Brass Tone Synthesis by Spectrum Evolution Matching with Nonlinear Functions

by

James Beauchamp
University of Illinois
at Urbana-Champaign

There has been ample evidence from analytical studies done in recent years that brass harmonic spectra (and probably those of other wind instruments) are somehow related to spectra which can be produced by the nonlinear distortion of a variable amplitude sine wave. It is the purpose of this paper to review this evidence and present some results of attempts to synthesize brass tones using this technique.

Nonlinear Synthesis Model

As shown in Figure 1 a sine wave of amplitude A is applied to a nonlinear processor with describing function F. The index A is controlled by an envelope generator. The nonlinear processor acts to distort the sine wave into a waveform containing several harmonics whose amplitudes depend on A(t) and F. The harmonic amplitudes are further altered by the filter having response H(f). The output amplitude may be further controlled by multiplication by $A_1(t)$, which is produced by a second envelope generator.

Compiler's note: A revised version of this paper will appear in *Computer Music Journal*, Vol. 3, No. 2. (1979)

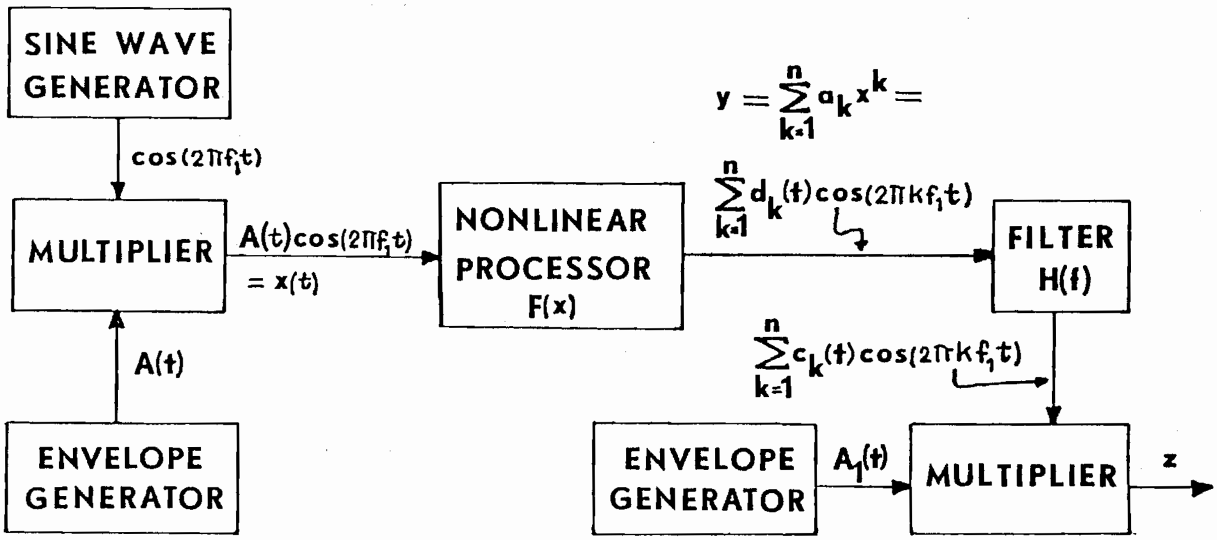

Figure 1

Nonlinear Synthesis Model

The essential feature of the nonlinear synthesis model is that as the index A increases, the evolution of the harmonic spectrum of the output is governed by the form of the function F. If we assume that F is "smooth," it can be well approximated by an nth degree polynomial

$$F(x) = a_0 + a_1 x + a_2 x^2 + . . . + a_n x^n. \tag{1a}$$

x is a sinusoid given by

$$x = A \cos(\Theta), \Theta = 2\pi f_1 t, \tag{1b}$$

where A is its amplitude and f_1 is its (fundamental) frequency. The action of F is such as to produce a harmonic series at the nonlinear processor output:

$$y = 1/2\ d_0 + d_1 \cos(\Theta) + d_2 \cos(2\Theta) + . . . + d_n \cos(n\Theta) \tag{2}$$

The harmonic amplitudes of the nonlinear processor output are the d_k coefficients and may be derived directly from the a_k coefficients and the amplitude A according to the formula

$$d_k = 2 \sum_{j=0}^{[(n-k)/2]} \frac{(k+2j)!}{j!(k+j)!} \frac{A}{2}^{k+2j} a_{k+2j}. \quad (3)$$

A detailed derivation of this formula is given by Suen (1970).

The filter acts to selectively modify the harmonic amplitudes according to its response characteristic H(f) which is "sampled" at the frequencies 0, f_1, $2f_1$, $3f_1$, . . . Thus, the output of the synthesis model (assuming H(0) = 0) is

$$z = A_1 \{ c_1\cos(\Theta) + c_2\cos(2\Theta) + . . . + c_n(n\Theta)\} \quad (4a)$$

where $c_k = H(kf_1).d_k$ and $\Theta = 2\pi f_1 t$. (4b)

For the purposes of most subsequent discussions we will assume that $A_1 = 1$.

The nonlinear processor provides a spectrum which changes with the intensity of the tone, and, in fact, is a one-to-one function of the intensity. One way to depict the way a spectrum evolves is by using "nonlinear interharmonic relationships" (NIHR). NIHR can be given as a set of plots of c_k vs. c_1 expressed in decibels. (For further discussion of NIHR see Beauchamp (1975).)

We define

$$C_k = 20 \log_{10}(c_k) \quad (5a)$$

$$D_k = 20 \log_{10}(d_k) \quad (5b)$$

$$H_{db} = 20 \log_{10}(H) \quad (5c)$$

as the logarithmic versions of c, d, and H, respectively, Equation 4b may then be converted to

$$C_k = H_{db}(kf_1) + D_k. \tag{6}$$

The NIHR curves for the output (z) are plots of C_k vs. C_1; at the output of the nonlinear processor (y) they are plots of D_k vs. D_1. It may be seen from Equation 6 that the *shapes* of the corresponding NIHR curves for z and y are identical; only their relative positions are different.

