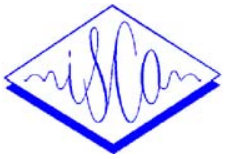
ISCA Archive
http://www.isca-speech.org/archive



INTERSPEECH 2004 - ICSLP
8th International Conference on Spoken Language Processing
ICC Jeju, Jeju Island, Korea
October 4-8, 2004

ANALYSIS OF HYPERNASALITY BY SYNTHESIS

P. Vijayalakshmi and M. Ramasubba Reddy

BioMedical Engineering Division, Dept. of Applied Mechanics
Indian Institute of Technology, Madras

Abstract

Speakers with velopharyngeal incompetence, produce hypernasal speech across voiced elements. During the production of vowels, due to the defective velopharyngeal mechanism, oral cavity gets coupled with the nasal cavity. This coupling may introduce resonances of nasal cavity in the resulting vowel sound. In the present study, an attempt is made to convert pure vowels in to its nasalized counterpart, by introducing nasal resonances in full/part, in addition to the oral tract resonances. Further, an analysis is done by varying the formant bandwidths of all the resonances. By considering nasalized vowel of normal speakers as reference, a perceptual experiment is conducted and validated. The speech data of 10 patients with cleft palate, who are expected to produce hypernasal speech were analyzed. The presence of nasal formants in their Linear Prediction spectra, validated the above results. Presence of nasal formants in the hypernasal speech, may be a cue for detection of hypernasality. Another experiment is also conducted in the time domain, to verify whether the widening of formant bandwidth is due to the distribution of energy between oral and nasal cavities.

10.21437/Interspeech.2004-211

1. Introduction

Speakers with a defective velopharyngeal mechanism, produce speech with inappropriate nasal resonances across voiced elements. A defective velopharyngeal mechanism, and the corresponding hypernasal speech, can be caused by anatomical defects (cleft palate), central nervous system damage (cerebral palsy or traumatic brain injury), peripheral nervous system damage (Moebius syndrome) [1] or hearing impairment [2].

Researchers have developed algorithms to detect hypernasality by analyzing (i) speech of hearing impaired, (ii) synthesized hypernasal speech and (iii) nasalized vowels. Stevens et al.[2] in their discussion on assessment of nasalization, in the speech of deaf children claim that, perception of nasality can be influenced primarily by inadequate or improper adjustment of velopharyngeal opening, and other factors like misarticulation, pitch variations and speech tempo. They have observed improper nasalization of certain vowels for the deaf children. Chen [3] has performed acoustic analysis of speech of hearing impaired children, and nasalized vowels of normal hearing speakers. From the acoustic analysis, he has observed an extra peak between the first and second formants around 950 Hz and a reduced first formant prominence for both the cases. From his study, the difference between the amplitude of the first formant and the amplitude of the extra peak is shown to be a promising measure of nasality. The author has concluded that, when the bandwidth of first formant widens, the perception of nasality increases. Cairns et al.[1] have utilized the sensitivity of Teager energy operator to multicomponent signals, to detect the hypernasality. As the nasalized speech is characterized by formants, antiformants and nasal formants, the authors have taken nasalized speech to be a multicomponent signal and normal speech as a single component signal. They have observed a measurable difference between the low pass and bandpass profile of Teager energy operator for the nasalized vowels, whereas, the normal speech which is assumed to be a single component signal, does not show any difference. Glass [4] has found that, the most noticeable difference between the nasalized and non-nasalized vowels, is in the low frequency regions of the magnitude spectrum. They have also suggested that, a measure of the spread of energy about the first resonance is found to be a best indication of nasalization.

