Published in final edited form as:

J Voice. 2017 July ; 31(4): 512.e17–512.e24. doi:10.1016/j.jvoice.2016.12.002.

Oscillatory onset and offset in young vocally healthy adults across various measurement methods

Rita R. Patel, Ph.D., CCC-SLP*, Reuben Walker, M.M.*, and Michael Döllinger, Ph.D.†

*Department of Speech and Hearing Sciences, Indiana University; Bloomington, Indiana, USA

†University Hospital Erlangen, Medical School, Division of Phoniatics and Pediatric Audiology, Department of Otorhinolaryngology, Head and Neck Surgery, Friedrich-Alexander University Erlangen-Nürnberg

Abstract

Objective—To investigate the relationship between (a) oscillatory onset/offset time across various approaches that use different measurement criteria and (b) oscillatory onset and offset times in vocally healthy young adults.

Method—Oscillatory onset/offset times were obtained from 71 vocally normal adults with high-speed videoendoscopy. Comparisons between the different onset methods involved measurement of the oscillatory onset time (OOT), voice initiation period (VIP), and the phonation onset method (POT) and for offset methods involved computation of the Oscillatory Offset Time (OOT_{off}) method and the phonation offset time (POT_{off}).

Results—Correlation of the OOT with the VIP was 0.240 ($p = .04$) & with the POT form glottal area waveform (GAW) was 0.248 ($p = .04$); however correlation between the VIP and POT GAW was 0.661 ($p < .001$). For offset there was a moderate correlation ($r_s = 0.503$, $p < 0.001$) across OOT_{off} and VOP_{off}. The onset time was longest for the OOT followed by the VIP and the POT. There was no correlation between onset and offset for all methods.

Conclusions—A framework for quantification of oscillatory onset/offset time was developed for /hi/ tasks, which can be used for future measurements of disordered voice. A positive relationship was observed between VIP & POT and between OOT_{off} and VOP_{off}. There was a nonlinear relationship between the OOT and the VIP, POT measures. Onset/offset times are strongly influenced by the calculation method used, the pros and cons of which are discussed in this paper. Vibratory onset and offset represent physiologically different phenomena.

Author to whom correspondence, reprint requests, and proofs should be sent: Rita R. Patel, Ph.D., Indiana University, Dept. of Speech & Hearing Sciences, 200 S. Jordan Avenue, C145, Bloomington, IN 47405-7002, patelrir@indiana.edu, Phone: 812-855-3886, Fax : 812-855-5561.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest: None

Financial disclosure: None

Presentation: Oral presentation at the 45th Annual Voice Foundation Symposium.

Keywords

oscillatory onset time; oscillatory offset time; high-speed videoendoscopy; glottal area waveform; voice onset

INTRODUCTION

Evaluation of vocal fold vibrations is fundamental for appropriate assessment and management of voice disorders. Clinically, vocal fold vibrations are typically evaluated using sustained phonations of varying pitch and loudness. There is widespread belief that evaluation of oscillatory onset and offset are useful for providing additional information regarding vocal fold pathology that is not obtained from sustained phonation; especially inferences concerning the tissue pliability of the vocal folds and vibratory asymmetries.^{1–5} In order to obtain insights into the tissue pliability and vibratory asymmetries from oscillatory onset and offset phenomena, measurement methods must take into account the entire duration of empirically associated variations in oscillatory onset and offset behaviors. However, studies quantifying oscillatory onset and offset vary considerably in terms of the selection criteria for start and end of onset and offset times, thereby resulting in inconclusive findings.^{6–9} How the various methods vary in terms of the onset and offset times is one of the questions of this study. This study also investigates the relationship between the oscillatory onset and offset times in large number of vocally healthy adults as limited investigations of oscillatory onset^{3,5–7,9–19} and particularly of oscillatory offset^{5,9,10} in vocally healthy adults, are on small number of subjects ranging from 1–23 subjects.

Vocal fold oscillatory onset and offset behavior comprises of various transitory segments. Broadly, oscillatory onset involves, (1) vocal fold adductory phase of approximately 203 ± 147 ms, followed by (2) small irregular vocal fold edge oscillatory motion for a brief duration of time (80 ± 43 ms) until steady-state cyclic oscillations are achieved.¹² In some instances of normal onset, small irregular vocal fold oscillatory motions are preceded by complete,^{6,14,20,21} or incomplete^{6,16,17,20} prephonatory glottal closure. The presence or absence of the prephonatory glottal closure is dependent on the nature of the task used to elicit the onset behavior. Various tasks have been used to elicit the oscillatory onset time. These include successive /e.e.e/,^{5,6} /mama/,²² and /he.he.he/.^{8,14} Similarly, oscillatory offset is broadly characterized by (1) brief duration of non-steady state oscillatory motion, followed by (2) complete cessation of oscillatory motion and vocal fold abduction (75 ms or 17 ± 2.5 cycles).^{6,10} Clearly, high temporal resolution as offered by high-speed video endoscopy^{23,24} and/or videokymography²⁵ is required to capture these transitory oscillatory segments of the vocal folds during onset^{6,12,14} and offset.⁶

