Journal of Oral Rehabilitation 2000 **27**; 723–730

Spectral characteristics of hypernasality in maxillectomy patients[1]

H. YOSHIDA*, Y. FURUYA*, K. SHIMODAIRA†, T. KANAZAWA†, R. KATAOKA‡ & K. TAKAHASHI‡ *Department of Oral and Maxillofacial Surgery, Institute of Clinical Medicine, University of Tsukuba, †Division of Dental Laboratory, Tsukuba University Hospital, and ‡First Department of Oral and Maxillofacial Surgery, School of Dentistry, Showa University, Japan*

SUMMARY **To reveal the acoustic characteristics associated with hypernasality and to ascertain their correlation to the severity of hypernasality, 30 speech samples produced by 15 maxillectomy patients were acoustically analysed with and without an obturator prosthesis in place. The isolated, sustained Japanese vowel /i/ was used as the stimulus for acoustic measurement and perceptual judgment to evaluate the severity of hypernasality. Normalized 1/3-octave spectral analysis demonstrated the spectral characteristics of hypernasality as a rise in amplitude between the first and second formants around the 1 kHz region, and a reduction in amplitude of the frequencies higher than the second formant. High correlation was shown between the perceptual ratings and the predicted values derived from stepwise regression analysis.**

Introduction

Most patients who have undergone resection of the maxillae for cancer in the palatomaxillary region suffer from dysfunctions such as disorders in speech, swallowing and mastication, because they cannot achieve effective separation of the oral and nasal cavities. Coupling of the oral and nasal cavities increases nasal resonance, resulting in hypernasality. Mild hypernasal speech sounds slightly distorted to a listener, but severe hypernasality can render a patients' speech virtually unintelligible, resulting in inhibition of patients from social contact. Prosthodontists and oral and maxillofacial surgeons are responsible for restoring sufficient oral–nasal partition in maxillectomy patients to permit correct articulation. Placement of obturator prosthesis or surgical closure with effective decoupling of oronasal cavities can reduce nasal resonance and improve speech intelligibility (Minsley, Warren & Hinton, 1987; Yoshida, Michi & Ohsawa, 1990; Yoshida *et al.*, 1993).

One approach to evaluating effective separation is to quantitatively examine whether the hypernasality is eliminated or reduced following treatments. Clinical evaluation of the degree of nasality has been dependent on the subjective judgment of individual listener in how they perceive the sounds. Although perceptual assessment has provided adequate reliability in quantitative measurements of hypernasality in research settings, particularly when the judges are experienced in evaluating nasality, it offers poor intra- and inter-judge reliability in rating hypernasality in clinical settings (Bradford, Brooks & Shelton, 1964). Consequently, instrumental measurements, such as those used in aerodynamic or acoustic analyses, have received much interest with regard to their potential for clinical use in quantitatively evaluating nasal resonance.

Few reports have instrumentally measured hypernasality in maxillectomy patients to objectively determine the effectiveness of maxillary obtutator prostheses. Aerodynamic measurements have shown

[1] This work has been carried out at the Department of Oral and Maxillofacial Surgery, Institute of Clinical Medicine, University of Tsukuba.

an increase of oral pressure and nasal airway resistance, and a decrease of nasal air flow rate, cross-sectional nasal airway area and respiratory volume after placement of prosthesis (Minsley *et al.*, 1987Leeper *et al.*, 1994). The only acoustic analysis of which we are aware has revealed that prosthodontic restoration significantly reduced nasal resonance by either eliminating the nasal resonance, reducing the amplitude of the resonance, or changing the frequencies of the formants toward the more normal frequency regions of the vowels (Tobey & Lincks, 1989). These studies have represented instrumentally measured variables associated with hypernasality, but have not yet determined the correlation between the severity of nasality and these variables.

Several acoustic measurements have been conducted to quantitatively evaluate hypernasality in cleft palate patients (Forner, 1983; Philips & Kent, 1984; Kataoka, 1988; Kataoka *et al.*, 1996). However, the cause of hypernasality in maxillectomy patients physiologically differs from that in cleft palate patients: hypernasality in the former results from ineffective oronasal separation, while in the latter it results from velopharyngeal inadequacy. Because of this difference, findings derived from cleft palate patients may not be applicable to maxillectomy patients.

