MOUNT SINAI SCHOOL OF MEDICINE

Significance of analysis window size in maximum flow declination rate (MFDR)

Linda M. Carroll, PhD
Department of Otolaryngology, Mount Sinai School of Medicine

Goal:

1. To determine whether a significant difference exists for mean MFDR across 4 different data extraction methods on the same data set.
2. To determine interaction between subject skill level and fundamental frequency on MFDR.

Background:

Examination of laryngeal aerodynamics remains crucial to our understanding of voice function in normal and non-normal subjects. Extensive research over the past 40 years has focused on subglottal pressure and transglottal flow, particularly as it relates to frequency and intensity control. More recently, the speed of closure at the maximal negative slope of the differentiated inverse-filtered waveform, or maximum flow declination rate (MFDR), has emerged as a valuable measure of laryngeal function [1-8]. Although subglottal pressure and transglottal flow have established measurement techniques for data extraction methods (e.g.: peak pressure value during [p] for subglottal pressure), such standards do not exist for MFDR. As such, it becomes difficult to compare results across studies which have used a wide range of measurement techniques.

Assumptions:

- MFDR is the point of sharpest change in the closing velocity of the vocal folds, and reflects the velocity when the vocal fold surfaces are nearly parallel and touching in the anterior (membranous) glottis [1-4, 9].
- It is hypothesized that a more rapid decrease (or stoppage) of the flow yields a more efficient and powerful glottal source, thereby allowing improved acoustic intensity [3,4,5,7,9].
- Previous investigators have reported MFDR values for speaking and singing using a range of 1-60 periods of analysis [3,6,7,10,11].

Experimental Design:

Subjects:

- Eight professional lyric sopranos employed as solo artists at international opera houses (N=4) or regional/national opera houses (N=4) served as volunteers in the IRB-approved study.

Group	A (international level singers) N=4	B (regional/national level singers) N=4
Age	34.5 years	36 years
Professional experience	8.25 yrs	8 yrs
Years Training	16.7 yrs	13.7 yrs

Tasks:

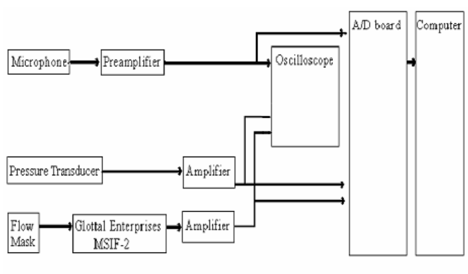
- Three tokens of a 7-syllable /pa/ train at progressively increasing and then decreasing intensity (messa di voce) in singing mode at two contrasting frequencies (F_01=330 Hz, F_02=660 Hz), with each /pa/ syllable lasting 1 second in duration.
- Intensity changes were not prescribed. The subject was instructed to sing a messa di voce as they typically would on the operatic stage.

Figure 1. Sample flow waveform for subject during 7-syllable /pa/ task.

Data Collection:

- The subject held a pneumotachograph mask firmly in place over her nose and mouth, with a pressure tube passing between the lips. A microphone was fitted in the mask handle.
- Signals from flow, pressure and microphone were digitized by a 12-bit analog to digital converter board with a sampling rate of 10 kHz per channel. Digitized signals were imported to the Alamed Voice Plus® and CSpeech® 3.1 analysis systems on a Pentium based computer. Waveforms were optimized by adjusting the amplifier gain to ensure optimum signal input for each subject prior to data collection, and were monitored by a Tektronix® TDS-420A 4-channel digitizing oscilloscope during computer data collection.
- Glottal velocity waveforms were recorded from two differential pressure transducers (Glottal Enterprises® PTL-2) mounted in a Rothenberg single-layer circumferentially vented pneuomotachograph mask, which was connected to a Glottal Enterprises® MSIF-2 inverse-filtering unit.
- Calibration for pressure (water u-tube manometer) and airflow (Matheson glass-float rotameter) was done immediately after each subject's data collection using known pressures and flow that produced output voltages that approximated those observed on the oscilloscope during data collection.

Figure 2. Block diagram of experimental instrumentation.