Various automated quantitative approaches involving measurement of when the onset of vocal fold edge oscillatory motion begins to when steady-state has been achieved, define steady-state as a fraction of the amplitude maximum. Mergell et al (1998)⁷ computed oscillatory onset as the duration of exponential growth from 32.2% to 67.8% of initial to steady-state amplitude derived from the mid-membranous vocal fold trajectories for left and right vocal folds and named this as Phonation Onset Time (POT). Kunduk et al (2006,

2013)^{6,10} calculated oscillatory onset as the time period from the first point the frame-by-frame differences in the glottal area waveform reach 5% of the maximum value to the point where the glottal area waveform reaches 90% of its maximum value and named this the Voice Initiation Period (VIP). The above approaches do not take into account the first oscillatory motion of the vocal fold and when the actual steady-state vibrations begin. The vocal fold onsets often demonstrate amplitude maxima beyond where oscillations could qualitatively be described as steady. Similarly, voice onsets in speech rarely, demonstrate a clean exponential approach to a consistent, steady-state value. Moreover, the above studies report findings on single subject (female). In order to encapsulate the entire growth in the amplitude waveform during onset, Patel, R. et al (2016)²⁶ and Patel R (2016)²⁷ obtained oscillatory onset time from the start of the first oscillation to when the amplitude periodicity was within two standard deviations of the individuals average steady state value and called it the vibratory onset time. In order to avoid confusion with the acronym VOT, since VOT is generally referred to as the 'Voice Onset Time', which is a well-established measure that is computed acoustically from a vowel following a voiceless plosive or fricative,²⁸ we denote the vibratory onset time as Oscillatory Onset Time (OOT).

Compared to oscillatory onset time, there are scant research studies related to the oscillatory offset time. The oscillatory offset time was first quantified as the duration from the first abductory motion of the arytenoids to the cessation of vocal fold vibration motion in the normalized glottal area difference waveform in one young and one old subject.⁶ Kunduk et al (2006)⁶ called it the Voice Offset Period (VOP). Like oscillatory onset time this does not take into account the time from the first instance of departure from steady-state to complete cessation of oscillation. In order to capture the entire offset behavior, Patel, R. (2016)²⁷ calculated the oscillatory offset time as the initial point at which the value of the glottal gap index or the open quotient exceeded two standard deviations of the individuals steady state average value and ending at the last detectable peak in the glottal cycle above zero. To keep terminology consistent with our onset measure, we refer to this as the Oscillatory Offset Time (OOT_{off}). To the best of our knowledge large group studies differentiating oscillatory onset and offset times in vocally normal subjects were not accessible after a thorough literature search. Given the wide range of possibilities of measuring oscillatory onset/offset time, it is critical to compare the various approaches for quantification of onset/offset in normal voices, as the latencies of oscillatory onset/offset would vary based on the choice of the waveforms, the landmarks for the start and endpoints, and the task used to elicit the onset/offset behavior. Further based on findings from studies incorporating mathematical modeling that has shown increased hysteresis,²⁹ reduced phonation threshold flow,³⁰ and reduced subglottal pressure at offset compared to onset, we examined the relationship between onset and offset within measurement methods. Using an oscillation hysteresis model, Lucero (1999)³¹ demonstrated that phonation threshold pressure at offset is 50 to 100% of the onset phonation threshold pressure. In excised canine larynges, Regner et al (2008)³⁰ reported that the phonatory threshold flow was consistently reduced at offset, with the ratio of offset-onset phonatory threshold flow values of 0.707, 1.0. Based on these findings we expect onset and offset to correlate highly within methods.

In summary the study proposes to evaluate the relationship between: (1) oscillatory onset times between the model based method, the Phonation Onset time (POT) versus the two

different empirical based methods: Voice Initiation Period (VIP), and the Oscillatory Onset time (OOT), and (2) oscillatory offset time between the two empirical based methods: Vocal Offset Period (VOP_{off}) and Oscillatory Offset Period (OOT_{off}), and (3) relationship between onset and offset within methods: $OOT \& OOT_{off}$ and $VIP \& VOP_{off}$.

METHOD

Participants

Human subjects data from our prior study (Patel et al, in press)³² was used for this current study. In our prior study we investigated the relationship between the onset and offset of the acoustic signal with various events associated with oscillatory onset and offset of the vocal folds. The current study is substantially different in that this study aims at evaluating the relationship between the various methods investigating onset and offset only from high-speed videoendoscopy. Additionally, the current study has more number of participants compared to our prior study with 56 vocally healthy adults (male = 21, female = 35).³² Ninety-two subjects (21–45 years) were recruited at the Vocal Physiology and Imaging Lab at the University of Kentucky after signing of the Institutional Review Board approved informed consent form. All participants were perceptually judged to have a normal voice on the GRBAS³³ scale by a speech language pathologist specializing in the area of voice. Participants with an overall grade of zero in the GRBAS scale were included in the study. Additionally, all participants had negative histories of vocal pathology, smoking, and professional voice use. Data from 21 subjects could not be obtained due to heightened gag reflex. Hence, high-speed videos from 71 subjects (Female = 44, Male = 27) were subjected to further analyses.