We previously reported the correlation of spectral properties and quantitative evaluation of hypernasality in acoustic analysis of cleft palate children (Kataoka *et al.*, 1996). This study was carried out to reveal acoustic characteristics associated with hypernasality in maxillectomy patients and to establish their correlation to the severity of hypernasality, by modification of the approach used for cleft palate patients.

Materials and methods

Subjects

Fifteen maxillectomy patients, consisting of four males and 11 females ranging in age from 38 to 78 years, were selected from a population of individuals receiving prosthetic treatment for coupling of the oral and nasal cavities resulting from various types of resection of palatomaxillary or maxillary sinus cancers. All patients had undergone surgery and fabrication of obturator prosthesis at the Department of Oral and Maxillofacial Surgery, Institute of Clinical Medicine, University of Tsukuba. Resected portions included the hard palate and the anterior segment of the soft palate in 12 patients, and only the hard palate in the other three patients. They had been normal speakers prior to the surgery, and suffered from nasal resonance attributed not to velopharyngeal inadequacy, because the middle and posterior segments of their soft palate remained intact, but to ineffective oronasal partition.

Speech samples

The Japanese vowel /i/ sound in isolation was used as the stimulus to evaluate hypernasality, because clinical judgments of hypernasality are based on vowels. Among the five Japanese vowels /a/, /i/, /u/, /e/ and /o/, the vowel /i/ was considered to be the most suitable for judging the severity of hypernasality, because nasal sound transmission can be appreciable even with a small degree of nasal coupling due to the narrow lingual constriction of this vowel (Fant, 1970).

The subjects were asked to produce the vowel /i/ for 0·5 s at an individually preferred pitch and loudness, with and without the obturator prosthesis in place. An electret condenser microphone (ECM-44S†) was placed approximately 15 cm from the lips of the speaker. Recordings were made using a digital audio tape-recorder (TCD-3D†) in a quiet room with background noise less than 18 dB.

Thirty speech samples were recorded: 15 pronounced with and the other 15 without the patients' prosthesis in place.

Acoustic measurements

The speech samples were acoustically analysed by a modification of the method previously used for assessment of hypernasality in cleft palate patients (Kataoka *et al.*, 1996). Briefly, each speech sample was low-pass filtered below 8 kHz with a roll-off of 84 dB per octave, and digitized at a sampling rate of 20 kHz using a 16-bit analog-to-digital converter (Sound Master‡) for storage on floppy disk. A whole segment of the time-domain waveform of the stored sample served as input data for spectral analysis. The data were transformed into its power spectrum by means of fast Fourier transform

† Sony, Tokyo, Japan.
‡ Canopus, Tokyo, Japan.

analysis, and then was 1/3-octave band-pass filtered in the frequency range of 125 Hz to 6·3 kHz. After the mean amplitude at every 1/3-octave band was calculated, a 1/3-octave spectrum of the speech sample was displayed as a frequency-versus-amplitude envelope on the screen of a computer (PC-9801 nx/c[§]). To minimize the effects of individual differences such as loudness of voice, the 1/3-octave spectrum was normalized by adjusting the amplitude of the band containing the fundamental frequency to 0 dB. The acoustic characteristics of hypernasality were assessed by comparing the normalized 1/3-octave spectra of the samples produced with and without prosthesis.

Perceptual judgment

Two speech pathologists participated as listeners for perceptual judgment of hypernasality. Each had received a professional education, and had 15 years' experience evaluating and treating hypernasality in cleft palate patients. The stored speech samples were read from the floppy disk, reconverted to analog signals, and standardized to the 60 dB sound pressure level. The samples were randomly ordered, and two sets of samples were created to eliminate ordering effects. The order of presentation of the two sets was randomized across listeners, and no information about the subjects or speaking conditions was disclosed to the listeners. The stimuli were presented over headphones (CD-550[†]) in a quiet room. The listeners were asked to rate the perceived severity of hypernasality on a 5-point equal-appearing interval scale with 0 representing no nasal resonance and 4 representing severe hypernasality. The listeners were encouraged to listen to the stimuli as many times as needed to feel confident about their ratings. The mean value of the ratings obtained by the two listeners was determined to constitute the perceptual severity of hypernasality.

