

Research Note: Vocal Attack Time—Extended Analysis

R.J. Baken and Ben C. Watson, *Valhalla, New York*

Summary: Normative data for vocal attack time (VAT) have previously been presented, but descriptive statistics and assessments of statistical significance of differences have previously been based on a data corpus that included both negative- and positive-valued VATs. Negative VAT values denote a glottal vocal attack, but, at the present time, the signification of the magnitude of a negative VAT value is unknown. The magnitude of a positive VAT value, on the other hand, conveys useful information about glottal behavior at the time of voice onset and is much more likely to be of use, especially in the clinical domain. We present descriptive statistics for the set of positive-valued VATs and demonstrate that the VAT differences between genders and among tone categories in Cantonese remain valid.

Key Words: Voice onset–Vocal attack–Voice initiation–VAT–Positive VAT.

Vocal attack time (VAT) is a measure of the duration of the onset phase of phonation. First proposed in 1998,^{1,2} VAT involves quantification of the time lag between the growth of the sound pressure and vocal fold contact (EGG Lx) signals at the start of phonation. Following validation studies,^{3,4} normative data for VAT were obtained for the onset of the isolated vowel /a/ in normal adult (age range: 22–50 years) speakers of American English⁵ and for the syllable /a/ at each of the lexical tone levels in normal adult speakers of Hong Kong Cantonese.⁶ Changes in VAT associated with vocal fundamental frequency⁷ and with the lexical tones of Mandarin Chinese,⁸ Amdo-Tibetan plosives,⁹ and the F_0 -vowel interaction in Mandarin Chinese¹⁰ have also been examined.

Details of the calculation of VAT values are reported in detail elsewhere.⁵ VAT represents the lag time of the peak in the cross-correlation function of the /a/ bandpass-filtered analytic sound pressure and electroglottographic signals during the period of vocal onset. It is possible for VAT to have a negative value and, even in normal speakers, it often does. This may be somewhat counterintuitive—how could onset time be less than zero?—but in fact is quite reasonable. Figure 1 demonstrates why. Here, simultaneous videokymographic, sound pressure, and vocal fold contact area (EGG) records are shown for the period of phonatory onset of two sustained vowels. In Figure 1A, phonation begins with small oscillations of the vocal folds that grow in amplitude until, finally, the vocal folds lightly contact each other at about 115 ms. As the oscillations grow stronger, the amount of vocal fold contact rapidly increases. As the EGG is primarily representative of vocal fold contact, the signal is extremely weak during the noncontact cycles but grows strongly and very rapidly once contacts are established. Hence, in this case, the VAT—the delay time between growth of the EGG and sound pressure signals—is positive: sound pressure precedes EGG.

Figure 1B shows a different situation: the vocal folds are in full contact before oscillation begins. The result is that there is no vocal sound pressure signal until the first glottal opening, and

even then, the sound pressure is very small, because most of the vocal acoustic energy is generated by vocal fold *closure*. The EGG signal therefore grows before the sound pressure signal—a situation that is signaled by a negative VAT. Thus, a negative VAT does signify something meaningful: vocal fold contact is effectively complete before phonatory onset.

It is clear that, although VAT values may span a continuous domain from some negative to some positive limit, negative and positive values represent two distinctly different physiological regimes. VAT values that are either zero or negative signal the probability of a “hard onset,” a “glottal attack,” or, in the older clinical terminology, “coup de glotte.”¹¹ It has long been believed that, at best, hard onset contributes to a “harsh” or otherwise unaesthetic voice,¹² and, at worse, is associated with vocal fold dysfunction or injury¹³ and thus is a target of therapy.^{14,15}

Unfortunately, although values of $VAT \leq 0$ signal something useful—that a hard glottal attack has occurred—it is not yet possible to specify what the *magnitude* of a negative VAT represents, except to say that there is some as-yet-unspecified prephonatory activity involving vocal fold contact. Pending further clarifying research, all $VATs \leq 0$ can only be taken to mean the same thing.

