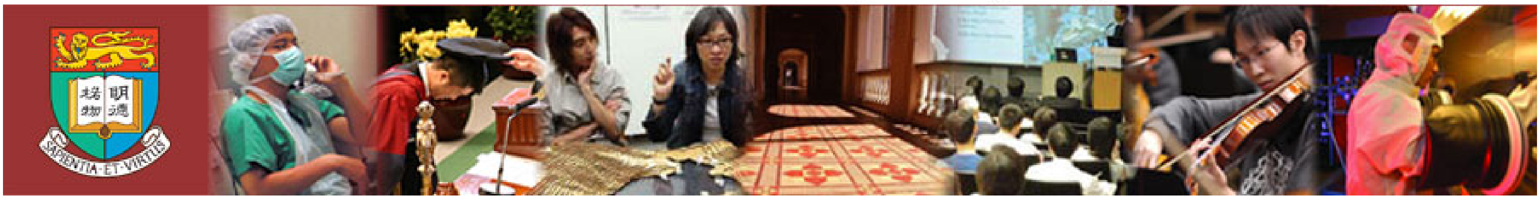
View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by HKU Scholars Hub

The HKU Scholars Hub The University of Hong Kong 香港大學學術庫



Title	Spectral correlates of hypernasality in consonants
Other Contributor(s)	University of Hong Kong
Author(s)	Chen, Pak-wei, Danny
Citation	
Issued Date	2005
URL	http://hdl.handle.net/10722/56196
Rights	Creative Commons: Attribution 3.0 Hong Kong License

Spectral correlates of hypernasality in consonants

Chen Pak Wei Danny

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, May 6, 2005.

Abstract

This study applied one-third-octave analysis for measuring spectral correlates of hypernaslity in consonants produced by speakers with a range of aetiologies (cleft palate, maxillectomy, and dysarthria). The speech materials were consonants /s/ and /n/ segmented from four Cantonese oral sentences and two single words. The results showed that the hypernasal speakers produced /s/ with flattened spectra that significantly lower amplitude at high frequencies above 2000 Hz, and significantly higher amplitude at low frequencies below 500 Hz and mid frequencies from 500 to 2000 Hz than those with normal resonance. Multiple regression analysis revealed a high correlation (R = 0.80) between a combination of the spectral scores for fricatives, nasals, and vowels with the perceptual ratings of hypernasality. These results suggested that these spectral measurements on consonants and vowels are reliable and valid to quantify hypernasality in connected speech.

Introduction

Hypernasality is a perceptual quality associated with excessive nasal resonance because of velopharyngeal incompetence (Kent, & Read, 2002). Subjective judgment of perception of sounds by individuals has been widely used as a standard for rating the degree of hypernasality in clinical settings (Yoshida et al., 2000). This is because nasality is a perceptual quality by definition. However, perceptual judgments are associated with poor intra- and inter-judge reliability due to individual differences of the listeners on rating hypernasal sounds (Bradford, Brooks, & Shelton, 1964; Bzoch, 1989; Counihan, & Cullinan, 1970). Instrumental measurements in evaluating nasal resonance have received much interest due to their potential for clinical use (Yoshida et al., 2000). The use of reliable and valid quantitative analysis could help plan the intervention goal and check the treatment progress of the client. Acoustic analysis is an instrumental technique that investigates directly the product of the speech mechanism. Moreover, this type of analysis is non-invasive, inexpensive, and can be applied to speakers with various age, gender and speech impairments with different etiologies (Kent, & Kim, 2003). However, no acoustic method has found to be effective for estimating the degree of hypernasality in clinical settings (Kataoka, Warren, Zajac, Mayo, & Lutz, 2001). The aim of this study is to apply the one-third-octave analysis for measuring the spectral correlates of hypernasality in speech. This could help to quantify hypernasality using an acoustic method.