Statistical analyses

One-factor analysis of variance was employed to compare the amplitude of each multiple between the normalized 1/3-octave spectra with and without prosthesis. To ensure the reliability of the perceptual assessments used as standards for rating the severity of hypernasality, intra- and inter-listeners reliabilities were evaluated using the Spearman rank correlation for the speech samples. The Mann–Whitney *U*-test was used to compare the perceptual nasality scores with and without prosthesis. The correlation between the amplitude of the multiples in the normalized 1/3-octave spectra and the perceptual assessments was measured by stepwise regression analysis. All statistical procedures were carried out using a statistical software package (SPSS 6.1[¶]).

[§] NEC, Tokyo, Japan.

Results

Acoustic characteristics of hypernasality

The fundamental frequency in the 1/3-octave spectrum of the vowel /i/ was displayed as the frequency band between 160 and 200 Hz for males, and between 200 and 250 Hz for females. The first and second formants were depicted as the frequency bands from 320 to 400 Hz and 2 to 2·5 kHz for males, and from 400 to 500 Hz and 2·5 to 3·2 kHz for females, respectively (Fig. 1). By normalization of the fundamental frequency, the third and eleventh multiples were represented as the bands containing the first and second formants for both males and females, respectively (Fig. 2).

To demonstrate the acoustic characteristics of hypernasality, the normalized 1/3-octave spectra of the vowel /i/ with and without the obturator prosthesis in the same patient were compared by superimposing the band containing the fundamental frequency (Fig. 3). Collectively comparing each set of 15 spectra, those without the obturator prosthesis showed a significant increase in the amplitudes of the sixth, seventh and eighth multiples, and a significant decrease in the amplitudes of the eleventh and twelfth multiples relative to those with prosthesis (6th: $F[1, 28] = 7{\cdot}7101$, $P = 0{\cdot}0097$; 7th: $F[1, 28] = 15{\cdot}9999$, $P = 0{\cdot}0004$; 8th: $F[1, 28] = 11{\cdot}5518$, $P = 0{\cdot}0020$; 11th: $F[1, 28] = 4{\cdot}3806$, $P = 0{\cdot}0455$; 12th: $F[1, 28] = 6{\cdot}5291$, $P = 0{\cdot}0163$). Accordingly, the acoustic properties of hypernasality in the maxillectomy patients were recognized as a rise in the amplitude between the first and second formants around the 1 kHz region, and a reduction in the amplitudes of the frequencies higher than the second formant (Fig. 4).

[¶] SPSS Inc., Chicago, IL, U.S.A.