One important feature of nonlinear processes is exhibited at *low amplitudes*, i.e., for small values of A. In this case Equation 3 reduces to

$$d_k \doteq 2\, a_k \left(\frac{A}{2}\right)^k \tag{7a}$$

with the special case

$$d_1 \doteq a_1\, A. \tag{7b}$$

Solving Equation 7b for A and substituting the result into Equation 7a yields the interharmonic relationship

$$d_k + 2\, a_k \left(\frac{d_1}{2a_1}\right)^k \tag{8a}$$

with its decibel equivalent

$$D_k = kD_1 + 20 \log_{10} [2a_k/(2a_1)^k]. \tag{8b}$$

An important conclusion to be drawn from Equation 8b is that for low amplitudes the slope of the kth harmonic NIHR curve is equal to k. Figure 2 illustrates the low amplitude NIHR curves for the case $y = x + x^2 + x^3 + \ldots$ Note that as the harmonic number increases the curves become steeper and shift to the right. This is a general principle of low amplitude nonlinear behavior.

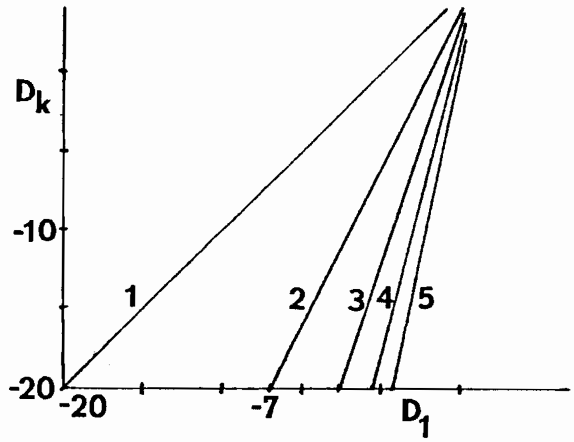

Figure 2

Low level Nonlinear Interharmonic Relationship (NIHR) curves for the function $y = x + x^2 + x^3 + \ldots$

The polynomial coefficients a_k can be derived from harmonic amplitudes d_k by means of a formula which is the inverse of Equation 3. If the d_k's vary with time, one specific time must be chosen for the calculation. Also, A_o is arbitrary and may be chosen to be equal to 1 without loss of generality. The formula is:

$$a_k = \frac{1}{2} \left(\frac{2}{A_o}\right)^k \sum_{j=0}^{[(n-k)/2]} \frac{(-1)^j (k+2j)(k+j-1)!}{k! \quad j!} \, d_{k+2j} \qquad (9)$$

Unfortunately, when the signal y(t) is resynthesized from Equations 2 and 3, the resulting spectrum d_k is guaranteed to match the original only for the d_k vector used in the calculation of Equation 9, i.e., for $A = A_o$. If the synthesized spectrum matches the original for other values of A, either this is a coincidence or the acoustics have dictated such a result.

The NIHR curves can be used to visually estimate the goodness of match between the synthesized spectra and the original ones over a range of A values. It is expected that matching the d_k's at some intermediate intensity would give the best overall match.

Brass Tone Analysis Results

Backus and Hundley (1971) outlined the mechanism for production of harmonics in trumpet tones. Their analysis was based on measurements of the nonlinear relationship between the acoustic impedance and size of the lip opening in the mouthpiece. The theory derived allowed a prediction of the shapes of trumpet mouthpiece pressure waveforms.

I have made some measurements of simultaneous trombone mouthpiece and output signals. Both types of waveforms for tones played pp, mf, and ff at 117 Hz are shown in Figure 3a (1-6). Note the tremendous range of output pressure. This was measured to be 67 dB to 103 dB SPL at the output vs. 152 to 169 dB in the mouthpiece for tones played between pp and ff. This fact may be explained as follows:

The trombone pipe acts as a high pass filter, severely attenuating the low frequencies while accentuating the high frequencies above 800 Hz. As the tone builds up in amplitude, the upper partials of the mouthpiece signal increase in strength much more rapidly than do the lower partials. Since the pipe-filter responds primarily to these upper partials, the output dynamic range is a magnified version of the input.

Spectra for six waveform cases (pp, mf, ff mouthpiece and output) are shown in Figure 3b (1-6). The two sets are related via the high pass characteristic shown in Figure 3c. (Measurement difficulties allow me only to give an approximate curve here.)

