Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization

Ingo R. Titze
National Center for Voice and Speech, 136 South Main Street, Suite 320, Salt Lake City, Utah 84101, USA

Ronald J. Baken
Department of Otolaryngology, New York Eye and Ear Infirmary of Mount Sinai, 310 East 14th Street, New York, New York 10003, USA

Kenneth W. Bozeman
Lawrence University Conservatory of Music, 711 East Boldt Way, Appleton, Wisconsin 54911, USA

Svante Granqvist
Royal Institute of Technology (KTH), School of Technology and Health (STH), Basic Science and Biomedicine, Campus Hamine, SE-136 40 Handen, Sweden

Nathalie Henrich
GIPSA-lab, Département Parole et Cognition, Domaine Universitaire, 11 rue des Mathématiques, BP 46 38402 Saint Martin d’Hères cedex, France

Christian T. Herbst
Laboratory of Biophysics, Department of Experimental Physics, Palacký University Olomouc, Czech Republic

David M. Howard
Department of Electronics, University of York, Heslington, York, United Kingdom

Eric J. Hunter
Communicative Sciences and Disorders, Michigan State University, 1026 Red Cedar Road, East Lansing, Michigan 48824, USA

Dean Kaelin
Dean Kaelin Vocal Studio, 2539 East 4430 Street, Holladay, Utah 84124, USA

Raymond D. Kent
The Waisman Center, University of Wisconsin–Madison, 1500 Highland Avenue, Madison, Wisconsin 53705, USA

Jody Kreiman
Head and Neck Surgery, University of California–Los Angeles School of Medicine, 1000 Veteran Avenue, Los Angeles, California 90095, USA

Malte Kob
Musikalische Akustik and Theorie der Musikübertragung Erich-Thienhaus-Institut Hochschule für Musik Detmold, Neustadt 22, 32756 Detmold, Germany

Anders Lüfqvišt
Haskins Laboratories, Yale University, 300 George Street, Suite 900, New Haven, Connecticut 06511, USA

I. INTRODUCTION

The study of vocalization brings together a long history of voice terminology from acoustics, linguistics, phonetics, speech pathology, laryngology, music, theater, biology, and speech technology. One challenge is to maintain consistency in symbolic representation of key variables used for resonant frequencies of the airways and the frequencies produced by sound sources. Scientists who use mathematical notation are encouraged to use single letters with subscripts for algebraic clarity (Cohen and Giacomo, 1987), whereas clinicians often prefer multiple-letter abbreviations without subscripts for ease of written and spoken communication. For example, the symbol $f_o$ as the fundamental frequency of oscillation of the vocal folds has been used in thousands of publications, both with upper case and lower case letters, and both with subscript and no subscript. If capitalized, the symbol is not clearly dissociated from formant frequencies $F_1, F_2, \ldots, F_n$. The subscript of fundamental frequency, if written as a “zero,” does not indicate a first harmonic, but rather a meaningless “zero” harmonic. If written as an alphabetic “o,” it can stand for “oscillation,” which is more meaningful. Some investigators have expressed a desire to abandon $f_o$ altogether, but such a dramatic shift would render a disservice to volumes of historic literature.

A new attempt at clarity has arisen, in which authors are beginning to identify harmonics of the sound source as $H_1, H_2, H_3, \ldots, H_n$. The problem with this notation is that the symbol “H” can refer to either the frequency or...
the amplitude of a harmonic. For example, when an $H_2/H_1$ ratio is computed for a spectral balance measure, an amplitude ratio is intended, not a frequency ratio. In other cases, when authors refer to $H_i/F_1$ or $H_1-F_1$ relations (especially in singing), they are talking about frequency ratios or differences. Thus, confusion has not been eliminated by introducing the $H$ symbol. We will show that its use is not necessary.

With regard to airway resonances, historical precedence and current usage of terminology are also slightly at odds. Joe Wolfe and colleagues suggest that the symbol $R$ be used to stand separate from the symbol $F$ for formant (Wolfe, 2014). The distinction is being made because a formant was originally defined as a peak in the output spectrum envelope radiated from the mouth (Hermann, 1894, 1895; Russell, 1929; Fant, 1960, p. 20). A similar definition appears in the current ASA standard of acoustic terminology (Acoustical Society of America, 2004), namely, that a formant is “a range of frequencies in which there is absolute or relative maximum in the sound spectrum. The frequency at the maximum is the formant frequency.” As such, a formant involves both the source and the filter. However, as speech analysis and synthesis have progressed in a half century, the definition has not been universally maintained. Fant (1960, pp. 20, 53) defined a formant frequency as the frequency at the maximum sound spectrum. The frequency at the maximum is the formant frequency.”

In analogy (Acoustical Society of America, 2004), namely, that a formant is “a range of frequencies in which there is absolute or relative maximum in the sound spectrum. The frequency at the maximum is the formant frequency.”

**III. A REASONABLE COMPROMISE FOR WRITTEN AND SPOKEN COMMUNICATION**

The present authors suggest the following notation to be used. Harmonic frequencies should be written as multiples of $f_0$, namely, $nf_0$. The letters $H$ and $P$ are not needed. Harmonic amplitudes should be written as $A_n$. The letter $R$ can be used as a word abbreviation for resonance, but two subscripts should be assigned to specify the resonance properties (see Table 1). The letter $F$ can be used as a word abbreviation for formant, but if only a single subscript is assigned, it must refer only to formant frequency (Table 1). Level and bandwidth of the formant should carry two subscripts to be distinguishable from those of resonances.