The same is not true for a $VAT > 0$, which does signal something that is physiologically interpretable. Clinical voice specialists commonly accept that a voice onset that is “soft,” but not *too* soft, can be a meaningful contributor to clinical dysphonia and is therefore a legitimate goal of therapy.^{13,16,17} Given that, it is likely that the magnitude of only positive VAT values will be useful in the clinical sphere. Our previous reports on VAT did not provide statistical data for the isolated corpus of positive-valued VAT. The purpose of the present report was to correct that lack.

Table 1 summarizes descriptive statistics for positive VAT values in normal native speakers of American English.^a Repeated measures analysis of variance revealed significant main effects for gender ($F(1, 502) = 5.08, P < 0.05$) and age group ($F(3, 502) = 7.02, P < 0.001$). For the complete data set, as well as for men considered separately, significant differences were found between the ≤ 24 -year-olds and the 30- to 39-year-olds, between the 25- to 29-year-olds and the 30- to 39-year-olds, and between

Accepted for publication December 22, 2017.

From the School of Health Sciences and Practice, New York Medical College, Valhalla, New York.

Address correspondence and reprint requests to Ben C. Watson, School of Health Sciences and Practice, New York Medical College, 40 Sunshine Cottage Road, Valhalla, NY 10595. E-mail: ben_watson@nymc.edu

Journal of Voice, Vol. 33, No. 3, pp. 258–262
0892-1997

© 2018 The Voice Foundation. Published by Elsevier Inc. All rights reserved.
<https://doi.org/10.1016/j.jvoice.2017.12.014>

^aTo mirror our original publication of these data, the mean positive VAT by gender and age is summarized in the Appendix.

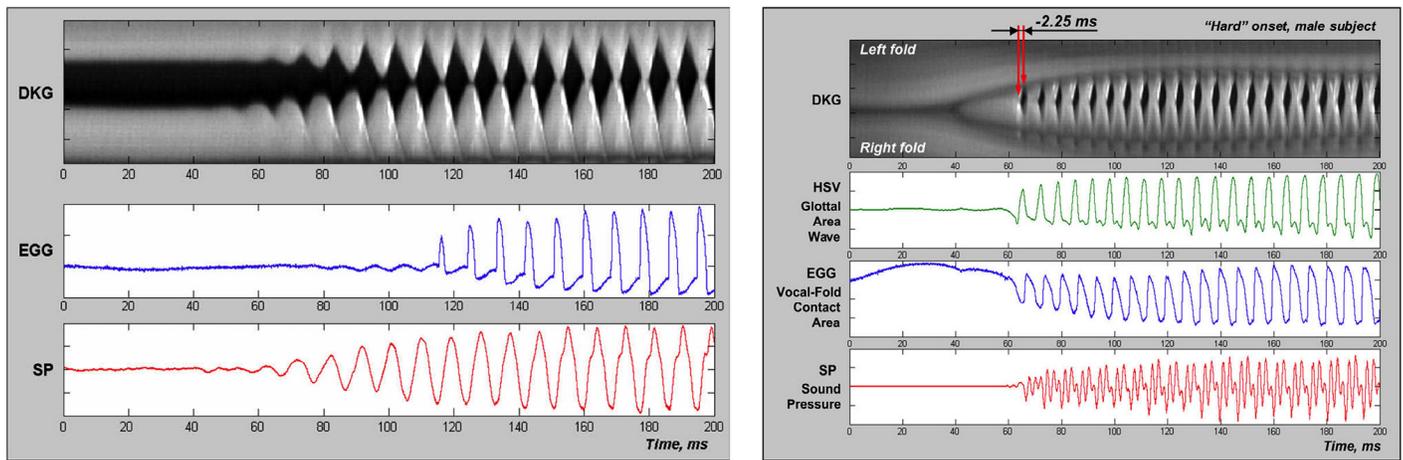


FIGURE 1. Two examples of vocalic phonatory onset. 1A (Left): Normal “soft” onset. Top to bottom: digital kymogram (DKG), vocal-fold contact area (EGG), and sound pressure (SP) signal. A sound pressure wave is present before the development of the EGG signal, resulting in a positive VAT. 1B (Right): Phonation beginning with a closed glottis (“hard attack”). The EGG signals grows very early in the onset phase compared to the sound pressure signal. The VAT is thus slightly negative. DKG: Digital kymogram; EGG: vocal fold contact area (EGG); SP: Sound pressure; HSV: High-speed video.

the 25- to 29-year-olds and the ≥40-year-olds. For women, considered separately, significant differences were found between the ≤24-year-olds and the 30- to 39-year-olds and between the 25- to 29-year-olds and the 30- to 39-year-olds.