Data Collection

High-speed videoendoscopic recordings were captured at 4000 frames per second with a spatial resolution of 512×256 pixels (PENTAX digital Model 9710, Montvale, New Jersey) for a duration of 4.094 seconds. The participants' produced a minimum of three repetitions of /hi.hi.hi/ at self-selected habitual pitch and loudness level. The combination of a voiceless fricative and a vowel were chosen to obtain the true physiologic onset of the vocal fold by preventing the potential for a hard glottal attack.¹⁴ The average fundamental frequency for females was 241.61 ± 29.51 Hz and for males was 149.74 ± 25.37 Hz.

Data Segmentation and Signal Processing

To prevent any unusual occurrences with the start and end of phonation, the middle /hi/ from a series of 3 consecutive /hi/ syllables was segmented for further analysis. The entire glottal area waveform was extracted using custom developed image processing software called Glottal Analysis Tools (GAT), developed at the Division of Phoniatrics and Pediatric Audiology, University Hospital Erlangen (Germany). The applied image processing algorithms are based on Lohscheller et al (2007).³⁴

Prior to calculation of the POT and the VIP, VOP, the steady-state fundamental frequencies of the trajectories and glottal area waveforms were obtained from the Fast Fourier transform of a 200 ms window centered at the midpoint of the section of the glottal area waveform

exhibiting glottal closure. This point of closure was chosen as a reflection of steady-state oscillation.

A fourth order Butterworth filter was next used to filter each signal between 80% and 120% of the calculated fundamental frequency for the VIP and the POT. Visual inspection of filtered samples showed that this frequency range consistently produced waveforms that accurately reflected the amplitude growth of the unfiltered glottal area waveform.

Computation of Onset Times

Three methods were used to calculate the vibratory onset times: (1) OOT, (2) POT, and (3) VIP.

The *OOT* was calculated in milliseconds from the entire glottal area waveform as the time interval between first detectable oscillatory motion from the zero baseline and the first point at which amplitude periodicity³⁵ was within two standard deviations of the subject's average steady state value (Figure 1). Although time periodicity³⁵ would seem to provide a strong alternative, it was reported to stabilize rapidly for soft and breathy onset qualities.¹⁶ However, amplitude periodicity, calculated from the ratio of neighboring peak maxima, is consistent for sustained phonation in normal voices³⁵ and presents a useful threshold for defining steady-state phonation.

The *POT* as defined by Mergell et al (1998)⁷ was calculated from the left and right mid-membranous trajectories (Figure 1). For comparison with other methods that use the glottal area waveform, the POT was also calculated from the glottal area waveform. In order to calculate POT, amplitude envelope of the initial 40% of the /hi/ sample, calculated via Hilbert transform, was compared to the growth function $r(t)$ provided in equation (1) in Mergell et al (1998)⁷

$$r(t) = r_0 * \left(\left(1 - \left(\frac{r_0}{r_f} \right)^2 \right) e^{-2at} + \left(\frac{r_0}{r_f} \right)^2 \right)^{-\frac{1}{2}} \quad \text{Equation 1}$$

r_0 and r_f are the initial and saturation amplitudes respectively, a is the exponential decay constant, and t is time. The initial 40% of the /hi/ sample was selected to exclude offset behavior. The measure of interest for this investigation is phonation onset time τ , calculated

$$\text{as } \tau = \frac{1}{a}.$$

The POT represents the interval corresponding to amplitude growth from 32.2% to 67.8% of the steady-state value. Solutions to the growth function $r(t)$ were calculated for all r_f between zero and the maximum value of the trajectory and for a (all natural numbers between 1 and 1000). Mean squared error between these solutions and the envelope were calculated for each possible combination of r_f and a and the smallest difference between envelope and r was taken as the best fit. POT was then calculated as the inverse of the best fit value a .

The VIP^{10} was calculated from the total glottal area waveform and is defined as the time in seconds between the 5% and 90% values of the difference between the initial minimum and initial maximum of the polynomial fit (Figure 1). Since the /hi/ task used for this study produces vocal fold oscillations prior to vocal fold closure and the complete adduction, oscillations were considered from the point, which they were first visually identified prior to contact. In order to calculate the 5% and 90% values of maximum glottal area for the VIP, the coefficients for a degree five polynomial fit on the glottal area waveform maxima were computed using MATLAB's polyfit function, which performs a least-squares fit of the data. The polynomial was calculated for the first 100 ms of the onset sample as the value was visually evaluated as sufficient to include the initial maximum of the glottal area waveform in all but one case and served to limit the inclusion of offset behavior. For the single exception, the sample was expanded to 200 ms to include the initial peak.

Computation of Offset Times

Two methods were used to calculate the vibratory offset times: (1) OOT_{off} and (2) VOP_{off} .

The OOT_{off} was calculated in milliseconds from the entire glottal area waveform as the time interval between the initial point at which the open quotient³⁶ or glottis insufficiency^{37,38} exceeded two standard deviations of the steady state average value and terminating at the peak of the last identifiable glottal cycle (Figure 2). The measure of open quotient was used when complete vocal fold closure was observed, whereas in cases of a posterior glottal gap, the glottal gap index^{37,38} was used, as the value of the open quotient would be equal to 1 for the duration of the phonation gesture in instances of incomplete glottal closure.

