



HAL
open science

Comparison of Nasalance Measurements from Accelerometers and Microphones and Preliminary Development of Novel Features

Nicolas Audibert, Angelique Amelot

► **To cite this version:**

Nicolas Audibert, Angelique Amelot. Comparison of Nasalance Measurements from Accelerometers and Microphones and Preliminary Development of Novel Features. INTERSPEECH 2011 - 12th Annual Conference of the International Speech Communication Association, Aug 2011, Florence, Italy. pp.2825-2828. halshs-00676877

HAL Id: halshs-00676877

<https://shs.hal.science/halshs-00676877>

Submitted on 6 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comparison of nasalance measurements from accelerometers and microphones and preliminary development of novel features

Nicolas Audibert¹, Angélique Amelot¹

¹Laboratoire de Phonétique et Phonologie
UMR 7018 CNRS/Université Sorbonne Nouvelle Paris 3, Paris, France
nicolas.audibert@gmail.com, angelique.amelot@univ-paris3.fr

Abstract

This study compares four nasalance measures computed as ratios between the amplitude of signals recorded with accelerometers and microphones. Two new measures based on RMS amplitude differences between the nasal signal and either the vocal folds vibration signal (LND) or the oral acoustic signal (OND) are introduced. Measures were compared on a total of 584 utterances produced by four native French speakers. Results show that (1) all measures separate nasal from oral consonants, (2) the different experimental setups cannot be considered equivalent, (3) difference-based measures appear to better describe the time course of nasality than ratio-based measures.

Index Terms: nasality, nasalance, accelerometer, microphone

1. Introduction

Acoustic correlates of nasality are complex [1]. It is therefore interesting to supply the acoustic signal by other types of signals, which give a more direct indication of nasality [2]. X-ray [3], EMA [4], Nasoendoscopes [5] or PNG [6] give a direct indication of the state of the velopharyngeal port, but these instruments are invasive. When the velum lowers, the nasal cavity is coupled with the oral cavity, and the excitation from the glottis propagates in the nose. The signal from the nasal cavity can be recorded by non-invasive instruments, either by measuring the vibration on the skin by an accelerometer or by recording the acoustic signal using a microphone.

The main limitation of acoustic recordings for the study of nasality is the difficulty of separating the nasal and the oral acoustic output, since the mouth and the nose are very close and both outputs are mixed. This issue has been addressed by inserting a probe microphone inside the nostril [7] or by physically separating the nasal and oral signals using a barrier [8, 9]. Moreover, the intensity of the nasal output is not only a function of the degree of nasal coupling: it also reflects the vocal effort, which cannot be easily controlled or kept constant through utterances by speakers [10].

A number of instruments and measures based on acoustic measures of nasality have been developed since the 70s. Only measures that enable a temporally accurate tracking of nasality are considered here. Those measures are based on combinations of RMS amplitude measures (hereafter Arms(signal)) of a nasal and a vocal signal. Four different signals, summarized in Table 1, were used in the computation of those measures.

Table 1: Signals used in the calculation of nasality measures

Nm	nasal acoustic signal measured by a nasal microphone
Vm	oral acoustic signal measured by an oral microphone
Na	nasal vibration measured by a nasal accelerometer
Va	laryngeal vibration measured by a throat accelerometer

Fletcher [11] defined The Oral Nasal Acoustic Ratio (TONAR) measured as an analog ratio, implemented on Kay elements, between Na and Vm:

$$\text{TONAR} = \frac{\text{Arms(Nm)}}{\text{Arms(Vm)}} \quad (1)$$

Still using Kay elements and the nasometer (which measures Nm and Vm), Dalston [12] introduced a measure of nasalance defined as the ratio between Nm and the sum of Nm and Vm:

$$\text{Nasalance} = \frac{\text{Arms(Nm)}}{\text{Arms(Nm)+Arms(Vm)}} \quad (2)$$

Since the vibration on the nasal surface can be detected, some authors [13, 14] used a nasal accelerometer placed on the lateral nasal cartilage (Na). However, Stevens et al. underline that the accelerometer not only captures variations in nasality, but also changes in intensity, as the acoustic measures do [15]. Horii [16] defines the Horii Oral Nasal Coupling (HONC) index as the ratio between the amplitudes of a nasal accelerometer (Na) and a glottal accelerometer placed on the neck surface (Va), normalized by a calibration constant k that represents the value of the ratio on maximally nasalized speech productions:

$$\text{HONC} = k \frac{\text{Arms(Na)}}{\text{Arms(Va)}} \quad (3)$$

The NAVI measure, equivalent to HONC, was developed by Redenbaugh and Reich to evaluate the link between physiology and perception of hypernasality [18].