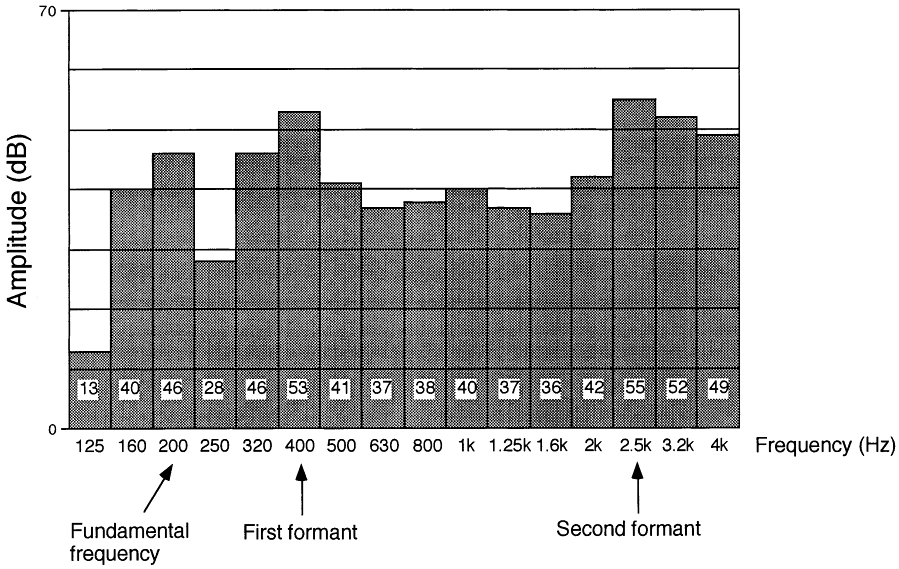

Fig. 1. A 1/3-octave spectrum displayed on the computer. The numbers within each column indicate the amplitude of each 1/3-octave frequency bandwidth.

Perceptual assessment of hypernasality

Intra-listener reliability was evaluated by comparing differences between each perceptual rating by the same listener for the same stimuli. Of the 30 speech samples, exact agreement and agreement within 1 scale value of the ratings by one listener were 20 (66·7%) and 30 (100·0%), and those by the other were 23 (76·7%) and 30 (100·0%), respectively. The correlation coefficients were 0·9094 ($P < 0{\cdot}001$) and 0·9416 ($P < 0{\cdot}001$), respectively. Inter-listener reliability was evaluated by comparing the mean perceptual ratings of each listener for the same speech samples, and found to be 19 (63·3%) for exact agreement and 27 (90·0%) for agreement within 1 scale value, with a correlation coefficient of 0·8729 ($P < 0{\cdot}001$) (Table 1).

The means and standard deviations (s.d.s) of nasality scores were 3·767 and 0·586 for the samples without prosthesis, and 1·433 and 0·980 for those with prosthesis, respectively. There was a significant difference in the nasality scores between samples produced with and without prosthesis ($P < 0{\cdot}001$).

Correlation of spectral analysis with perceptual assessment

The following equation derived from stepwise regression analysis demonstrated the correlation between spectral characteristics and perceptual ratings of hypernasality:

$$y = 3{\cdot}43 + 0{\cdot}12\mathrm{F7} - 0{\cdot}08\mathrm{F12},$$

where y represents the predicted value, and F7 and F12 signify the amplitudes of the seventh and the twelfth multiples in the normalized 1/3-octave spectra, respectively (Fig. 5). There was high correlation between the perceptual ratings and the predicted values ($r = 0{\cdot}8419$, adjusted $r^2 = 0{\cdot}6872$, $F[2, 27] = 32{\cdot}8480$, $P < 0{\cdot}001$).

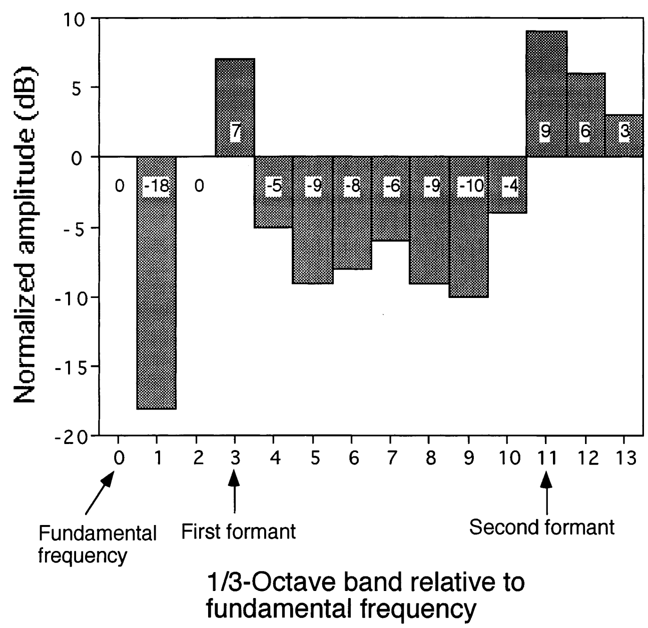

Fig. 2. A 1/3-octave spectrum normalized by adjusting the amplitude of the band containing the fundamental frequency to 0 dB. The third and eleventh multiples contain the first and second formants. The numbers within each multiple indicate the normalized amplitude of each 1/3-octave frequency bandwidth.