The VOP_{off} as defined by Kunduk et al (2013)¹⁰ was calculated from the glottal area waveform and is defined as beginning when the difference in adjacent glottal minima crossed a threshold of ten pixels (Personal Communication with M. Kunduk on 2/27/15) and ending with the cessation of oscillation (Figure 2). The final 50% of glottal area waveform minima were selected to isolate offset behavior. Differences between glottal minima were calculated backward with respect to time in order to exclude cases where glottal area increased and then decreased before the final offset gesture. Calculated backward, the final glottal area minima difference greater than 10 pixels was identified. The difference in time between this point and the last visually identified oscillation represents the voice offset time. Each VOP_{off} was visually inspected to ensure that no small, final oscillations had been misclassified as the beginning of the offset gesture. Beginning the calculations at a threshold of seven cycles prior to cessation of oscillation was sufficient to prevent the algorithm from such a misclassification.

Reliability

A total of 20 participants (males = 10, females = 10) were randomly selected for intra rater reliability for the OOT and OOT_{off} measures. One rater performed the analysis for intra-rater reliability. Intra-rater cronbach's alpha (α) was calculated for OOT and OOT_{off} . Oscillatory onset time was calculated as the time between a visually identified onset oscillation landmark and an amplitude periodicity landmark and showed good internal consistency, $\alpha = 0.880$. Oscillatory offset time was calculated as the time between an open quotient or glottal

gap index landmark and a visually identified offset oscillation landmark, and showed excellent internal consistency, $\alpha = 0.931$.

Statistical Analysis

A Shapiro-Wilk test was used to test the normality distribution of all five investigated measures. Correlations between various onset/offset methods were calculated using Spearman's rho to examine shared variance.

In order to evaluate the relationship between OOT and OOT_{off} and VIP and VOP_{off} Spearman's rho correlational analysis was conducted. Significance level was set at $p = 0.05$. The interquartile ratio (Interquartile range / Median) was calculated for all measures to evaluate the variability across the measures. All analyses were performed using IBM SPSS Statistics software Version 19.0.

RESULTS

Comparison of Onset Time Across Measurement Methods

Normality was violated among female subjects under all measures with the exception of VOP_{off}, while normality was violated among male subjects for VIP and VOP_{off} (Table 1). All non-normal measures exhibited a positive skew. Due to the violations of normality, Spearman's rho was used to calculate bivariate correlations between the OOT, VIP, and POT calculated using the glottal area waveform, left and right vocal fold trajectories (Table 1). The mean and standard deviations of the various onset methods across gender are provided in Table 2. As expected, all correlations were positive and significant for pairwise comparisons (Table 3). The correlation between OOT and VIP/POT is weak however, the correlation between the VIP and POT left/right and the POT derived from the glottal area waveform is moderate (Table 3).

Comparison of Offset Time Across Measurement Methods

Due to violations of normality, bivariate correlations were calculated for the OOT_{off} and VOP_{off} using Spearman's rho. A moderate positive correlation was observed between the OOT_{off} and VOP_{off} ($r_S = 0.503$, $p < 0.001$). The mean and standard deviations of the various offset methods across gender are provided in Table 2.

Relationship between onset and offset

There was no correlation between subjects OOT and OOT_{off}, $r_S = 0.007$, $p = 0.951$ and VIP and VOP_{off}, $r_S = -0.116$, $p = 0.337$.

DISCUSSION

Clinically, evaluation of vocal fold vibrations from techniques of laryngeal imaging is crucial for the assessment and treatment of dysphonia. Vibratory onset and offset times can provide insights into the efficiency of voice production with respect to the required subglottal pressure, myoelastic aerodynamic forces acting on the vocal folds, and the elasticity of the vocal folds.^{3,39-42} In the current study, we investigate the relationship

between, oscillatory onset and offset time in vocally healthy young adults and across/within various measurement approaches for the /hi/ task obtained from HSV recordings during typical phonation.

Comparison of onset and offset times across measurement methods

The most important finding is that the three investigated methods (OOT, VIP, and POT) correlate positively and significantly when calculated with the waveform extraction method for which they were designed. This shows that the three measurements from these methods are in fact describing similar behavior. The primary difference between the three investigated methods can be explained through the differences in amplitude growth they describe. The OOT method attempts to estimate the entire duration of onset behavior, from first oscillation to steady-state. The VIP¹⁰ method is expected to be shorter than the semi-automated method because the initial measurement occurs after 5% of steady state amplitude has developed and the final measurement occurs at 90% of the maximum value. The POT⁷ is likewise expected to be significantly shorter than both values because it corresponds to the amplitude growth from 32.2% to 67.8% of a steady-state value. However, a comparison of means shows that the relationship is not simply a linear relationship. Although the VIP and POT reflect 90% and 35.6% of the total amplitude growth respectively, the mean time values for the VIP reflect 63% of the mean OOT measurement for males and 70% for females while the means for POT reflect 15% of the OOT for males and 14% for females. These findings show that the VIP and especially the POT reflect regions of accelerated amplitude growth. For the /hi/ task, changes in vocal fold adduction certainly have an effect on the onset of vibrations, as the vocal folds will only start periodic vibrations when the vocal folds have achieved a sufficient level of adduction. It is possible that future comparisons of differences between these three methods could begin to help identify this adductory threshold in the onset process where such accelerated regions of amplitude growth occur. It remains to be empirically investigated how the VIP varies, when calculated from the median values of the Root-Mean-Square (RMS) amplitudes of bandpass filtered glottal width waveforms as in Ikuma et al (2015).⁵

