

Mask-Based Nasometry A New Method for the Measurement of Nasalance

ABSTRACT

The term nasalance has been proposed by Fletcher and his associates (Fletcher and Frost, 1974) for a measure of velopharyngeal closure during voiced speech in which nasally emitted acoustic energy is compared to the orally emitted energy. In this paper, a nasalance measure referred to as F0-nasalance is defined in which the amplitudes of only the fundamental frequency components of the nasal and oral acoustic energy are used for this comparison.

When derived from the nasal and oral airflows, as by using a dual chamber circumferentially vented mask, F0-nasalance offers several advantages over a previously proposed measure of nasalance, termed here F1-nasalance, which is derived from acoustic energies in the approximate frequency range of the first formant, as recorded by pressure-sensitive microphones on either side of a finite width sound barrier held against the upper lip. F0-nasalance is also shown to have advantages over methods that compare nasal and oral average or low-pass filtered airflows during voiced speech.

F0-nasalance is a precisely specified measure of velopharyngeal acoustic leakage that is less sensitive to vowel value and voice pitch than is F1-nasalance, and less sensitive to articulatory movements than are methods comparing the low-pass filtered airflows. A system for measuring F0-nasalance using a dual-chamber CV wire-screen mask can be readily extended to the recording of unvoiced nasal emission during consonants by coupling low frequency pressure transducers to the mask chambers. Other advantages and limitations of this new method are described and illustrated.

I. Velar Control and Oronasal Valving in Speech

During speech or singing, it is necessary to open and close the passageway connecting the oral pharynx with the nasal pharynx, depending on the specific speech sounds to be produced. This is accomplished by lowering and raising, respectively, the soft palate, or velum. Raising the velum puts it in contact with the posterior pharyngeal wall, to close the opening to the posterior nasal airflow passages.

This velopharyngeal (or oronasal) passageway must be opened when producing nasal consonants, such as /m/ or /n/ in English, and is generally closed when producing consonants that require a pressure buildup in the oral cavity, as for stops (such as /p/ and /b/ in English), or approximants (as /s/ and /z/). During vowels and sonorant consonants (such as /l/ or /r/ in English), the oronasal passageway must be closed or almost closed for a clear sound to be produced, though in some languages an appreciable oronasal opening during a vowel is occasionally required for proper pronunciation, as during the first vowel in the French words "français" or "manger".

There are many disorders that result in inappropriate oronasal valving, usually in the form of a failure to sufficiently close the oronasal passageway during non-nasal consonants or non-nasalized vowels. Such disorders include a cleft palate, a hearing loss sufficient to make the nasality of a vowel not perceptible, and many neurological and developmental disorders.

The effect on speech production of insufficient oronasal closure is usually separated into the 'nasal emission' effect, which limits oral pressure buildup in those speech sounds requiring an appreciable oral pressure buildup, and the spectral distortion in vowels and sonorant consonants that is often referred to as 'nasalization'. (Baken 1987, Chapter 10). The terminology used here is that suggested by Baken, who also prefers to reserve the term 'nasality' for the resulting perceived quality of the voice.)

The action of the velum is not easily observed visually, and there is little proprioceptive feedback associated with velar movements. In addition, the acoustic effects of improper velar action are sometimes difficult to monitor auditorally. Therefore, there is a need in the field of speech pathology for convenient and reliable systems to monitor velar action during speech, both to give the clinician a measure of such action and to provide a means of feedback for the person trying to improve velar control.

II. Previous Methods for Measuring Velar Function

Methods for instrumentally monitoring velopharyngeal closure during speech have been reviewed extensively by Baken (1987, Chapter 10). The less invasive methods described by Baken generally fall under the following four categories:

1. Measuring the low frequency, primarily subsonic and including zero frequency, components of the airflow through the nose or through the nose and mouth simultaneously, often with a measure of the intraoral pressure. (Baken 1987, pages 416-421; McLean, et al. 1987)
2. Placing an accelerometer (vibration detector) on the nose to detect sound passing through the nose. (Baken 1987, pages 404-407)
3. Measuring the sound (acoustic pressure waveform) emitted from the nose and mouth, respectively, usually in conjunction with the placing of a solid sound barrier against the upper lip to improve the separation of the nasal and oral sounds, with microphones placed above and below the barrier, respectively. (Baken 1987, pages 401-404; Nasometer literature)
4. Analyzing the acoustic properties of the radiated speech to detect the acoustic properties associated with nasalization. (Baken 1987, pages 398-401)

These various methods can generally be divided into two types, according to the aspect of velar control being measured: (a) those that measure velar control during those consonants requiring an oral pressure buildup (e.g., stops and approximants), and (b) those that measure velar control during vowels and sonorant consonants. Methods of type (b), namely for measuring the nasalization of vowels and sonorant consonants, have been more difficult to implement successfully (Baken 1987, page 393).

