Sensitivity of odd-harmonic amplitudes to open quotient and skewing quotient in glottal airflow (L)

Ingo R. Titze^{a)}

National Center for Voice and Speech, 136 South Main Street, Suite 320, Salt Lake City, Utah 84101

(Received 30 September 2014; revised 13 November 2014; accepted 25 November 2014)

It is well known that a half-sinusoid has no odd harmonics other than the fundamental. If glottal flow in phonation were to approximate this exact waveshape, which is generally unlikely, some misperception of pitch and loss of vowel intelligibility would occur. The sensitivity of the glottal waveshape to this special shape is explored by systematically varying two parameters, open quotient and skewing quotient. Mild asymmetry (open quotient below 0.45 or above 0.55 and/or skewing quotient greater than 2.0) equalizes the odd-even harmonic series. Singers and speakers avoid the exact symmetry by skewing the flow pulse with source-filter interaction.

© 2015 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4904539>]

[JW]

Pages: 502–504

I. INTRODUCTION

It is generally understood that pulsatile glottal airflow in phonation (speaking or singing) has a harmonic frequency spectrum with an approximate -12 dB/octave spectral slope. The time-derivative has a -6 dB/octave slope. In most textbook and tutorial descriptions of this source spectrum (Fant, 1960, p. 19; Sundberg, 1977; Stevens, 1998, p. 68; Titze, 2000, p. 130), little differentiation is made between odd and even harmonics amplitudes. Both odd and even harmonics are present in a typical airflow waveform, and they generally decay in amplitude at the same slope, even in excised larynges (Alipour *et al.*, 2012). However, there are special cases in which the odd harmonic series differs greatly from the even harmonic series. Symmetry in the rise and fall of the flow pulse (quantified by a skewing quotient Q_s) and symmetry in the duty ratio (quantified by the open quotient Q_o) produce these special cases. Odd and even harmonic series must then be treated separately. This letter clarifies the conditions under which such a distinction would need to be made and why it is generally not a major concern.

II. GLOTTAL PULSE SHAPE AND ITS MAGNITUDE SPECTRUM

A typical caricature of a glottal airflow shape is a half-sinusoid, shown in Fig. 1(a). It is conceptually and mathematically simple, but illustrates the problem. This pulse shape has no odd harmonics other than the fundamental [Fig. 1(b)], as a simple Fourier series calculation shows (see, for example, Brandwood, 2011). In terms of voice science terminology, it has an open quotient Q_o equal to 0.5 (also known as duty ratio) and a skewing quotient Q_s equal to 1.0 (rise time of pulse equals fall time of pulse). The non-existence of harmonics 3, 5, 7, ... is due to phase cancellations of frequency components when perfect symmetry exists between rise/fall duration and flow/no-flow duration.

If such a symmetric flow pulse were to be used in natural or synthetic phonation, some obvious acoustic and perceptual effects would occur. From a production point of view, the odd harmonics would not be available to excite formants. Consider an example where a vowel has a 600 Hz first formant. With a 200 Hz fundamental, harmonic energy above f_o would be available at 400 Hz (second harmonic) and at 800 Hz (fourth harmonic). The critical 600 Hz third harmonic at or near F_1 would be non-existent. Since vowels are perceived by the combination of F_1 and F_2 , lack of excitation of F_1 would make it difficult to judge the correct vowel. As a second example, a singer producing A_4 (440 Hz fundamental) would have little excitation of a second formant around 1500 Hz. The critical third harmonic (1320 Hz in this

FIG. 1. (Color online) (a) Half-sinusoid as a glottal flow waveform, (b) Fourier series spectrum.

^{a)} Author to whom correspondence should be addressed. Also at: Department of Communication Sciences and Disorders, the University of Iowa, Iowa City, IA 52242. Electronic mail: ingo-titze@uiowa.edu

case) would be missing. It has been shown that classical singers rely heavily on third harmonic energy at pitches near A_4 (Miller, 2008), both for timbre and for vowel distinction.