While all the three measures (OOT, VIP, and POT) correlated positively and significantly, the correlation was moderate only between the POT and the VIP measures, suggesting that both the POT and the VIP methods are positively related, i.e. values of POT increase, the values for VIP would also increase. The POT and the VIP measure growth in the amplitude as a percentage of steady-state value rather than reflecting the entire duration of the transitory behavior like the OOT method. Even though the measurement of the VIP is derived from the entire glottal area waveform and the POT from the left and the right vocal fold trajectories, the correlation between the measures was moderate due to the growth function it captures. Unlike the POT and the VIP methods, the measurement of OOT are derived from the actual glottal area waveform without fitting a curve but using amplitude periodicity as an indicator of steady-state function. The correlation between the POT and the VIP was best when the POT was derived from the glottal area waveform, suggesting that the glottal area waveform could be preferable for measuring oscillatory onset time compared to trajectories due to higher pixel resolution. The amplitude of a single vocal fold trajectory can be less than ten pixels while the amplitude of a glottal area waveform can be two orders of

magnitude higher. This potentially allows for a clearer recognition of initial vocal fold oscillations and increased accuracy of cycle-to-cycle measurements such as amplitude periodicity, open quotient, and glottal insufficiency.

The choice of the best method for measurement of oscillatory onset time will depend on the nature of the research question, type of the task, and the type of the waveform used. Overall, for male and female subjects, the VIP resulted in the least dispersion as determined by the interquartile ratio and hence could be the preferred methods for evaluation of onset in males and females respectively when investigating the rate of growth of amplitude as a function of steady-state value. Measures such as the POT and the VIP, however, are intended for implementation with onset samples of duration long enough for the subject to sustain a steady target pitch and loudness. It is important to remember that such samples are not always possible in severely dysphonic subjects. In such instances, the OOT would be the likely choice of method, as the OOT measure most closely resembles the pre-phonatory durations from video observations. Additionally it should be noted that pass band filters applied to any waveform not beginning with a steady-state level of adduction would remove any evidence of adductory effects. The presence or absence of prephonatory adductory closures will have a significant effect on vocal fold dynamics and will likely influence amplitude growth. When investigating subtle differences in onset time such as adduction gesture, individual variability, or dysfunction, it is possible that an empirical measure such as the OOT method would provide information missed by a mathematically derived measure such as POT, which was developed for a perfect model. It is also worth noting that some subjects' onsets exhibited a peak prior to steady state phonation, which was not predicted by the model used to develop POT (Figure 3). Although POT is based on a mathematical model, it is possible that the modeled amplitude envelope is not capturing all behavior present in voice onset. When present, this peak also served as the maximum for VIP and may have served to produce shorter times compared to the steady-state landmark that was observed in the video. Future investigations of onset using automatic extraction techniques should continue to incorporate both quantitative and qualitative measurement of adductory behavior.

There was a significant moderate positive correlation between the two offset methods of OOT_{off} and VOP_{off} . The primary difference between the two methods is how the two methods measure the amplitude decay. The OOT_{off} method attempts to capture the entire duration of the offset, from the first instance of departure from the steady-state to complete cessation of vocal fold vibration. The VOP_{off} method is expected to be shorter as the initial measurement takes place at first minimum that is 10 pixels larger than the previous minimum. Overall, the OOT_{off} resulted in the lowest interquartile ratio compared to the VOP_{off} method and hence would be the preferred method for measuring oscillatory offset time. A possible weakness of both the OOT_{off} and VOP_{off} is that both are based on empirical observations and could fail to take into account yet-unobserved offset behavior.

In an effort to move measurement of onset and offset times as a standardized clinical measure for evaluation of vocal pathology and change in vocal function following treatment, these study findings present robust data for future studies. As researchers continue to refine measurement methods for automatic calculation of oscillatory onset and offset, it is

important to ensure that such methods continue to reflect empirically observed variations in oscillatory onset and offset.

Relationship between oscillatory onset and offset time

Even though the onset durations were shorter than the offset, the study did not reveal statistically significant correlation between onset and offset times within the two measurement methods (OOT_{on}/OOT_{off} and VIP_{on}/VOP_{off}). The lack of a significant correlation between onset and offset times is likely due to variation in measured values as well as the number of factors underlying thresholds for sustained oscillation in onset and offset. One factor could be the nature of the task used in the study. The vowel /i/ following an /h/, while a good task for onset, results in an unconstrained offset whose duration is largely dependent on rate of abduction. This impedes conclusions about vibratory information from offset measurements. Hence, in order to evaluate the relationship between the onset and offset, it would be advisable to constrain the offset with the presence of a stop consonant, e.g. 'heet,' and a metronome. However, the total absence of an observable relationship suggests that onset and offset durations reflect different physiological mechanisms. Onset and offset are both transient behaviors suggested to provide information not observable during steady-state phonation. This finding suggests that studies interested in transient oscillatory behavior should carefully consider whether onset or offset better reflects the behavior under investigation as onset and offset may reflect different behaviors.