Data Analysis:

- The most negative value from the first derivative of the inverse-filtered waveform (MFDR) was extracted using CSpeech® 3.1 for each cycle at $F_0$1 and $F_0$2.
- Some subjects' differentiated flow waveforms had two negative peaks, which were often reduced to one negative peak 20 ms later, as a result of unexpected, intermittent presence of formant energies from high voice quality. Because of errors in peak detection for automated MFDR computation, hand cycle-by-cycle determination of the MFDR value for each cycle was used for the eight subjects.
    - Raw flow signal was compared with the inverse-filtered flow signal during MFDR detection at $F_0$1 (330 Hz) and $F_0$2 (660 Hz) for most negative point within each cycle.

Figure 3. Sample flow signal (A) and inverse filter of signal (B).

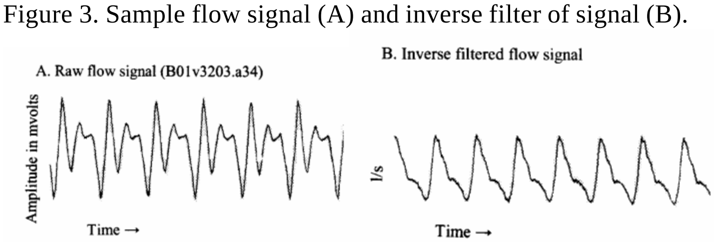
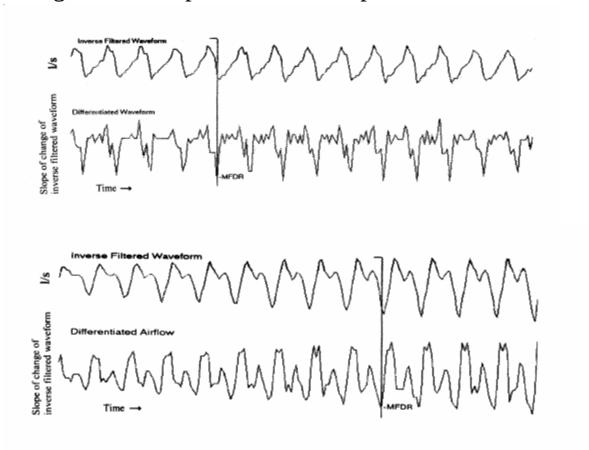


Figure 4. Sample of Inverse-filtered flow signal with differentiated waveform for MFDR for one subject. Tracings show easy marking of MFDR point for upper trace, and need for hand-marking of MFDR point in lower trace with change in cursor position within /pa/ from 256 ms into /pa/ (upper trace) to 321.7 ms into /pa/ (lower trace).



Subject performance was compared from 4 different extraction windows within each /pa/ for the 7-syllable train at $F_0 1$ and $F_0 2$.

- Method A: mean MFDR from *middle* 1000 ms for each /pa/ segment (if less than 1000 ms available in /pa/, then 20 ms excluded from onset/offset)
    - analysis of 330 cycles for $F_0 1$, 660 cycles for $F_0 2$ at mid-portion
- Method B: mean MFDR from *middle* 100 ms of /pa/ segment, with center at mid-portion of entire /pa/ segment
    - analysis of 33 cycles for $F_0 1$, 66 cycles for $F_0 2$ at mid-portion
- Method C: mean MFDR for -/+ 50 ms from *greatest value* of MFDR from entire /pa/ segment
    - analysis of 33 cycles for $F_0 1$, 66 cycles for $F_0 2$ at greatest value
- Method D: mean MFDR for -/+ 10 cycles from *greatest value* of MFDR from entire /pa/ segment
    - 20 cycles for $F_0 1$, 20 cycles for $F_0 2$ at greatest value

Statistical Analysis

SPSS®, with overall $\alpha=0.05$, with each /pa/ studied as unique variables. Each subject's mean MFDR (and sd-MFDR) was a composite of three trial tokens at each pitch condition.

