Sources and interactions of long-term phonatory instabilities

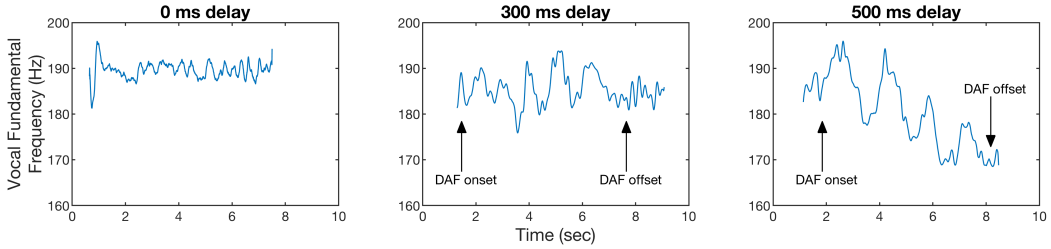
François-Xavier Brajot
Communication Sciences and Disorders
Ohio University. Athens, OH 45701.

Long-term phonatory instabilities are slow oscillations of vocal fundamental frequency (f_o) or intensity ranging from 0.1 to upwards of 25 Hz. They are posited to originate from various sources, including neurological, biomechanical, aerodynamic and acoustic. I review recent empirical and theoretical work that focuses on three such sources: tremor, wow, and a vocal "beat" due to cardiovascular pulsation. I argue that these phonatory instabilities provide a unique window into speech sensorimotor mechanisms and are of significant clinical value.

I take as starting point the reflex resonance model of vocal vibrato proposed by Titze et al. (2002). The model consists of two negative feedback loops that alter descending cortical activation to cricothyroid and thyroarytenoid muscles based on changes in vocal fold tension (a laryngeal reflex). The principle control parameters are the gain and conduction times of sensory feedback. When gain is large enough, vocal f_o oscillates around 4 to 7 Hz. The model accounts for a number of observations, including differences in modulation depth and frequency associated with practice in singers, changes with age, or pathology resulting in vocal tremor.

In contrast to tremor, the 0.2-2 Hz wow has been associated with auditory feedback (Ternström & Friberg, 1989). Support for this is provided in Brajot and Lawrence (2018), in which f_o wow extent and period could be reliably increased by delaying auditory feedback (Figure 1).

Figure 1. Example f_o time series for different delays of auditory feedback (DAF) in one subject.

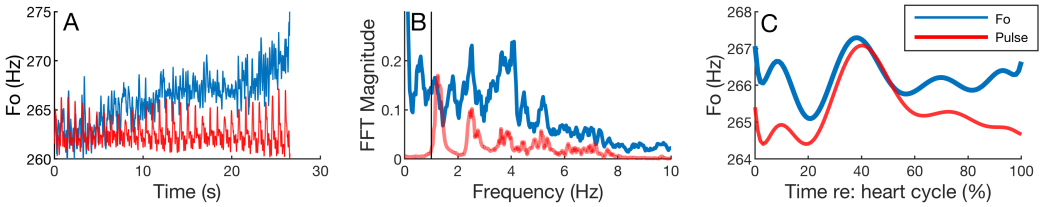


Incorporating an auditory feedback loop into the reflex resonance model described above, a colleague and I show that wow can be elicited from otherwise overdamped dynamics as auditory feedback parameters reach critical values described by an Andronov-Hopf bifurcation (Brajot & Neiman, forthcoming). As with tremor, modulation depth and frequency depend primarily on feedback gain and delay, respectively. We also successfully derive a parametric formula to estimate delays in the feedback loop from wow frequency.

Evaluating the expanded model with both laryngeal reflex and auditory feedback loops in place, we find that the two systems interact. Increasing reflex gain, for example, can lower the threshold at which auditory feedback gain may induce wow. As reflex gain approaches the bifurcation line, f_o oscillations become quasiperiodic (tremor superimposed on wow) for certain intervals of auditory delay. For reflex gains well above bifurcation, tremor predominates. Such interactions have only been identified by numerical analysis and have yet to be verified experimentally.

Another set of oscillations occurs roughly between 1 and 5 Hz, which we have identified as the "phonovascular effect" described by Orlikoff & Baken (1989). This is considered a phonatory instability of biomechanical origin, as the stiffness of the vocal folds varies with the pressure pulse from each heartbeat. Normalizing f_o to heartbeat intervals, we successfully replicate the original findings and further determine that harmonics of the fundamental pulse frequency also contribute to f_o oscillations (Figure 2; Goff & Brajot, submitted). Our data show that these higher harmonics can extend to 6 or 8 Hz, overlapping and possibly interacting with vocal tremor. There also appears to be an interaction with the lower frequency wow, in which the amplitude of this vocal beat, but not its frequency, varies with auditory feedback delay and concomitant wow.

Figure 2. Example f_o and pulse times series (A), corresponding spectrum (B), and heartbeat-normalized average (C) for sustained /a/ in one subject. Pulse was recorded at the earlobe.



Supporting claims by Buder and Strand (2003), these findings suggest that long-term phonatory instabilities provide information relevant for the differential diagnosis of voice disorders. Vocal tremor may point to a disease process that impacts somatosensory function (e.g. Parkinson disease), whereas wow points to deficits in auditory processing (e.g. sensorineural hearing loss). The ability to estimate feedback parameters from the acoustic signal provides a means of quantifying the extent to which a disease has impacted sensorimotor function (e.g. delays in sensory conduction due to demyelination in multiple sclerosis). Interactions across systems may also help explain certain clinical observations. The role of vocal beat in generating tremor may explain the moderate effectiveness of beta-blockers for treating tremor, for example. There is also the question as to whether these observations can generalize to speech oromotor control, which presumably makes use of negative feedback as well. To date, we have not identified a clear oscillatory response to auditory feedback perturbation from analysis of articulatory acoustics alone (Brajot & Lawrence, 2018).

References

Brajot FX, Lawrence D (2018). Delay-induced low-frequency modulation of the voice during sustained phonation. JASA 144: 282-291.

Brajot FX, Neiman AB (forthcoming). An adapted reflex resonance model of vocal wow. JASA.

Goff AE, Brajot FX (submitted). Speaking from the heart: A tutorial. Perspectives ASHA SIGs.

Buder EH, Strand EA (2003). Quantitative and graphic acoustic analysis of phonatory modulations: The modulogram. JSLHR 46: 475-490.

Orlikoff R, Baken R (1989). Fundamental frequency modulation of the human voice by the heartbeat: Preliminary results and possible mechanisms. JASA 85: 888 -893.

Ternström S, Friberg A (1989). Analysis and simulation of small variations in the fundamental frequency of sustained vowels. STL-QPSR 3: 1-14.

Titze IR, Story B, Smith ME (2002). A reflex resonance model of vocal vibrato. JASA 111: 2272-2282.