Table 2 summarizes descriptive statistics for positive VAT values for a similar group of native Cantonese speakers. Table 3 summarizes the descriptive statistics for positive VAT values in

Cantonese speakers by contour vs level tone pattern. Repeated measures analysis of variance revealed significant main effects for only tone pattern (contour vs level) ($F(1, 57) = 53.12, P < 0.0001$).

The frequency distributions for American English male and female speakers’ VAT are presented in Figure 2 and, for the entire corpus, in Figure 3.

TABLE 1.
Descriptive Statistics for Positive VAT Values (ms) of Normal Native Speakers of American English

Group	Number of Tokens	Mean (SD)	95% Confidence Interval	Median	25th–75th Percentiles
All subjects	728	4.77 (3.58)	4.50–5.02	3.99	2.06–6.71
Women	304	4.08 (3.56)	3.68–4.48	3.22	1.51–5.33
Men	424	5.26 (3.51)	4.92–5.59	4.55	2.49–7.44
Both genders					
≤24 YO	240	4.60 (3.93)	4.09–5.09	3.51	1.63–6.43
25–29	275	5.56 (3.62)	5.13–5.99	5.06	2.74–7.87
30–39	121	4.01 (2.86)	3.49–4.52	3.65	2.02–5.13
≥40	92	3.83 (2.82)	3.25–4.41	3.22	1.87–5.59
Women					
≤24 YO	139	4.08 (3.72)	3.45–4.70	3.22	1.52–4.83
25–29	90	4.63 (3.69)	3.85–5.39	3.67	1.77–6.42
30–39	26	2.23 (1.46)	1.64–2.81	2.34	0.95–3.81
≥40	49	4.08 (3.40)	3.09–5.05	3.45	1.50–6.01
Men					
≤24 YO	101	5.31 (4.12)	4.49–6.12	4.35	2.04–7.28
25–29	185	6.02 (3.50)	5.50–6.52	5.71	3.51–8.48
30–39	95	4.50 (2.96)	3.89–5.10	4.19	2.27–5.62
≥40	43	3.56 (1.95)	2.95–4.15	3.18	1.99–4.92

Abbreviations: SD, standard deviation; YO, year old.

TABLE 2.
Descriptive Statistics for Positive VAT Values (ms) of Normal Native Speakers of Cantonese by Tone