- Analyses of variance (ANOVA) were used to test whether a significant difference exists for MFDR across the four different measurement techniques. Statistical adjustment was made for pitch, group, and all interactions. These analyses were repeated for each /pa/ during the 7-syllable train.
- In the ANOVAs, pitch, group and window were fixed factors, and subjects within groups was a random factor. We used a full ANOVA model that included all interactions. A significance level of 0.05 was used for each analysis. Marginal significance was defined as a p-value between 0.05 and 0.10. Contrasts were performed to compare the four measurement techniques for each /pa/. Bonferroni adjustments were used for these pairwise comparisons.

Results:

- Mean MFDR
    - A significant main effect was found for pitch condition ($F_0 1$, $F_0 2$) at [pa3], [pa4] and [pa5]
    - A significant main effect was found for window (method 1, 2, 3, 4) at [pa1] through [pa6], with marginal main effect at [pa7]
    - A significant pitch condition by group (A, B) interaction was found at [pa1] through [pa6]
    - A significant pitch condition by window interaction was found at [pa4] and [pa5], with marginal 2-way interaction for [pa3]
- Standard deviation of MFDR (sd-MFDR)
    - A significant main effect was found for pitch condition throughout the 7-syllable /pa/ train
    - A significant main effect was found for window at [pa1] through [pa6]
    - A marginal main effect was found for group at [pa2]
    - A significant pitch condition by group interaction was found throughout the 7-syllable /pa/ train
    - A significant pitch condition by window interaction was found at [pa3] through [pa5], with marginal 2-way interaction at [pa2]
- No significant difference between groups for MFDR or sd-MFDR at any [pa]
- Pairwise comparisons
    - Window 1 vs. window 2
        - No significant difference between mean MFDR
        - Significantly different sd-MFDR for window 2 at [pa1], [pa2], [pa5] and [pa6]
    - Window 1 vs. window 3
        - Significantly greater mean MFDR for window 3 at [pa2], [pa5] and [pa6], with marginal significance at [pa1] and [pa3]
        - Marginally different sd-MFDR for window 3 at [pa5]
    - Window 2 vs. window 3
        - No significant difference for mean MFDR or sd-MFDR
    - Window 3 vs. window 4
        - No significant difference for mean MFDR

- Significant difference for sd-MFDR at [pa2], [pa3], [pa5], and [pa6], with marginal significance at [pa4]

MFDR means (and standard deviations) and maximum MFDR for $F_0$1 (pitch=1) for four different data extractions

	Group	Peak Value	M1	M2	M3	M4
Pa1	A	488	79 (25.71)	83 (10.33)	105 (18.51)	120 (13.67)
	B	264	75 (29.64)	83 (11.23)	113 (16.72)	118 (15.91)
	A	547	122 (30.56)	129 (13.31)	155 (14.46)	158 (14.1)
Pa2	B	166	67 (19.16)	74 (7.17)	91 (9.81)	93 (9.06)
	A	313	145 (22.97)	150 (17.76)	154 (19.36)	173 (16.44)
Pa3	B	283	97 (23.63)	108 (10.18)	125 (11.26)	128 (12.56)
	A	430	157 (21.48)	159 (17.52)	175 (19.88)	181 (21.07)
Pa4	B	264	114 (22.14)	121 (10.66)	133 (11.4)	135 (11.44)
	A	254	124 (18.97)	132 (11.24)	143 (13.49)	146 (12.29)
Pa5	B	596	109 (30.23)	129 (12.58)	147 (17.22)	151 (16.54)
	A	195	67 (11.49)	74 (6.75)	79 (7.22)	81 (6.78)
Pa6	B	127	52 (19.23)	54 (7.8)	76 (7.98)	79 (7.49)
	A	88	36 (8.2)	37 (4.75)	43 (5.22)	44 (4.8)
Pa7	B	195	27 (10.61)	31 (6.68)	41 (7.9)	43 (7.78)

MFDR means (and standard deviations) and maximum MFDR for $F_0$2 (pitch=2) for four different data extractions

