Bermúdez Cervantes, Eric; García Quintana, David, dir. Acoustical analysis of hypernasal voice : towards a clinical prediction test. 2019. 30 pag. (889 Grau en Logopèdia)

A study develops a clinical test to evaluate hypernasality in voice

Barcelona, May 17, 2019. Graduation research work by a Speech-Language Pathology student at Universitat Autonoma de Barcelona (UAB) provides a novel tool for the clinical evaluation of so-called *hypernasal voice*, the result of a frequent alteration in the regulation of the soft palate during speech.

Hypernasal voice is less intelligible, as a result of attenuated volume and distorted acoustical clues. Speech-Language therapists have intervention strategies to improve soft palate function and reduce hypernasality. However, hypernasal voice is often difficult to assess in clinical practice, as is currently based on perceptual evaluation by the therapist. Subtle cases pass undetected.

Previous studies on nasalized vowels showed little coincidence on the acoustical traits of nasality. The experimental approach of the UAB team compared pairs of vowels from 40 different volunteers, each pair consisting of a purely oral production and a nasalized production. The recordings were evaluated by 5 different expert judges for the degree of perceived nasality. The 80 voice productions were thus ranked according to the average score of perceived nasality. The samples were then thoroughly analysed using advanced acoustic analysis tools. In this manner, the study identifies two significant and consistent acoustic alterations due to nasal resonances, robustly specific for hypernasal vowels.

The investigators finally went on to explore whether the two acoustic parameters might provide a test for the reliable, objective evaluation of hypernasality. They generated a simple, automatized index that indeed shows a robust clinical value, with high sensibility -the vast majority of healthy cases are identified as such-, and high specificity -the vast majority of healthy cases are identified as such. Furthermore, the test is able to grade the degree of severity.

In all, the study provides for the first time an objective test with a robust clinical value. The scientist expects to distribute the test among the clinical community in the foreseeable future.

For more information, contact Eric Bermudez at Eric.Bermudez@e-campus.uab.cat

Executive summary:

Voice problems are a demanding issue that affects an important sector of the population every year. Some authors state that the amount of people that can have a voice problem at least one time in their lives are up to 29,9%.

This gives importance to the accurate diagnosis of the voice, in order to give response to the issues when the problem is starting, to avoid the consequences that involve a voice problem (psychological, emotional, economic and communicative).

Currently, there are some voice problems that are diagnosed perceptively. This is done that way because the lack of objective tools designed for Speech and Language Pathologists. But the issue with this kind of diagnose is that is subjective and thus, there is always a certain degree of variability between professionals, making difficult to arrive to any consensus.

One of the problems that are diagnosed perceptually is hypernasality, understood as the excessive amount of nasal coupling during normal speech. This issue alters the acoustic characteristics of the speech and makes it more difficult to be understood at the same time it makes the voice less efficient, with the risk of forcing the voice in order to have a better performance.

The current research has been done with the purpose of check if there was possible to give to the agents involved a tool to assess hypernasality.

The agents involved can be from a wide spectrum of disciplines that go from Speech and Language Pathologists that need a tool to diagnose voice problems, researchers that want to do their job in the field of nasality acoustics, voice teachers that want a tool in order to teach hypernasality for scenic reasons or singers who want to check when they suspect their voice is not behaving as usual.

The objectives were:

1. Find the acoustical keys that make a sound hypernasal.
2. Check if those keys correlate with the hypernasality perceptions.
3. If they correlate, check if there is possible to develop a tool to assess nasality.

In order to find the acoustical keys, voices from 47 healthy subjects were recorded. For each subject a pair were recorded, one sample with the sound [a] oral (with little or 0 degree of nasality) and one sample with the sound [a] nasalized (at the same pitch and intensity), making a total sample size of 94 elements.

Those samples were anonymized and scrambled, with no information about which sample was nasal or oral. Then 5 judges listened the voices and gave them a total score of perceived nasality, on a scale from 0 (no nasality) and 10.

With the samples scored, the acoustical analysis could be done and then only the voices that fit the inclusion and exclusion criteria were picked.

The inclusion and exclusion criteria were:

1. If one sample of each pair had only 1 individual score that was +/- 3 points away of its mean, was assumed conflictive and **all the pair were excluded.**
2. For a pair to be included, the nasal sample of the pair needed to have a mean score at least 4 points higher than its oral cognate.

The acoustical analysis consisted on a comparison of the relative energy in the region between 0 and 3,300 Hz, checking the differences and their significance between the nasal and the oral group.

The results show there are statistically differences between the groups in the areas of the low nasal murmur, on frequencies between 100 and 500 Hz; and in the first formant region, on frequencies between 700 and 1,300 Hz. Using the critical points given by these differences, there are different clinical performances, ranging from 65% to 80%.

Evaluating these results, a combined index has been designed, using a multiple linear regression analysis in order to design it.

The data analysed with the combined index shows that there are differences between the nasal and the oral groups of samples with a significance up to 99.9%.

		mean	**standard deviation**	**t value**	**significance**
Combined index	nasal	6.265	2.592	5.331	99.9%
	oral	2.915	1.087		

Table I. *Summary of the results of the t-test on the combined nasality index values for the nasalized and the oral vowels.*

The combined algorithm also correlates with the perceptual degree of hyper nasality, showing a strong correlation with a Pearson Correlation Coefficient up to 0.74.

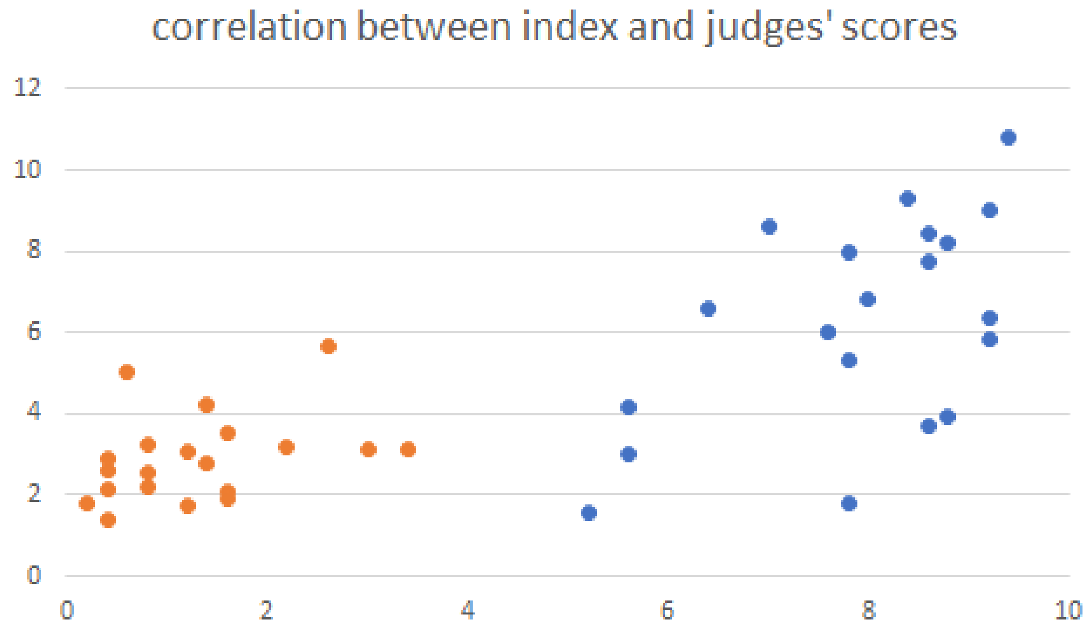

Figure 1. *Scatterplot showing the two groups of values for the* ***combined nasality index*** *of oral (orange) and nasalized (blue) vowels. Y axis, combined nasality index; X axis, 0-10 mean perceptual evaluation scores.*

Using the combined index, an overall clinical efficiency of 85% is scored, scoring both 85% in specificity and sensitivity.