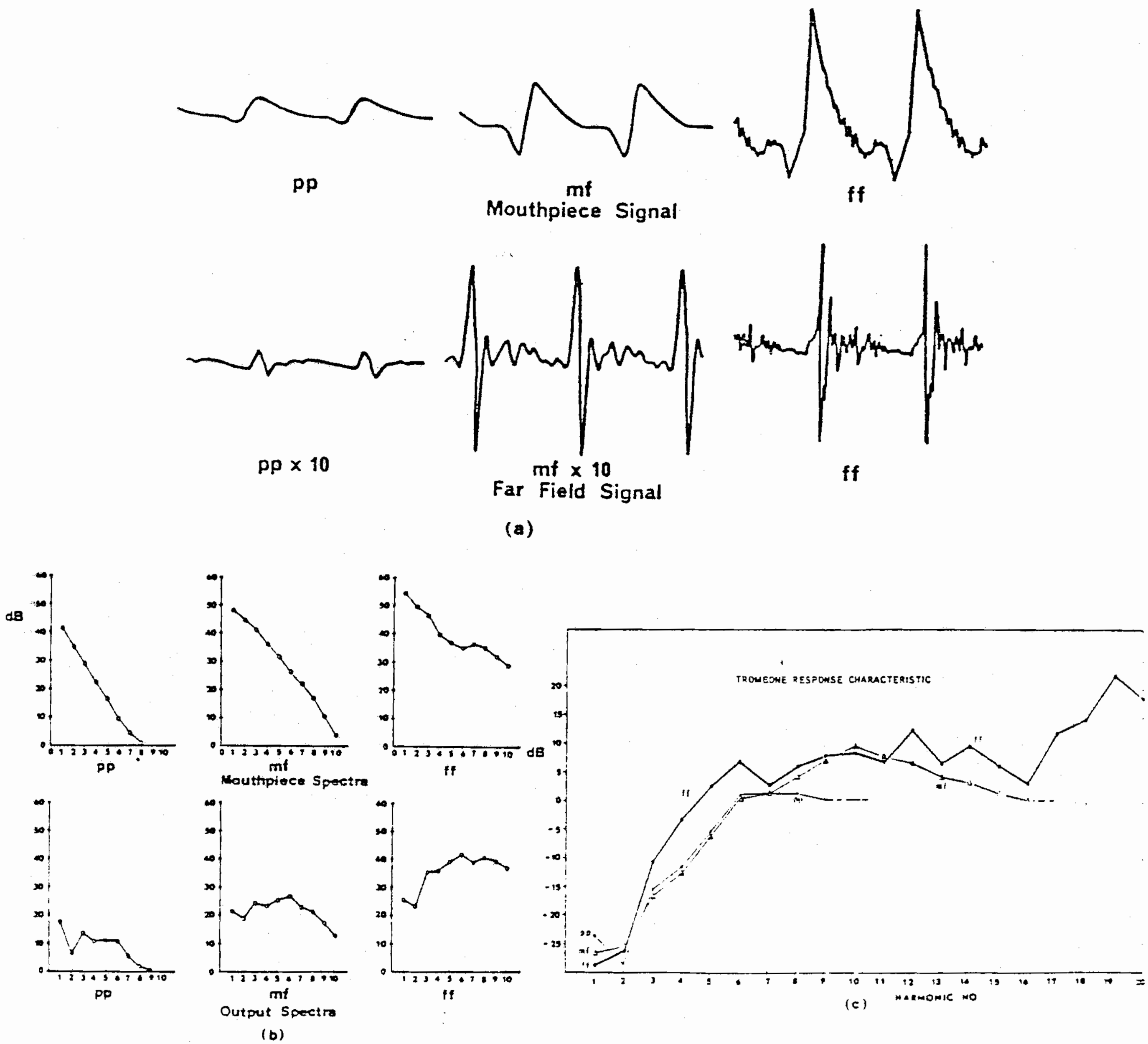

Figure 3

Waveforms and corresponding steady-state spectra for three 117 Hz trombone tones a) Waveforms (as labeled). Note that there is phase distortion due to tape recorder. b) Spectra (as labeled). c) Transmission response characteristic.

At this point I would like to make the assumption that the mouthpiece waveform corresponds to the signal x in the synthesis model of Figure 1 and that the output waveform corresponds to the signal y of the same figure. Arbitrarily taking $A_o = 1$ for the mf case we can compute the polynomial coefficients using Equation 9 and use Equation 1a to plot the nonlinear processor function shown in Figure 4a. Using this function to distort a sine wave of unit amplitude will produce a spectrum exactly matching the one shown in Figure 3b(2). Interestingly, the function shape is remarkably similar to a diode I vs. V curve!

To be effective the function should also work for the pp and ff cases. Near-optimum values of A for these two cases were determined by trial and error to be 0.87 and 1.13, and the resulting spectra (together with the originals) are plotted in Figure 4b(1-3). We see that the computed spectrum of x closely matches the original for the pp case but not so well for the ff case.

I have synthesized tones with a computer using this model. The tones

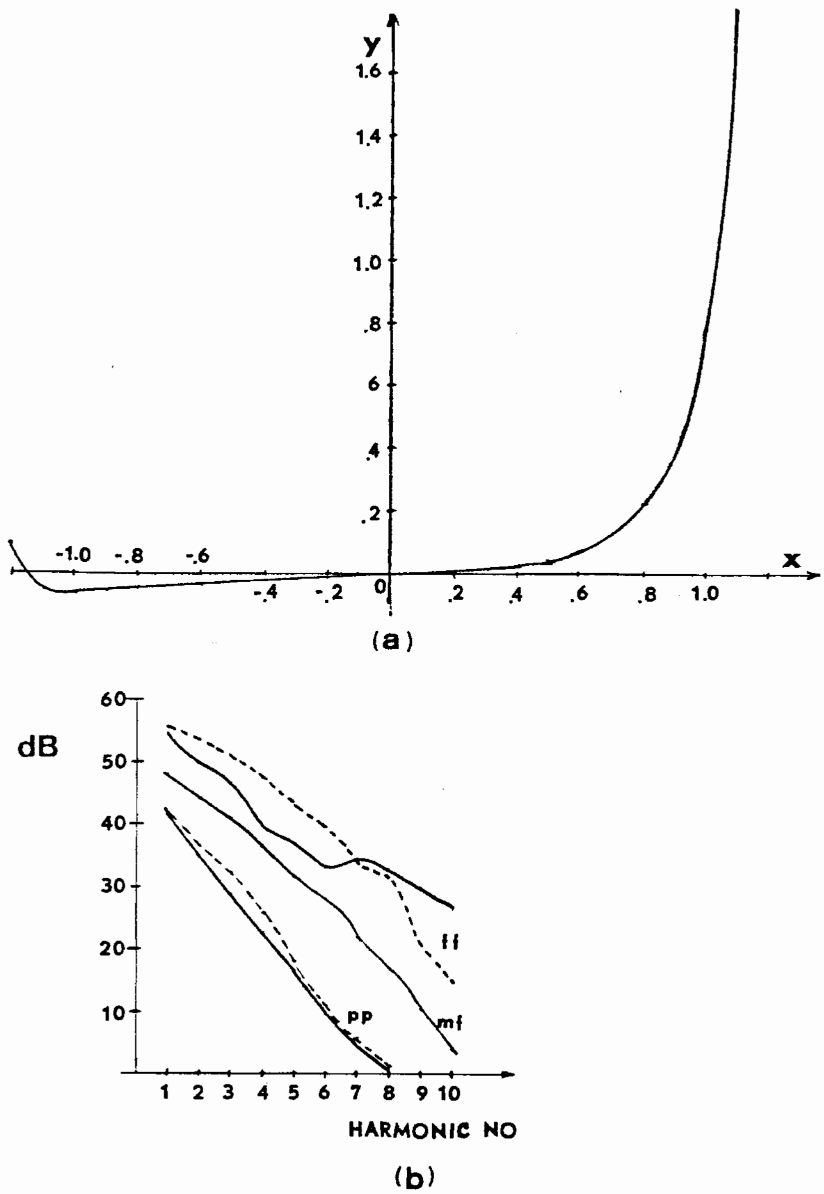

Figure 4

Nonlinear processor characteristic and corresponding synthesized mouthpiece spectra for trombone tones. a) y vs. x characteristic. b) Synthesized (--) and original (-) mouthpiece spectra (as labeled).