	Group	Peak value	M1	M2	M3	M4
Pa1	A	1436	167 (114.74)	210 (94.82)	315 (121.77)	422 (50.86)
	B	264	48 (49.38)	50 (20.03)	85 (32.8)	116 (9.79)
	A	908	178 (120.43)	242 (91.88)	310 (115.32)	396 (50.39)
Pa2	B	352	44 (30.82)	50 (23.53)	85 (36.67)	122 (12.84)
	A	1221	233 (116.1)	250 (95.3)	345 (109.87)	442 (52.73)
Pa3	B	986	91 (65.81)	105 (46.47)	170 (77.86)	240 (39.13)
	A	1084	238 (130.74)	264 (126.2)	320 (124.77)	445 (51.48)
Pa4	B	811	112 (72.39)	154 (57.64)	199 (75.34)	274 (38.93)
	A	1191	201 (119.57)	266 (111.25)	344 (92.44)	415 (42.52)
Pa5	B	908	87 (50.12)	108 (39.65)	152 (52.89)	204 (28.76)
	A	430	78 (40.63)	92 (40.41)	118 (45.71)	160 (13.49)
Pa6	B	352	47 (30.86)	54 (18.67)	79 (28.94)	109 (11.51)
	A	264	45 (23.65)	47 (20.97)	66 (25.98)	89 (11.05)
Pa7	B	186	28 (13.69)	35 (11.32)	43 (13.05)	56 (7.37)

Main effect and interaction effects of pitch condition (frequency $F_0 1$, $F_0 2$), window size (method 1, 2, 3, 4) and group (A, B) on mean MFDR and sd-MFDR for each individual /pa/ during the 7-syllable train .*

	Source	F	Sig. α=0.05		Source	F	Sig. α=0.05
Pa1	Pitch	0.507	ns	sdPa1	Pitch	14.875	0.000***
	Window	5.212	0.004**		Window	4.479	0.008**
	Group	0.782	ns		Group	0.960	ns
	Pitch x window	0.638	ns		Pitch x window	1.370	ns
	Pitch x group	22.006	0.000***		Pitch x group	16.192	0.000***
	Window x group	0.005	ns		Window x group	0.194	ns
Pa2	Pitch	0.667	ns	sdPa2	Pitch	70.859	0.000***
	Window	7.607	0.000***		Window	7.915	0.000***
	Group	4.449	0.079†		Group	4.350	0.082†
	Pitch x window	1.725	ns		Pitch x window	2.610	0.064†
	Pitch x group	20.071	0.000***		Pitch x group	15.129	0.000***
	Window x group	0.031	ns		Window x group	0.149	ns
Pa3	Pitch	18.287	0.000***	sdPa3	Pitch	166.142	0.000***
	Window	7.123	0.001***		Window	5.139	0.004**
	Group	2.204	ns		Group	2.288	ns
	Pitch x window	2.748	0.055†		Pitch x window	3.054	0.039*
	Pitch x group	11.566	0.001***		Pitch x group	5.941	0.019*
	Window x group	0.174	ns		Window x group	0.488	ns
Pa4	Pitch	33.673	0.000***	sdPa4	Pitch	82.902	0.000***
	Window	6.407	0.001***		Window	3.468	0.024*
	Group	1.514	ns		Group	1.789	ns
	Pitch x window	3.868	0.016*		Pitch x window	3.054	0.039*
	Pitch x group	6.326	0.016*		Pitch x group	7.646	0.008**
	Window x group	0.168	ns		Window x group	0.564	ns
Pa5	Pitch	8.496	0.006**	sdPa5	Pitch	242.386	0.000***
	Window	10.896	0.000***		Window	12.910	0.000***
	Group	1.836	ns		Group	0.947	ns
	Pitch x window	3.080	0.038*		Pitch x window	5.054	0.004**
	Pitch x group	65.663	0.000***		Pitch x group	57.965	0.000***
	Window x group	0.254	ns		Window x group	1.245	ns
Pa6	Pit ch	0.382	ns	sdPa6	Pitch	42.081	0.000***
	Window	6.612	0.001***		Window	6.594	0.001***
	Group	0.771	ns		Group	0.552	ns
	Pitch x window	1.321	ns		Pitch x window	2.206	ns
	Pitch x group	7.631	0.008**		Pitch x group	13.064	0.001***
	Window x group	0.156	ns		Window x group	0.595	ns
Pa7	Pitch	0.154	ns	sdPa7	Pitch	10.222	0.003**
	Window	2.702	0.058†		Window	1.792	ns
	Group	0.768	ns		Group	0.314	ns
	Pitch x window	0.271	ns		Pitch x window	0.568	ns
	Pitch x group	1.123	ns		Pitch x group	7.678	0.008**
	Window x group	0.030	ns		Window x group	0.349	ns

*Note: † $p \leq 0.10$ (marginal significance), and *$p<0.05$, **$p<0.010$, and *** $p<0.001$.

