THE ROLE OF THE SINUS CAVITIES IN THE PRODUCTION OF NASAL VOWELS

Shinji Maeda
Centre National d'Études des Telecommunications, 22301 Lannion, France
and
Bell Laboratories, Murray Hill, New Jersey, 07974 U.S.A.

ABSTRACT

The acoustic wave propagation inside the vocal tract was simulated assuming plane waves. When an acoustic tube having appropriate area function but no side branching cavities was used as the nasal tract, the low frequency nasal formant below the first formant of low vowels as observed in natural speech could not be simulated. The relatively short nasal tract with both ends open cannot produce the low nasal formant. We therefore employed a side cavity for the nasal tract system, which gave rise to the nasal formant in the expected frequency region. The computed output signals for different nasalized vowels sounded very naturally nasal.

INTRODUCTION

This paper describes a study of acoustic modeling of the vocal tract, in particular, with appropriate nasal passages, by means of computer simulation. The model of the entire vocal tract is shown in Fig.1.

The vocal tract system consists of the pharyngeal, oral and nasal cavities. The salient feature of this model is that a side cavity, representing the nasal sinuses, is connected to the main nasal passage. The system is excited by varying the glottal area as an appropriate function of time, thereby, without resorting to the source-filter concept of the speech production (1). The speech signals are obtained as the sum of the acoustic pressure at the mouth opening and at the nostrils.

In the preliminary stage of this study, the nasal tract was represented by a simple acoustic tube without the side cavity. The eleven different French vowels synthesized were of high quality, in particular, in terms of naturalness. In the synthesis of nasal sounds, however, most of the vowels could not be nasalized properly by coupling this simple tube.

In our subjective judgement, the high vowels, such as [i] and [u], were well nasalized, and the degree of nasalization increased as the magnitude of the coupling increased. For the mid and low vowels, for example [a] or [oe], on the contrary, the vowel quality was modified, but the vowels did not sound nasal.

The sweep-tone experiment of Fujimura and Lindqvist (2) has shown that the nasal pole-zero pair appears near and often below the first formant peak, especially for low vowels. On the contrary, in the simulation, the pair was located above the first shifted vowel formant peak (see Fig.6, for example). The value of the lowest resonance frequency of the nasal tract in our simulation was apparently too high. Our nasal model was based on data reported by Fant (3) and by House and Stevens (4).

One possible way of lowering the nasal resonance frequency is, of course, to modify the area function of the nasal tract. The value of the lowest resonance of the nasal tract proper was 670 Hz, when the inner end of the nasal tract was closed. To lower this value to, say, 450 Hz, which seems to produce the observed nasal formant, requires rather considerable modification of the tract shape. If the length alone were modified, then it would be necessary to increase roughly 40% of its original length, resulting in a 16 cm long nasal passage.

Fig.1 The model of the vocal tract.

911
A more attractive alternative is to connect a side cavity to the existing nasal tract. Fujimura and Lindqvist (5) have suggested that the acoustic effect of the "subsidiary cavities" within the nose and/or of the asymmetry of the nose may explain the occurrence of a zero at low frequencies as observed in their sweep-tone data for the consonant [ ]. More direct evidence has been shown in the study of the acoustic measurement of the nasal transfer function, by Lindqvist-Gauffin and Sundberg (6). It is quite plausible, therefore, as suggested by Fant (7), that the sinus cavities may play an important role in shaping appropriate nasal spectra.

A NASAL SYSTEM WITH THE SIDE CAVITY

We have postulated, thus, a simple tube having a single side cavity as an acoustic model of the nasal tract. Its area function is specified as shown in Fig. 2.

![Fig. 2 The area function of the nasal system](image)

There are at least three or four pairs of the sinuses in the nose. We consider only the maxillary sinus pair that is largest in volume, about 25 cu.cm with the two cavities together, and has fairly large opening, about 0.1 sq.cm, to the nasal tract proper, according to the medical literature. We also assumed that the two nasal passages including the maxillary sinuses are symmetric. Its configuration was estimated, by a try-and-error procedure, at 0.1 sq.cm in the opening area (that is the sinus coupling area), 0.5 cm in the length of the coupling section, and 20.8 cu.cm in volume of the cavity. Because of the narrow passage to the main nasal tract, the cavity function as a Helmholtz resonator. The shape of the area function is not so important in terms of acoustic characteristics, and it was determined arbitrarily.

It may be interjected here that the manner how the nasal coupling is implemented is also crucial for achieving high nasal quality.

The first three sections of the nasal area function (indicated by the dashed line in Fig. 2) are varied depending on the degree of the nasal coupling. The coupling is defined by the cross-sectional area of the innermost section, i.e., the first section of the nasal tract. The area of the two remaining variable sections are calculated by a linear interpolation (in area) between the coupling section and the first fixed section of the nasal tract, that is the fourth section. Then the amount of the increased area for each of the three variable sections due to velum lowering is subtracted from the corresponding section within the oral tract, to represent the concomitant reduction of the distance between the tongue surface and the soft palate as the result of nasalization. This subtraction was very important in improving the nasal quality, especially, for mid and low vowels.