Since the throat accelerometer may be influenced by factors such as the vowel quality and the mode of phonation, Horii [17] examined the correlation between the N/V ratio, V being obtained from an oral microphone (Vm) and N from a nasal accelerometer (Na), and perceived hypernasality in speech:

$$N/V = \frac{\text{Arms(Na)}}{\text{Arms(Vm)}} \quad (4)$$

All measures presented above were mainly developed in order to describe the speech of English-speaking patients suffering from hypo- or hyper-nasality. The first purpose of this study is to compare the different instruments on healthy speakers of a language with both nasal vowels and nasal consonants such as French. The second purpose of this study is to determine the best minimal experimental setup that can be used on the field. We also introduce new measures computed from a nasal accelerometer and either a glottal accelerometer (LND) or an oral microphone (OND), which appear to bring important information to study the time course of nasality for the nasal consonants.

2. Methodology

2.1. Subjects

Four native speakers of standard French (two males, two females, aged 30 in average) participated in this study. None of them had any diagnosed language trouble.

2.2. Speech material

The recorded corpus can be subdivided into three parts. The first part is composed of isolated sustained vowels. All French oral (/i,y,u,e,ø,o,ε,œ,ɔ,a/) and nasal (/ã,ẽ,õ/) vowels were included. The second part is composed of 30 CVCVCVC logatoms, with $C=\{t,d,n,s,z\}$ and $V=\{a,i,u,ã,ẽ,õ\}$. The third part of the corpus is composed of 30 French words including a nasal consonant with another consonant as left of right context. All these words have a CN, NC or NN sequence, where $C=\{p,t,k,b,d,g,f,s,ʃ,v,z,ʒ,l,ʁ\}$ and $N=\{m,n\}$, for instance *apnée* (/apne/), *amena* (/amna/) or Parkinson (/paʁkɪnsɔ̃/). All the utterances were produced twice by every subject.

2.3. Recording setup

All data were recorded in a soundproof room. Two microphones and two pairs of accelerometers were used, as illustrated in Figure 1. The oral acoustic signal was recorded using an AKG C520 headset microphone. A double piezoelectric accelerometer (K&K Sound System Hot Spot) was stuck to the speaker's lateral nose bridges using double-sided adhesive tape [19]. A Panasonic reversible electrodynamic transducer was inserted into the speaker's right nostril to be used as a nasal microphone. This transducer is equivalent to a high acoustic pressure (120 dB_{SPL}) electret microphone with a frequency response range from 20 Hz to 20 kHz [20]. A double piezoelectric accelerometer (K&K Sound System Hot Spot) was placed on the speaker's throat at the larynx level using double-sided adhesive tape. These accelerometers have a frequency response range from 20 Hz to 15 kHz.

The nasal microphone and the double piezoelectric accelerometers were connected to preamplifiers with 30 dB gain. These preamplifiers and the headset microphone were connected to a multichannel recording device (Rack Flight Pro Mixer 12U) and linked up to a Mac computer. The recording process was monitored by the Pro Tools LE software v7.4. All signals were recorded at a sampling rate of 48 kHz.

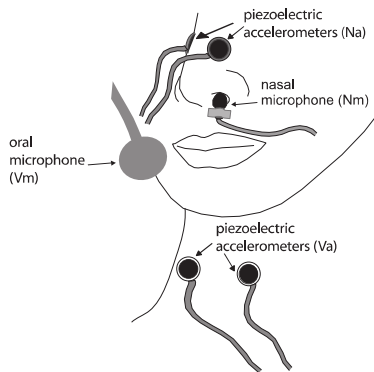


Figure 1: Location of the two microphones and the two pairs of piezoelectric transducers on the speaker's face and neck (adapted from [21]).