resembled those of a trombone but were somehow lacking in quality. One problem was that since the measurements were steady-state, I could only guess at reasonable envelope functions to use.

Benade has published the spectra of trumpet mouthpiece pressures produced by a professional musician and has found them to be very repeatable (1976a). He also gives a transmission characteristic which can be used to calculate the output spectrum (1976b), and the mouthpiece spectra are plotted in the form of NIHR curves (1976c). The salient feature of these curves is that for low levels the curves obey "Worman's Law," a formula derived by Walter Worman (1971) which can be written as follows:

$$D_k = k(D_1 - D_{1o_k}) \qquad D_1 < 30 \text{ dB}^* \tag{10}$$

Note that this is precisely the form of Equation 8b, which is the low level interharmonic relationship for the output of a generalized smooth nonlinear processor. From Benade's data for a 233 Hz trumpet tone $D_{1o_0} = 0$, $D_{1o_1} = 18$, $D_{1o_2} = 23$, and $D_{1o_3} = 26$ dB*. A Bode plot analysis of his transmission function shows that it can approximated by

$$H_{dB} = 10 \log_{10}\left(\frac{1 + (300/f)^4}{1 + (1500/f)^4}\right) \tag{11}$$

which resembles the frequency response of a treble boost circuit.

The other interesting aspect of the NIHR curves obtained by Benade is

* These values of dB are not absolute; they are relative to an arbitrary reference.

that for values of D_1 greater than 36 dB, the NIHR curves are approximately parallel. I. e., $D_k \doteq D_1 - 2k$, $D_1 > 36$ dB. In between $D_1 =$ 30 and $D_1 = 36$ there is a transition region. It is as if above a certain point the spectrum "freezes" into a certain position, and it only increases in overall amplitude. This can be effected by the model of Figure 1 by increasing A_1 while keeping A constant.

Analysis and Synthesis of Cornet Tones

I have measured the time-variant output spectrum of several cornet tones using a computer and plotted the resulting NIHR characteristics. (See Beauchamp (1975).) Figure 5 shows the harmonic amplitude curves for a single 350 Hz mf cornet tone. Figure 6 shows the NIHR's for the same tone with the attack portion (the first 0.1 sec) omitted. In 1975 I derived approximate, smooth curves to fit the NIHRcharacteristics, and the results indicated crossover slopes less than those given by Worman's Law. However, the somewhat murky nature of the data below 26 dB contributed to an ambiguity of the crossover slopes. The dashed lines in Figure 5 indicate slopes of 1, 2, 3, ..., 9 for the corresponding harmonics, and show that Worman's Law is, indeed, plausible for $C_1 < 24$ dB. For $C_1 > 26$ dB, we again note that the NIHR curves are parallel, as in the case of Benade's trumpet data.