Mean MFDR and sd-MFDR pairwise comparisons of window sizes (where W1=method 1; W2=method 2; W3=method 3; W4=method 4) for individual /pa/ during 7-syllable train.

	Pairwise Comparison	Mean difference	Sig.		Pairwise Comparison	Mean difference	Sig.
Pa1	W1 x W2	-0.068	ns	sdPa1	W1 x W2	-.0.673	0.011*
	W1 x W3	-0.446	0.023†		W1 x W3	0.223	ns
	W2 x W3	-0.377	ns		W2 x W3	-0.450	ns
	W3 x W4	-0.191	ns		W3 x W4	0.586	ns
Pa2	W1 x W2	-0.114	ns	sdPa2	W1 x W2	0.609	0.001*
	W1 x W3	-0.452	0.004*		W1 x W3	0.308	ns
	W2 x W3	-0.339	ns		W2 x W3	-0.301	ns
	W3 x W4	-0.160	ns		W3 x W4	0.468	0.009*
Pa3	W1 x W2	-0.113	ns	sdPa3	W1 x W2	0.294	Ns
	W1 x W3	-0.263	0.015†		W1 x W3	0.153	ns
	W2 x W3	-0.150	ns		W2 x W3	-0.141	ns
	W3 x W4	-0.188	ns		W3 x W4	0.425	0.008*
Pa4	W1 x W2	-19.448	ns	sdPa4	W1 x W2	8.685	Ns
	W1 x W3	-51.394	ns		W1 x W3	3.839	ns
	W2 x W3	-31.946	ns		W2 x W3	-4.846	ns
	W3 x W4	-52.310	ns		W3 x W4	27.117	0.014†
Pa5	W1 x W2	-0.167	ns	sdPa5	W1 x W2	0.392	0.001*
	W1 x W3	-0.352	0.000*		W1 x W3	0.245	0.023†
	W2 x W3	-0.184	ns		W2 x W3	-0.147	ns
	W3 x W4	-0.146	ns		W3 x W4	0.386	0.001*
Pa6	W1 x W2	-0.086	ns	sdPa6	W1 x W2	0.507	0.009*
	W1 x W3	-0.362	0.010*		W1 x W3	0.271	ns
	W2 x W3	-0.276	ns		W2 x W3	-0.237	ns
	W3 x W4	-0.168	ns		W3 x W4	0.519	0.008*
Pa7	W1 x W2	-0.075	ns	sdPa7	W1 x W2	0.447	Ns
	W1 x W3	-0.330	ns		W1 x W3	0.086	ns
	W2 x W3	-0.255	ns		W2 x W3	-0.362	ns
	W3 x W4	-0.149	ns		W3 x W4	0.316	ns

*Note: † $p \leq 0.0.025$ (marginal significance), and *$p<0.0125$

Discussion:

MFDR was found to be significantly greater for louder intensities (during a messa di voce task), and greater for the more elite (level A) singers throughout a messa di voce. The value of MFDR was significantly higher for the louder portion of the messa di voce task. MFDR was found to be more variable (higher sd-MFDR) among the more elite singer, which suggests a more *reactive* relationship for source and filter for those subjects during the sung task (Titze, 2004).

More detailed examination of transglottal flow and subglottal pressure from the raw data had revealed greater variability (higher sd-flow) among the B level singers, but no significant difference in mean flow rate (even with change in frequency). There was a higher correlation of subglottal pressure to frequency for the A group singers in the lower register transition (Carroll, 2001).

This suggests that the elite singer (A group) and regional singer (B group) balance source and filter characteristics differently. First, the elite singer monitors use of support (reflected in subglottal pressure-frequency interaction) at both the upper and lower register transition, while the regional singer monitors support in the higher frequency, not the lower frequency. Second, the elite singer reacts and adjusts MFDR throughout sung events, while the regional singer maintains status quo.