	Combined index > 3.61 dB	**Combined index < 3.61 dB**	**totals**			
nasals	17	3	20	85%	% sensibility	
orals	3	17	20	85%	% specificity	
totals	20	20	40		**Positive predictive value =**	85%
					Negative predictive value =	85%
					Overall Clinical Efficiency =	85 %

Table II. *Confusion matrix using the critical point obtained as a result of the combined index.*

After reviewing the results, it can be concluded that there are some hypernasality markers that are possible to find in an acoustic study, being the ones exposed before.

With these findings, there is an opportunity to carry on further research in order to precise the results, consider narrower bands to avoid interferences and develop and validate a clinical test that could be used to assess nasality objectively, eliminating the perceptual gap that exists currently.

Acoustical analysis of hypernasal voice: Towards a clinical prediction test

Hypernasalization of oral voice sounds results in a decrease of energy, due to nasal attenuation, and in loss of acoustical clues with the subsequent loss of intelligibility. Whereas severe and moderate hypernasality is easily perceived auditorily, mild cases are difficult to assess. Therefore, there is a need for an objective, acoustic test for clinical evaluation of hypernasal voice. In this research work we identify two acoustic markers that specifically and consistently correlate with nasalization of oral vowels. We have used these two markers to generate a robust combined acoustic index that predicts nasalization with high sensitivity and specificity, and an overall clinical efficiency of 85%.

Graduation Thesis
Eric Bermúdez Cervantes
Speech and Language Pathology Degree
Universitat Autònoma de Barcelona
2019

1. Introduction:

1.1. Motivation:

Voice dysfunctions affect a significant portion of population every year. According to Roy et al. (1), 29,9% of people will endure a voice disorder in any moment of their life. Because of the complexity of anatomic structures and movements that intervene in the voice production process, such disorders are caused by very different reasons.

A group of dysphonic alterations are the consequence of the a misfunction of the soft palate. Such alterations may result in either an excess of nasality (when the soft palate doesn't close well), or a decrease of nasality when it fails to open adequately. Depending if there is an excess or a decrease of nasality, those voices are called hypernasal or hyponasal voices, respectively (2).

When it comes to hypernasal voices, the inefficient closing of the soft palate, results in an oral sound affected by some degree of nasality, which has two relevant consequences. First, a **loss of energy** as passage through the nasal tract greatly attenuates the sound with the subsequent efficiency loss (3)**.** Second, a **loss of the definition of the acoustical clues** due to the interference of the oral resonances with the nasal resonances, leading to a loss of intelligibility (4)**.**

Worldwide, oral clefts in any form occur in about one in every 700 live births (World Health Organization [WHO], 2001). Furthermore, this figure excludes under – clinical cases, therefore underestimating the prevalence of the alteration.

Speech and Language therapists evaluate nasality perceptively. To this respect, studies stress the importance of perceptive training, exposing differences in diagnosis among professionals, depending on their training (5). In addition, whereas severe or moderate hypernasal voice is easily evaluated perceptually by Speech and Language Therapists, evaluation fails in subtle cases. It turns even more difficult when additional alterations are present in the evaluated voice.

.

With the above considerations in mind, the **motivation** of this research is to identify objective acoustical parameters that allow **the objective clinical evaluation of the nasality in voice**.

1.2. Introduction and current state of the research:

The vocal tract is a variable resonator that selects particular frequencies in the glottal sound, to produce different voice sounds (6). The vocal tract is a resonator tube with a closed end at the glottis and an open end. The open end is the mouth opening for oral sounds such as [a], and the nostrils for nasal sounds such as [m]. Acoustic analysis of nasal sounds reveals specific nasal formants (poles) and zeroes (anti-formants) due to resonances and anti-resonances in the nasal cavity (7). Whereas no consensus is found for the frequency of the nasal poles for a given nasal sound among different individuals, a characteristic low frequency resonance, called nasal murmur, is found in all nasal sounds in the 40 – 270 Hz region. (7).

During the production of nasal sounds the velopharyngeal passage is voluntarily open. However, in some anatomical alterations, such as cleft palate, different degrees of oral voice nasalization results from undue opening of the velopharyngeal passage. In 1993, Warren et al. determined that openings as little as 0.10 cm^2 were enough for listeners to perceive nasality: (8). **Nasalized voice can be distinguished perceptively and therefore acoustically from oral voice.**

Information relevant to pathological hypernasal voice may be drawn from acoustic studies focused in naturally occurring **nasal vowels** (such as [ɛ̃] in the French word *train*) and in **nasalized vowels**, i.e. vowels modified by coarticulatory interference of a neighboring nasal sound, as is the case of the [ɪ] in the English word *train*, compared with oral vowels (such as the [ɪ] in the English word *trait*). Different works report the occurrence of nasal poles and zeros in nasalized vowels with respect to oral vowels (9), (10), (11), (12), (13). However, there is little coincidence among pole frequencies reported by different authors, except for the lowest pole, compatible with the nasal murmur in nasal sounds (Table I).

	Feng & Castelli (11)	**Chen** (9)	**Styler et al.** (10)
Pole 1 (P0)	250 – 300 Hz	250 – 450 Hz	250 – 450 Hz
Pole 2 (P1)	400 – 450 Hz	~1000 Hz	790 – 1100 Hz
Pole 3	650 – 700 Hz		~1250 Hz
Pole 4	850 – 900 Hz		
Pole 5	1200 – 1300 Hz		
Pole 6	1800 – 2200 Hz		
Pole 7	2800 – 3300 Hz		
Pole 8	~4000 Hz		

***Table I**, Frequencies of the nasal poles in three different studies.*

In addition, Feng and Castelli describe a zero in the 1,200 Hz region and a loss of energy above 2,000 Hz in their measures of French nasal vowels (11).

Interestingly, despite the lack of consistency, in frequencies of poles and zeros in nasalized vowels, **the nasal coupling appears to cause more definite modifications in the vowel spectrum**, such as **reduction in amplitude of the first vocal formant (F1),** broadening the bandwidth of F1, **shifting F1 upwards in frequency**, and a relative **strengthening of the spectrum around 250 Hz**, the region of the nasal murmur mentioned above (12–17).

Finally, a limited number of studies have been carried on the acoustics of pathologically nasalized (hypernasal) voice. A seminal acoustic study with pathological hypernasal voices by Kataoka et al. (18) reported areas with decreased and areas with increased energy (18). although the frequency regions were poorly defined. More recently, a more detailed study by Lee et al. (19) unveil that hypernasal voices in patients with a cleft palate show increased amplitude in frequency bands centered at 630 Hz, 800 Hz and 1,000 Hz, and a decrease in the frequency band centered at 2,500 Hz, again likely to be the result of interference of the oral signal with the nasal poles and zeros.