Acoustic correlates of hypernasality

Nasalization is associated with a set of acoustic features that can severely compromise the acoustic analysis of a speech signal. The speech signal can be influenced by a high degree of damping, and by antiformants whose potential effects on the acoustic signal are numerous and complex (Kent, & Read, 2002). Kent, Liss and Philips (1989) found that nasalized vowels have acoustic correlates in spectrograms such as an increase in formant bandwidth, a decrease in the overall energy of the vowel, a slight increase of the F1 frequency, and a slight lowering of the F2 and F3 frequencies, and an introduction of a low frequency nasal formant with a center frequency of about 250 to 500 Hz for adult males. Another study by Maeda (1993) claimed that nasalization of the vowel /i/ leads to an increase in amplitude in first formant, and a small trough in the spectrum at around 1000 Hz. Recent studies on the spectral characteristics of hypernasal sounds have used one-third-octave analysis to evaluate hypernasality in children and young adults with cleft palate and cleft lip (Kataoka, Michi, Okabe, Miura, & Yoshida, 1996; Kataoka et al., 2001), and in adults following maxillectomy (Yoshida et al., 2000). One-third octave interval was selected because it compared well with the critical bandwidth of analyzing mechanism utilized by the ears (Kataoka et al., 1996). The power spectrum obtained from digitalized samples was evaluated at every one-third-octave band to calculate the mean power level of each band. These levels were then normalized relative to the amplitude of the band that contained the fundamental frequency (Kataota et al., 1996). Normalization could minimize the effects of individual differences such as the loudness of voice (Yoshida et al., 2000). In the study by Kataoka et al. (2001), an amplitude adjustment was made such that the average amplitude of all one-third-octave bands was subtracted from the amplitude of each band. The spectral characteristics associated with hypernasal sounds were an increase in power level between the first and second formant, and a reduction in power in the second and third formants region (Kataota et al., 1996; Kataota et al., 2001; Yoshida et al., 2000).

Lee, Ciocca and Whitehill (2003) and Lee (2004) replicated the studies by Kataoka and colleagues. They extended the use of one-third-octave analysis to the speech of adults with hypernasality due to different aetiologies (cleft palate, maxillectomy and dysarthria). Lee et al. (2003) found that hypernasality was associated with an increase in amplitude between F1 and F2, and a decrease in amplitude at F2 in the vowel /i/. These findings were consistent with that of the Maeda's (1993) study.

In all the studies on the spectral characteristics of the hypernasality, vowels were used, especially /i/. Vowel /i/ is considered to be ideal for listeners to judge severity of hypernasality among other vowels because even with low degrees of nasal coupling, the narrow tongue passage of the vowel /i/ should lead to noticeable nasal sound transmission (Fant, 1970). Sustained phonation of vowel /i/ was used in the studies by Kataoka and colleagues (Kataota, et al., 1996; Kataota et al., 2001; Yoshida et al., 1996). The vowel /i/ segmented from single words was used in the study by Lee et al. (2003). In the study by Lee et al. (2003), the results showed that speakers with resonance disorders had significantly higher amplitude for the bands centered at 630, 800 and 1000 Hz, and significantly lower amplitude for the band centered at 2500 Hz than speakers with normal resonance. The difference in the mean amplitudes at each of these adjusted frequency bands was only about 5 dB or smaller. Although the spectral characteristics related to hypernasality found in the above studies were consistent in general, the difference between the mean amplitude of the one-third-octave bands of the hypernasal and control group was relatively small. Only patients with quite pronounced degree of hypernasality could be detected acoustically and visually from the spectra. In this way, acoustic analysis of hypernasality did not have a good diagnostic value.

While previous studies focused on vowels, especially /i/, the present study focused on the spectral characteristics of nasalization on consonants. Production of oral consonants requires velopharyngeal closure to separate oral and nasal cavities. In contrast, nasal consonants involve velopharyngeal opening that sound energy propagated to the nasal cavity (Pickett, 1999). Therefore, stops and fricatives would be affected by velopharyngeal incompetence as it prevents the speaker from building up sufficient pressure in the oral cavity.

Resulting escape of air from nose would result in nasal resonance (Borden et al., 2003). It is then suggested that nasalized oral consonants might show up with spectral features of nasal consonant counterparts with the same place of articulation. Fricatives are sounds involving noise generation, and a restricted airflow characterized by noise spectra. Stops are characterized by complete blockage of airstream for a brief period of time followed by a sudden release of constriction. The acoustic signal of a typical stop consists of a short closure interval, a release burst and formant transitions (Borden et al., 2003; Kent, & Read, 2002). When comparing with stops, fricatives have relatively longer duration of noise with lengthy interval aperiodic energy. Therefore, fricative is chosen in this study for its salient feature in the spectra with a longer noise segment presented. In Cantonese, /s/, /f/ and /h/ are the only sounds in the class of fricative (So, & Dodd, 1995). For the labiodental fricative /f/, the fricative energy is relatively low because there is no appreciable resonating cavity anterior to the point of constriction with spectral peaks around 1500 to 8500 Hz; for the lingua-alveolar fricative /s/, most of its high sound energy is above 4000 Hz with spectral peaks around 5000 to 8000 Hz; for the glottal fricative /h/, there is no particular spectral peak with diffused spectrum (Kent, & Read, 2002). Fricative /s/ is then used in this study due to its prominent peak at high frequencies and the presence of a nasal counterpart /n/. The presence of nasalization could then be easier to be measured in the one-third-octave bands spectra with reference to the nasal counterpart.