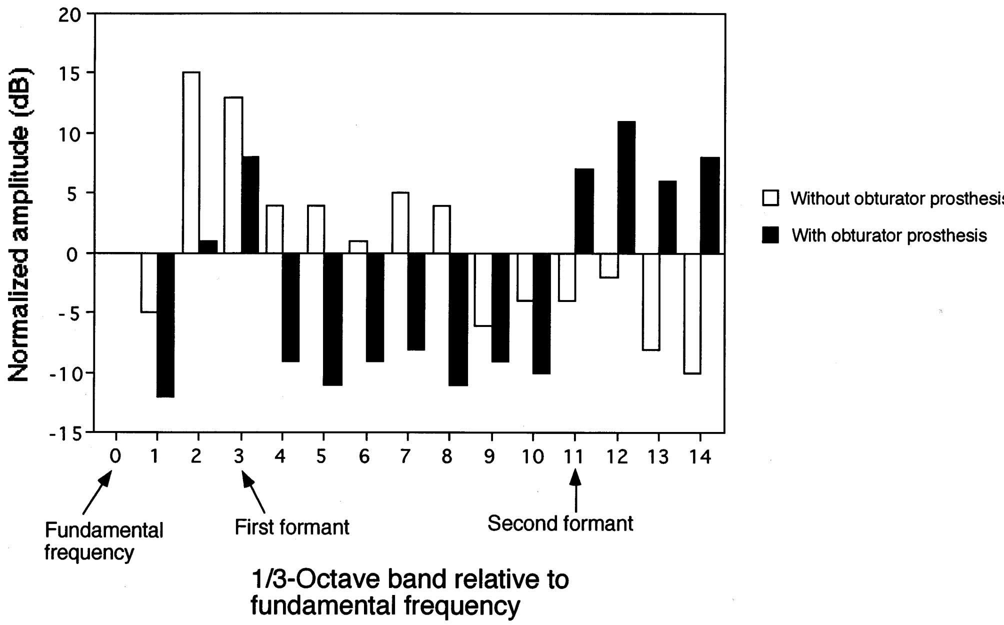

Fig. 3. Superimposition of a normalized 1/3-octave spectra of the vowel /i/ produced with and without the obturator prosthesis in the same patient.