SIMULATION RESULTS

When the side cavity is connected to the nasal tract, theoretically, the "sinus pole" should appear always below the original first resonance in the nasal-tract transfer. This is the case, indeed, as shown in Fig. 3.

![Fig. 3 The transfer function of the nasal tract with or without the sinus coupling. The innermost section is closed.](image)

Note that the original first resonance peak, shown by the solid line, is split into two, as indicated by the dashed line, when the sinus is coupled. With the sinus side cavity coupled, the first resonance occurs at 446 Hz and the original first resonance at 670 Hz is shifted up to 817 Hz. The two poles at low frequencies in the nasal transfer function can affect the first or the second formant, or both, of the vowel, depending on the particular vowel and on the degree of nasal coupling.

The transfer function of the vowel [ ] and its nasalized version resulting from this nasal system are shown in Fig. 4.
Fig. 4 The transfer function, under the closed glottis condition, of the vowel \[ \text{a} \] and that with nasal coupling, NC=0.5 sq cm. The sinus cavity is connected to the main nasal tract.

In the case of weak coupling, as shown by the dash-dot line in Fig. 4, the first nasal formant (marked by \( N_1 \)) with a small peak appears below the shifted first formant \( (F_1') \), and also the second nasal formant \( (N_2) \) appears between the two vowel formants. In Fig. 5, the degree of coupling is increased to 2.0 sq cm;

Fig. 5 The same as Fig. 4, except NC=2.0 sq cm.

The magnitude of the first nasal formant peak \( (N_1) \) now compares well with that of the first shifted vowel formant \( (F_1') \), and the second nasal formant \( (N_2) \) appears very close to the second vowel formant \( (F_2') \). Note that the first formant-antiformant cluster is weaker than the second cluster. Such qualitative transfer characteristics of the nasalized vowel agree well with those of the sweep-tone data. The noticeable downward shift in each of the higher formants are not due to the acoustic effect of the nasal tract, but rather due to the change in the area function of the oral cavity introduced by the lowering of the velum.

**DISCUSSION**

Using the new nasal-tract configuration, the eleven vowels were synthesized with various degree of nasal coupling. An informal listening of the sounds has indicated that the low vowels, such as \[ \text{a} \], require a greater velum lowering in comparison with high or mid, especially front vowels for similar degree of nasalization in terms of the perception. This may be explained by assuming that the spectral modification of the first formant region, due to the occurrence of the nasal pole and zero in that zone, is primarily responsible for the perception of nasalization. The pole-zero pair would appear at low frequency when the coupling is small. As the amount of coupling increases, the pair moves into higher frequencies. Thus, in order to produce such spectral modification, the higher the first formant frequency, the greater shift of the pair, and then the greater velum lowering is needed.

This hypothesis also explain why the simple acoustic tube without the side cavity as the nasal tract cannot produce good nasal sounds for mid or low vowels. The transfer function of the vowel \[ \text{a} \] and its nasalized version by employing the simple tube is shown in Fig. 6.

Fig. 6 The transfer function of the vowel \[ \text{a} \] and that with NC=0.5 sq cm. The sinus cavity is disconnected.

In the case of this weak coupling, the nasal pole-zero pair appears between the first and second vowel formant, as indicated by \( N_1 \) and \( A_1 \) in Fig. 6. If the degree of coupling is increased, the pole-zero pair would shift upward, and eventually the second vowel formant and the nasal zero would cancel out each other, leaving only the first vowel formant and the nasal formant in that region. In such case a listener may interpret the signal as a different vowel rather than the nasalized vowel.
The assumption we postulated, however, seems to contradict with our observation in the case of the vowel [i], where the nasal pole-zero frequencies are always higher than the first formant frequency, and the pair shifts toward the second formants, departing from the first formant, when the coupling increases, as shown in Fig. 7.

Note that the side cavity is not connected in this particular simulation. According to the hypothesis, the nasality should lessen in such case, which is the opposite of our listening impression. The modification of the wide valley between the first and the second formant, in this case, seems to contribute to the perception of nasality (8). This issue must be investigated yet on the basis of more formal perceptual tests, however. In any case, the qualities of all nasalized vowels seems to be far more satisfactory than those synthesized without the sinus cavity.

CONCLUSION

The simulation experiment has suggested strongly that the nasal pole-zero pair must appear in the vicinity of the first vowel formant for mid and low vowels to be nasalized appropriately. The nasal side cavity may be regarded as an effective means for lowering the nasal pole, which is otherwise too high, to the first formant or lower region. In this sense, this study has made the participation of the nasal sinuses even more plausible. However, to answer the question whether or not the sinuses actually play a relevant role in the production of nasal sounds, we need further corroborations based on acoustic and anatomical data. Finally, it is emphasized that the simulation experiment such as described here would provide a powerful tool for studying the perceptual cues of nasal and nasalized sounds.

ACKNOWLEDGEMENTS

The author wishes to thank Osamu Fujimura for helpful discussion and advice during the course of this study.

REFERENCES