2.4. Calibration

Since the four signals are recorded in independent channels as described supra, the gain of each channel and their relative values had to be carefully controlled for each speaker.

The retained calibration procedure was to set the gain levels from a sustained nasal consonant /m/, supposed to be produced with the highest nasal amplitude [17], and a sustained open vowel /a/, which is known to the vowel with the highest oral amplitude. Levels were first adjusted on a sustained /a/ to

obtain the same volume V_{max} for the headset microphone and the laryngeal accelerometer on this vowel, used as a reference. Secondly, the gain values of the nasal microphone and the nasal accelerometers were adjusted to obtain the volume V_{max} on a sustained /m/. The purpose of this gain calibration process is to obtain higher amplitude on nasal signals (N_m and N_a) compared to vocal signals (V_m and V_a) during the production of nasal segments.

2.5. Implemented measures

The four measures based on ratios of signal amplitudes and presented in section 1 were tested: TONAR, Nasalance, HONC and N/V. The calibration constant k used in the computation of the HONC measure is supposed to represent the value of the ratio on maximally nasalized speech productions. For each speaker, the value of k was therefore defined as the ratio $Arms(N_a)/Arms(V_a)$ during the production of a sustained /m/.

2.6. Measurements based on RMS amplitude difference

A drawback of ratios between oral and nasal signals is that they exhibit high values on segments with very low energy on both signals used in the ratio calculation, which is typically the case in voiceless plosives (see Figure 4). Such segments as well as silent parts can therefore be incorrectly interpreted as nasalized.

Some authors indicate that the nasality ratio should be applied only to voiced segments [18]. Indeed, the unwanted inclusion of low-energy parts of signal may have drastic consequences. Since nasality indices based on amplitude differences are expected to be less sensitive to such cases, we introduce and evaluate the new measures LND (Laryngeal Nasal Difference) and OND (Oral Nasal Difference).

The LND difference is computed from the RMS amplitudes of the nasal and laryngeal accelerometers, as:

$$LND = Arms(kV_a \cdot N_a) - Arms(V_a) \quad (5)$$

In order to compensate for the possible imprecision of the gain calibration process, the nasal signal is rescaled to the level of the laryngeal signal. The constant kV_a is therefore set to the ratio between the amplitude of the laryngeal accelerometer during the production of a sustained /a/ (considered as the most oral segment) and that of the nasal accelerometer during the production of a sustained /m/ (considered as the most nasal segment).

The OND difference is computed from the RMS amplitudes of the nasal accelerometers and that of the oral microphone, as:

$$OND = Arms(kV_m \cdot N_a) - Arms(V_m) \quad (6)$$

Similarly to the kV_a constant used in LND calculation, kV_m is set to the ratio between the amplitude of the oral microphone during the production of a sustained /a/ and that of the nasal accelerometer during the production of a sustained /m/. Note that since the units of the N_a and V_m signals are not homogenous, kV_m is expressed in $m^3 \cdot Pa^{-1} \cdot s^{-2}$.

An expected advantage of LND/OND over ratio-based measures is that it makes the definition of a nasality threshold more straightforward. Indeed, values of LND/OND are expected to be positive on nasal segments. LND/OND are supposed to be less sensitive to the inclusion of low-energy parts of signal such as voiceless consonants, compared to ratio-based measures (see Figure 4). However, it should be noted that the interpretation of the value zero as mid-oral or mid-nasal would not be fully consistent with the notion of nasality. In order to retain only values corresponding to nasalized portions of the signal, negative values are therefore set to 0.

2.7. Phonetic labeling

The phonetic labels and boundaries were manually determined by a trained phonetician, for all 146 sustained vowels and utterances produced by each of the 4 speakers.