In the current data, the latencies for oscillatory onset were shorter than the oscillatory offset for typical phonation. These data reinforces the hysteresis hypotheses from computational modeling and excised larynx experiments.^{31,41,43} The hysteresis hypothesis from studies on computational modeling reveals that when the glottal half-width is held constant, the phonation threshold pressure and the subglottal pressure required for the onset of oscillation is greater than the pressures at which oscillations will cease.^{31,41}

Conclusions and Future Directions

This study presents normative ranges for oscillatory onset time and oscillatory offset time for normal male and female voices using high-speed endoscopy of /hi/ syllable trains and presents suggestions for future measurement derived from relationships between empirical and mathematical models. A positive relationship was observed between the VIP and the POT measures. There was a nonlinear relationship between the OOT and the VIP, POT measures. For offset methods a positive relationship was observed between the OOT_{off} and the VOP_{off} measures. Overall, vibratory onset and offset represent physiologically different phenomena.

Onset and offset times are strongly influenced by the calculation method used and hence the preferred choice of the method is dependent on the nature of the research question. For measurement of the rate of the growth of the amplitude as a percentage of steady-state phonation the VIP resulted in the lowest normalized dispersion (as measured by interquartile ratio) among male and female subjects and should therefore be considered the preferred method for evaluation of onset. The values produced by the OOT method better reflect the complete duration of transitory behavior. For offset the OOT_{off} appears to be the preferred

method compared to the VOP_{off} . The conclusions here are however, restricted to normal subjects and may not hold true for future investigations in individuals with various types of dysphonia. There was no relationship between onset and offset between OOT vs. OOT_{off} and VIP vs. VOP_{off} suggesting that the onset and offset represent different behaviors. The use of /hi/ syllables allowed for consistency of onset gestures and could be instructive if replicated among subjects with voice disorders. The findings from this study can serve as basis for future studies investigating the sensitivity and specificity of oscillatory onset and offset time in differentiating various pathologies. It is possible that there exists, for offset a direct mathematical analog to the phonation onset time measure, which could be based on the Hopf bifurcation or a similar mathematical theory of bifurcation as the system passes back into instability. Such a measure would allow for a direct quantitative comparison between vocal fold dynamics during onset and offset with the same method, within and across the samples. Future investigations should continue to develop mathematical models that fully capture the variability present in onset behavior including level of adduction and the presence of amplitude peaks prior to steady-state vocal fold vibrations.

Acknowledgments

Sources of financial support or funding: This research was supported by NIH/NIDCD R03DC11360 awarded to Indiana University. Prof. Döllinger's contribution was supported by Deutsche Forschungsgemeinschaft (DFG), Grant no. DO 1247/8-1.

REFERENCES

1. Verdonck-de Leeuw IM, Festen JM, Mahieu HF. Deviant vocal fold vibration as observed during videokymography: the effect on voice quality. *J Voice*. 2001; 15:313–322. [PubMed: 11575628]
2. Goberman AM, Blomgren M. Fundamental frequency change during offset and onset of voicing in individuals with Parkinson disease. *J Voice*. 2008; 22:178–191. [PubMed: 16950600]
3. Braunschweig T, Flaschka J, Schelhom-Neise P, Döllinger M. High-speed video analysis of the phonation onset, with an application to the diagnosis of functional dysphonias. *Medical Engineering & Physics*. 2008; 30:59–66. [PubMed: 17317268]
4. Patel, Pickering J, Stemple J, Donohue KD. A Case Report in Changes in Phonatory Physiology Following Voice Therapy: Application of High-Speed Imaging. *J Voice*. 2012
5. Ikuma T, Kunduk M, Fink D, McWhorter AJ. A Spatiotemporal Approach to the Objective Analysis of Initiation and Termination of Vocal-fold Oscillation With High-speed Videoendoscopy. *J Voice*. 2015
6. Kunduk, Yan Y, McWhorter AJ, Bless D. Investigation of voice initiation and voice offset characteristics with high-speed digital imaging. *Logoped Phoniatr Vocol*. 2006; 31:139–144. [PubMed: 16966156]
7. Mergell P, Herzel H, Wittenberg T, Tigges M, Eysholdt U. Phonation onset: vocal fold modeling and high-speed glottography. *J Acoust Soc Am*. 1998; 104:464–470. [PubMed: 9670538]
8. Patel, Walker R, Sivasankar M. Spatiotemporal quantification of vocal fold vibration following exposure to superficial laryngeal dehydration: A preliminary study. *Journal of voice*. In press.
9. Patel R. Vibratory onset and offset time in children: A laryngeal imaging study. *Int J Pediatr Otorhi*. in press.
10. Kunduk, Chen JH, McWhorter A, Gunturk B. Effects of pitch, loudness, and phonation types on voice initiation and offset investigated with high-speed imaging. *Proceedings of the 10th International Conference on Advances in Quantitative Laryngology, Voice, and Speech Research*. 2013:39–40.
11. Wittenberg T, Tigges M, Mergell P, Eysholdt U. Functional imaging of vocal fold vibration: Digital multislice high-speed kymography. *Journal of Voice*. 2000; 14:422–442. [PubMed: 11021509]