Methods in each of the four categories of methodology described above have one or more serious inherent drawbacks.

Methods measuring low frequency (or low-pass filtered) volume airflow (in Category 1.) can show well the oronasal valving patterns during voiced or unvoiced consonants requiring a strong oral pressure buildup (measurement type (a)). However, because these methods rely on low frequency airflow components, during vowels and sonorant consonants they yield readings contaminated with significant low frequency artifacts due

to lip and jaw motion and soft palate deflection. These methods also require a well-fitting mask over both nose and mouth or nasal plugs and an oral mask. The mask used can also cause a muffling of the voice (McLean 1997), though such muffling can be greatly reduced by use of a circumferentially vented mask (see below), or by using a mask incorporating one or more acoustically transparent diaphragms in the mask walls to allow the higher frequency components in speech to be more effectively radiated and also reduce deleterious acoustic loading of the vocal tract caused by the mask. (reference to Rothenberg mask patent). The principles of the circumferentially vented mask and the diaphragm mask can also be combined for minimal voice muffling in low frequency airflow measurements.

The other categories of methods focus on measurements of voiced sounds:

Accelerometer methods (Category 2) generally require adhering a small accelerometer or vibration detector to the side of the nose, and yield a measurement that is highly dependent on the vowel being spoken, the voice pitch, nose geography and the consistent placement of the accelerometer.

The oral/nasal sound-pressure-ratio methods (Category 3) are highly dependent on the precise geometry of the oral-nasal sound barrier used, the placement and directivity characteristics of the microphones, and the frequency range over which energy in each channel is measured. The choice of frequency range is especially problematic, since the spectral distribution in the oral and nasal channels can differ greatly, with the sound emitted from the nose consisting primarily of energy at the lower voice harmonics. Thus if too wide a bandwidth is used, such a system would be comparing the energy in mostly lower frequency voice harmonics emanating from the nose with the energy of mostly higher frequency harmonics from the mouth. For a popular commercial version of this method, the Nasometer (Kay Elemetrics), as well as its previous research version, TONAR II, this frequency range has been empirically chosen to be roughly 300 Hz to 750 Hz, with half-power points at 350 and 650 Hz (Baken 1987; Nasometer Manual). This frequency range was presumably chosen to emphasize the lower frequency harmonics that predominate in the nasal emissions, while capturing the energy of the first formant (the lowest vocal tract resonance) for most vowels and sonorant consonants in the oral channel. However, since the directivity of even a directional microphone at the lower frequencies of this range is limited by the long wavelengths (approximately 3.3 feet at 300 Hz), there is necessarily some appreciable sound crossover between the oral and nasal channels, assuming reasonable proportions for the sound barrier against the upper lip. Thus, a nasal consonant would be expected to register appreciable oral pressure, even in the presence of complete oral closure. There is also a strong dependency in versions of this method on the voice pitch and on the vowel or consonant being spoken.

In the fourth category of methods, the spectrum of the radiated pressure waveform during voiced speech is analyzed to determine the degree of nasalization. However, in attempts to do this it has been difficult to obtain meaningful quantitative results (Baken 1987). The effect of incomplete velopharyngeal closure on the spectrum of a voiced speech sound is highly variable between speech sounds and is highly dependent on the acoustic properties of the nasal passages. For example, consider the great changes in speech quality produced when the nasal passages are partially occluded by nasal congestion during a cold. Thus readings for the same level of velar control could vary greatly from day-to-day, even for the same subject.