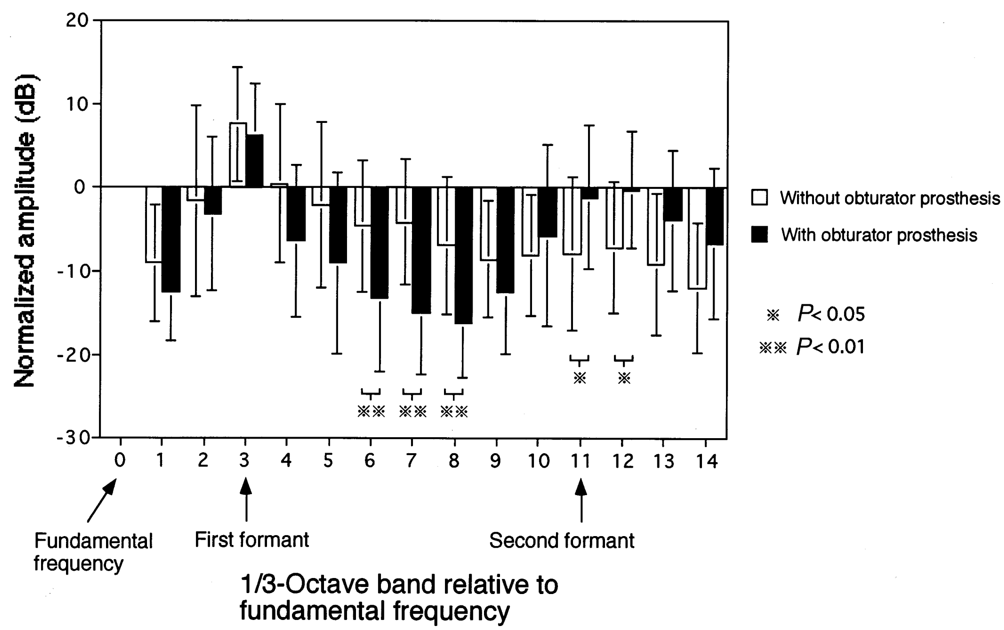

Fig. 4. Collective superimposition of each set of 15 normalized 1/3-octave spectra of the vowel /i/ produced with and without the obturator prosthesis. Values represent mean and 1 s.d. Comparing the two speech conditions, there are significant increases in the amplitudes of the sixth, seventh and eighth multiples, and significant decreases in the amplitudes of the eleventh and twelfth multiples in spectra without prosthesis, relative to those with prosthesis. These results indicate that the acoustic characteristics of hypernasality in maxillectomy patients were a rise in the amplitude between the first and second formants around the 1 kHz region, and a reduction in the amplitude of the frequencies above the second formant.

Table 1. Rating reliability for hypernasality (%)

	Intra-listener reliability		
	Listener A	Listener B	Inter-listener reliability
No. of exact agreements	20 (66·7)	23 (76·7)	19 (63·3)
No. of agreements within 1 scale value	30 (100·0)	30 (100·0)	27 (90·0)
Correlation coefficient (r)	0·9094 ($P<0·001$)	0·9416 ($P<0·001$)	0·8729 ($P<0·001$)

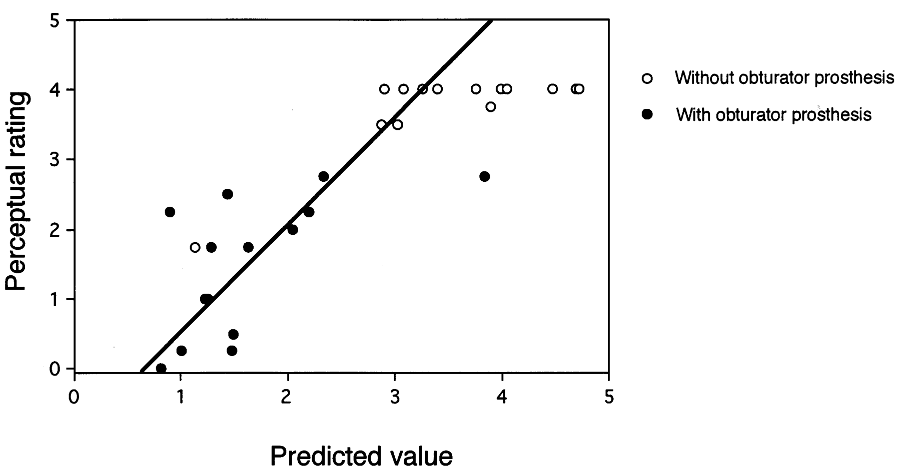

Fig. 5. Scattergram showing high correlation between the perceptual ratings and the predicted values calculated from the amplitudes of the seventh and twelfth multiples by stepwise regression analysis. The multiple regression equation is: $y = 3·43 + 0·12F7 - 0·08F12$ (y represents the predicted value, and F7 and F12 signify the amplitudes of the seventh and the twelfth multiples in the normalized 1/3-octave spectra, respectively).

Discussion

The speech stimuli presented to listeners for rating hypernasality have typically been of one or more of four types: isolated vowels, consonant–vowel–consonant syllables, connected speech played forward, and connected speech played backward (Counihan & Cullinan, 1970). Use of isolated vowels has been preferred to the other three types with consonants, because the production of isolated vowels can be more easily controlled and standardized than that of connected speech, and because isolated vowels are less affected by confounding articulatory information than are connected speech fragments (de Krom, 1994). Listeners can focus more closely on the characteristics of the actual voice quality of sustained vowels. Accordingly, the sustained vowel was employed as the stimuli in this study.

The reliability of perceptual evaluations of hypernasality has commonly been calculated with both intra- and inter-listeners' consistency. Although the rating methods and subjects were different, the reliability coefficients for our subjects were comparable to those of previous studies (Bradford *et al.*, 1964Counihan & Cullinan, 1970). Consequently, the perceptual evaluations of hypernasality obtained from this study were considered to be reliable.