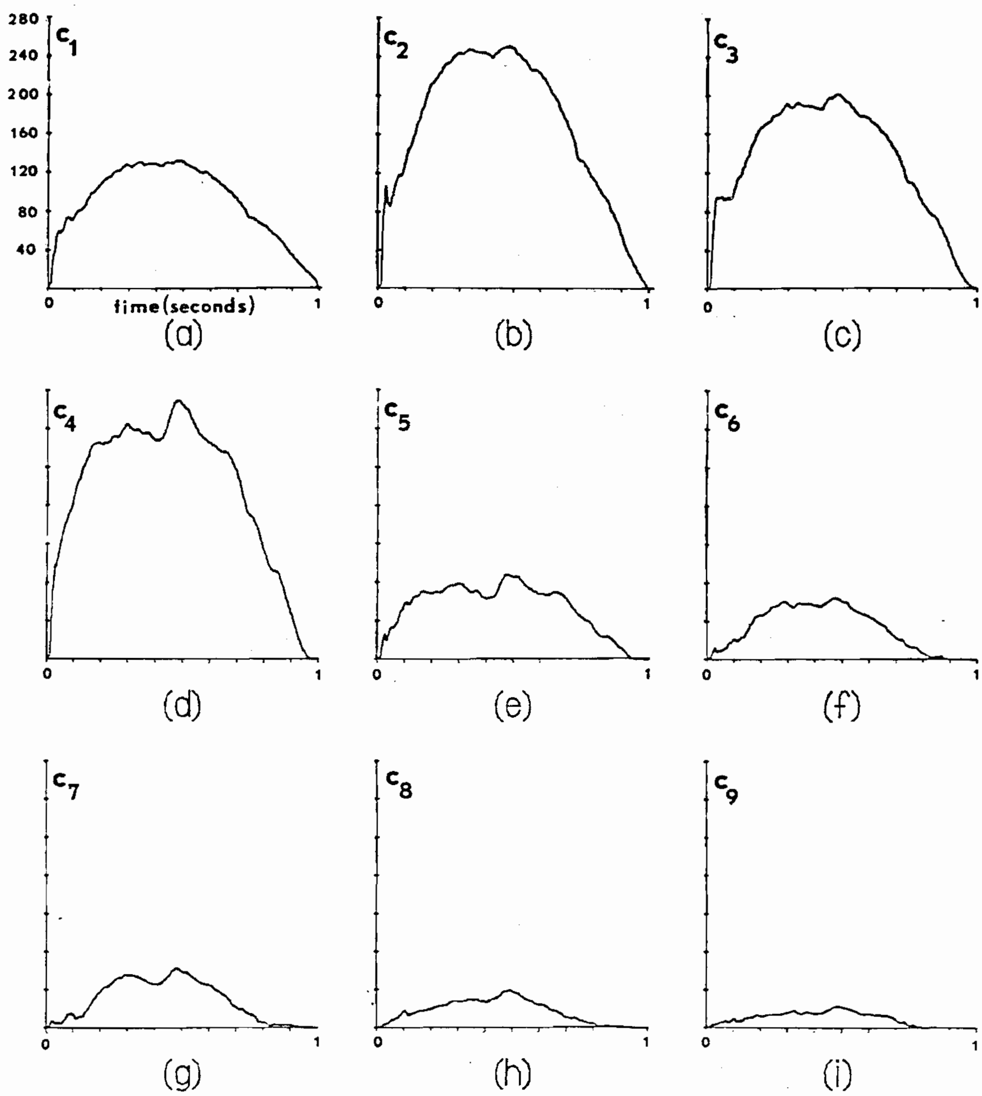

Figure 5

Harmonic amplitudes (c_1 - c_9) for original cornet tone (played 350 Hz (F4), mf). Duration = 1 second.

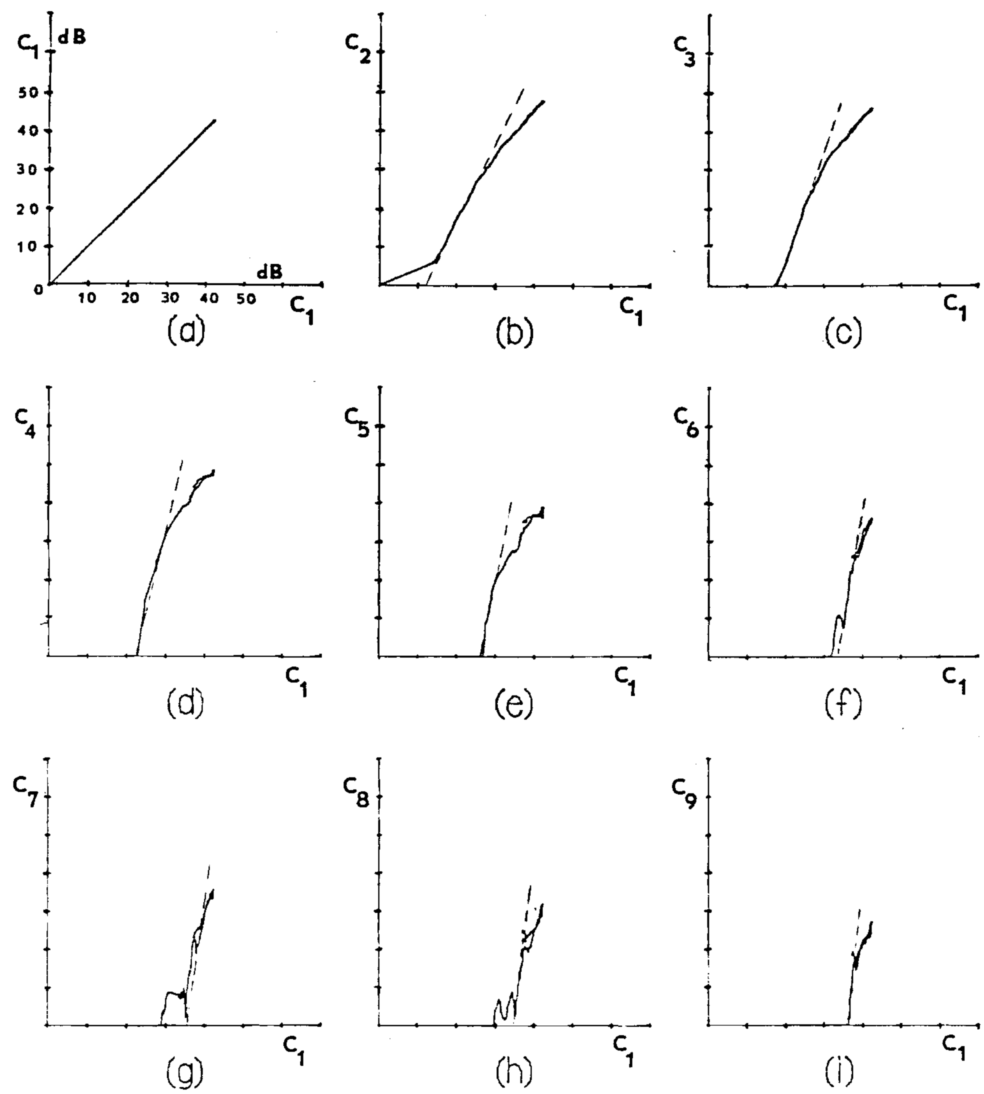

Figure 6

NIHR curves for original cornet tone. Dashed straight lines have slopes of 1, 2,...,9 corresponding to the harmonic number