III. Definitions of Nasalance as a Measure of Nasalization

Fletcher and his associates (1974) have coined the term 'nasalance' to describe various measures of the balance between the acoustic energy at the nares, A_n , and the acoustic energy at the mouth, A_o , during voiced speech. This balance between A_n and A_o can be expressed as a simple ratio, A_n/A_o , to yield a measure that can be referred to as a "Nasalance Ratio" (NR) or it can be expressed as a percentage, $A_n / (A_o + A_n)$, to yield a

measure that can be referred to as "% Nasalance" (%N). Each measure contains the same information but with a different scale. Most recent measurements of nasalance have been reported in the % Nasalance form.

There can be numerous measures of nasalance, as defined by Fletcher, depending on manner in which the nasal and oral energies are measured. For example, in the original TONAR system, Fletcher allowed the user to set the parameters of a bandpass filter in each channel, so as to select a nasalance measure best for the application at hand.

In recent practice, clinicians and researchers have been reporting numbers generated by the particular combination of microphone type, microphone placement, separator dimensions, and bandpass filter parameters in the Kay Elemetrics Nasometer. According to its manual, the Nasometer measures the amplitudes of band-pass filtered oral and nasal radiated sound pressure waveforms, as transduced by means of directional microphones mounted in metal blocks attached to the separator and located approximately 5 cm from the lips and nares. The two band-pass filters each consist of cascaded low pass and high pass 4-pole Butterworth filters, with -3dB points of 350 Hz and 650 Hz, respectively. Thus energy below about 300 Hz and above about 750 Hz would be significantly attenuated. Attenuated components would therefore include the voice fundamental frequency component (especially for adult male voices) and formant energy above the first formant for most vowels.

To the extent that the microphones at their locations on the Nasometer have a flat frequency response when inferring pressure at the lips and nares (perhaps a questionable assumption when a directional microphone is used so close to the sound source), and to the extent that the partition between the microphone separates the oral and nasal sounds in the 300 to 750 Hz range, the Nasometer can be considered to be measuring what might be termed first-formant nasalance, or F1-nasalance. In addition, since its measurements are derived from radiated sound pressure (as differentiated from particle velocity or volume velocity, as described below), the Nasometer can be more completely considered to be measuring 'pressure-derived F1-nasalance'. However for brevity we only refer to this method as measuring F1-nasalance.

IV. F0-Nasalance

A. F0-nasalance defined

A second commercially available system for measuring nasality, the Glottal Enterprises OroNasal system, derives a measure of nasalance from the ratio of nasal-to-oral airflow volume-velocity at the voice fundamental frequency, F0, yielding what might be termed 'flow-derived F0-nasalance'. In the OroNasal system, the nasal and oral airflow are recorded from a circumferentially vented (CV) wire screen mask that is separated into nasal and oral chambers by a separator within the mask that rests against the upper lip. A CV mask records airflow by putting a small flow resistance (in this case, a fine-mesh wire screen) in the air path and recording the resulting pressure drop (Rothenberg 1973, 1977). In the OroNasal system, these pressure waveforms are recorded by means of two matched omnidirectional microphone elements selected to have a linear response over the pressure ranges found within the mask chambers.

Though F0-nasalance can be theoretically derived from either the airflow (volume-velocity) or the radiated pressure waveforms, the form derived from airflow is easier to specify unambiguously. The volume velocity, being the total flow from the respective orifice at any point in time, can be measured by summing flow components over any surface in space enclosing that orifice, while the radiated pressure will vary with the distance from the orifice, orientation of the microphone with respect to the orifice, and the size of the orifice (e.g., lip opening).

By measuring airflow instead of radiated pressure, and limiting the measurements to primarily the fundamental frequency component, the OroNasal system attains many of the advantages of both the low frequency flow systems and the systems measuring wide-band acoustic pressure. The reasoning supporting this claim follows.

B. Flow-derived F0-nasalance vs. low frequency airflow

The ratio of F0 flow at the nostrils to F0 flow at the mouth reflects well the ratio of low frequency flows at these locations, since the amplitude of the F0 component, A_{f0} , for normal nonbreathy voiced speech is strongly correlated with the average or low frequency airflow, A_{av} . This conclusion follows from the observation that most of the periodic energy in the airflow pulses through the glottis is contained in the F0 component. It follows from this that the F0 component is similar in shape to the entire waveform. This is illustrated in Rothenberg 1977, in which it is shown that the shape of the glottal airflow pulses is well represented by the lowest few harmonics.