There does *not* appear to be a significant difference in overall data from a 1000 ms analysis window to a smaller 100 ms analysis window. However, the location of the 100 ms segment *does* appear to alter the mean MFDR value. A greater mean MFDR was found when centered on the peak MFDR for the utterance. MFDR was found to be significantly greater at the higher fundamental frequency during the middle of a messa di voce task in the peak window analysis segment (method 3) and higher among elite singers (group A).

There is *no difference* in MFDR data from a 100 ms analysis segment vs. a 20 cycle analysis segment for medium low pitch ($F_0$1=330 Hz) or medium high pitch ($F_0$2=660 Hz) among professional female singers for mean MFDR. If variability is of interest (sd-MFDR), then 100 ms is a <u>better</u> analysis segment when compared to 20 cycles.

It is suggested that window extraction specifics be included in future research to allow closer comparison of mean MFDR. As analysis moves to nonlinear aspects of the voice, data analysis segments should have a minimum of 100 ms.

Summary:

A moderate sized window segment appears to be sufficient for determining mean MFDR. There does not appear to be a significant advantage to using a large (1000 ms) analysis window. There does appear to be a loss of data when the analysis window is reduced from moderate (100 ms) to small (20 cycles).

Among the professional singer population, there does appear to be a difference at the glottal level in management of airflow shut-off when fundamental frequency increases among subjects who are employed in regional/national level opera companies vs. those employed at international level opera companies. Both groups were found to increase MFDR as fundamental frequency increased, and greater MFDR for louder portions of the messa di voce task.. During sustained phonation, the elite singer appears to use a more inertive vocal tract and more nonlinear productions.

References:

1. Titze IR (1986). Mean intraglottal pressure in vocal fold oscillation. *Journal of Phonetics* 14: 359-364.
2. Titze IR (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America* 85(4): 1699-1707.
3. Holmberg EB, Hillman RE, Perkell JS (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal and loud voice. *Journal of the acoustical Society of America* 84: 511-529.
4. Perkell JS, Hillman RE, Holmberg EB (1994). Group differences in measures of voice production and revised values of maximum flow declination rate. *Journal of the Acoustical Society of America* 96(2 Part 1): 695-698.
5. Stathopoulos ET, Sapienza C (1993). Respiratory and laryngeal function of women and men during vocal intensity variation. *Journal of Speech and Hearing Research* 36: 64-75.
6. Peterson KL, Verdolini-Marston K, Barkmeier JM, Hoffman HT (1994). Comparison of aerodynamic and electroglottographic parameters in evaluating clinically relevant voicing patterns. *Annals of Otolaryngology, Rhinology and Laryngology* 103: 335-346.
7. Sulter AM and Wit HP (1996). Glottal volume velocity waveform characteristics in subjects with and without vocal training related to gender, sound intensity, fundamental frequency, and age. *Journal of the Acoustical Society of America* 100(5): 3360-3373.
8. Sundberg J, Andersson M, Hultqvist C (1999). Effects of subglottal pressure variation on professional singers' voice source. *Journal of the Acoustical Society of America* 105(3): 1965-1971.
9. Holmberg EB, Hillman RE, Perkell JS (1989). Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal and high pitch. *Journal of Voice* 3: 294-305.
10. Södersten M, Hertegård S, Hammarberg B (1995). Glottal closure, transglottal airflow, and voice quality in healthy middle-aged women. *Journal of Voice* 9(2): 182-197.
11. Homberg EB, Hillman RE, Perkell JS, Gress C (1994). Relationship between intraspeaker variation in aerodynamic measures of voice production and variation in SPL across repeated recordings. *Journal of Speech and Hearing Research* 37: 484-495.
12. Titze IR (2004). Nonlinear source-filter interaction in singing. Presented at The Second International Physiology and Acoustics of Singing Conference, Denver, CO, October 7, 2004.
13. Carroll LM (2001). Airflow characteristics among highly skilled professional singers. Dissertation for PhD at Columbia University, New York, NY.