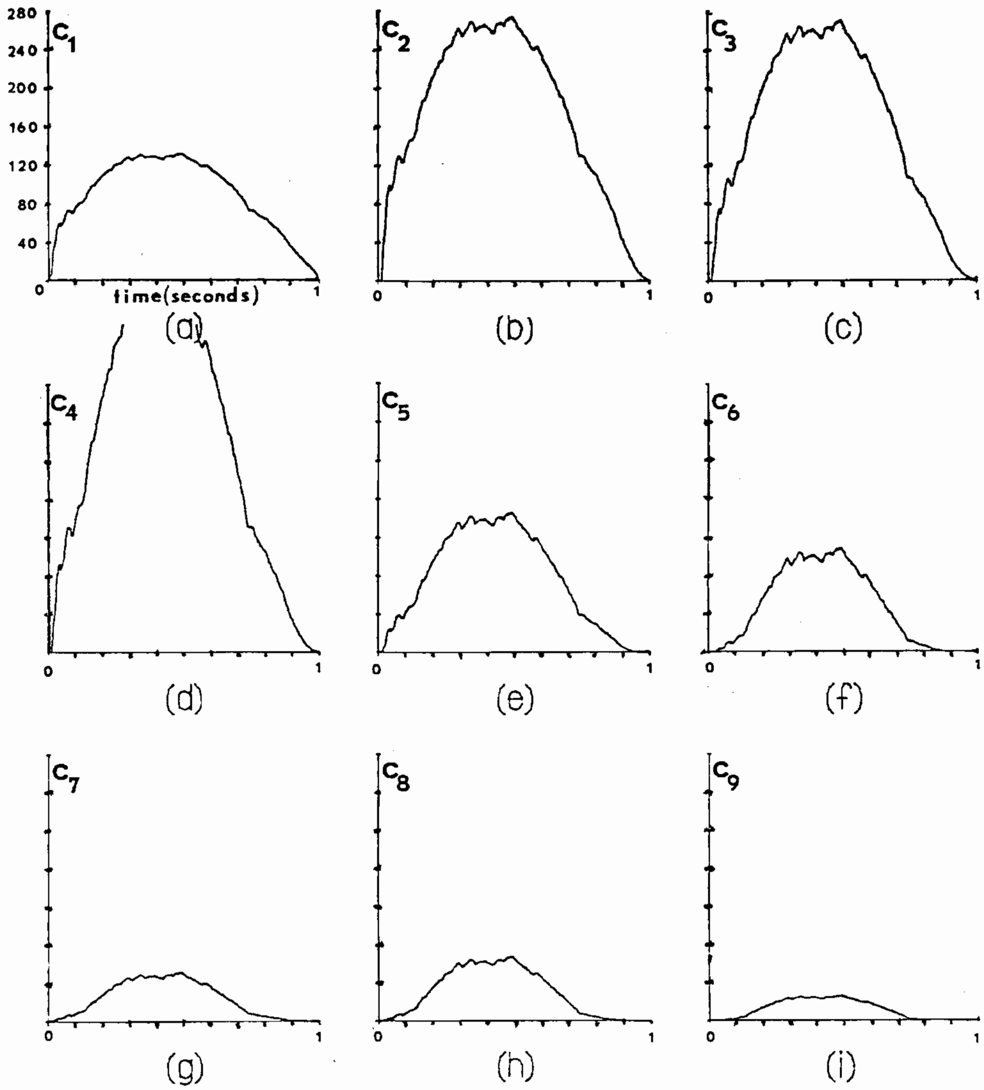

Figure 8

Harmonic amplitudes of synthesized cornet tone.
Compare with Figure 5.

The synthesis procedure can be summarized as follows:

1. From the NIHR curves of Figure 6 (for the mf case) we select a value of C_1 and corresponding values of C_2, C_3, ...

2. Using one of the transmission curves of Figure 7 (the p_o/p_m curve seemed to work best) we derive the D_k vector and from that the corresponding d_k vector using the equations

$$D_k = C_k - H_{dB}\ (350k) \tag{13a}$$

$$d_k = 10^{Dk/20} \tag{13b}$$

3. We use Equation 9 to compute the polynomial coefficients a_k with A_o set to unity.

4. Using Equations 3 and 4b we derive c_1 as a function of A. This allows us to obtain the inverse function to get A from c_1.

5. From the original first harmonic envelope for each tone (pp, mf, ff) we use the inverse function obtained in step 4 to generate A(t).

6. The A(t) functions as derived by step 5 are used with Equation 3 to generate $d_k(t)$ functions for the pp, mf, and ff cases.

7. We use Equation 4b to compute the $c_k(t)$ functions and Equation 4a to compute the output waveform z(t) (We take $A_1 = 1$.) for each of the three cases.

8. The NIHR curves for the synthetic case are also generated by variation of A and use of Equations 3 and 4b.

The synthesis procedure just described is an additive synthesis equivalent of the direct nonlinear synthesis technique illustrated by Figure 1. Of course, the direct nonlinear technique using Equation 1 is the preferred one of the two for reasons of computational efficiency. It involves a table-lookup to get $A(t)$, increment and table-lookup to get $\cos(2\pi f_1 t)$, a multiplication to get $A(t)\cos(2\pi f_1 t)$, a table-lookup to get $F[A(t)\cos(2\pi f_1 t)]$, and a final digital filter operation. The filter can be approximated by a second order high pass type requiring three adds and two multiplies.

The harmonic amplitude curves resulting from the procedure described are shown in Figure 8. These should be compared with the original curves of Figure 5. The match was selected to be perfect for all harmonics for the times at which $c_1 = 63.1$, i.e., for $C_1 = 36$ dB. Note that the $c_1(t)$ curves match perfectly, whereas the other curves match only at one amplitude value. The NIHR cuves for the synthetic tone are shown in Figure 9. These should be compared with the original NIHR curves of Figure 6.

Despite some visible differences between the original and synthesized envelopes, the perceptual difference between the synthetic and original tones is very small. This is true of the pp and mf cases and to a lesser extent the ff case.

In 1973 I visited Arthur Benade at his laboratory at Case Institute and obtained from him measurements of the input impedance (mouthpiece pressure divided by mouthpiece velocity) and transmission response (both output pressure divided by mouthpiece pressure and output pressure divided by mouthpiece velocity) for the same Conn 80A cornet which was used for the tones analyzed in the 1975 paper. The original graphs of p_o/p_m and p_o/v_m for no valves depressed are shown in Figure 7.