Nasal consonants produce the nasal murmur is characterized by a low frequency spectral prominence below 500 Hz, and low amplitude peaks above 500 Hz due to presence of the antiformants (Pickett, 1999). Nasal /n/ would have the antiformants between 1450 and 2200 Hz; these antiformants are capable of absorbing acoustic energy and lead to the reduction of amplitude at frequencies of antiformants (Kurowski, & Blumstein, 1993). If a fricative consonant is nasalized, it may contain i) a nasal murmur that shows up as a low frequency peak, and ii) lower amplitude at high frequencies due to the presence of antiformants. Since both /s/ and /n/ have the same place of articulation, lingua-alveolar, it was hypothesized that production of a nasalized fricative /s/ due to the velopharyngeal incompetence would have similar spectral characteristics of that of a nasal /n/. It was hypothesized that there would be much higher mean amplitude of a nasalized /s/ in the lower frequencies below about 500 Hz, and much lower mean amplitude of the nasalized /s/ in the higher frequencies above about 2000 Hz compared with a non-nasalized one. Therefore, the difference in mean amplitude of fricative /s/ and nasal /n/ on certain sets of bands at both high, mid and low frequencies produced by the hypernasal group should be much smaller than that of the control group. Greater significant differences in the mean amplitude on certain sets of one-third-octave bands of consonants between both groups of speakers might help to characterize the presence of nasalization better than that of vowel /i/.

Perceptual judgments of hypernasality

Whitehill, Lee and Chun (2002) and Zarick and Liss (2000) suggested that direct magnitude estimation (DME) would be more valid method to evaluate hypernasality when comparing with the equal-appearing interval scale. In using DME for perceptual judgments, a standard speech sample is provided and the listeners were then free to assign a number to represent the level of nasality to the modulus. Direct magnitude estimation with practice and feedback during training would be preferable due to its improvement on intra- and inter-judge reliability (Lee, 2004). The correlation between the results of one-third-octave analysis in the present study and the perceptual ratings on the connected speech samples produced by the same set of hypernasal speakers done in study of Lee (2004) was measured. This correlation might reveal an association between the spectral characteristics and the perceptual ratings on the hypernasality of clients with different aetiologies.

Research aims

First, Lee et al. (2003) applied one-third-octave analysis in vowel /i/ and the speakers with hypernasality showed significant higher differences in only four one-third-octave bands than those with normal resonance; and the difference of mean amplitude on these bands was small at around 5 dB. Therefore, the first research aim was to extend the use of one-third-octave analysis in consonants. It was hypothesized that the difference of mean amplitude between the oral /s/ and nasal /n/ consonants at high, mid and low frequency bands would be much smaller than in the experimental group. The spectral characteristic of nasalized fricative /s/ would help determine the presence and the degree of hypernasality.

Second, the spectral characteristics of consonants on each type of aetiologies and their correlations with the perceptual ratings were investigated. The generalizability of this measure to a range of aetiologies in consonants was tested.

Third, the relationship between the spectral analysis on consonants in this study and on vowels (Lee et al., 2003; Lee, 2004), and the perceptual judgments (Direct Magnitude Estimates: Practice-Feedback) of hypernasality was examined by using of multiple regression analysis. This might further reveal that these acoustic parameters were reliable to quantify hypernasality.

The overall goal of the study was to develop a more reliable and valid way of quantifying perceived hypernasality acoustically using fricative /s/, and nasal /n/ together with vowels samples for use in clinical settings for making diagnosis and charting the prognosis of the clients with hypernasality.