In summary, despite the poor consensus seen among different studies, the available data indicate that vowel nasalization results in marked changes in energy in particular frequency regions. **Such regions could perhaps serve as nasality markers**.

2. Objective and Hypothesis:

The objective of this research is to identify acoustical markers for the objective evaluation of pathological hypernasal voice.

The hypotheses driving this work are:

- Among the diverse acoustic alterations in hypernasal voice, it will be possible to identify one or more markers common to hypernasal vowels among different individuals.
- Some of those markers will correlate with perceived hypernasality.
- Among those markers, some will be of use to build an objective evaluation test of clinical value.

3. Methodological proposal:

Samples and recording

Voices from 47 healthy subjects were obtained. Subjects were selected among volunteers able to produce nasalized vowels on demand. Thus, pairs of nasalized and oral vowels were recorded.

Voices were recorded in a silent environment, using a plain response microphone SAMSOM C01UPRO, placed in front of the subject, at a distance between 10 and 20 cm, half-way from mouth and nostrils to record the sound resulting from nasal and oral resonance.

Software Praat (23), 64-bits version 6.0.46, 2019, was used as recording software, at a sampling frequency of 44,100 Hz.

After some rehearsals, volunteers were requested to produce an oral [a] vowel (closed soft palate), then a nasalized [a] vowel with the soft palate half-opened.

Afterwards, all the samples were leveled for intensity at 70 dB. The recordings were anonymized and scrambled, with no information on whether the sample was either nasalized or oral: *Voice1, Voice2, ...*

Perceptual evaluation

The samples were independently evaluated by 5 judges, all of them well-regarded professionals with expertise in voice clinics or voice research. Three of the judges are Speech and Language Pathologists, all with clinical trajectory and teaching voice. One of the judges is a Professor in Phonetics with expertise in clinical phonetics. One of them is a Professor in Medical Biophysics with expertise in voice acoustic analysis.

The judges were requested to evaluate the perceptual degree of nasality in the samples within a 0 – 10 scale, where 0 is least nasality perception and 10 is maximum nasality perception. The samples were evaluated in the same order in all cases. The judges had no contact with each other.

Inclusion and exclusion criteria

Upon evaluation by the judges, the voices underwent the following inclusion and exclusion criteria. The criteria were designed to ensure that the voice pairs **robustly and consistently** contained one vowel perceived as more oral and one vowel more nasal by all the judges.

1. Samples which one individual judge's score was more than +/-3 points away from the mean, were assumed to be conflictive and the corresponding oral/nasalized pair was **excluded**.
2. For a pair to be **included**, the nasal sample of the pair must have a mean score at least 4 points higher than its oral cognate.

In all, 20 out of the initial 47 pairs of voices were included in the study.

Acoustic analysis

The samples were analyzed with the acoustical analysis software Praat (23), 64-bits version 6.0.46, 2019.

In nasalized vowels the nasal coupling is expected to cause spectral alterations in the oral wave, resulting in increases and decreases in energy at different frequency regions. Thus, strengthening of the spectrum around 250 Hz due to interference with nasal murmur, a reduction in amplitude of the first vocal formant (F1) due to interference with nasal zeroes, and an upward frequency shift of F1 due to interference with nasal poles, have been described for nasalized vowels in speech samples (11-17, 19).

Based on the above information, we investigated which of these alterations were **consistently** and **significantly** present in our nasalized pairs with the goal to identify markers of hypernasal voice. We developed algorithms (Praat scripts) to analyze energy gains or losses in different frequency regions in nasalized vowels with respect to their oral pair. In all cases, the algorithms subtract the Intensity Level (dB) in the region of interest to the Intensity Level in a broader region of reference. Because dB are logarithmic, subtraction is indeed a **ratio** between the energies of the two regions, which makes it independent of the intensity of the emission.

- One of the algorithms explores whether there is increased energy in the 100 – 500 Hz region in nasalized vowels compared with their oral counterpart, as suggested in some of the above referenced works with nasalized vowels, likely due to the nasal murmur resonance (also termed pole P1 or P0 depending on the authors). The index in the algorithm results from subtracting the Intensity Level in this region to the Intensity Level in the 0 – 2,000 Hz region. Therefore, the prediction would be that the index should be lower in nasalized vowels than in their oral counterparts.

- Another algorithm explores whether there is a loss of energy in the 700 – 1,300 Hz region in nasalized vowels compared with their oral counterpart, as suggested in some of the above referenced works with nasalized vowels, likely due to a nasal zero in this region. The index in the algorithm results from subtracting the Intensity Level in this region to the Intensity Level in the 0 – 2,000 Hz region. Therefore, the prediction would be that the index should be higher in nasalized vowels than in their oral counterparts.

- Another of the algorithms explores whether there is a gain of energy in the 1,600 – 2,700 Hz region in nasalized vowels compared with their oral counterpart, as suggested in some of the above referenced works with nasalized vowels, likely due to the contribution of nasal poles in this region. The index in the algorithm results from subtracting the Intensity Level in this region to the Intensity Level in the 0 – 3,300 Hz region. Therefore, the prediction would be that the index should be lower in nasalized vowels than in their oral counterparts.

- One last algorithm explores whether there is a loss of energy in the 2500 – 3300 Hz region in nasalized vowels compared with their oral counterpart, as suggested in some of the above referenced works with nasalized vowels, likely due to nasal zeros in this region. The index in the algorithm results from subtracting the Intensity Level in this region to the Intensity Level in the 0 – 3,300 Hz region. Therefore, the prediction would be that the index should be higher in nasalized vowels than in their oral counterparts.

Statistical analysis

The acoustic data generated by the analysis algorithms were analyzed with the Kolmogorov – Smirnov test in order to determine whether they follow a **normal distribution**, so that parametric statistical analysis can be performed (www.socscistatistics.com).

Values lower than Q1 – 1.5·IQR or higher than Q3 + 1.5·IQR were assumed to be **outliers** and discarded from the analysis (Q1 is the value of the 1st quartile, Q3 is the value of the 3rd quartile, IQR = Q3 – Q1).

In order to check whether the perceptual scoring among the different judges was consistent, the **inter-rater reliability** was evaluated by means of an ICC (Intraclass Correlation Coefficient) test. ICC measures which fraction of the total variance is due to differences between judges. First, a two-factor (scores / judges) ANOVA test without replication (one only score per voice per judge) was carried out on the 5 series of perceptual scores. From the ANOVA results an ICC was calculated assuming Two-way random effects, absolute agreement, single rater/measurement: ICC(2,1,absolute agreement) = (MS voices – MS error)/((MS voices + ((k – 1) * MS error) + ((k/n)*(MS judges – MS error)), where MS = Mean Squares, an estimate of the variance, k is the number of judges, and n is the number of voices evaluated in each series.

The **significance of the differences** for a given acoustic parameter between the oral and the nasalized groups was evaluated statistically with a t-test (www.mathportal.org).

To combine more than one predictor in a single test, **multiple linear correlation analysis** was used. The data from two algorithms that individually correlate linearly with perceived hypernasality were entered an independent variable, and the perceptual evaluation mean score as dependent variable (www.socscistatistics.com).