The harmonic notation is tied to the Fourier series expansion of an acoustic pressure

$$P(t) = A_n \sin(2\pi f_0 t + \phi_n).$$

(1)

The parentheses (1) for the first harmonic in Table I is generally not written or spoken, but always implied. This is important to point out so that the harmonic integer series is complete. The subscript for fundamental frequencies is an “$o$,” not a zero to emphasize “oscillation.” The letter $L$ for resonance level is used because we usually express relative formant peaks in dB. $L_i$ is generally assumed to be 0 dB, thereby using the amplitude of the fundamental as the reference amplitude.

For harmonic source frequencies, the symbols $f_1$, $f_2$, $f_3$, …, can be used without reference to any harmonic index or $f_0$. It is then important to speak “source frequency $f_1$,” “source frequency $f_2$,” etc. For subharmonic frequencies, $nf_0/i$ will identify the period-i subharmonic series. Consensus on symbols for amplitudes and levels of subharmonics has not yet been discussed.

A little training will be needed for people to say, “two $f_0$,” “three $f_0$,” and so on, for harmonic frequencies. The beauty of that training, however, is that the harmonic relationship with $f_0$ will always be kept in mind. Also, for subharmonics, “one-half $f_0$” or “one-third $f_0$” is an easy extension. Speaking the extra letters in $f_i/k_1$, …, $f_i/k_n$, $f_i/L_1$, …, $f_i/L_m$, and $f_i/B_1$, …, $f_i/B_m$ will also be an immediate reminder of “resonance” rather than “source” or “formant.”

With this nomenclature, an $nf_0/k_m$ ratio or an $nf_0 - f_m$ difference describes a source-resonance frequency relation. Likewise, an $nf_0/A_n$ or an $nf_0 - F_m$ describes a source-formant relation. The ratio $A_n/F_m$ describes a harmonic relation (linear scale), $L_m - L_n$ describes a logarithmic (dB) source-harmonic relation, $L_m - L_{k_m}$ describes a logarithmic source-resonance amplitude relation in dB, and $L_{f_m} - L_{f_1}$ describes a formant level relation in dB.

The classical equation for the resonance frequency of an idealized, uniform closed-open tube

$$f_m = \left(\frac{m-1}{4L}\right)^2 \frac{c}{4L}$$

remains a benchmark of comparison between resonances and formants, inasmuch as no resonance coupling occurs to other airway structures. For this idealized airway structure, as well as for closed glottis vowels, $f_{k_m} = F_m$.

Syntactic notation for subglottal resonances has not been addressed here, nor for resonances of side branches of the airways (nasal tract, sinuses). Some precedence exists for labeling subglottal resonance and formant frequencies with a “prime” superscript (e.g., $f_0/k_1$ and $F_1$). Subscripts “sg” have

| TABLE I: Harmonic, resonance, and formant symbols for quantitative relations. |
|-------------|-------------|-------------|-------------|-------------|-------------|
| **Harmonics** | **Resonances** | **Formants** |
| **Frequency (Hz)** | **Amplitude (Pa)** | **Level (dB)** | **Frequency (Hz)** | **Level of peak (dB)** | **Bandwidth (Hz)** | **Frequency (Hz)** | **Level of peak (dB)** | **Bandwidth (Hz)** |
| $(1)f_o$ | $A_1$ | $L_1$ | $f_{R1}$ | $L_{R1}$ | $B_{R1}$ | $F_1$ | $L_{F1}$ | $B_{F1}$ |
| $2f_o$ | $A_2$ | $L_2$ | $f_{R2}$ | $L_{R2}$ | $B_{R2}$ | $F_2$ | $L_{F2}$ | $B_{F2}$ |
| $3f_o$ | $A_3$ | $L_3$ | $f_{R3}$ | $L_{R3}$ | $B_{R3}$ | $F_3$ | $L_{F3}$ | $B_{F3}$ |
| … | … | … | … | … | … | … | … | … |
| $nf_o$ | $A_n$ | $L_n$ | $f_{Rn}$ | $L_{Rn}$ | $B_{Rn}$ | $F_m$ | $L_{Fm}$ | $B_{Fm}$ |


Titze et al.: Symbols for harmonics, resonances, and formants
also been used in presentations, but these additional subscripts are unappealing due to overuse of subscripts.

IV. CONCLUSION

A compromise has been reached between preserving historical nomenclature and symbols for source harmonics, vocal tract resonances, and formants while providing clarity for speaking the symbols and assigning numbers and units to them. Little re-training is needed. One extra subscript is added for resonance characteristics and for formant levels and bandwidths. The harmonic number is explicitly written and spoken together with the fundamental frequency. Authors who are heavily invested in formant frequency analysis are encouraged to be as clear as possible about the relation between a peak in the output spectrum and a presumed resonance of the vocal tract. Likewise, those who describe airway resonances are encouraged to be as clear as possible about their manifestation in the output spectrum. It is important to clarify what the boundaries of the resonator are. In some cases, only the supraglottal vocal tract is described as a resonator (with the glottis closed), in other cases the resonance includes the interaction with the glottis, and in yet other cases resonance includes the entire airway, lungs to lips. As benchmarks are being developed for characteristic frequencies and bandwidths of vowels and consonants, across species, genders, age and cultures, it becomes ever more important to define the exact geometry and boundary conditions of the portion of the airway under investigation.

ACKNOWLEDGMENT

Support for this project was provided by NIH NIDCD Grant No. 1 R01 DC012045-01.
