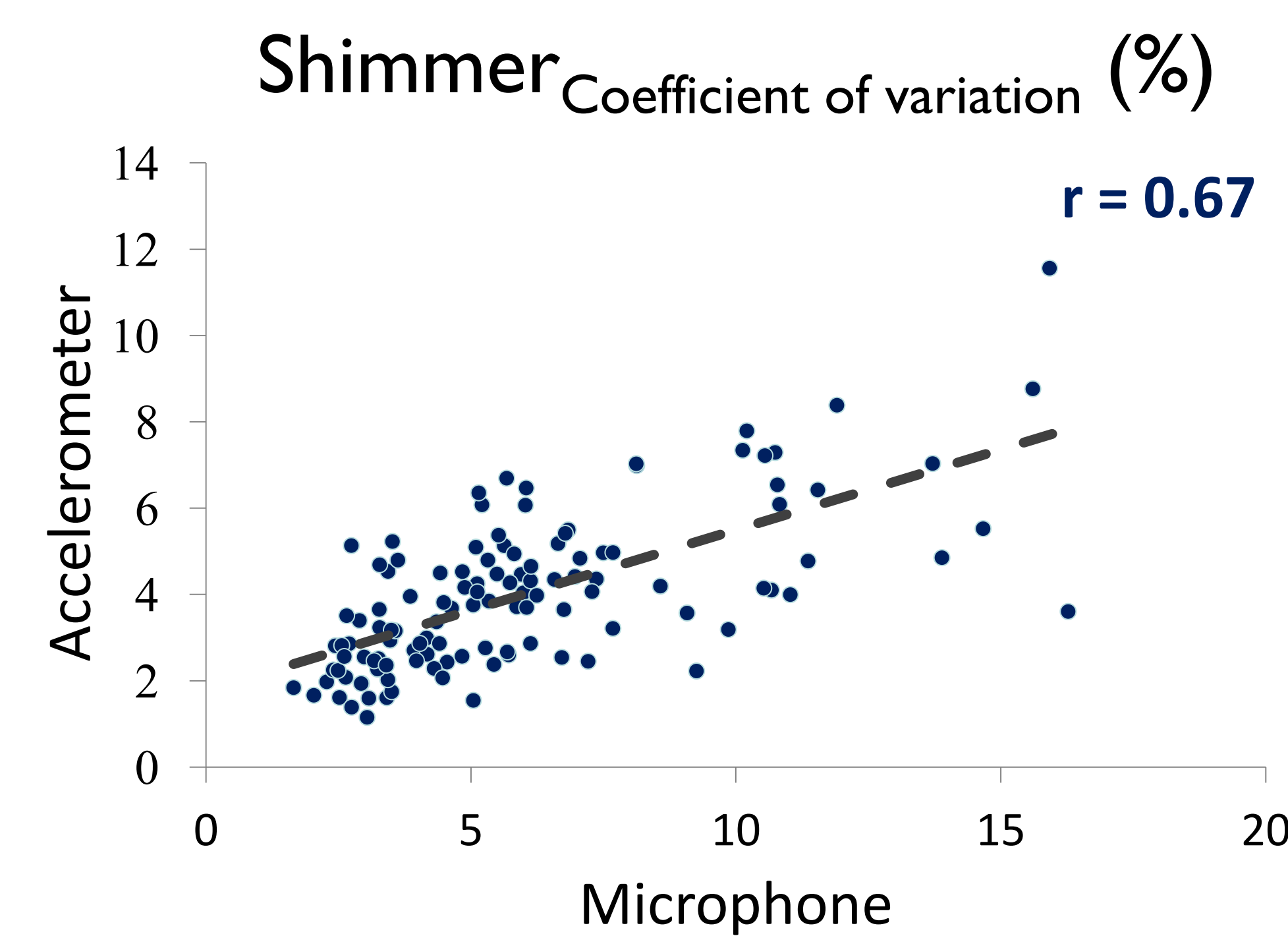