Though A_{av} is only coarsely represented by A_{f0} , it offers a better representation of airflow for the purpose of measuring relative velopharyngeal airflow during voiced speech, since it is not affected by airflow components generated by articulatory movements. Articulatory movements have a spectrum in the range of approximately 0 to 10 Hz, as limited by the dynamic constraints in these movements. Though this range overlaps with the frequency range for average airflow measurement, usually 0 to 20 or 30 Hz, it is well below the range for F0 values in speech or singing. Therefore, measurements of F0-component amplitude are not significantly affected by articulatory movement. This can be readily illustrated by comparing the low frequency oral airflow and the amplitude of the F0 component in a syllable sequence having a large amount of jaw movement, such as /wawawa. (That the syllabic variation in the low frequency flow trace is primarily caused by the movement of the mandible, can be verified by making a similar periodic jaw movement with the glottis closed.)

A second advantage of using the amplitude of the F0 flow instead of the low frequency flow is that F0 flow is much less sensitive to mask air leakage. This is because the inertive component of the flow impedance of a narrow flow path, which increases proportional to frequency, impedes the airflow at the frequency of F0.

C. F0-nasalance vs. F1-nasalance

A problem in obtaining definitive comparisons between F0-nasalance and F1-nasalance is the lack of a theoretically sound definition for the latter. F0-nasalance has been defined here independent of a particular mechanism for its measurement, i.e., as the ratio of the amplitudes of the F0 components of the nasal and oral volume airflows. Thus a particular system for its measurement can be evaluated, in principle, for its accuracy in recording the theoretically correct value. However, no such precise standard exists for F1-nasalance. F1-nasalance has been defined instead in terms of a particular commercial device commonly used for its measurement, the Kay Nasometer. In the following discussion we will assume that the Nasometer parameters define F1-nasalance, though for convenience in comparing measurements of the same speech sample, test comparisons will be made using the CV mask emulations of the Nasometer.

As with F0-nasalance, F1-nasalance is insensitive to articulatory movements and circumferential mask air leakage if a mask is used. However, F1-nasalance will be more influenced by the specific vowel being produced than is F0-nasalance. This is because F1-nasalance is basically comparing the amplitudes of two different types of spectra. It compares the F1 energy radiated from the lips with that part of the low-frequency dominated nasal murmur that overlaps with the overall F1 frequency range.

There would also be expected to be a stronger pitch dependence for F1-nasalance when the value of F0 is close to the lower band limit of the F1 filter used (350 Hz in the Nasometer). The reasoning behind this conclusion is as follows. The nasal audio has comparatively less F1 energy and a comparatively stronger F0 component than does the oral audio. Therefore if the value of F0 rises to approach the lower bandwidth limit, even if there has been no increase in velopharyngeal opening, the energy passed by the bandpass filter in the nasal channel will rise in comparison to the energy passed by the bandpass filter in the oral channel, and the value of nasalance displayed will increase.

Related to the greater pitch dependence of F1-nasalance is the inter-subject variation that may be induced by the fact that the oral channel amplitude is dependent on the oral formant energy while the nasal channel amplitude is must less so. Thus users having a 'strong' voice (stronger formant excitation for the same average air flow and fundamental frequency component energy) would be expected to have a lower nasalance reading for the same degree of velopharyngeal opening.

D. Sources of error for flow-derived F0-nasalance

Though F0-nasalance derived from airflow offers a number of important advantages over other previous methods for measuring the degree of velopharyngeal closure during speech, there are some limitations associated with this technique.

An obvious disadvantage of the method is the need for a facemask, with attendant voice distortion and muffling. However, a CV design for the mask can keep this voice distortion and muffling to a level acceptable for most clinical applications.

Another limitation is the possible leakage of sound between the oral and nasal chambers. Assuming a good mask seal to the face, this leakage can result from at least three factors: vibration of the interchamber mask barrier, radiated sound from one chamber reentering the other chamber, and vibration of the soft palate.

Of the three factors, the first, vibration of the mask barrier, can be made small by thickening or stiffening the barrier, and is generally not significant. This can be verified by recording Nasalance Ratio during an alveolar nasal consonant while the ports in the oral chamber are occluded or covered (to eliminate the reentrant acoustic energy).