Method

Subjects

The speakers of the experimental and control groups in this study have been previously described (Lee, Coicca, & Whitehill, 2003; Lee, 2004). In the experimental group, the speakers were 20 adults with hypernasality (ten males and ten females) aged from 21 to

65 years with a mean age of 40 years (*SD* = 12). They had different etiologies that six had dysarthria, two had maxillectomy, and twelve had cleft palate. The speakers with dysarthria were recruited from the Speech and Hearing Language Clinic of The University of Hong Kong, the Spastic Association of Hong Kong, and the Tung Wah Hospital. The speakers with cleft palate were recruited from the Cleft Lip and Palate Centre of the Prince Philip Dental Hospital, The University of Hong Kong. Those with maxillectomy were recruited from the Oral and Maxillofacial Unit, Faculty of Dentistry, The University of Hong Kong.

All speakers passed screening tests on hearing, speech and language. These tests were (a) a pure-tone audiological screening at 35 dB HL at the octave frequencies of 250 to 4000 Hz for the better ear (except two patients who only had octave frequencies of 500 to 4000 Hz tested); and (b) an aphasia/apraxia screening test derived from the Cantonese Aphasia Battery (Yiu, 1992). All the speakers were judged to have hypernasality by the first and third authors who are qualified speech-language therapists and specialized in resonance disorders (Lee et al., 2004). An informal severity rating ("mild", "moderate", and "severe") was made by these two authors based on sentences with no nasal consonants that were used in this study (Appendix A). All those recruited as speakers were judged to have resonance disorders with hypernasal speech without hearing impairments, aphasia and apraxia of speech.

Twenty age- and gender-matched Cantonese speakers with no history of neurological disorders or facial anormalies were recruited. All were independently judged by authors Lee and Whitehill to have normal resonance with the speech samples in Appendix A. Screening tests on hearing, speech and language were carried out. These speakers passed pure-tone audiological screening at 25 dB HL at the octave frequencies of 250 to 4000 Hz for the better ear (except two speakers who had 35 dB HL and 40 dB HL at 4000 Hz for the better ear), the aphasia/apraxia screening and oral-motor examination. However, only 12 (nine males and three females) of the speakers in the control group with a mean of 39 years (*SD* = 12) were used in this study, because only these control speakers in the study of Lee (2004) had recorded the oral sentences in Appendix A. These oral sentences contained words with consonants /s/ that were targeted in this study.

Materials and procedures

The speech materials for one-third-octave analysis in this study were productions of /n/ within two words loaded with nasal stops at initial and final positions (NVN); the two words were /mɐn_{22}/ and /min_{22}/. Each speaker repeated the words five times in randomized order within a list containing other 16 Cantonese single words with consonant-vowel consonant structure. Another set of speech materials were four oral sentences (Sentence 4, 5, 6, and 9 in Appendix A). In these four sentences, there were six Cantonese single words with /s/ at the initial position in the structure of CVV, CV and CVC (two occurrences /sei_{33}/, /sik_{2}/, /si_{21}/, /sik_{5}/, and /sy_{55}/). The speech samples of the single words and oral sentences were recorded on to a DAT tapes using a digital tape recorder (TASCAM DA-30 MK II) and a Bruel and Kjaer Type 4003 microphone with mouth-to-microphone distance of about 10 cm in a soundproof booth or in a quiet room.

Spectral Analysis

For spectral analysis, the speech samples were digitalized at a sampling rate of 44.1 kHz with 16-bit resolution using the Sound Designer II software (version 2.8, Degidesign Inc.). The middle 50 or 100 ms portion of /s/ or /n/ was segmented and analyzed in forms of one-third-octave interval bands using Praat version 4.0.1 (Boersma and Weenink, 1992-2001). The amplitude of each of the 21 one-third-octave bands from 80 to 12600 Hz (80 to 125 Hz, 126 to 160 Hz, 161 to 200 Hz, 201 to 250 Hz, 251 to 320 Hz, 321 to 400 Hz, 401 to 500 Hz, 501 to 630 Hz, 631 to 800 Hz, 801 to 1000 Hz, 1001 to 1250 Hz, 1251 to 1600 Hz, 1601 to 2000 Hz, 2001 to 2500 Hz, 2501 to 3200 Hz, 3201 to 4000 Hz, 4001 to 5000 Hz, 5001 to 6300 Hz, 6301 to 8000 Hz, 8001 to 10000 Hz, and 10001 to 12600 Hz) of the extracted portions were measured. The amplitude of the extracted segment was normalized following the procedures by Kataoka et al. (2001), Lee et al. (2003) and Lee (2004). The amplitude value of each one-third-octave band was deducted by the average amplitude values of all the bands. The mean amplitude was calculated for each adjusted one-third-octave band from segmented /s/ and /n/ of each speaker in the experimental and control groups.