3. Results and discussion

3.1. Distinctiveness of nasality measures

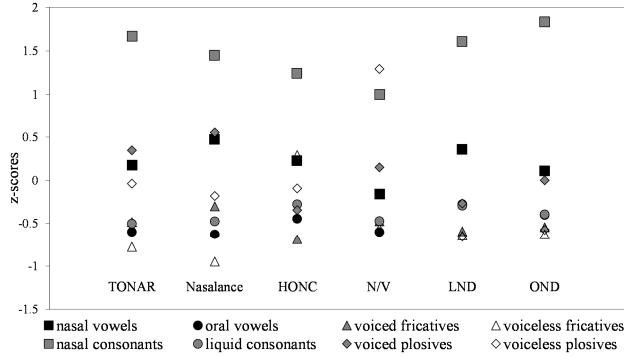


Figure 2: Distribution of the mean z-score values of the 6 calculated nasality measures on phonetic classes (all 4 speakers pooled).

Nasality measures values can be mainly interpreted relatively to a nasality threshold depending on the speaker. In order to account for this variability, raw values of these measures were transformed into z-scores for each speaker.

As illustrated on Figure 2, the distribution of nasality values on phonetic classes varies according to the nasality measure considered. For every measure except N/V, nasal consonants are attributed the highest values. The particularly high values attributed to voiceless plosives by N/V underline the urge to apply ratios only on voiced segments, as recommended by [18]. Though less obvious, such overestimation of nasality in voiceless plosives can be observed in the three other ratios. Figure 2 suggests that only LND and OND attribute a high degree of nasality for nasal consonants and nasal vowels, while voiceless consonants are ranked as the less nasalized. The distribution of the degree of nasality on phonetic classes for LND and OND is therefore more consistent with the sonority scale of nasality [22].

Figure 3 shows the average z-score values of the six measures for nasal vs. oral consonants produced by the four speakers.

The ability of each measure to distinguish nasal vs. oral segments was evaluated separately for consonants and vowels by ANOVAs with the z-transformed measure as dependent variables, and the speaker and the nasality status of the segment as fixed factors. The effect of nasality was found highly significant ($p < .001$) for all six measures, for both consonants and vowels.

Since the speaker*nasality interaction was also found highly significant, separate analyses were performed for each speaker, using ANOVAs with only nasality as fixed factors. For consonants, the effect of nasality was found highly significant ($p < .001$) for all measures and all speakers. For vowels, though the effect of nasality was found highly significant for most speakers and measures, LND and OND do not distinguish oral vs. nasal vowels produced by speaker M2.

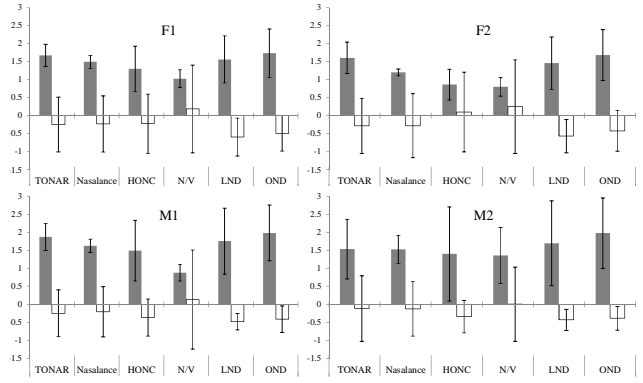


Figure 3: Z-transformed values of the six measures for nasal (grey) vs. oral (white) consonants produced by the four speakers. Error bars = standard deviation.

3.2. Relationships between nasality measures

Table 1 summarizes the mean correlation coefficients computed on voiced segments between the six nasality measures for all the four speakers. Unsurprisingly, the highest correlation is found between TONAR and nasalance, which are calculated from the same signals. Measures computed from the nasal and laryngeal accelerometers appear to be moderately correlated to measures computed from the nasal accelerometer and the oral microphone (HONC vs. N/V and LND vs. OND). The rather high correlations between N/V and measures obtained from both microphones show that the nasal microphone and the nasal accelerometer capture similar phenomena. Indeed, the correlation between the RMS amplitudes of Na and Nm is $r = .874$ on voiced segments. Comparatively, the correlation between the RMS amplitude of the oral microphone signal and that of the laryngeal accelerometer is only $r = .676$.