Cantonese Tone Number	Homophone	Mean (SD)	Overall Mean (SD)
Women			
Level (F_0 pattern) tones			
1 (High)	丫叉 /a ₁ ts ^h a ₁ /	3.99 (2.87)	4.22 (3.49)
	鴉片 /a ₁ p ^h in ₃ /	4.47 (4.08)	
3 (Mid)	亞洲 /a ₃ tsɛu ₁ /	5.25 (3.15)	4.71 (3.36)
	亞軍 /a ₃ kwɛn ₁ /	4.17 (3.49)	
6 (Low)	硬幣 /an ₆ pɛi ₆ /	5.78 (3.20)	5.19 (3.53)
	掙扎 /a ₆ tsa ₆ /	4.48 (3.80)	
Contour (F_0 pattern) tones			
2 (High rising)	啞巴 /a ₂ pa ₁ /	5.60 (4.74)	5.82 (4.22)
	啞拑 /a ₂ liŋ ₄ /	6.03 (3.66)	
4 (Low falling)	芽菜 /a ₄ ts ^h ɔi ₃ /	5.53 (3.16)	5.63 (3.32)
	牙齒 /a ₄ ts ^h i ₂ /	5.71 (3.47)	
5 (Low rising)	瓦片 /a ₅ p ^h in ₂ /	5.87 (3.02)	5.77 (3.16)
	雅典 /a ₅ tin ₂ /	5.67 (3.29)	
Men			
Level (F_0 pattern) tones			
1 (High)	丫叉 /a ₁ ts ^h a ₁ /	5.32 (3.10)	5.13 (3.73)
	鴉片 /a ₁ p ^h in ₃ /	4.92 (4.28)	
3 (Mid)	亞洲 /a ₃ tsɛu ₁ /	5.48 (3.97)	5.48 (3.94)
	亞軍 /a ₃ kwɛn ₁ /	5.48 (3.92)	
6 (Low)	硬幣 /an ₆ pɛi ₆ /	5.35 (3.19)	5.56 (3.55)
	掙扎 /a ₆ tsa ₆ /	5.76 (3.88)	
Contour (F_0 pattern) tones			
2 (High rising)	啞巴 /a ₂ pa ₁ /	6.68 (3.90)	6.58 (3.76)
	啞拑 /a ₂ liŋ ₄ /	6.48 (3.63)	
4 (Low falling)	芽菜 /a ₄ ts ^h ɔi ₃ /	6.26 (4.47)	6.20 (4.10)
	牙齒 /a ₄ ts ^h i ₂ /	6.14 (3.73)	
5 (Low rising)	瓦片 /a ₅ p ^h in ₂ /	5.78 (3.30)	6.16 (3.55)
	雅典 /a ₅ tin ₂ /	6.59 (3.76)	

Abbreviation: SD, standard deviation.

DISCUSSION

A negative-valued VAT signals a hard (“glottal”) voice onset, for which the vocal folds are in contact before airflow begins. The physiological signification of the magnitude of negative values

is uninterpretable at the present time, and thus, beyond signaling a glottal attack or coup de glotte, their magnitude is likely to be of limited use to the clinician. The magnitude of positive VAT values, on the other hand, does quantify behavior which

TABLE 3.
Descriptive Statistics for Positive VAT Values (ms) of Cantonese Speakers for Production of Contour vs Level Tone Patterns

Tone Pattern	Number of Tokens	Mean (SD)	95% Confidence Interval	Median	25th–75 Percentiles
All patterns	1653	5.59 (3.69)	5.41–5.77	5.10	2.90–7.80
Contour	883	6.02 (3.69)	5.78–6.26	5.60	3.30–8.10
Women	440	5.72 (3.58)	5.39–6.06	5.50	3.10–8.10
Men	443	6.32 (3.79)	5.96–6.67	5.70	3.70–5.70
Level	770	5.10 (3.63)	4.84–5.36	4.50	2.20–7.40
Women	337	4.73 (3.46)	4.36–5.10	4.10	1.80–4.10
Men	433	5.40 (3.74)	5.04–5.74	4.70	2.60–7.60

Abbreviations: SD, standard deviation.