Perceptual analysis

Listeners for perceptual ratings of Direct Magnitude Estimation (DME) on hypernasality speech in this study were 36 Cantonese native listeners (four males and thirty two females) (Lee, 2004). They were students in the Division of Speech & Hearing Sciences with limited exposure to hypernasality speech in the one-hour nasometry lecture or no previous experience with hypernasality speech. These listeners had normal hearing and passed a pure-tone audiological screening at 20 dB HL at the octave frequencies of 250 to 4000 Hz for the better ear. They were given short training sessions with audiosamples of varying degree of hypernasality, and were randomly assigned to three listeners group: Exposure, Practice, and Practice-Feedback groups. Feedback was given to the listeners when practicing rating hypernasality in the Practice-Feedback group in the training sessions. Both feedback and practice led to significant increase in inter- and intra-reliability for perceptual rating (Lee, 2004). They gave perceptual ratings of the speakers in the experimental group that based on the oral sentences in Appendix A. Listeners heard the stimuli through the headphones (AKG K141 Monitor); the speech samples were presented using a HyperCard program running on a Macintosh G3 ibook. The listeners were instructed to listen to a modulus and assign any positive integer to it. For example, nasality of a modulus was assigned a value “100”. If the first sample was as twice nasalized as the modulus, a rating of “200” would be given to the first sample. The order of presentation of stimuli was randomized and the modulus was repeated after each speech sample. Listeners could hear the stimuli more than once, and rated each speech sample by assigning a numerical value to present the nasality level relative to the modulus.

Results

Intra- and inter-judge reliability of spectral analysis on both /s/ and /n/ was measured using the Pearson product-moment correlation. Fourteen percent of the total samples (one sample of /s/ and one sample of /n/ in each subject) were randomly selected and analyzed again to test the intra-judge reliability. A year four student in the Division of Speech & Hearing Sciences, who had basic knowledge of acoustics of speech, randomly selected one /s/ sample and one /n/ sample from each subject for estimating inter-judgment reliability. Pearson product-moment correlations showed high inter-judge (r = 0.97) and inter-judge reliability (r = 0.96).

Comparison of the spectral scores between the experimental and control groups

Articulation of consonants /s/ and /n/ of each sample in both groups was judged by the author in this study and other two listeners who were year four students from the Division of Speech & Hearing Sciences through listening to the whole sentence and word, but not the extracted segments. Articulations of /s/ in all the speech samples in the control group were perceived as normal. However, articulation errors were made in the experimental group and summarized in Appendix B. Some speech samples /s/ perceived as affricate /ts/ and stop /t/ respectively. Since this study explored the spectral characteristics of the hypernasality of consonants, all substitution errors made in the speech sample by speakers in the experimental group were discarded in this study. Moreover, the speech samples for which segments of 50 ms or longer could not be extracted, and those with deleted consonants were also discarded. Therefore, nineteen speech samples of /s/ were discarded. The number of subjects in the experimental group for /s/ measurements became 19 (n = 19). No articulation errors were noted in the speech samples of /n/ in both the experimental and control groups.

The mean amplitude of the 21 adjusted one-third-octave bands for /s/ and /n/ samples were calculated and plotted as a line graph. The adjusted one-third-octave spectra of /s/ and /n/ for both the experimental and control groups were shown in Figure 1. The pattern of spectra for /n/ in both control and experimental groups were similar. However, in the spectra of /s/, the experimental group showed lower mean amplitude at high frequencies above 2000 Hz, and higher mean amplitude at low frequencies below 500 Hz and mid frequencies from 500 to 2000 Hz than that of the control group.

As number of subjects in the control (n = 12) and experimental (n = 20) groups were different for /n/ measurements, Mann-Whitey U tests were conducted to compare the mean amplitude of each of the 21 one-third-octave bands in both groups. The normal group had significantly higher amplitude ($p < 0.05$) only for 3201 to 4000 Hz band.