12. Wittenberg T, Moser M, Tigges M, Eysholdt U. Recording, processing, and analysis of digital high-speed sequences in glottography. *Machine Vision and Applications*. 1995; 8:399–404.
13. Eysholdt U, Tigges M, Wittenberg T, Proschel U. Direct evaluation of high-speed recordings of vocal fold vibrations. *Folia Phoniatr Logop*. 1996; 48:163–170. [PubMed: 8823984]
14. Koster O, Marx B, Gemmar P, Hess MM, Kunzel HJ. Qualitative and quantitative analysis of voice onset by means of a multidimensional voice analysis system (MVAS) using high-speed imaging. *J Voice*. 1999; 13:355–374. [PubMed: 10498052]
15. McDonnell M, Sundberg J, Westerlund J, Lindestad PA, Larsson H. Vocal fold vibration and phonation start in aspirated, unaspirated, and staccato onset. *J Voice*. 2011; 25:526–531. [PubMed: 20951547]
16. Werner-Kukuk E, von Leden H. Vocal initiation. High speed cinematographic studies on normal subjects. *Folia Phoniatr (Basel)*. 1970; 22:107–116. [PubMed: 5486597]
17. Freeman E, Woo P, Saxman JH, Murry T. A comparison of sung and spoken phonation onset gestures using high-speed digital imaging. *J Voice*. 2012; 26:226–238. [PubMed: 21256709]
18. Hong KH, Kim HK, Niimi S. Laryngeal gestures during stop production using high-speed digital images. *J Voice*. 2002; 16:207–214. [PubMed: 12150373]
19. Lindestad PA, Sodersten M, Merker B, Granqvist S. Voice source characteristics in mongolian "throat singing" studied with high-speed imaging technique, acoustic spectra, and inverse filtering. *Journal of Voice*. 2001; 15:78–85. [PubMed: 12269637]
20. Wittenberg, T., Mergell, P., Tigges, M., Eysholdt, U. IEEE International Conference on Acoustics, Speech, and Signal Processing. Washington D.C., USA: 1997. Quantitative Characterization of Functional Voice Disorders Using Motion Analysis of Highspeed Video and Modeling; p. 1663
21. Titze IR. On the relation between subglottal pressure and fundamental frequency in phonation. *The Journal of the Acoustical Society of America*. 1988; 85:901–906.
22. Petermann S, Kneisburges S, A Z, Schutzenberger A, Dollinger M. Evaluation of analytical modeling functions for the phonation onset process. *Computational and Mathematical Methods in Medicine*. 2016
23. Patel, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol*. 2008; 117:413–424. [PubMed: 18646437]
24. Deliyski DD, Hillman RE, Mehta DD. Laryngeal High-Speed Videoendoscopy: Rationale and Recommendation for Accurate and Consistent Terminology. *J Speech Lang Hear Res*. 2015; 58:1488–1492. [PubMed: 26375398]
25. Švec JG, Schutte HK. Videokymography: high-speed line scanning of vocal fold vibration. *J Voice*. 1996; 10:201–205. [PubMed: 8734395]
26. Patel RR, Walker R, Sivasankar PM. Spatiotemporal Quantification of Vocal Fold Vibration After Exposure to Superficial Laryngeal Dehydration: A Preliminary Study. *J Voice*. 2016; 30:427–433. [PubMed: 26277075]
27. Patel R. Vibratory onset and offset times in children: A laryngeal imaging study. *Int J Pediatr Otorhi*. 2016; 87:11–17.
28. Klatt DH. Voice onset time, frication, and aspiration in word-initial consonant clusters. *J Speech Hear Res*. 1975; 18:686–706. [PubMed: 1207100]
29. Lucero JC. Optimal glottal configuration for ease of phonation. *J Voice*. 1998; 12:151–158. [PubMed: 9649070]
30. Regner MF, Tao C, Zhuang P, Jiang JJ. Onset and offset phonation threshold flow in excised canine larynges. *Laryngoscope*. 2008; 118:1313–1317. [PubMed: 18401267]
31. Lucero JC. A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offset. *J Acoust Soc Am*. 1999; 105:423–431. [PubMed: 9921668]
32. Patel R, Forrest K, Hedges D. Relationship between acoustic voice onset and offset and selected instances of oscillatory onset and offset in young healthy males and females. *Journal of Voice*. In Press.
33. Hirano, M. Clinical examination of voice. New-York: Springer-Verlag; 1981.