The mean perceptual evaluation of 1·433 for the samples produced with prosthesis indicated that nasal resonances were not completely eliminated, but were significantly reduced with the prosthesis. Consequently, the stimuli produced with prosthesis were not always non-nasal. However, considering that the only difference in vowel production was the presence of oronasal separation, the comparison of the samples with and without prosthesis was considered to explicitly demonstrate acoustic characteristics that closely correlated with the severity of hypernasality.

Acoustic analysis has provided valuable information regarding hypernasality (Forner, 1983; Philips & Kent, 1984) and synthesized nasal speech (Kingston & Macmillan, 1995), but involved difficulties in quantifying spectral characteristics. One method to simply demonstrate the spectral characteristics of hypernasality is to measure the amplitude of the spectra at 1/3-octave intervals (Kataoka, 1988; Kataoka *et al.*, 1996).

This specific bandwidth was selected because, over a broad frequency range, it compares well with the critical bandwidth of the analysing mechanism utilized by the ear (Pols, van der Kamp & Plomp, 1969). However, the amplitude of formant frequencies is related to the loudness of voice, and varies among speakers, even for individual production of the same vowels. Therefore, to permit assessment of the severity of hypernasality regardless of age, sex or loudness, each 1/3-octave spectrum was normalized to the amplitude and frequency of the band containing the fundamental frequency (Kataoka, 1988; Kataoka *et al*., 1996). The resulting normalized 1/3-octave spectrum revealed that the bands containing the first and second formant frequencies were represented as the third and eleventh multiples, respectively.

A comparison of the spectra revealed a significant increase in the amplitudes of the sixth, seventh and eighth multiples and a significant decrease in the amplitudes of eleventh and twelfth multiples in the spectra without the obturator prosthesis compared to those with the prosthesis. These results indicate that the acoustic characteristics of hypernasality in maxillectomy were a rise in the amplitude between the first and second formants around the 1 kHz region, and a reduction in the amplitude of the frequencies above the second formant. In the only known report to have analysed acoustic properties in maxillectomy patients, Tobey & Lincks (1989) described significant changes that occurred in the second and third formant frequencies in five maxillectomy patients while speaking with prosthesis, and that prosthodontic restoration significantly reduced nasal resonances depicted between the first and second formants, by either completely eliminating the resonances, reducing the amplitude of the resonances, or changing the frequency of the resonances to regions closer to the vowels. Comparing the results of our study to theirs is difficult because of the different methods of assessing hypernasality. However, the presence of nasal resonances can be presumed to be a rise in amplitude between the first and second formants, and changes in the second and third formant frequencies can be regarded as changes in the amplitudes of these regions. Based on these assumptions, the spectral characteristics of hypernasality in maxillectomy patients are considered to be similar in the results of both studies.

The spectral characteristics of hypernasality demonstrated in the current study were nearly identical to the results obtained from the study of cleft palate patients (Kataoka *et al*., 1996), in which almost the same spectral analysis was used. Consequently, the spectrum of hypernasality is considered to be consistent, even though the causes of nasal resonances are physiologically different.

Stepwise regression analysis specified the seventh and twelfth multiples as the acoustic variables associated with the judgment of the severity of hypernasality in maxillectomy patients, and showed high correlation between the predicted values calculated with these two variables and the perceptual evaluations. The utility of the instrumental measurements is demonstrated by the correlation between measured values and ratings of perceived voice quality (Hillenbrand, 1988; Klatt & Klatt, 1990). If the instrumentally measured values are closely related to listeners' judgments, the measurements are considered to be clinically useful (Gerratt *et al*., 1993). Although the 1/3-octave spectral analysis is suggested to be effective as an instrumental measurement of hypernasality, further research on a larger sample of maxillectomy patients may be necessary to establish the validity of this approach and to determine its efficacy as a definitive tool for clinical use.

Acknowledgments

The authors thank speech pathologists Y. Yamashita and S. Imai, First Department of Oral and Maxillofacial Surgery, School of Dentistry, Showa University for serving as expert listeners. The authors also thank Mr. Edward J. Coan for editorial assistance. This research was supported by Grant-in-Aid for Scientific Research of The Ministry of Education, Science and Culture of Japan (Grant-in-Aid for Developmental Scientific Research, 04557092).