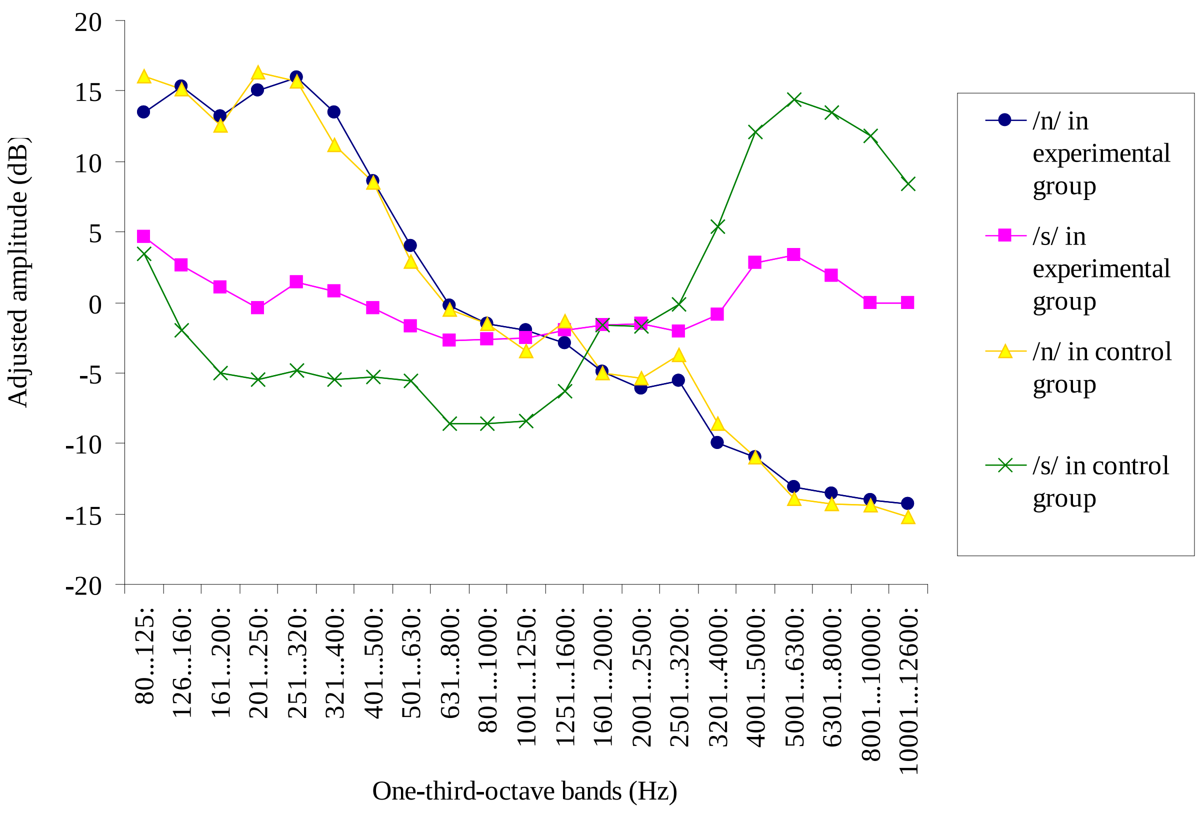


Figure 1. Adjusted one-third-octave spectra of /s/ and /n/ of the experimental and control groups

Visual inspection on the spectra of the consonants revealed discrepancies between the two speaker groups for these consonants. At high frequency bands above 2000 Hz, five bands of both /s/ and /n/ (4001 to 5000 Hz, 5001 to 6300 Hz, 6301 to 8000 Hz, 8001 to 10000 Hz, and 10001 to 12600 Hz) were selected. For each subject, the average of the mean amplitude of the five bands of /s/ was calculated, and minus that of the average of the 5 bands of /n/. This score was taken for each of the subject and named “High5”. At low frequencies bands below 500 Hz, five bands (126 to 160 Hz, 161 to 200 Hz, 201 to 250 Hz, 251 to 320 Hz, and 321 to 400 Hz) of both /s/ and /n/ were selected. For each subject, the average of the mean amplitude of the five bands of /n/ was calculated, and minus that of the average of the five bands of /s/. This score was taken for each of the subject and named “Low5”. Another score was generated for each sample of /s/. Visual inspection revealed that the slope of the spectra of /s/ in the experimental group was flattened. Three bands (4001 to 5000 Hz, 5001 to 6300 Hz, and 6301 to 8000 Hz) at high frequencies and the other three bands (631 to 800 Hz, 801 to 1000 Hz, and 1001 to 1250 Hz) at mid frequencies were selected. The difference of the averaged mean amplitude of the two sets of bands was calculated in each sample and this score was named “High-Mid3”. The mean score of “High-mid3” on samples /s/ was calculated for each subject.