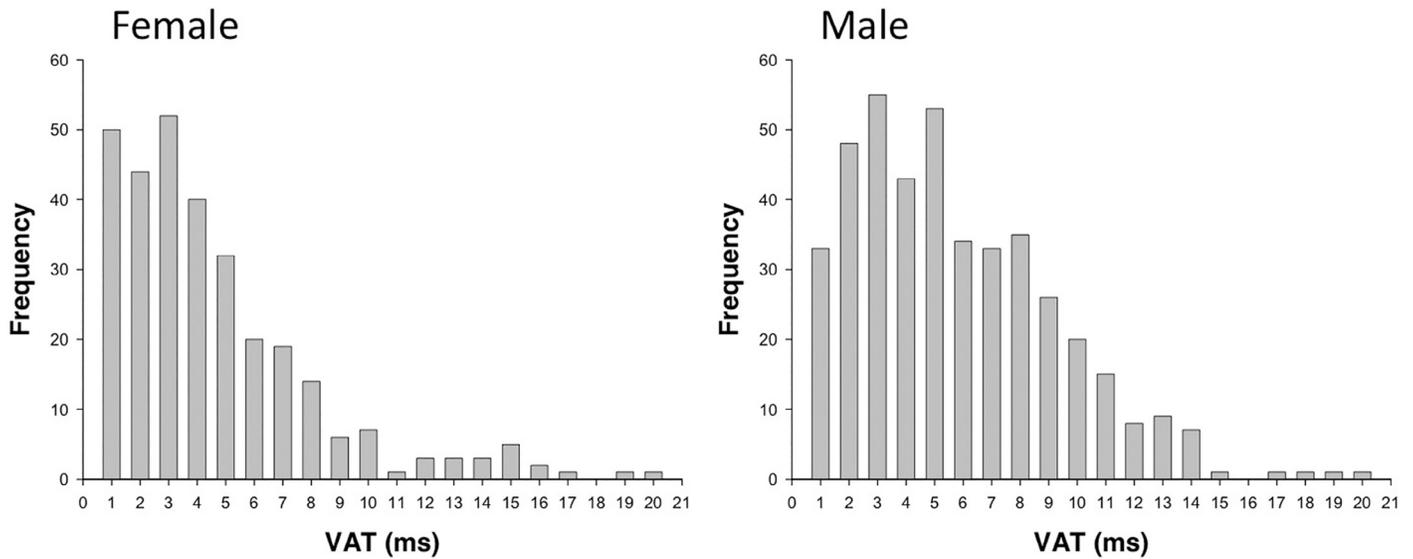


FIGURE 2. Frequency distributions of VAT for American English male and female speakers.

is physiologically relevant: the “softness” of voice onset, a quality that is generally agreed to be a legitimate focus of clinical intervention. The descriptive statistics for the set of positive-valued VATs that we present may therefore be useful.

Certain characteristics of voice initiation that were observed in the total (combined negative- and positive-valued) data sets

retain their significance in the isolated positive-valued set. In particular, among both American English and Cantonese speakers, men have longer VATs than women. Moreover, for Cantonese speakers, contour tones are associated with longer VATs than level tones. Both of these observations have plausible explanations that merit further exploration.

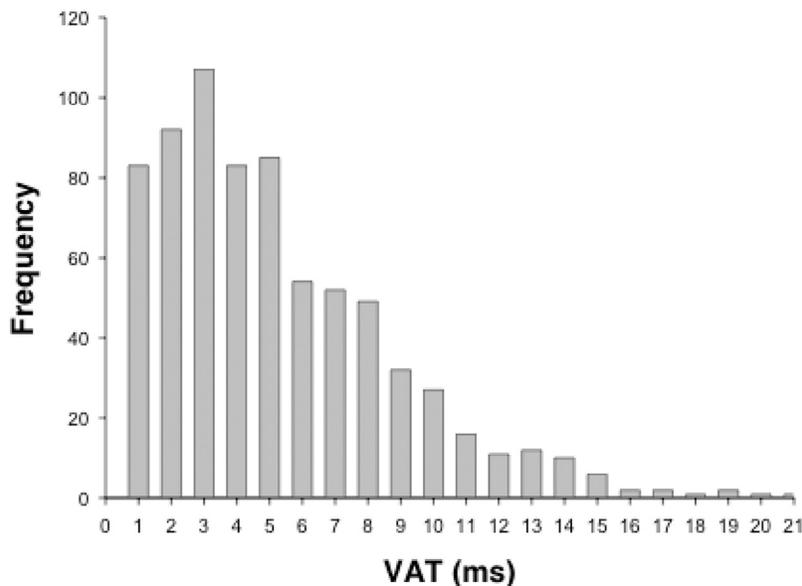


FIGURE 3. Frequency distributions of VAT for the entire data corpus.