Perceptually, there would also be a problem because the second harmonic is likely to determine the pitch. This is explained in the time domain by a double excitation of the vocal tract, one at glottal opening and one at glottal closing. Since the two excitations are equal in magnitude, even though opposite in phase in terms of their time derivatives, pitch perception is based on the half-period rather than the whole period. Pitch doubling can also be explained spectrally by the fact that only an even-harmonic series exists. The auditory system extrapolates the harmonic series downward (by autocorrelation) to an apparent $2f_o$ fundamental rather than the f_o fundamental (Licklider, 1951). Thus, the perceived pitch is an octave higher, and f_o becomes a subharmonic.

Speakers and singers have no doubt learned how to avoid this problem by staying clear of perfect symmetries in the flow waveform (Holmberg *et al.*, 1988). Women generally increase the open quotient to 0.6 or greater, while men often stay between 0.55 and 0.6 in speech, but also produce open quotients around 0.4 or less, especially in singing (Miller, 2008). More importantly, both genders increase the skewing quotient (Holmberg *et al.*, 1988), either by nonlinear source-filter coupling (Titze, 2004) or by accelerating the

closure time of the glottis (Granqvist, 2003). Figure 2 shows a flow waveform and its spectrum with an open quotient of 0.55 and skewing quotient of 2.0. Note the presence of all harmonics with an approximate -10 dB/octave slope (a little less than typical because the flow-shutoff is quite abrupt).

III. SIMULATION OF HARMONIC AMPLITUDE VARIATION NEAR WAVEFORM SYMMETRY

A simple formula was used to generate glottal flow waveforms with a variety of open quotients and skewing quotients,

$$u(t) = \max[0, e^{\alpha t}(\beta + \sin \omega t)], \quad (1)$$

where α and β are parameters that produce variations in Q_o and Q_s . The closing and opening events on the waveform (due to truncation at zero flow) were calculated as

$$t_o = (T/2\pi) \sin^{-1}(-\beta), \quad (2)$$

$$t_c = T/2 - t_o, \quad (3)$$

where T is the period ($2\pi/\omega$). The location of the peak was detected numerically as t_m . The open quotient and skewing quotient were then computed as

$$Q_o = (t_c - t_o)/T, \quad (4)$$

$$Q_s = (t_m - t_o)/(t_c - t_o). \quad (5)$$

The time of opening, t_o , is negative if the sinusoid is raised, as in Fig. 2(a). The exponential term $e^{\alpha t}$ in Eq. (1) was applied only in the open portion, $t > t_c$. Once an exact cycle of the waveform was simulated with 1000 points, a MATLAB (The Mathworks, 2014) Fourier series routine was used to calculate the amplitudes A_n of the first 6 harmonics ($n=1,2,\dots,6$) in the spectrum. Approximately 100 waveform variations, with Q_o ranging between 0.4 and 0.6 and Q_s ranging between 0.4 and 3.5, were used to plot A_n/A_1 as a function of open quotient and skewing quotient.

Figure 3(a) shows how third, fifth, and seventh harmonic amplitudes drop sharply when $Q_o=0.5$ is approached from either direction while Q_s is constant at the symmetry value 1.0. The amplitude ratio A_n/A_1 is expressed in dB. Solid lines show the relative amplitude of the odd harmonics and dashed lines show the relative amplitudes of the even harmonics. Note the lesser sensitivity of the even harmonics to Q_o when the critical value $Q_o=0.5$ is approached. In fact, harmonic amplitudes 4 and 6 rise while 3, 5, and 7 fall sharply.

Figure 3(b) shows a similar set of curves for variation of Q_s around its critical value 1.0, with Q_o held constant at 0.5. Again, even harmonics show a lesser sensitivity than odd harmonics to variation in Q_s , but they also fall slightly.

When either Q_o or Q_s is kept away from its critical symmetry value, the sharp loss of even harmonics never appears. Figure 3(c) shows variation with Q_o when Q_s is held constant at 2.0, while Fig. 3(d) shows variation with Q_s when Q_o is held constant at 0.55. There are still minor dips in harmonic amplitudes, but the overall spectral slope is on the order of -10 dB/octave. Figure 2 (shown earlier) is the waveform

FIG. 2. (Color online) (a) Skewed sinusoid slightly raised to produce an open quotient of 0.55, (b) corresponding Fourier series spectrum.