Confusion matrixes were generated to analyze the clinical value of the findings, assessing the sensibility and the specificity of the potential predictors, as well as the clinical efficiency.

For a test to be of clinical validity it must accomplish a minimum set of conditions. First, it must be **Sensitive**, detecting as may true positives as such and not missing them as false negatives. Second, it must be **Specific**, not scoring true negatives as (false) positives. From these data, three additional parameters can be calculated, the **Positive Predictive Value** (PPV), which is the percentage of true positives that the test identifies as such; the **Negative Predictive Value** (NPV), which is the percentage of true negatives that the test identifies as such; and the **Overall Clinical Efficiency**, the percentage of true positives and true negatives that the test identifies as such.

Ideally, a perfect clinical test should be 100% sensitive and specific, resulting in PPV, NPV and Overall Clinical Efficiency of 100%. In reality, though, a compromise threshold must be selected, based on the confusion matrix, that maximizes the Overall Clinical Efficiency and one of the two: either Sensitivity or specificity, depending on what is more relevant for the clinical case of interest.

	The test predicts as nasal	**The test predicts as oral**	**Totals**		
Real nasals	Number of True Positives	Number of False Negatives	Number of real nasals	**Sensitivity =**	TP / real nasals
Real orals	Number of False Positives	Number of True Negatives	Number of real orals	**Specificity =**	TN / real orals
	TP+FP	FN+TN	Number of all samples	**Positive Predictive Value =**	TP / (TP+FP)
				Negative Predictive Value =	TN / (TN+FN)
				Overall Clinical Efficiency =	(TP+TN) / all samples

4. Ethical considerations:

According to the current law (*Real Decreto 1716/2011 de 18 de noviembre*, and *Real Decreto 1720/2007 de 21 de diciembre*), regarding personal information treatment and the storage of personal and clinical data, the voice donors participating in the study were informed on the scope of the study, what would their contribution be, how their samples would be handled anonymously, and how would they be analyzed and by whom. In addition, they were informed they could withdraw from the study at any time and request their recordings to be deleted. Because of the low ethical implications of the study, the agreement was oral.

All the samples were labelled with a code consisting of the first letter of the name and surname of the participants, plus a suffix indicating whether it was an oral sample or a nasalized sample. No personal information regarding the participants was recorded.

All the recordings were saved in a personal computer, with a backup copy in a personal Google Drive account. The recordings were embedded in a Google form for the judges to listen without downloading. The files were anonymized as described above. The judges were asked not to attempt to download the recordings.

5. Results:

The raw results with the perceptual mean scores and the data from the 4 acoustic analysis scripts for the 40 included voice pairs are shown in the Annexes section.

5.1. Inter-rater reliability

ANALYSIS OF THE VARIANCE						
SOURCE OF THE VARIATIONS	***Sum of the squares***	***Degrees of freedom***	***Mean of the squares***	***F***	***P - value***	***Critical value for the F***
ROWS	2402.38	39	61.59948718	31.01145642	1.10452E-55	1.476986487
COLUMNS	2.13	4	0.5325	0.268080163	0.898143347	2.429625042
Error	309.87	156	1.986346154			
Total	2714.38	199				
ICC	0.85943543					

Table II. *Results of the ANOVA test carried out with a significance level of 95%.*

As can be seen in Table II, from the results from the ANOVA, an ICC = 0.859 is calculated, indicating an **excellent inter-rater reliability** in the perceptual evaluations.

Table 2 summarizing the ANOVA test carried out with a significance level of 95%. As can be seen, the results from the ANOVA show an interclass correlation (ICC) of 0.859, that means a strong correlation between the scores applied by the judges.

5.2. Normality of the data

Normality of the data series was assessed by means of a Kolmogorov – Smirnov test. Mean perceived nasality scores and the results from each of the 4 acoustical analyses were analyzed. The test indicates that the 5 data series fit a normal distribution. Therefore, parametric statistical tests can be used for subsequent analysis.

5.3. Analysis of the acoustical parameters as potential markers of nasalization

The results from the 4 acoustic analyses described above, were compared in the oral vs the nasalized voices. As summarized in Table III, two of the acoustical alterations analyzed are significantly, confidently different between nasalized and oral samples. Nasalized vowels consistently show increased energy in the 100 – 500 Hz region and attenuated energy in the 700 – 1,300 Hz, compared with their oral cognates. To the best of our knowledge, this is the first time ever that potential acoustic markers of nasalization have been identified with this degree of reliability.

None of the two other parameters shows significant differences.

		mean	standard deviation	t value	significance
acoustic index 1 (algorithm 1)	nasal	7.315	4.870	-2.067	95.0%
	oral	10.025	3.265		
acoustic index 2 (algorithm 2)	nasal	6.095	2.874	5.277	99.9%
	oral	2.395	1.253		
acoustic index 3 (algorithm 3)	nasal	14.605	7.564	-0.884	not significant
	oral	16.390	4.929		
acoustic index 4 (algorithm 4)	nasal	24.59	7.952	1.912	not significant
	oral	20.900	3.350		

Table III. *Summary of the t-test analyses on the different acoustic regions analyzed, comparing nasalized vs oral productions.*

5.4. Identification of discrimination thresholds

So far, two acoustic markers of nasalization have been identified. We next investigated whether discrimination threshold values could be identified with a clinical test as goal.

The data were represented as box plots with cat whiskers. In the case of excluding boxes, the value of Q1 of one group, Q3 of the other group, or an intermediate value, provide robust discrimination thresholds.

This is the case of the acoustic index 2, energy attenuation in the 700 – 1,300 Hz region, calculated by algorithm 2, as shown in Figure 1. Therefore, **this acoustic index is a robust marker for nasalized vowels**. From now on, we term it **nasality acoustic marker 2**.

As discrimination threshold values either 3.650 dB (the Q1 value of the nasal box), 2.875 dB (the Q3 value of the oral box), or an intermediate value, can be selected.

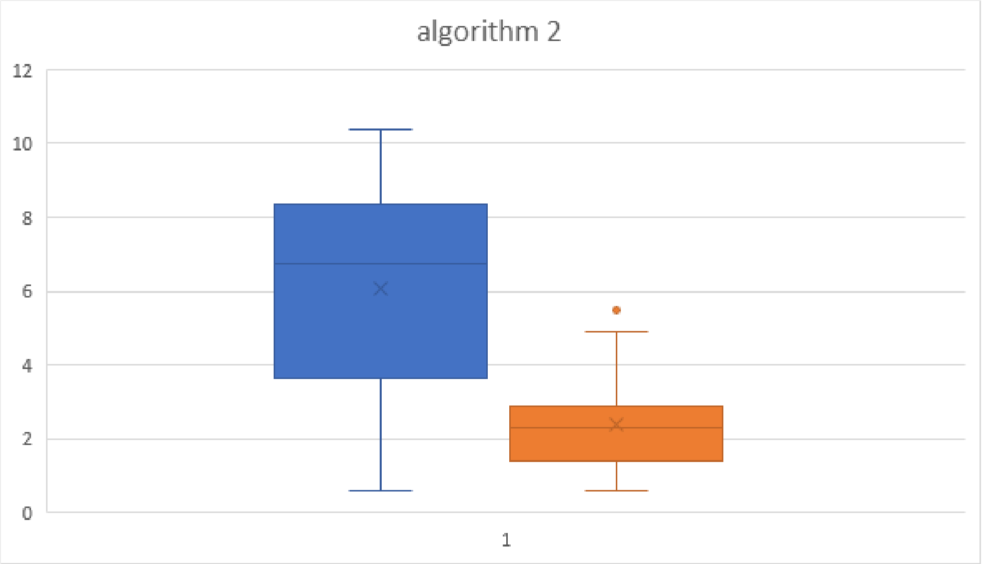

Figure 1. *Box plots and cat whiskers of **nasality acoustic index 2**, the relative energy in the 700 – 1,300 Hz region in nasalized (blue) vs oral (orange) vowels. The Y axis are dB as calculated by algorithm 2.*