Table 1: Mean correlations between the six nasality measures on voiced segments (four speakers pooled)

	TONAR	Na.	HONC	N/V	LND
Nasalance	.990				
HONC	.531	.530			
N/V	.822	.811	.669		
LND	.478	.474	.707	.534	
OND	.672	.657	.606	.719	.797

Figure 4 shows the acoustic signal recorded from the oral microphone and the variations of the six measures for the word 'cornflakes' (borrowed from English and produced as /kɔʁnflɛks/) uttered by speaker F1. The LND and OND measures seem to best describe the time course of nasality in the production of nasal segments compared to ratio-based measures. This observation is consistent with the qualitative analysis of measures variations on other utterances and speakers. Positive LND and negative OND values (before negative values cancellation) observed for the oral vowel /ɛ/ illustrate the difficulty of appropriately defining the nasality threshold from calibration values. Indeed, this vowel is detected as nasal in the first case, and as non-nasal in the second case. Defining nasality ranges depending on the phonetic class of analyzed segments might therefore be a more realistic objective than the definition of a unique multipurpose nasality threshold.

4. Conclusions

All evaluated nasality measures from accelerometers and microphones successfully separate nasal vs. oral consonants. However, voiceless segments have to be discarded before applying ratio-based measures. New difference-based measures LND and OND require an accurate definition of the calibration values, which have a direct impact on the nasality thresholds or ranges. Indeed, in our preliminary evaluation of these new measures, the chosen calibration values did not enable an appropriate separation of nasal vs. oral vowels for every speaker. Calibration values will have to be fine-tuned by matching difference-based measures with more direct physiological nasality data (velum opening measured using a PNG [23] and nasal pressure measures).

Correlations between measures indicate that different experimental setups can hardly be considered as equivalent. In order to determine minimal setups enabling a reliable measure of nasality, measures obtained from different setups have to be compared to physiological measures.

According to the current state of knowledge on nasality phenomena, difference-based measures appear to better describe the time course of nasality, compared to ratio-based measures. Reference physiological measures will also enable an objective evaluation of the appropriateness of nasality measures to describe the time-course of nasality.

5. Acknowledgements

The authors would like to thank Shinji Maeda and Jacqueline Vaissière for fruitful discussions and advice.

6. References

[1] Fujimura, O., "Analysis of nasal consonants." JASA 34(12), 1865-1875, 1962.
 [2] Amelot, A., "Etude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français". PhD dissertation, Phonetics, University Sorbonne Nouvelle, 2004.
 [3] Vaissière, J., "Prediction of Velum Movement from Phonological Specifications", *Phonetica*, 45, pp. 122-139, 1988.
 [4] Rossato, S., "Du son au geste, inversion de la parole : le cas des voyelles nasales", PhD dissertation, Signal Image Parole Télécoms, Institut National Polytechnique de Grenoble, 2000.
 [5] Benguerel, A.P., Hirose, H., Sawashima, M., and Ushijima, T., "Velar Height and its Timing in French: a Fiberscopic Study", *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, 9, pp. 67-78, 1975.
 [6] Ohala, J.J., "Monitoring soft-palate movements in speech", *Journal of the Acoustical Society of America*, 50, (140), 1971.