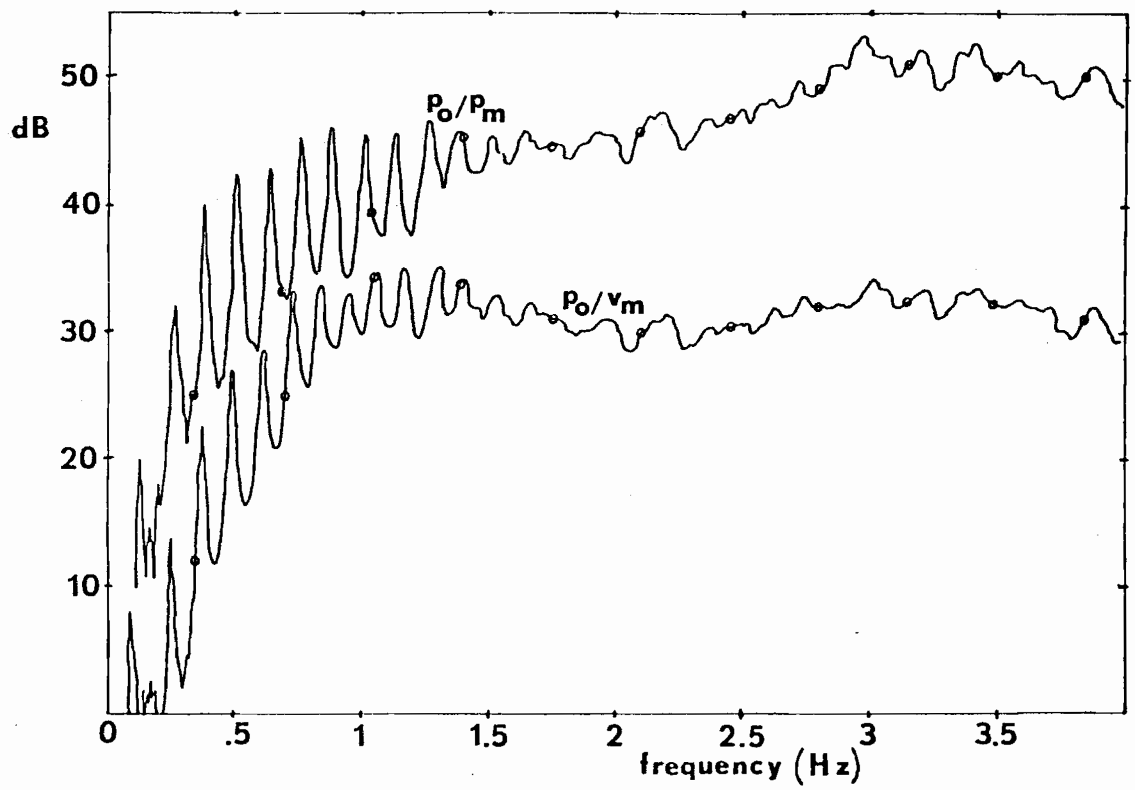

Figure 7

Transmission response graphs for cornet (conn 80A, no valves depressed). Output pressure over mouthpiece pressure (top curve); output pressure over mouthpiece velocity (bottom curve). Responses are sampled at harmonics of 350 Hz.

The transmission curves of Figure 7 were sampled at multiples of 350 Hz to produce values of $H_{db}(kf_1)$. This allowed us to deduce the d_k harmonic amplitudes from a specific C_k vector, i.e., for a particular value of C_1. We then used Equation 9 to compute the nonlinear processor coefficients a_k. They would, hopefully, allow a reasonable replication of the original tone by means of Equations 3 and 4 with appropriate variation of the amplitude A.

A(t) can be chosen to match the original $c_1(t)$ envelope or the envelope for any original harmonic for that matter. (It could also be chosen to match the rms amplitude of the tone.) This is a matter of constructing a tabular function for c_1 vs. A and inverting the function to get A from c_1. The original $c_1(t)$ function becomes the "driver" of the synthesis model, and it is guaranteed that the amplitude of the output signal's first harmonic will match this exactly.

There were some problems in applying the data obtained in Benade's laboratory to the data reported in my 1975 paper. The most difficult problem was that the transmission curves contain several resonances, and harmonic samples which occur on the steep slopes of these resonances are subject to significant changes as the frequency of the tone and other factors such as temperature and humidity change. Another problem was that the microphone position used in Benade's lab was about 6 inches from the bell, as opposed to the 66 in. used in the other data (which was actually collected about 1967); this was only expected to affect frequencies below the radiation cutoff though (about 1500 Hz). Nevertheless, I decided to apply the data to see if it would give a useful result.