In the case of the acoustic index 1, energy increase in the 100 – 500 Hz region, calculated by algorithm 1, as shown in Figure 2, the two boxes are overlapped. However, a threshold can still be defined that excludes the boxed values for oral vowels. Despite such threshold will miss a fraction of true positives, it will be specific, as true negatives will not be scored as (false) positives, and to some degree sensitive, picking a fraction of true positives as such. From now on, we term it **nasality acoustic marker 1**.

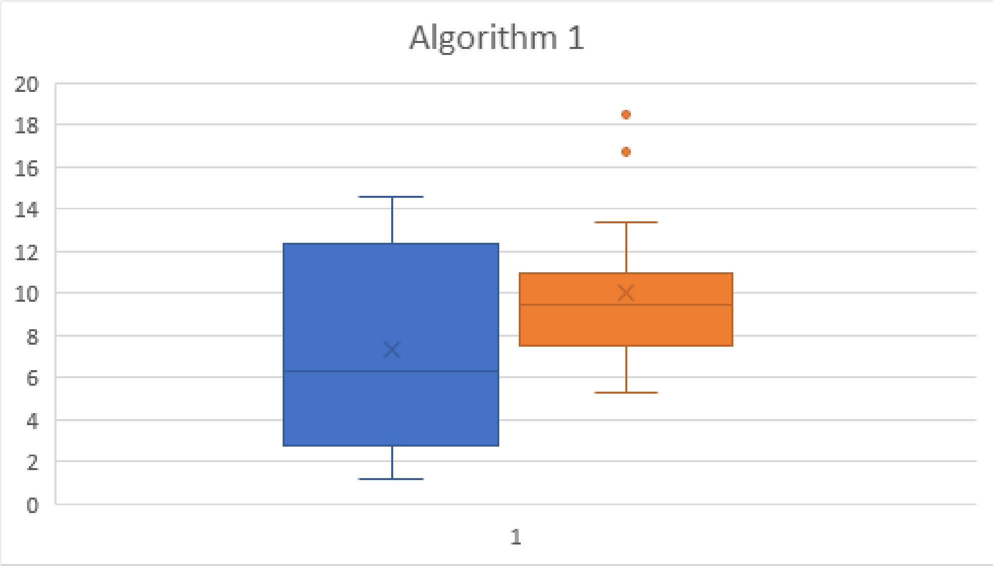

Figure 2. *Box plots and cat whiskers of **nasality acoustic index 1**, the relative energy in the 100 – 500 Hz region in nasalized (blue) vs oral (orange) vowels. The Y axis are dB as calculated by algorithm 1.*

In this case, two options for discrimination threshold values are either 7.525 dB (the Q1 value of the oral box), or 5.300 dB (the tip of the oral lower whisker).

5.5. Analysis of clinical prediction value

We next explored the clinical validity of the two nasality acoustic markers, using the two different possible discrimination thresholds for each.

In the case of nasality acoustic marker 1, the energy increase in the 100 – 500 Hz region, calculated by algorithm 1, the two possible thresholds in the confusion matrixes in Table IV indicate that none of them on its own provides a powerful clinical prediction value, although it may be tested in a combined test together with a second marker.

	marker1 < 5.300 dB	**marker1 > 5.300 dB**	**totals**			
nasals	9	11	20	45%	% sensitivity	
orals	0	20	20	100%	% specificity	
totals	9	31	40		**Positive predictive value =**	100%
					Negative predictive value =	65%
					Overall Clinical Efficiency =	72 %

	marker1 < 7.525 dB	**marker1 > 7.525 dB**	**totals**			
nasals	11	9	20	55%	% sensitivity	
orals	5	15	20	75%	% specificity	
totals	16	24	40		**Positive predictive value =**	69%
					Negative predictive value =	62%
					Overall Clinical Efficiency =	65 %

Table IV. *Confusion matrixes for the two discrimination thresholds with nasality acoustic marker1.*

In the case of nasality acoustic marker 2, the energy attenuation in the 700 – 1,300 Hz region, calculated by algorithm 2, the two possible thresholds in the confusion matrixes in Table V reach an identical, satisfactory Overall Clinical Efficiency of 80%. One of the options provides better specificity, and Positive Predictive Value, at the expense of sensitivity and Negative Predictive Value. The other option does the other way around.

	marker2 > 3.650 dB	**marker2 < 3.650 dB**	**totals**			
nasals	15	5	20	75%	% sensitivity	
orals	3	17	20	85%	% specificity	
totals	18	22	40		**Positive predictive value =**	83%
					Negative predictive value =	77%
					Overall Clinical Efficiency =	80 %

	marker2> 2.875 dB	**marker2 < 2.875 dB**	**totals**			
nasals	17	3	20	85%	% sensitivity	
orals	5	15	20	75%	% specificity	
totals	22	18	40		**Positive predictive value =**	77%
					Negative predictive value =	83%
					Overall Clinical Efficiency =	80 %

Table V. *Confusion matrixes for the two discrimination thresholds with nasality acoustic marker2.*

5.6. Generation of a combined clinical test by means of Multiple Linear Regression Analysis

Because the two nasality acoustic markers identified in this study were only partially satisfactory, we investigated whether a combined index might provide a more powerful clinical predictive value. As seen above, the nasality acoustic marker 1 provides maximum specificity, but with a low sensitivity. nasality acoustic marker 2 provides a better sensitivity / specificity ratio and a good overall clinical efficiency. We investigated whether a combined acoustic index might provide a more powerful clinical test for nasalization.

In order to develop the combined index, a linear regression was performed. Before, the two variables need to be confirmed that have linear correlation with the scores of the judges.

Pearson correlation between judges and the algorithm 1: -0.354559134
Pearson correlation between judges and the algorithm2: 0.73394854

The values of acoustic markers 1 and 2 were fed as independent variables, whereas the mean perceptual scores for nasalization were entered as dependent variable. The result of the regression analysis is shown in Table VI.