[7] Hyde, S.R. "Nose trumpet: apparatus for separating the oral and nasal outputs in speech", *Nature*, 219, (5155), pp. 763-765, 1968.
 [8] Feng, G., and Kotenkoff C. "Vers un nouveau modèle acoustique des nasales basé sur l'enregistrement bouche - nez séparé", 25e JEP, Fez (Morocco), pp. 213-216, 2004.
 [9] Hirano, M., Takeuchi, Y., and Hiroto, I. "Intranasal sound pressure during utterance of speech sounds", *Folia Phoniatrica*, 18 (5), pp. 369-381, 1966.
 [10] Schwartz, M.F. "The acoustics of normal and nasal vowel production", *Cleft Palate J*, 5, pp. 125-140, 1968.
 [11] Fletcher, S.G. "Theory and instrumentation for quantitative measurement of nasality", *Cleft Palate J*, 7, pp. 601-609, 1970.
 [12] Dalston, R.M., Warren, D.W., and Dalston, E.T. "Use of nasometry as a diagnostic tool for identifying patients with velopharyngeal impairment", *Cleft Palate Craniofacial J.*, 28, (2), pp. 184-188; discussion 188-189, 1991.
 [13] Stevens, K.N., Kalikow, D.N., and Willemain, T.R. "A miniature accelerometer for detecting glottal waveforms and nasalization", *J. Speech Hearing Res.*, 18, (3), pp. 594-599, 1975.
 [14] Tronnier, M., "Nasals and Nasalization in Speech Production with Special Emphasis on Methodology and Osaka Japanese", S-22100 Lund, 1998.
 [15] Stevens, K.N., Nickerson, R.S., Boothroyd, A., and Rollins, A.M. "Assessment of nasalization in the speech of deaf children", *J. Speech Hearing Res.*, 19 (2), pp. 393-416, 1976.
 [16] Horii, Y. "An accelerometric approach to nasality measurement: a preliminary report", *Cleft Palate J*, 17, (3), pp. 254-261, 1980.
 [17] Horii, Y. "An accelerometric measure as a physical correlate of perceived hypernasality in speech", *J. Speech Hearing Res.*, 26, (3), pp. 476-480, 1983.
 [18] Redenbaugh, M.A., and Reich, A.R. "Correspondence between an accelerometric nasal/voice amplitude ratio and listeners' direct magnitude estimations of hypernasality", *J. Speech Hearing Res.*, 28 (2), pp. 273-281, 1985.
 [19] Lippmann, R.P. "Detecting nasalization using a low-cost miniature accelerometer", *J Speech Hear Res*, 24, (3), pp. 314-317, 1981.
 [20] Montagu, J. "Analyse acoustique et perceptive des voyelles nasales et nasalisées du français parisien". PhD dissertation, Phonetics, University Sorbonne Nouvelle, 2007.
 [21] Vaissière, J., Honda, K., Amelot, A., Maeda, S., Crevier-Buchman, L. "Multisensor platform for speech physiology research in a phonetics laboratory", *J. Phonetic Society of Japan*, in press.
 [22] Clements, G.N., and Osu, S. "Nasal Harmony in Ikwere, a Language with no Phonemic Nasal Consonants", *J. African Lang. and Linguistics*, 26, (2), pp. 165-200, 2005.
 [23] Amelot, A., "Dispositifs d'imagerie pour l'observation de l'activité vélo-pharyngée". In Marchal, A.; Cavé, C. (Ed.), *L'imagerie médicale pour l'étude de la parole*, Cachan: Lavoisier, pp. 65-82, 2009.

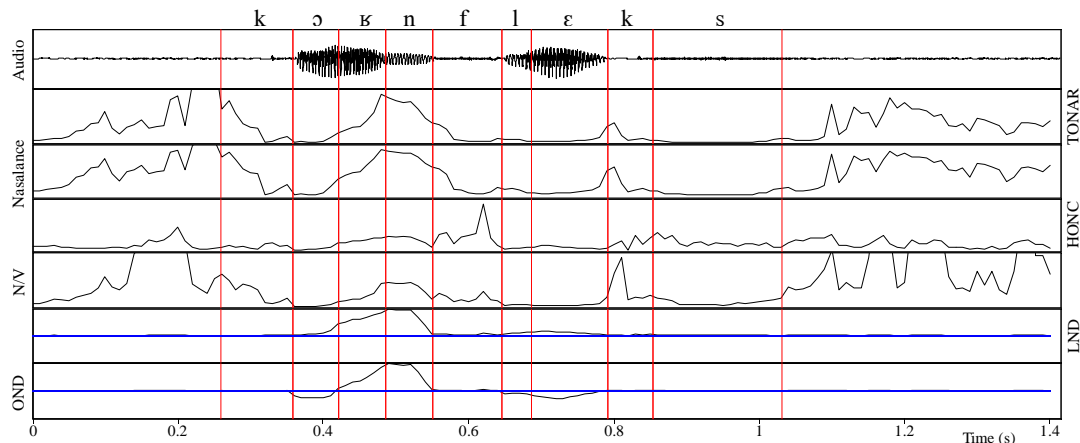


Figure 4: Variation of the values of the six nasality measures during the production of the word cornflakes (borrowed from English and produced as /kɔʁnflɛks/) by speaker F1. LND and OND values are plotted without negative values removal.