APPENDIX

Positive VAT Values by Gender and Age

Group	Number of Tokens	Mean (SD)	95% Confidence Interval	Median	25th–75th Percentiles
Both genders					
≤24 YO	240	4.60 (3.93)	4.09–5.09	3.51	1.63–6.43
25–29	275	5.56 (3.62)	5.13–5.99	5.06	2.74–7.87
30–39	121	4.01 (2.86)	3.49–4.52	3.65	2.02–5.13
≥40	92	3.83 (2.82)	3.25–4.41	3.22	1.87–5.59
Women					
≤24 YO	139	4.08 (3.72)	3.85–5.39	3.22	1.52–4.83
25–29	90	4.63 (3.69)	1.64–2.81	3.67	1.77–6.42
30–39	26	2.23 (1.46)	3.45–4.70	2.34	0.95–3.81
≥40	49	4.08 (3.40)	3.09–5.05	3.45	1.50–6.01
Men					
≤24 YO	101	5.31 (4.12)	4.49–6.12	4.35	2.04–7.28
25–29	185	6.02 (3.50)	5.50–6.52	5.71	3.51–8.48
30–39	95	4.50 (2.96)	3.89–5.10	4.19	2.27–5.62
≥40	43	3.56 (1.95)	2.95–4.15	3.18	1.99–4.92

Abbreviations: SD, standard deviation; YO, year old.

REFERENCES

- Baken RJ, Orlikoff RF. Vocal fold adduction time estimated from glottographic signals. Presented at: the 25th Mid-Winter Meeting of the Association for Research in Otolaryngology; February 1998; St. Petersburg, FL. 29.
- Baken RJ, Orlikoff RF. Estimating vocal fold adduction time from EGG and acoustic records. In: Schutte HK, Dejonckere P, Leezenberg H, et al., eds. *Programme and Abstract Book: 24th IALP Congress*. Amsterdam: 1998:15.
- Orlikoff RF, Deliyski DD, Baken RJ, et al. Validation of a glottographic measure of vocal attack. *J Voice*. 2009;23:164–168.
- Roark RM, Watson BC, Baken RJ. A figure of merit for vocal attack time measurement. *J Voice*. 2012;26:8–11.
- Roark RM, Watson BC, Baken RJ, et al. Measures of vocal attack time for healthy young adults. *J Voice*. 2012;26:12–17.
- Ma EP-M, Baken RJ, Roark RM, et al. Effect of tones on vocal attack time (VAT) in Cantonese speakers. *J Voice*. 2012;26:670.e1–670.e6.
- Watson BC, Baken RJ, Roark RM, et al. Effect of fundamental frequency at voice onset on vocal attack time. *J Voice*. 2013;27:273–277.
- Kong JP, Zhang R. A VAT study of the four tones on Mandarin Chinese. In: Kong JP, ed. *International Symposium on Voice: Physiology, Language, and Cultures*. Beijing: 2013:12. October 9–11.
- Sangta. VAT measurement of Amdo Tibetan plosives. In: Kong JP, ed. *International Symposium on Voice: Physiology, Language, and Cultures*. Beijing: 2013:17. October 9–11.
- Zhang R, Kong JP, Baken RJ. A study on VAT of different pitch levels and vowels in Mandarin. In: Kong JP, ed. *International Symposium on Voice: Physiology, Language, and Cultures*. Beijing: 2013:30. October 9–11.
- Colton RH, Casper JK, Leonard R. *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment*. 4th ed. Philadelphia: Williams and Wilkins; 2011.
- Rees M. Harshness and glottal attack. *J Speech Hear Res*. 1958;1:344–349.
- Andrade D, Heuer R, Hockstein NE, et al. The frequency of hard glottal attacks in patients with muscle tension dysphonia, unilateral benign masses and bilateral benign masses. *J Voice*. 2000;14:240–246.
- Mathieson L. *Greene and Mathieson's The Voice and Its Disorders*. 6th ed. London: Whurr; 2001.
- Froeschels E. Hygiene of the voice. *Arch Otolaryngol Head Neck Surg*. 1943;38:122–130.
- Luchsinger R, Arnold GE. *Voice-Speech-Language: Clinical Communicology: Its Physiology and Pathology*. Belmont CA: Wadsworth; 1965.
- Harris T, Harris S, Rubin JS, et al. *The Voice Clinic Handbook*. London: Whurr; 1998.