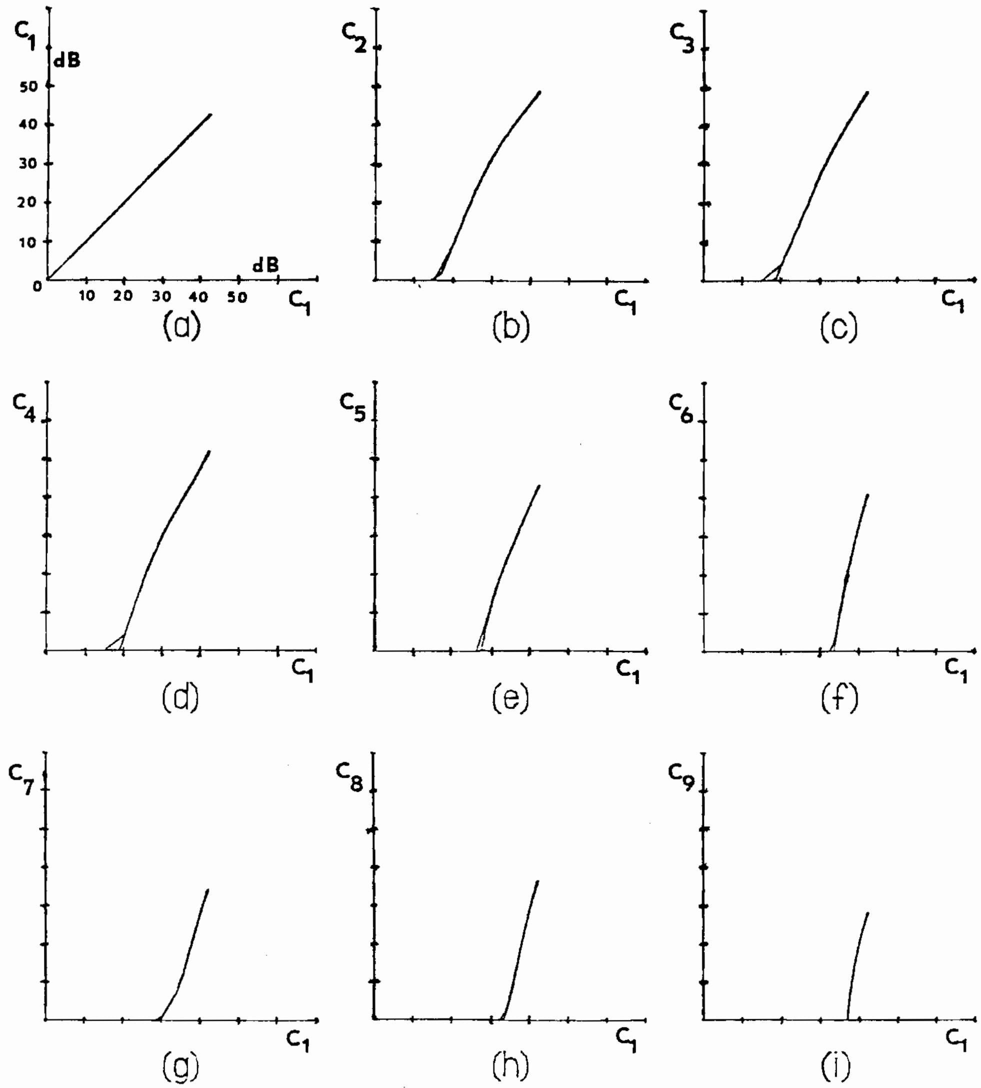

Figure 9

NIHR for synthesized cornet tone. Compare with Figure 5. Match points are for C_1 = 36 dB.

The entire experiment was repeated using a second order Butterworth high pass filter with - 3 dB cutoff at 1800 Hz. The resulting synthetic tones were not as close to the originals as those using the data of Figure 7, but they were still close.

The data derived for the nonlinear processor (the a_k coefficients), the harmonic amplitudes which are matched (C_{k_o}), and the corresponding filter samples (H(350k)) are given in Table 1 for the two filter responses used. Also given are the maximum values of A required for the three dynamic levels.

Summary and Conclusions

The acoustics of the lip-mouthpiece interaction leads one to conclude that a nonlinear process is responsible for tone formation in the mouthpiece. Any smooth nonlinear processes can be represented by a polynomial, and its effect in distorting a sine wave can be predicted in terms of the harmonic spectrum it produces for changing values of the sine wave's amplitude.

Given a harmonic spectrum the coefficients of the polynomial needed to produce it can be computed. A brass tone spectrum evolves in a particular way as the intensity changes. We can match the spectrum at one point and attempt to imitate the spectral evolution by using the matching nonlinear polynomial to distort an amplitude-varying sine wave. However, the spectrum is only guaranteed to match at the matched point. A closer imitation can be obtained by using a filter at the output of the nonlinear pro-

		Filter Data from Figure 7		Second Order Butterworth Filter Approximation	
k	c_{k_o}	H(350k)	a_k	H(350k)	a_k
1	64.0	.0562	489.1	.0378	1023.1
2	109.8	.1380	193.5	.1488	-170.4
3	85.4	.2985	646.4	.2810	690.7
4	104.8	.5888	1873.1	.5159	2114.7
5	28.5	.5370	-959.6	.6853	-996.3
6	5.69	.6095	-2151.9	.8046	-2194.6
7	4.67	.6839	2878.4	.8800	2820.7
8	3.68	.8913	3786.4	.9242	3802.3
9	0.15	1.12	-2483.9	.9506	-2533.5
10	0.65	1.00	-3300.7	.9668	-3346.7
11	0.89	1.00	915.5	.9769	935.7
12	0.59	1.00	1210.5	.9835	1229.4

	A max	A max
pp	1.10	1.21
mf	1.21	1.33
ff	1.46	1.46

Table 1

Data for Nonlinear Synthesis of Analyzed Cornet Tone cessor which simulates the transmission response of the brass pipe. The "goodness" of the imitation can be estimated by visual comparison of the nonlinear interharmonic relationship (NIHR) curves for the original and synthetic cases.

Once the parameters of the nonlinear synthesis model are established, tones can be synthesized with different articulations by supplying different A(t) functions which describe the envelope of the input sine wave. The A(t) functions can be derived from some parameter of the original tone such as the amplitude of the first harmonic; it is guaranteed that this parameter will be matched exactly, and the other parameters will hopefully "follow along" close enough for a reasonable imitation of the original sound.

When this procedure was tried for actual cornet tones using a filter response derived from actual measurement the synthetic tones sounded very close to the originals.

The nonlinear synthesis model is a practical one for synthesis using a computer music program such as Music 5 and Music 4BF. In cost efficiency it compares favorably with FM and has the advantage of a strong theoretical basis for matching spectra.

References

1. Suen, C. Y. (1970), "Derivation of Harmonic Equations in Nonlinear Circuits," J. Audio Engr. Soc., 18(6), 675-676.

2. Backus, J. and J. C. Hundley (1971), "Harmonic Generation in the Trumpet," J. Acoust. Soc. Am., 49, 509-519.

3. Benade, A. H., (1976a), "Fundamentals of Music Acoustics, Oxford University Press, Fig. 20.13, p. 419.

4. Benade, A. H. (1976b), F. M. A., Fig. 20.14, p. 421.

5. Benade, A. H. (1976c), F. M. A., Fig. 21.6 (A), p. 443.

6. Beauchamp, J. W. (1975), "Analysis and Synthesis of Cornet Tones Using Nonlinear Interharmonic Relationships," J. Audio Engr. Soc., 23(10), 778-795.

7. Arfib, D. (1978), "Digital Synthesis of Complex Spectra by Means of Multiplication by Nonlinear Distorted Sine Waves," Audio Engr. Soc. preprint No. 1319 (C-2).