References

Bradford, L.J., Brooks, A.R. & Shelton, R.L. Jr. (1964) Clinical judgment of hypernasality in cleft palate children. *Cleft Palate Journal*, **1,** 329.

Counihan, D.T. & Cullinan, V.L. (1970) Reliability and dispersion of nasality ratings. *Cleft Palate Journal*, **7,** 261.

Fant, G. (1970) Nasal sound and nasalization. In: *Acoustic Theory of Speech Production* (ed. G. Fant), pp. 148–161. Mouton, Hague.

FORNER, L.L. (1983) Speech segment durations produced by five and six year old speakers with and without cleft palate. *Cleft Palate Journal*, **20,** 185.

GERRATT, B.R., KREIMAN, J., BARROSO, N.A. & BERKE, G.S. (1993) Comparing internal and external standards in voice quality judgments. *Journal of Speech and Hearing Research*, **36,** 14.

HILLENBRAND, J. (1988) Perception of aperiodicities in synthetically generated voices. *Journal of Acoustical Society of America*, **83,** 2361.

KATAOKA, R. (1988) Quantitative evaluation of hypernasality – relation between spectral characteristics and reception of hypernasality. *Journal of Japanese Cleft Palate Association*, **13,** 204 (in Japanese).

KATAOKA, R., MICHI, K., OKABE, K., MIURA, T. & YOSHIDA, H. (1996) Spectral properties and quantitative evaluation of hypernasality in vowels. *Cleft Palate-Craniofacial Journal*, **33,** 43.

KINGSTON, J. & MACMILLAN, N.E. (1995) Integrality of nasalization and F1 in vowels in isolation and before oral nasal consonants: a detection-theoretic application of Garner paradigm. *Journal of Acoustical Society of America*, **97,** 1261.

KLATT, D.H. & KLATT, L.C. (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of Acoustical Society of America*, **87,** 820.

DE KROM, G. (1994) Consistency and reliability of voice quality ratings for different types of speech fragments. *Journal of Speech and Hearing Research*, **37,** 985.

LEEPER, H.A., JANZEN, V., SILLS, P.S. & AHMAD, D. (1994) Technique for assessing nasal airway resistance in patients treated prosthetically. *Journal of Prosthetic Dentistry*, **72,** 210.

MINSLEY, G.E., WARREN, D.W. & HINTON, V. (1987) Physiologic responses to maxillary resection and subsequent obturation. *Journal of Prosthetic Dentistry*, **57,** 338.

PHILIPS, B.J. & KENT, R.D. (1984) Acoustic-phonetic descriptions of speech production in speakers with cleft palate and other velopharyngeal disorders. In: *Speech and Language: Advances in Basic Research and Practice* (ed. N.J. Lass), pp. 132–160. Academic Press, Orlando.

POLS, L.C., VAN DER KAMP, L.J.TH. & PLOMP, R. (1969) Perceptual and physical space of vowel sounds. *Journal of Acoustical Society of America*, **46,** 458.

TOBEY, E.A. & LINCKS, J. (1989) Acoustic analyses of speech changes after maxillectomy and prosthodontic management. *Journal of Prosthetic Dentistry*, **62,** 449.

YOSHIDA, H., MICHI, K. & OHSAWA, T. (1990) Prosthetic treatment for speech disorders due to surgically acquired maxillary defects. *Journal of Oral Rehabilitation*, **17,** 565.

YOSHIDA, H., MICHI, K., YAMASHITA, Y. & OHNO, K. (1993) A comparison of surgical and prosthetic treatment for speech disorders attributable to surgically acquired soft palate defects. *Journal of Oral and Maxillofacial Surgery*, **51,** 361.

Correspondence: Dr Hiroshi Yoshida, Department of Oral and Maxillofacial Surgery, Institute of Clinical Medicine, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba-shi, Ibaraki-ken 305-8575, Japan. E-mail: hyoshida@md.tsukuba.ac.jp