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN OF SQUARES
MODEL	262.312306	2	131.156153
RESIDUAL	218.163696	37	5.8931611
TOTAL	480.476002	39	12.3198975

Number of observations	**40**
F (2, 37)	22.24
Prob > F	0.000
R - squared	0.5459
Adjusted R - squared	0.5214
Root MSE	2.4282

Judge	Coefficients	Standard Error	t -value	P > |t|	[95% Confidence Interval]
algorithm 2	0.9676714	0.1653631	5.85	0	0.3626138 - 1.302729
algorithm 1	0.0849097	0.1103773	0.99	0.447	- 0.138736 - 0.3085554
_cons	-0.2539319	1.52719	-0.17	0.869	- 3.348313 - 2.840449

Table VI. *Summary of the results of a Multiple Linear Regression calculation, with a confidence level of 95%*

And the resulting equation for the combined index is:

Combined acoustic nasality index (dB) = 0.08491· *marker1* + 0.96767· *marker2* - 0.25393

We next investigated whether the **combined acoustic nasality index** significantly discriminates oral and nasalized vowels. Results of a t-test, summarized in Table VII, indicate that values calculated for oral and nasalized vowels are significantly different with a confidence at least of 99.9%.

		mean	**standard deviation**	**t value**	**significance**
Combined index	nasal	6.265	2.592	5.331	99.9%
	oral	2.915	1.087		

Table VII. *Summary of the results of the t-test on the combined nasality index values for the nasalized and the oral vowels.*

In addition, the **combined acoustic nasality index** shows a strong linear correlation with the scores done by the judges with a Pearson Correlation Coefficient r = 0.7389. Figure 3 shows the corresponding scatter plot.

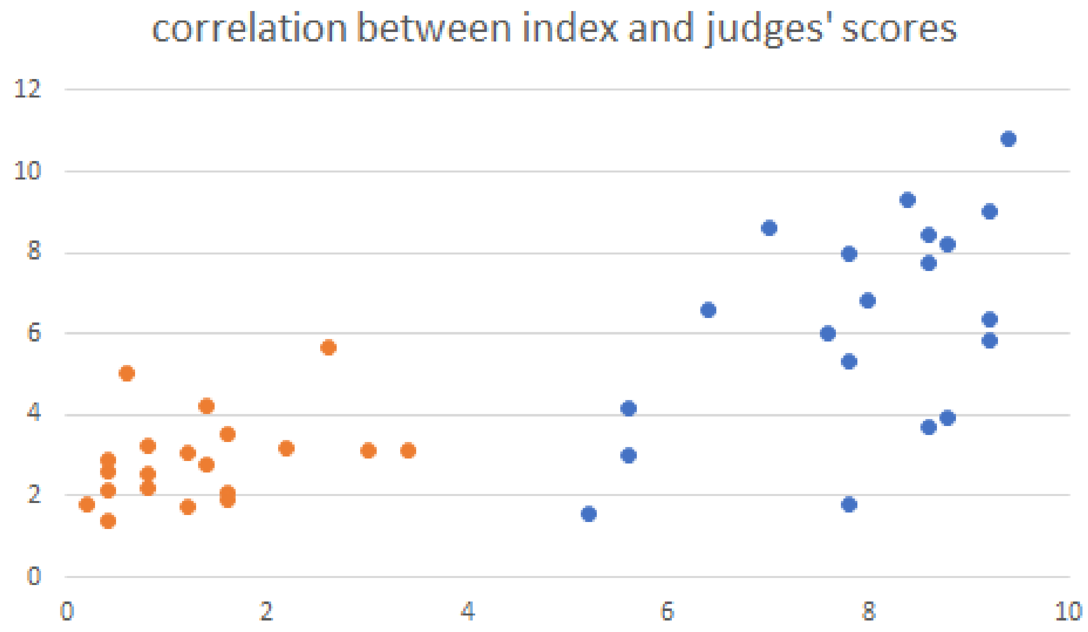

Figure 3. *Scatterplot showing the two groups of values for the* ***combined nasality index*** *of oral (orange) and nasalized (blue) vowels. Y axis, combined nasality index; X axis, 0-10 mean perceptual evaluation scores.*

Identification of discrimination thresholds for the combined acoustic nasality index

To investigate the clinical value of the combined acoustic nasality index, results were box-plotted for nasalized and oral vowels. As shown in Figure 4, the two boxes are robustly separated, pointing to a strong discrimination potential.

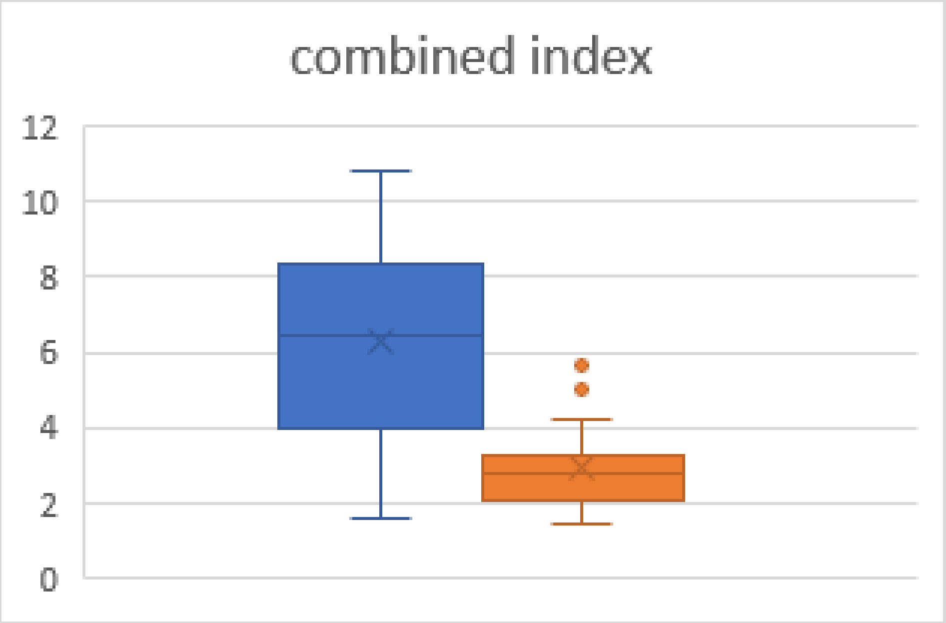

Figure 4. *Box plots and cat whiskers of the* ***combined nasality acoustic index*** *(Y axis) for nasalized (blue) and oral (orange) vowels.*

As discrimination threshold values either 3.982 dB (the Q1 value of the nasal box), 3.242 dB (the Q3 value of the oral box), or an intermediate value, can be selected.

Analysis of clinical prediction value for the combined acoustic nasality index

We next explored the clinical validity of the **combined acoustic nasality index** using the two different possible discrimination thresholds described above, in confusion matrixes. As shown in Table VIII, the results are no better than those using acoustic marker 1 alone.

	Combined index > 3.98 dB	**Combined index < 3.98 dB**	**totals**			
nasals	15	5	20	75%	% sensibility	
orals	3	17	20	85%	% specificity	
totals	18	22	40		**Positive predictive value =**	83%
					Negative predictive value =	77%
					Overall Clinical Efficiency =	80 %

	Combined index > 3.24 dB	**Combined index < 3.24 dB**	**totals**			
nasals	17	3	20	85%	% sensibility	
orals	5	15	20	75%	% specificity	
totals	22	18	40		**Positive predictive value =**	77%
					Negative predictive value =	83%
					Overall Clinical Efficiency =	80 %

Table VIII. *Confusion matrixes for the two discrimination thresholds with the* ***combined acoustic nasality index.***

We therefore explored the intermediate value as discrimination threshold. As shown in Table IX, when the critical point is set at 3.61, all relevant clinical parameters improve dramatically, reaching values well above many routine medical tests[1].