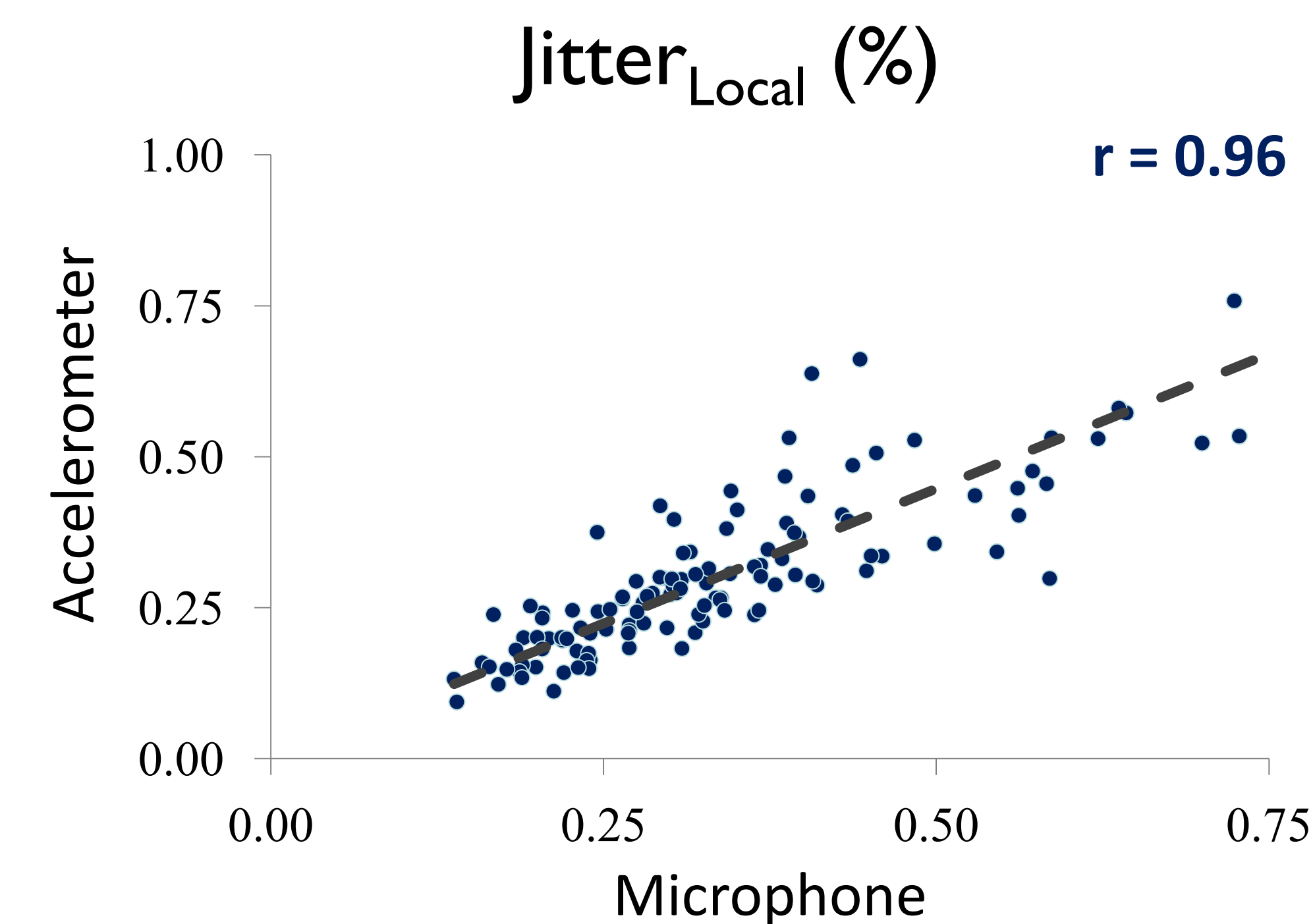
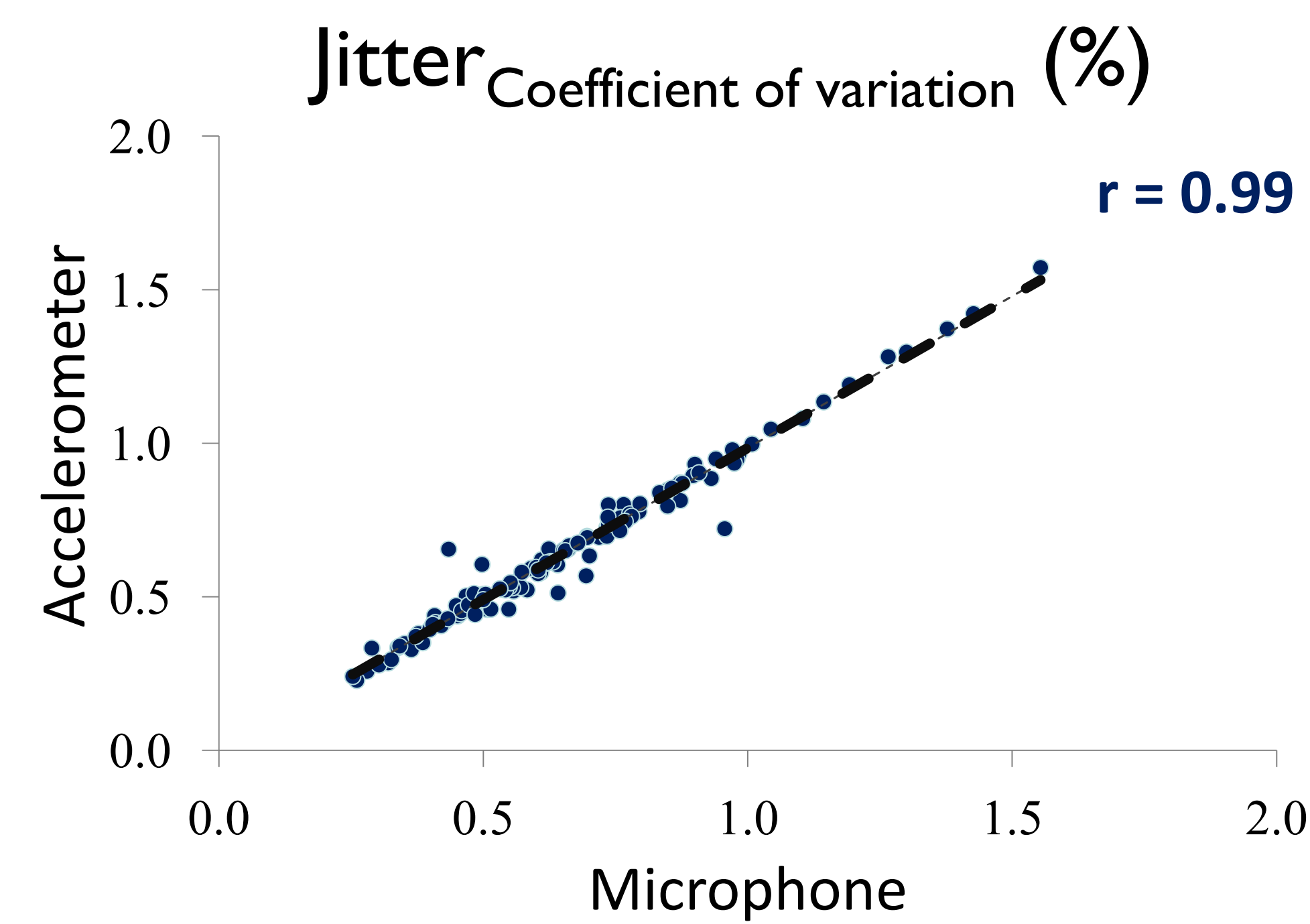
Microphone segment with pitch marks on /a/ vowel