The second factor, reentrant sound, can be made smaller by raising the flow resistance of the wire screen, at the expense of an increase in voice distortion and muffling and an increased perturbation of the oral-nasal flow balance. (With a very low screen flow resistance, the interchamber crosstalk due to this factor approaches that of the acoustic barrier method commonly used for measuring F1-nasalance.) The magnitude of the error caused by reentrant sound can be estimated by recording % Nasalance during an alveolar nasal consonant. Instead of a value close to 100%, the theoretically expected value, we generally see values of about 90%, indicating that roughly 10% of the nasal energy at F0 is reentering the oral chamber.

The third factor, vibration of the soft palate, is highly variable, since it depends on the vowel or consonant being produced, the value of F0, the degree of velopharyngeal leakage, and the acoustic compliance of the soft palate. This leakage is from the oral chamber (the sound source) to the nasal chamber. It is maximum when there is both a complete velopharyngeal closure and a strong oral constriction anterior to the velum, as during a non-nasalized tense /u/ or /l/, and would tend to raise the measured value of nasalance during such sounds.

The effect of vibration of the soft palate can be separated from the effect of reentrant sound by adding a large oral-nasal acoustic barrier to the mask, external to the mask, and recording nasalance for vowels produced with

a complete velopharyngeal closure (as when produced between two stop consonants by a normal speaker). However, we have not tried this experiment.

An increase in measured nasalance caused by velar vibration cannot be properly referred to as an error, since it results from actual acoustic energy at the nares. However, it does present to the user a bias that must be disregarded when using nasalance to judge the degree of velopharyngeal closure. In practical terms, we have found that the effects of the sum of reentrant sound and vibration of the soft palate will cause a bias of between 0.05 and 0.15 in measurements of NR made during a complete velopharyngeal closure.

V. Measuring Unvoiced Nasal Emission

A potential deficiency of all nasalance-based methods for measuring the degree of velopharyngeal closure is that they only function during voiced speech sounds. The occurrence of 'nasal emission' (using Baken's terminology cited above) during unvoiced consonants is not detected. However, nasalance measurement methods that employ a dual-chamber CV wire-screen mask can be readily adapted for the simultaneous detection of nasal emission.

To measure low frequency airflow using a wire screen mask, the pressure transducers that detect the pressure variations across the wire screen, and their associated electronics, must have a frequency response that extends down to zero frequency or constant pressure. In addition, to receive these signals, current microcomputers must have an A-D capacity added that extends down to zero frequency, since the signal capture capabilities of the computers themselves only include the audio range.

However, such added transduction and A-D capabilities are rapidly becoming less expensive. It is not difficult to conceive of mask-based nasalance monitoring systems for use on any general-purpose microcomputer marketed commercially in the near future with the added capability of recording nasal emission, and in a price range that would make the them accessible to a speech-communication-impaired person for home use.

Note: since the original drafting of this paper a system for nasometry that measures both nasalance and nasal emission has been marketed by [Glottal Enterprises](#), which refers to this system as the Nasality Visualization System (NVS).

References

- Baken, R. (1987). *Clinical Measurement of Speech and Voice*, Little, Brown & Co. - College Hill Press.
- Fletcher S.G. and Frost, S.D. (1974). Quantitative and graphic analysis of prosthetic treatment for "nasalance" in speech, *J. Prosthet. Dent.* **32**, No. 3, pp. 284-291.
- McLean, C.C., et al. (1997). An instrument for the non-invasive objective measurement of velar function during speech, *Med. Eng. Phys.* **19**, No.1, pp. 7-14.
- Nasometer Manual, Kay Elemetrics, Pine Brook, New Jersey, 1999 edition.
- Rothenberg, M. (1973). A new inverse-filtering technique for deriving the glottal airflow waveform during voicing, *J. Acoust. Soc. Amer.* **53**, No. 1, pp. 1632-1645.
- Rothenberg, M. (1977). "Measurement of Airflow in Speech", *J. Speech Hear. Res.* **20**, No.1, pp. 155-176.

Rothenberg, M. (1995). "Pneumotachograph Mask or Mouthpiece Coupling Element for Airflow Measurement During Speech or Singing", U.S. Patent No. 5,454,375, Oct. 3,1995

[Home](#)

[Publications](#)

[E-mail](#)

[Papers online](#)

glottal enterprises

Instruments for Voice Measurement