34. Lohscheller J, Toy H, Rosanowski F, Eysholdt U, Döllinger M. Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital high-speed videos. *Medical image analysis*. 2007; 11:400–413. [PubMed: 17544839]
35. Qiu Q, Schutte HK, Gu L, Yu Q. An automatic method to quantify the vibration properties of human vocal folds via videokymography. *Folia Phoniatr Logop*. 2003; 55:128–136. [PubMed: 12771464]
36. Timcke R, Von Leden H, Moore P. Laryngeal vibrations: measurements of the glottic wave. I. The normal vibratory cycle. *AMA Arch Otolaryngol*. 1958; 68:1–19. [PubMed: 13544677]
37. Kunduk, Döllinger M, McWhorter AJ, Lohscheller J. Assessment of the variability of vocal fold dynamics within and between recordings with high-speed imaging and by phonovibrograph. *Laryngoscope*. 2010; 120:981–987. [PubMed: 20422695]
38. Patel, Dubrovskiy D, Dollinger M. Characterizing vibratory kinematics in children and adults with high-speed digital imaging. *J Speech Lang Hear Res*. 2014; 57:S674–S686. [PubMed: 24686982]
39. Jiang JJ, Tao C. The minimum glottal airflow to initiate vocal fold oscillation. *J Acoust Soc Am*. 2007; 121:2873–2881. [PubMed: 17550186]
40. Lucero JC, Koenig LL. On the relation between the phonation threshold lung pressure and the oscillation frequency of the vocal folds (L). *The Journal of the Acoustical Society of America*. 2007; 121:3280–3283. [PubMed: 17552679]
41. Lucero JC, Lourenco K, Hermant N, Van Hirtum A, Pelorson X. Effect of source-tract acoustical coupling on the oscillation onset of the vocal folds. *J Acoust Soc Am*. 2012; 132:403–411. [PubMed: 22779487]
42. Zhang Z. Characteristics of phonation onset in a two-layer vocal fold model. *J Acoust Soc Am*. 2009; 125:1091–1102. [PubMed: 19206884]
43. Lucero JC, Koenig LL. Phonation thresholds as a function of laryngeal size in a two-mass model of the vocal folds. *J Acoust Soc Am*. 2005; 118:2798–2801. [PubMed: 16334896]

Figure 1. Comparisons of the boundaries between the Oscillatory Onset Time (OOT), Voice Initiation Period (VIP), and Phonation Onset Time (POT) of a 28-year-old female subject.

Figure 2.
Comparison of the boundaries of the Oscillatory Offset Time (OOT_{off}) and the Voice Offset Period (VOP_{off}) of a 28-year-old female subject.

Figure 3.
Amplitude peak prior to steady-state phonation of a 22-year-old female subject.

Table 1

Shapiro-Wilk test of normality for onset

Measure(s)	Method(s)	Male			Female		
		s-w	df	p	s-w	df	p
OOT	GAW	0.928	27	0.063	0.820	44	<0.001
VIP	GAW	0.743	27	<.001	0.922	44	<0.01
POT	L Traj	0.959	27	0.355	0.820	44	<0.001
POT	R Traj	0.966	27	0.505	0.819	44	<0.001
	GAW	0.966	27	0.500	0.881	44	<0.001
OOT _{OFF}	GAW	0.994	27	0.999	0.948	44	<0.05
VOP _{OFF}	GAW	0.886	27	<0.01	0.954	44	0.076

s-w = Shapiro-Wilk statistics, df = degrees of freedom, p = p value, OOT = semiautomated method onset, POT = Phonation Onset Time, VIP = Voice Initiation, L = left vocal fold, R = right vocal fold, GAW = Glottal Area Waveform, Traj = Trajectory.

Table 2

Summary statistics for oscillatory onset time, voice initiation period, phonation onset time, oscillatory offset time and the voice offset period. All values are presented in milliseconds.

Method	Waveform	Male (n = 27)			Female (n = 44)		
		Mean(SD)	Median (IQR)	Range	Mean(SD)	Median (IQR)	Range
Oscillatory Onset Time	GAW	40(18)	38 (27)	15–87	37(21)	31 (20)	13–121
Voice Initiation Period	GAW	25(10)	25(9)	10–68	26(7)	24 (6)	4–47
Phonation Onset Time	Left Trajectory	6(2)	6(4)	2–10	5(3)	5 (3)	1–14
	Right Trajectory	6(2)	6(3)	3–10	5(2)	5 (2)	2–14
Oscillatory Offset Time	GAW	7(2)	7(4)	2–11	6(2)	5(3)	3–12
	GAW	91(28)	91 (40)	23–151	93(28)	91 (26)	7–167
Voice Offset Period	GAW	83 (39)	74 (42)	38–151	79(35)	83 (41)	2–163

GAW = Glottal Area Waveform, SD = Standard Deviation, IQR = Interquartile range

Table 3

Pairwise Spearman Rank Coefficients for onset measures: Oscillatory Onset Time (OOT), Voice Initiation Period (VIP) and Phonation Onset Time for the left (POT_L), the right vocal fold (POT_R), and the glottal area waveform (POT_{GAW}); and offset measures Oscillatory Offset Time (OOT_{off}) and Voice Offset Period (VOP_{off}).

		OOT	VIP	POT _L	POT _R	POT _{GAW}	OOT _{off}	VOP _{off}
Onset Measures	OOT	1	.240 *	.390 **	.257 *	.248 *	.007	-.019
	VIP		1	.502 **	.573 **	.661 **	.143	-.116
	POT _L			1	.529 **	.519 **	-.013	-.186
	POT _R				1	.620 **	.037	-.092
	POT _{GAW}					1		
Offset Measures	OOT _{off}						1	.503 **
	VOP _{off}							1

**
 $p < .01$,

*
 $p < .05$