FIG. 3. (Color online) Sensitivity of harmonic amplitudes to open quotient near 0.5 (left panels) and skewing quotient near 1.0 (right panels).

and spectrum that corresponds to the value $Q_o = 0.55$ in Fig. 3(c) and to the value 2 in Fig. 3(d). With reference to Eq. (1), the value of α was 745 and the value of β was 0.16 for this case.

IV. SUMMARY

A small amount of asymmetry in the glottal airflow waveform is essential for a balanced odd-even harmonic spectrum. A simple Fourier series analysis demonstrated that open quotients below 0.45 and above 0.55 nearly equalize the odd-harmonics with the even ones without much waveform skewing. Waveform skewing alone (with a 0.5 open quotient) nearly equalizes the spectrum when $Q_s \geq 2.0$. When both asymmetries are utilized together, values of Q_o above 0.55 or below 0.45 (with $Q_s \geq 2$) are sufficient to eliminate all odd-even peaks and valleys in the spectrum.

In terms of a practical application, a case of interest may be when the glottal area skews to the left and vocal tract inductance skews the flow pulse to the right. The two skewing factors could cancel each other, leaving the flow pulse nearly symmetric. Then, if $Q_o \approx 0.5$, some un-clarity in pitch and vowel may arise. Future studies will explore this possibility.

Limitations of this brief letter are the lack of local roundness at opening and closing with a simple truncated sinusoidal model. In natural phonation, the flow is usually not cut off sharply, or initiated with a step-wise slope change. This limitation would affect mainly the higher harmonics and broader peaks and valleys in the spectral envelope (Ananthapadmanabha and Fant, 1982). For this reason, this study was focused only on the lower six harmonics. Future

studies can address oversimplification of the flow model with sophisticated fluid-structure interaction models.

ACKNOWLEDGMENT

Support for this research comes from Grant No. 1 R01 DC012045-01A1 by the National Institute on Deafness and Other Communication Disorders.

- Alipour, F., Scherer, R. C., and Finnegan, E. (2012). "Measures of spectral slope using an excised larynx model," *J. Voice* **26**(4), 403–411.
- Ananthapadmanabha, T. V., and Fant, G. (1982). "Calculation of true glottal flow and its components," *Speech Commun.* **1**(3), 167–184.
- Brandwood, D. (2011). *Fourier Transforms in Radar and Signal Processing*, 2nd ed. (Artech House, Norwood, MA), pp. 88–91.
- Fant, G. (1960). *The Acoustic Theory of Speech Production* (Moulton, The Hague), p. 19.
- Granqvist, S., Hertegard, S., Larsson, H., and Sundberg, J. (2003). "Simultaneous analysis of vocal fold vibration and transglottal airflow: Exploring a new experimental setup," *J. Voice* **17**(3), 319–330.
- Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (1988). "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice," *J. Acoust. Soc. Am.* **84**(2), 511–529.
- Licklider, J. C. R. (1951). "A duplex theory of pitch perception," *J. Acoust. Soc. Am.* **23**(1), 147–147.
- Miller, D. G. (2008). *Resonance in Singing: Voice Building Through Acoustic Feedback* (Inside View, Princeton, NJ), Chap. 1.
- Stevens, K. (1998). *Acoustic Phonetics* (MIT Press, Cambridge, MA), p. 68.
- Sundberg, J. (1977). "The acoustics of the singing voice," *Sci. Am.* **236**, 82–91.
- The Mathworks, 2014. MATLAB, version 8.3.0.532 (R2014a).
- Titze, I. R. (2000). *Principles of Voice Production*, 2nd printing (National Center for Voice and Speech, Salt Lake City, UT), p. 130.
- Titze, I. R. (2004). "Theory of glottal airflow and source-filter interaction in speaking and singing," *Acta Acust. Acust.* **90**(4), 641–648.