To summarize, there were three scores: “High5”, “Low5”, and “High-Mid3”. Results of Mann-Whitney *U* tests were conducted to compare these scores of the experimental group (n = 19 as all speech samples /s/ of S19 were discarded) with those of the control group (n = 12) shown in Table 1. Non-parametric tests were then applied in this study because of the violation of the homogeneity of variance assumption for *t*-tests. The differences in mean scores of “High5”, “Low5”, and “High-Mid3” were around 10, 5, and 16 dB respectively as shown in Table 1.

Table 1

Means and variances on the spectral scores of the experimental (n = 19) and control (n = 12) groups

	Mean (in dB)		Variance	
	Experimental	Control	Experimental	Control
Score	(n = 19)	(n = 12)	(n = 19)	(n = 12)
High5	15.10	25.78	41.72	7.70
Low5	13.76	18.72	29.55	9.30
High-Mid3	5.83	21.84	70.6	18.72

Note. All aetiologies were included in the experimental group (11 had cleft palate, 6 had dysarthria, and 2 had maxillectomy).

The Mann-Whitney *U* tests revealed significant differences in “High5” ($p < 0.01$), “Low5” ($p < 0.01$), and “High-Mid3” ($p < 0.01$) between the experimental and control groups. Mann-Whitney *U* tests were also conducted for different aetiologies (cleft palate, oral cancer and dysarthria). There were four sub-groups in the experimental group: 1) all aetiologies included (n = 19); 2) cleft palate (n = 13); 3) cleft palate and maxillectomy (n = 13); and 4) dyarthria (n = 12). These revealed that significant differences ($p < 0.01$) in the scores of “High5” and “High-Mid3” between the control group and experimental sub-groups of all aetiologies included, cleft palate, cleft palate and maxillectomy, and dysarthria. Moreover, there were statistically significant differences between the control and experimental groups on the score of “Low5” in all aetiologies included ($p < 0.01$), cleft palate and maxillectomy ($p < 0.05$), and dysarthria ($p < 0.01$), except for cleft palate.

Correlations between the spectral scores and perceptual ratings on hypernasality

Pearson product-moment correlation between each score in each sub-group of the experimental speakers and the geometric means of Direct Magnitude Estimation: Practice-

Feedback (DME-PF) was conducted. For all the etiologies included, significant correlations ($p < 0.05$) were found for all three scores: “High5” (r = - 0.60), “Low5” (r = - 0.46), and “High-Mid3” (r = - 0.62). The lower the value in these three scores, the higher the perceptual ratings on hypernasality as shown in the negative sign in the correlation coefficients. The results for the sub-groups were summarized in Table 2. For both the scores of “High5” and “High-Mid3”, there were significant correlations ($p < 0.05$) with the DME-PF in all sub-groups, except for dysarthria. In the sub-group of cleft and maxillectomy, stronger correlations were noted on the scores of “High5” (r = - 0.74) and “High-Mid3” (r = - 0.67). However, for the score of “Low5”, significant correlations ($p < 0.05$) were only shown in two subgroups: all etiologies included (r = - 0.46), and cleft and maxillectomy (r = - 0.56). Low correlation (r = - 0.28) for “Low 5” was noted in the cleft palate sub-group. In the subgroup of dysarthria, no significant correlation was shown in each of the three scores.

Table 2

Pearson product-moment correlations between each individual score (“High5”, “Low5”, and “High-Mid3”) and the geometric means of Direct Magnitude Estimation: Practice-Feedback in the sub-groups of the experimental speakers

	Pearson r			
Score	All aetiologies included (n = 19)	Cleft palate (n = 11)	Cleft palate and Maxillectomy (n = 13)	Dysarthria (n = 6)
High5	- 0.60**	- 0.62*	- 0.74**	- 0.47
Low5	- 0.46*	- 0.28	- 0.56*	- 0.46
High-Mid3	- 0.62**	- 0.61*	- 0.67**	- 0.46

Note. *$p < 0.05$, **$p < 0.01$.