Accelerometer segment with pitch marks on /a/ vowel



Results



Discussion

Accelerometer-Based Voice Quality

This study provides evidence supporting the **derivation of measures related to acoustic voice quality using an accelerometer sensor** placed on the subglottal surface of the neck above the sternal notch:

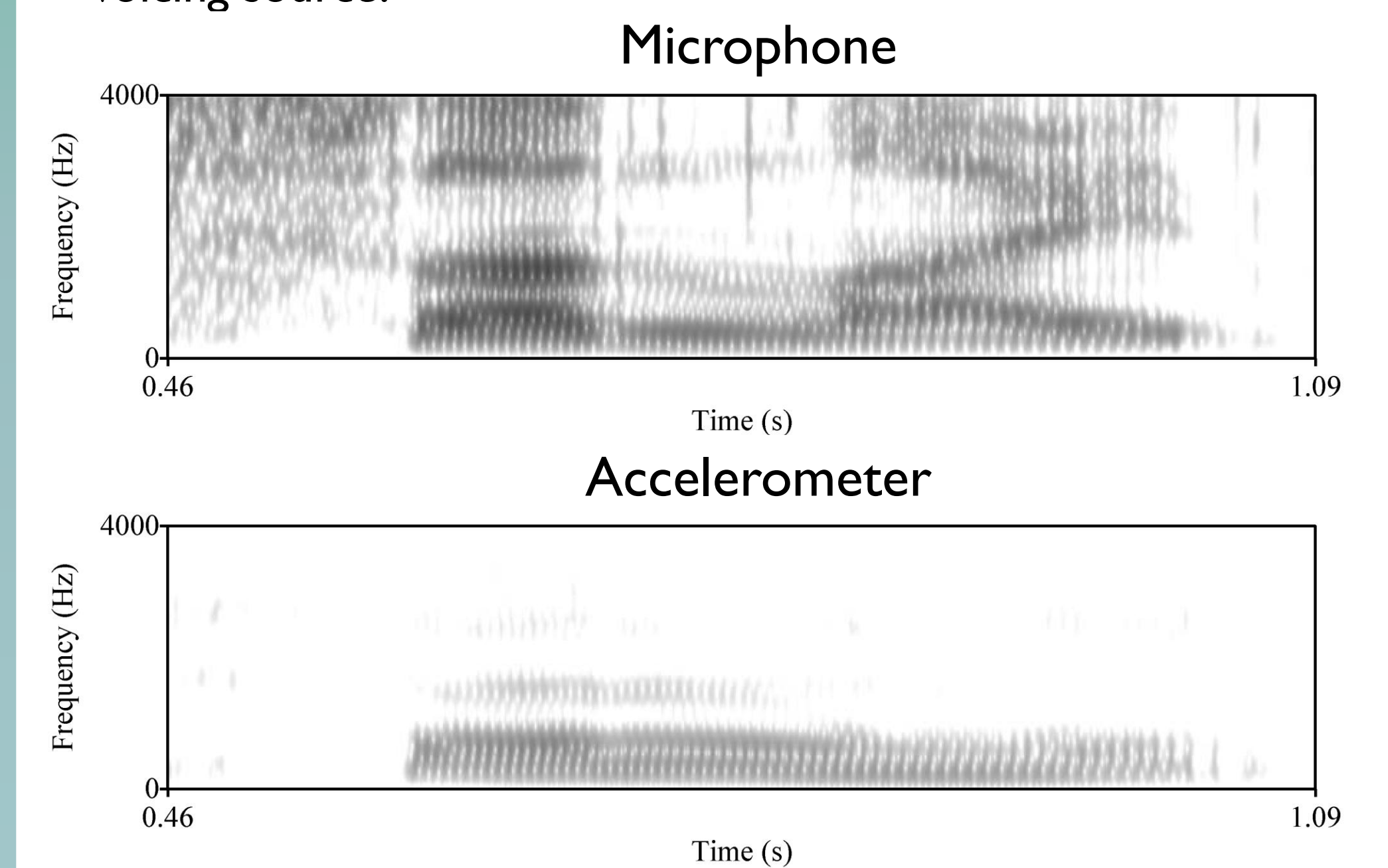
Jitter ($r \leq 0.99$): Strong relationship between waveform periods of neck surface vibration and radiated acoustic sound pressure.

Shimmer ($r \leq 0.67$): Subglottal filtering and vocal fold tissue collision potential have a different impact on the acceleration signal as compared to the radiated acoustic signal.

Harmonics-to-noise ratio ($r \leq 0.51$): Glottal aspiration noise energy appears to be significantly reduced in the subglottal neck-surface acceleration signal.

Spectral tilt ($r = 0.33$): Expected weak relationship between supraglottal and subglottal filtering of the voice source.

Cepstral peak prominence ($r = 0.89$): The accelerometer and acoustic signals reflect similar levels of periodic energy in the voicing source.



Subglottal Neck-Surface Acceleration: Aerodynamics vs. Collision Forces

Amplitude-based and spectral measures correlated less strongly than time-based perturbation metrics, warranting future investigation into the **relative contributions of vocal fold tissue contact and acoustic radiation in the subglottal neck-surface acceleration signal**. Although, auditory perception of the acceleration signal exhibits a muffled (lowpass-filtered) acoustic-like signal, it is unclear whether a mechanical skin surface wave due to vocal fold collision forces significantly affects acceleration properties.

Conclusions

The moderate-to-strong correlations between accelerometer-based and acoustic-based measures of perturbation and cepstral peak prominence provide the potential for **adding voice quality estimation to ambulatory monitoring**. Although accelerometer-based measures of harmonics-to-noise ratio and spectral tilt display weaker correlations with acoustic versions of these measures, the potential of subglottal neck-surface acceleration to contribute additional information regarding **glottal source aerodynamics and vocal fold collision forces** merits further study.

Acknowledgments

This project is supported by a grant from the **NIH National Institute on Deafness and Other Communication Disorders** (R33 DC011588). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.