Hypernasal speech occurs typically in conjunction with abnormalities in pitch, loudness, voice quality and/or articulation [5]. These co-existing features, affect the perception of hypernasality [6]. This aspect of hypernasal speech has necessitated the analysis of hypernasal speech by synthesis. In our present study, an attempt is made to synthesize hypernasal speech, by introducing nasal resonances in full/part, in addition to the resonances of the vowels (/a/, /i/, /u/) and also by reducing the radii of the resonances of vowels and nasals. The advantage of synthetic speech is that, it becomes possible to control and artificially manipulate acoustic variables, related to a specific perceptual feature, without confounding effects from other co-occurring problems, often found in disordered speech. A perceptual experiment is conducted by taking nasalized vowels (from the CVC units (/mam/, /min/, /mun/) as reference and is validated. An acoustic analy-

I thank Dr. Savithri of All India Institute of Speech and Hearing for providing me the hypernasal speech data

sis of hypernasal speech is also conducted. The presence of nasal formants in the hypernasal speech, validated the results obtained from nasalized vowels and synthesized speech. To prove that, the widening of formant bandwidth is due to the distribution of energy between oral and nasal cavities, vowel and its nasalized counterpart are constructed, keeping a single formant frequency. This time-domain analysis shows that, signal corresponding to nasalized vowel dies out at a faster rate than that of the vowel.

The rest of the paper is organized as follows. Section II presents acoustic analysis on nasalized vowels and nasal consonants. Section III describes the synthesis of hypernasal speech and the perceptual study. In Section IV, the acoustic analysis of hypernasal speech are discussed. Section V discusses on time-domain analysis, to give an evidence for the widening of formant bandwidth.

2. Acoustic analysis

The speech data analyzed in this study, is collected from 10 patients with cleft palate. The speech of 6 healthy speakers are also collected for analysis on nasalized vowels. For synthesis, phonations from 10 healthy speakers are collected. All of the recorded words are digitized at 8 kHz.

2.1. Acoustic analysis on nasalized vowels

The transition from one oral vowel to nasal consonant or vice versa, may involve an intermediate nasalized vowel sound, as the velum lowers before or raises after oral closure[7]. The data analyzed in this study, are collected from six (3 male and 3 female) healthy speakers. This data is composed of words with /m/ and /n/ as initial and final nasal consonants, with an intermediate vowel (/mam/, /min/, /mun/). The formant frequencies of pure vowels, [8] (see Table 1) is considered as reference. The nasalized vowels are analyzed at 20 ms intervals, to extract the formants and are tabulated (Table 2). The additional formants (around 250 Hz for /a/, around 1000 Hz for /i/ and /u/) found (refer Table 2) in the nasalized vowels, may be because of the adjacent nasal consonants. To confirm this, an analysis is done on formant frequencies of nasal consonants, which is discussed in the following Section.

Table 1: Formant frequencies(in Hz) of vowels

phone	F1	F2	F3
/a/	730	1090	2440
/i/	270	2290	3010
/u/	300	870	2240

Table 2: Formants (mean, minimum, and maximum in Hz) of nasalized vowels

words	mean	min	max
man	236	180	266
/a/	746	711	766
	1119	1055	1180
	1950	1732	1984
	2577	2328	2741
min	239	172	297
/i/	524	436	555
	1044	945	1165
	2236	1992	2454
	2728	2538	2894
mun	241	164	305
/u/	563	461	711
	1042	914	1195
	1885	1847	2007
	2496	2386	2592

2.2. Acoustic analysis on nasal consonants

The speech data analyzed in this study, are collected from 10 (5 male and 5 female) healthy speakers. The speech data is composed of words *summer, sunny* and *singing*. An acoustic analysis is performed on the nasal portions of the data, to extract formant frequencies and the results are tabulated (Table 3). It is noticed that, the additional formants introduced in the nasalized vowels, are almost same as the formants of the nasal consonants. To cross verify this, synthesis of hypernasal speech is performed by introducing the nasal formants, in addition to the vowel formants, as explained in Section 3.

Table 3: Formants (mean, minimum, and maximum in Hz) of nasal consonant /m/

formants	mean	min	max
F1	219	200	244
F2	947	777	1111
F3	2421	2355	2533
F4	3315	3155	3511

3. Synthesis

Synthesis of hypernasal speech is carried out by introducing resonances of nasals, in addition to the resonances of vowels. From the phonations /a/, /i/, /u/ of 10 healthy speakers, Linear Prediction (LP) coefficients, LP residual and gain are extracted. The speech signal of the vowel sounds (/a/, /i/, /u/), are modified by introducing

- each of the nasal resonances separately.
- each pair of nasal resonances.
- all the nasal resonances.
- all the nasal resonances, with widened formant bandwidths of nasal as well as oral formants.