---

[1] For comparison purposes, diabetes mellitus prediction tests based on fasting plasma glucose levels have been shown to be in some cases as little specific as 68% and as little sensitive as 40% (24).

	Combined index > 3.61 dB	**Combined index < 3.61 dB**	**totals**			
nasals	17	3	20	85%	% sensibility	
orals	3	17	20	85%	% specificity	
totals	20	20	40		**Positive predictive value =**	85%
					Negative predictive value =	85%
					Overall Clinical Efficiency =	85 %

Table IX. *Confusion matrix for the* ***combined acoustic nasality index****, using an intermediate discrimination threshold 3.61 dB.*

6. Discussion and conclusions:

Our research identifies two robust acoustic markers of nasalization.

Firstly, there is a consistent energy decrease in the region between 700 and 1,300 Hz in the nasalized samples. Such attenuation is likely the result of the interference of the oral signal with an anti-resonance in the nasal signal. This finding is consistent with the previous phonetic works studying nasalized vowels in a speech context (11-17).

Secondly, there is an energy increase in the region 100 – 500 HZ in the nasalized samples. Such increase is likely to result from the interference of the oral signal with the nasal murmur formant in the nasal signal, described by previous phonetic works studying nasalized vowels in a speech context (9–11).

Combination of the two acoustic markers of nasalization provides a robust combined **acoustic nasality index** that shows a strong correlation with the perceived degree of nasality (r = 0.74).

In addition, the combined **acoustic nasality index** shows a strong clinical value for the objective evaluation of hypernasal voice in patients. The index shows a sensitivity of 85% (percentage of nasalized voices identified as such) and a specificity of 85% (percentage of not nasalized voices identified as such), which results in an overall clinical efficiency of 85% (percentage of correctly identified nasalized voices and not nasalized voices). Such figures are better than some of the medical tests currently in use in the clinic.

Limitations of the study and future research prospects:

The **acoustic nasality index** created in this research work is based on 40 selected pairs of samples of voluntarily generated oral and nasalized vowels. Future work should aim at: (1) validate the index with a larger sample; (2) further study the correlation of the index with the degree of severity; (3) validate the index as a screening tool when it comes to real patients, e.g. with different degrees of diagnosed cleft palate.

7. References:

1. Roy N, Merrill RM, Gray SD, Smith EM. Voice disorders in the general population: Prevalence, risk factors, and occupational impact. Laryngoscope. 2005;115(11):1988–95.
2. Boone DR, McFarlane SC, Von Berg SL, Zraick RI, Boon DR, McFarlane SC. The voice and voice therapy [Internet]. 9th ed. Allyn & Bacon. Boston: Pearson Higher Ed; 2014 [cited 2018 Dec 5]. 361 p. Available from: https://cataleg.uab.cat/iii/encore/record/C__Rb1756631__Svoice and voice therapy__Orightresult__U__X7;jsessionid=2CDF3369BA97208883A4E711D901ED4C?lang=cat
3. Schwartz MF. The acoustics of normal and nasal vowel production. Cleft Palate J. 1968;5(2):125–40.
4. Chen MY. Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers. J Acoust Soc Am [Internet]. 1995;98(5):2443–53. Available from: http://asa.scitation.org/doi/10.1121/1.414399
5. Lee A, Whitehill TL, Ciocca V. Effect of listener training on perceptual judgement of hypernasality. Clin Linguist Phon [Internet]. 2009 Jan 1;23(5):319–34. Available from: https://doi.org/10.1080/02699200802688596
6. Titze IR, Martin DW. Principles of Voice Production. J Acoust Soc Am [Internet]. 1998;104(3):1148. Available from: http://scitation.aip.org/content/asa/journal/jasa/104/3/10.1121/1.424266
7. Fant G, Erikson Y. Analysis of Nasal Consonants. J Acoust Soc Am. 2005;52(1A):182–182.
8. Warren DW, Dalston RM, Mayo R. Aerodynamics of Nasalization [Internet]. Vol. 5, Nasals, Nasalization, and the Velum. ACADEMIC PRESS, INC.; 1993. 119-146 p. Available from: http://linkinghub.elsevier.com/retrieve/pii/B9780123603807500090
9. Chen MY. Acoustic correlates of English and French nasalized vowels. J Acoust Soc Am [Internet]. 1997;102(4):2360–70. Available from: http://asa.scitation.org/doi/10.1121/1.419620
10. Styler W. On the acoustical features of vowel nasality in English and French. J Acoust Soc Am. 2017;142(4):2469–82.
11. Feng G, Castelli E. Some acoustic features of nasal and nasalized vowels: a target for vowel nasalization. J Acoust Soc Am [Internet]. 1996;99(6):3694–706. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8655801
12. Hattori S, Yamamoto K, Fujimura O. Nasalization of vowels in relation to nasals. J Acoust Soc Am. 1958;30(4):267.
13. Hawkins S, Stevens KN. Acoustic and perceptual correlates of the non-nasal–nasal distinction for vowels. J Acoust Soc Am [Internet]. 1985;77(4):1560–75. Available from: http://asa.scitation.org/doi/10.1121/1.391999
14. House AS, Stevens KN. Analog Studies Of The Nasalization Of Vowels. J Speech Hear Disord. 1956;21(2):218–32.
15. Fant G. Acoustic Theory of Speech Production (Mouton, The Hague, The Netherlands). 1960;

16. Fujimura O, Lindqvist J. Sweep-Tone Measurements of Vocal-Tract Characteristics. J Acoust Soc Am [Internet]. 1971;49(2B):541–58. Available from: https://doi.org/10.1121/1.1912385

17. MAEDA S. ACOUSTICS OF VOWEL NASALIZATION AND ARTICULATORY SHIFTS IN FRENCH NASAL VOWELS. Nasals, Nasalization, and the Velum [Internet]. 1993 Jan 1 [cited 2019 May 8];147–67. Available from: https://www.sciencedirect.com/science/article/pii/B9780123603807500107

18. Kataoka DR, Michi DK-I, Okabe DK, Miura DT, Yoshida DH. Spectral Properties and Quantitative Evaluation of Hypernasality in Vowels. Cleft Palate-Craniofacial J [Internet]. 1996;33(1):43–50. Available from: https://doi.org/10.1597/1545-1569_1996_033_0043_spaqeo_2.3.co_2

19. Lee AS, Ciocca V, Whitehill TL. Acoustic correlates of hypernasality. Clin Linguist Phon [Internet]. 2003;17(4–5):259–64. Available from: https://doi.org/10.1080/0269920031000080091

8. Annexs:

Table with the scores and the scores of all the voices:

Voice code		1st judge	2nd judge	3rd judge	4th judge	5th judge	Mean
10	nasal	8	10	6	10	9	8.6
11	oral	1	1	2	1	3	1.6
86	nasal	9	9	10	10	8	9.2
61	oral	0	0	0	4	0	0.8
89	nasal	5	8	8	7	10	7.6
21	oral	0	1	4	2	0	1.4
90	nasal	8	9	5	5	8	7.0
50	oral	1	0	1	2	0	0.8
91	nasal	8	8	8	7	8	7.8
75	oral	4	2	4	4	3	3.4
62	nasal	8	9	8	9	9	8.6
35	oral	4	0	4	6	1	3.0
60	nasal	8	9	8	7	8	8.0
70	oral	3	1	2	1	4	2.2
53	nasal	8	10	8	8	10	8.8
14	oral	0	0	2	0	4	1.2
68	nasal	8	3	4	7	4	5.2
76	oral	0	1	0	1	2	0.8
27	nasal	10	8	8	6	7	7.8
84	oral	0	0	1	1	0	0.4
16	nasal	10	10	9	9	8	9.2

92	oral	2	3	6	0	2	2.6
88	nasal	9	10	10	8	9	9.2
83	oral	0	1	2	0	0	0.6
32	nasal	4	8	4	4	8	5.6
66	oral	2	1	3	1	1	1.6
81	nasal	10	9	9	9	10	9.4
65	oral	0	0	0	2	0	0.4
44	nasal	9	10	8	9	8	8.8
33	oral	4	1	1	0	2	1.6
56	nasal	8	9	9	7	9	8.4
38	oral	0	1	4	1	0	1.2
47	nasal	7	9	7	8	8	7.8
29	oral	0	1	0	0	0	0.2
82	nasal	10	9	5	9	10	8.6
31	oral	0	0	0	2	0	0.4
63	nasal	2	6	3	9	8	5.6
15	oral	1	2	1	1	2	1.4
45	nasal	6	4	8	7	7	6.4
64	oral	0	1	1	0	0	0.4

Table with the scores of the 4 algorithms

Number		**Mean**	**Algorithm 1**	**Algorithm 2**	**Algorithm 3**	**Algorithm 4**
10	nasal	8.6	3.1	3.8	27.8	35.4
11	oral	1.6	10.7	1.5	13.9	20.7
86	nasal	9.2	1.3	9.5	28.6	37.8
61	oral	0.8	8.3	2.9	21.3	27
89	nasal	7.6	3.4	6.2	13.5	19.4
21	oral	1.4	7.1	2.5	23.1	22.2
90	nasal	7	5.5	8.7	7.3	25.1
50	oral	0.8	12	1.5	10.4	21.6
91	nasal	7.8	7.2	7.9	17.6	33.2
75	oral	3.4	8.8	2.7	11.4	22.4
62	nasal	8.6	5.5	8.5	10.2	29.5
35	oral	3	8.2	2.8	10.8	25.9
60	nasal	8	1.2	7.2	21.8	34.4
70	oral	2.2	5.3	3.1	14.3	22.4
53	nasal	8.8	12.8	3.2	11	21.6
14	oral	1.2	16.7	0.6	16.6	24.5
68	nasal	5.2	14.6	0.6	15.7	21.5
76	oral	0.8	11	1.9	21.8	23
27	nasal	7.8	12	4.7	3.2	16.1
84	oral	0.4	13.4	1.3	18.3	16.6
16	nasal	9.2	12.5	5.2	2.4	20.8
92	oral	2.6	6.9	5.5	8.3	16.1

88	nasal	9.2	2.7	6.6	15.8	16.8
83	oral	0.6	6.6	4.9	11.7	20.4
32	nasal	5.6	7.8	2.7	14.8	25.4
66	oral	1.6	9.3	1.4	17.2	21.7
81	nasal	9.4	11.6	10.4	2.8	19.2
65	oral	0.4	10.7	2.3	13.9	15.1
44	nasal	8.8	14.3	7.5	8.2	11.5
33	oral	1.6	18.5	2.3	21.2	15.8
56	nasal	8.4	2.1	9.7	19.2	21.7
38	oral	1.2	7.3	2.8	20.2	18.6
47	nasal	7.8	12.9	1	20.7	11.2
29	oral	0.2	10.2	1.2	17	17.2
82	nasal	8.6	3.4	8	14.2	34
31	oral	0.4	9.6	2.1	22.2	22.5
63	nasal	5.6	10.7	3.6	15.8	26
15	oral	1.4	10.5	3.7	10.4	21.8
45	nasal	6.4	1.7	6.9	21.5	31.2
64	oral	0.4	9.4	0.9	23.8	22.5

Algorithm combined:

```
clearinfo


### print heading

writeInfoLine: "Nasality Index Script (0-100)"

appendInfoLine: "  Derives from two independent acoustic parameters showing correlations found:"

appendInfoLine: "    Relative energy gained by nasal resonance (100 - 500 Hz, nasal murmur)"

appendInfoLine: "    Relative energy lost by  nasal attenuation (700 - 1,300 Hz)"

appendInfoLine: "  Clinical discrimination threshold value = 22/100"

appendInfoLine: ""

### Script by David Garcia-Quintana, PhD, Associate Professor in Acoustic Physics, Biophysics Unit, School of Medicine, and Eric Bermudez, SLP senior student, Universitat Autonoma de Barcelona, 2019.




### batch processing loop

n_batch = numberOfSelected("Sound")

for i_batch to n_batch

      bsel'i_batch' = selected("Sound", i_batch)

endfor

for i_batch to n_batch

      select bsel'i_batch'

      name$ = selected$ ("Sound")


### regnr calculation

Filter (pass Hann band)... 100 500 1

Rename... oralnasalmurmurregion

select Sound oralnasalmurmurregion

To Intensity... 100 0

slporalnasalmurmurregion = Get mean... 0 0 dB

# SPL in the phonetic region
```

```
select bsel'i_batch'
Filter (pass Hann band)... 0 2200 1
Rename... phoneticregion
select Sound phoneticregion
To Intensity... 100 0
slpphoneticregion = Get mean... 0 0 dB
# calculation
regnr = slpphoneticregion - slporalnasalmurmurregion

### relna calculation
select bsel'i_batch'
Filter (pass Hann band)... 700 1300 1
Rename... slporalnasalattenuationregion
select Sound slporalnasalattenuationregion
To Intensity... 100 0
slporalnasalattenuationregion = Get mean... 0 0 dB
# SPL in the phonetic region
# select bsel'i_batch'
# Filter (pass Hann band)... 0 2200 1
# Rename... phoneticregion
# select Sound phoneticregion
# To Intensity... 100 0
# slpphoneticregion = Get mean... 0 0 dB
# calculation
relna = slpphoneticregion - slporalnasalattenuationregion

### multiple linear regression calculation, mlrc
mlrc = 0.96767*relna + 0.08491*regnr - 0.25393
if mlrc >= 3.61
      mlrcstar$ = " -> nasality detected!"
else
      mlrcstar$ = " -> no nasality detected"
```

```
endif


###  index education

mlrcindex = (mlrc - 1.4)*100/10.2


### print report

appendInfoLine: name$, "   ", " Nasality Index (dB) = ", fixed$ (mlrcindex,0), mlrcstar$


### clean and restart to loop

removeObject: "Sound oralnasalmurmurregion", "Sound phoneticregion", "Sound slporalnasalattenuationregion", "Intensity oralnasalmurmurregion", "Intensity phoneticregion", "Intensity slporalnasalattenuationregion"


endfor
```