For each case, separate speech files are generated after synthesis and used for perceptual analysis. 5 female and 3 male speakers participated in the perceptual experiment. They are with ages under 30. None of the subjects had known speech or hearing impairment. All are naive listeners, without any past training in perception of nasality. The listeners are asked to rate the modified speech signals, for the degree of nasalization and the results are tabulated. The test is conducted for three different phonations, leading to 24 tests for each case. Each entry in the Table 4, indicates the number of tests in favor of a particular rating.

Table 4: Perceptual Analysis

Formants (in Hz)	Degree of nasalization			
	No	Mild	Moderate	High
2500	22	2	-	-
1000	-	15	9	-
250	-	4	20	-
250,1000	-	-	16	8
All	-	-	16	8
All (inc. BW)	-	-	3	21

From Table 4, it is observed that, introduction of the nasal formant

- at 2500 Hz, does not have any influence on nasality.
- at 250 Hz or 1000 Hz, has moderate influence on nasalization, in particular for the nasal formant at 250 Hz.
- at 250 Hz and 1000 Hz or at all nasal resonant frequencies, have more influence than the previous case.
- at all nasal frequencies with increased bandwidth, shows high degree of nasalization.

In the present study, inclusion of zeros is not considered, which may be taken for future study. From this perceptual study, we may conclude that, all the formants of nasal consonants have influence on nasalization of vowels and we can expect the same set of formant frequencies in the hypernasal speech also.

Table 5: Formants (mean, minimum, and maximum in Hz) of hypernasal speech

words	mean	min	max
/a/	242	203	297
	703	512	900
	1062	859	1234
	2242	2133	2391
	2711	2569	2890
/i/	249	224	265
	556	483	574
	857	756	969
	1809	1617	1881
	2531	2400	2708
/u/	255	238	271
	566	484	734
	1084	906	1352
	1894	1720	2087
	2385	2191	2542

4. Acoustic analysis on hypernasal speech

The acoustic parameters (Formant frequencies and Formant bandwidths), that serve as an estimate of nasalization, are analyzed in the speech data of patients with cleft palate, who are expected to produce hypernasal speech. The LP (Linear Prediction) spectra of speakers with cleft palate, are examined closely for nasal formants or a widening of formant bandwidth.

LP spectra of phonations (/a/, /i/, /u/) of 10 speakers with cleft palate, are analyzed at 20 ms intervals throughout the vowel, to extract the formants and the average formants over the entire interval for each vowel are found (refer Table 5).

From the Table 5 we observe that,

- invariably for all the phonations, there exists a formant frequency around 250 Hz (see Figure 1 for the phonation /a/).
- in many cases, for the phonations /i/ and /u/, there is a broadening of F1 bandwidth. This widening of formant bandwidth may be because of the inclusion of one more frequency which is very close to F1, which is resolved clearly for the higher LP order.
- there exists an additional formant frequency around 1000 Hz for the phonations /i/ and /u/, as realized in the perceptual study and acoustic analysis of nasalized vowels (refer Table 2).

From this, we can conclude that, presence of two or more formants of nasal consonants, may be a cue for detection of hypernasality.

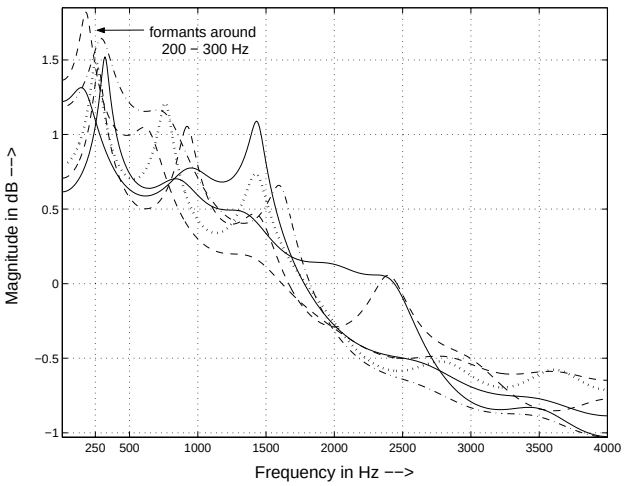


Figure 1: LP spectrum for the phonation /a/ uttered by speakers with hypernasality (see formants around 200 Hz

5. Widening of formant bandwidths

Widening of formant bandwidths of some of the formants of hypernasal speech, may be due to the energy distribution between oral and nasal cavities. To prove this, the time domain signal of vowel /a/ and its nasalized counterpart, are normalized to a same value. Second formants of both the signals are extracted by taking the roots corresponding to it and the time domain signal containing only this formant is constructed, for both the vowel and its nasalized counterpart. The signal of nasalized vowel, dies out (refer Figure 2(a)) at a faster rate compared to that of the pure vowel(see Figure 2(b)). This may be because of the distribution of energy between oral and nasal cavities, which in turn implies that, the corresponding formant in the spectrum will have a wider bandwidth.

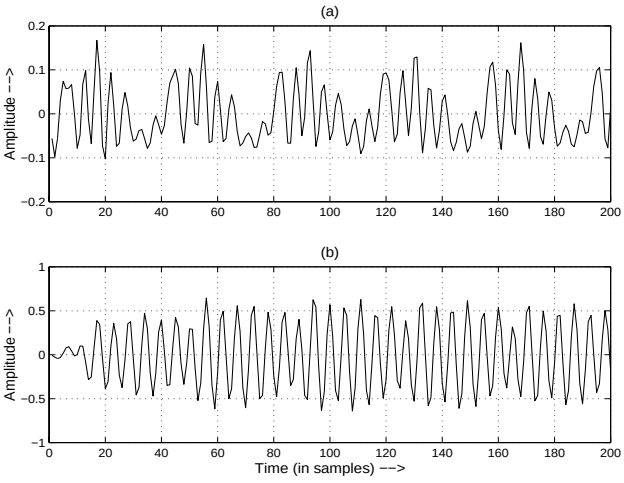


Figure 2: Time domain signal of F2 alone. (a) Nasalized vowel. (b) Pure vowel

6. Conclusions

The perceptual study, and acoustic analysis of hypernasal speech and nasalized vowels reveal that, nasal formants around 250 Hz and 1000 Hz play a major role in the nasalization of vowels. This may be due to the high concentration of energy, at the lower formants than the higher formants. For the vowel /a/, since it does not have a formant in the low frequency region, visibility of formant in the low frequency region, is a clear indication of hypernasality.

7. References

[1] Cairns, D.A., J.H.L. Hansen and J.E. Riski, "A noninvasive technique for detecting hypernasal speech using a nonlinear operator," *ieeebm*, vol. 43, no. 1, pp. 35–45, Jan. 1996.

[2] Stevens, K.N., R.S.Nickerson, A.Boothroyd, and A.M. Rollins , "Assessment of nasalization in the speech of deaf children," *jshr*, vol. 19, pp. 393–416, 1976.

[3] Chen, M.Y., "Acoustic parameters of nasalized vowels in hearing-impaired and normal hearing speakers," *jasa*, vol. 98, no. 5, pp. 2443–2453, Nov. 1995.

[4] Glass, J.R., and V.W. Zue, "Detection of nasalized vowels in american english," in *icassp*, 1985, pp. 1569–1572.

[5] Spriesterbach, D.C.,, "Assessing nasal quality in cleft palatespeech of children," *jshd*, vol. 20, pp. 266–270, 1955.

[6] Moll, K.L., *Speech characteristics of individuals with cleft lip and palate. In cleft palate and communications, D. C. Spriesterbach and D. Sherman, editors*, 1968.

[7] O'Shaughnessey, D.,, *Speech Communications*, Universities press, 2001.

[8] Rabiner, L.R., and B.H.Juang, *Fundamentals of Speech Recognition*, Prentice Hall, New Jersy, 1993.