Relationship between the spectral measurements and perceptual rating with a range of aetiologies included

Stepwise multiple regression analysis was then conducted to determine which spectral measures on consonants in the present study and vowels in study by Lee (2004) best predicted the direct magnitude estimates of the hypernasality with a range of aetiologies included by the Practice-Feedback group. Lee (2004) found that there were statistically significant correlations (p < 0.05) between the DME-PF and each of the four one-third-octave bands in vowels: the bands centered at 1000 Hz for /i/ (r = 0.61), the 250 Hz (r = 0.46) and the 800 Hz (r = - 0.48) bands for /ɔ/, and the 2000 Hz band (r = 0.43) for /u/. These four bands were used as a predictive variables in stepwise multiple regression analysis. The result of stepwise multiple regression analysis showed that these bands had a moderate correlation as shown in the multiple correlation coefficient (R = 0.72) with the perceptual ratings.

While from the results in this study, they showed that there were statistically significant correlations (p < 0.05) between the DME-PF and each of the three spectral scores on consonants: “High-Mid3” (r = - 0.62), “High5” (r = - 0.60), and “Low5” (r = - 0.46). As the scores “High-Mid3” and “High5” were highly correlated, only the “High-Mid3” and Low5” were entered as the predictive variables. The three bands centered at 1000 Hz in vowel /i/ and 200 and 800 Hz in vowel /ɔ/ together with the spectral scores of “Low5” in consonants /s/ and /n/ and “High-Mid3” in consonant /s/ in this study were then entered as the predictive variables to predict the DME-PF through stepwise multiple regression analysis. The result of the stepwise multiple regression analysis was shown in Table 3. Variables “High-Mid3”, “Low5”, the bands centered at 1000Hz in vowel /i/ and at 250Hz in vowel /ɔ/ accounted for 64% of the variance in the DME-PF of hypernasality. There was one outlier in S12 while the other residuals were located within $\pm$ 2 standard deviations. A higher correlation (*R* = 0.80) was shown between the perceptual ratings and the predicted values derived from the one-third-octave bands in vowels together with the spectral scores on consonants compared with that of vowels (*R* = 0.72) in the study of Lee (2004). Therefore, the spectral scores on consonants further increase the correlation with the perceptual ones. Independent variables "High-Mid3" ($p < 0.01$) and the band centered at 1000 Hz in vowel /i/ ($p < 0.05$) made significant contribution in the multiple regression model, while "High-Mid3" have the relatively strongest contribution ($\beta = -0.94$) in the prediction of the perceptual ratings.

Table 3

Stepwise multiple regression analysis between Direct Magnitude Estimates: Practice-Feedback of hypernasality and the scores derived from the one-third-octave bands in consonant /s/ and / or /n/ ("High5", "Low5", and "High-Mid3") & the three one-third-octave bands in vowels (centered at 1000 Hz in vowel /i/ and 200 and 800 Hz in vowel /ɔ/)

Variable	*β* coefficient	*p* level
High-Mid3	- 0.94	0.00**
1000Hz for /i/	0.40	0.04*
250Hz for / ɔ /	0.34	0.06
Low5	0.41	0.16

Note. Multiple $R = 0.80$; R-squared = 0.64; Adjusted R-squared = 0.53; $df = 4,14$; $p = 0.0045$
*$p < 0.05$, **$p < 0.01$.

Discussion

Spectral characteristics of hypernasality in consonants

In this study, the spectral characteristics of consonants in hypernasal speech with a range of aetiologies (dysarthria, maxillectomy, and cleft palate) were investigated using procedures used by Kataoka et al. (2001), Lee et al. (2003), and Lee (2004). The results showed that hypernasal speakers had significantly higher amplitude in frequency bands below 500 Hz and significantly lower amplitude in frequency bands above 2000 Hz than speakers with normal resonance. The slope of spectrum of /s/ produced by speakers with hypernasality was flattened as shown in the significant smaller value in “High-Mid3” when compared with that of the control group. These findings were consistent with the hypothesis that the difference of mean amplitude between the oral /s/ and nasal /n/ consonants at high and low frequency bands would be smaller in the experimental group. The hypernasal /s/ was characterized by the flattened slope with increase in amplitude at low and mid frequencies, and decrease in amplitude at high frequencies.

For the consonant /n/, the normal group had significantly higher amplitude at the band 3201 to 4000 Hz only. No significant difference in the mean amplitude in the other 20 bands of both groups was revealed. The spectra were characterized by prominence peak and trough at low frequencies below 500 Hz and high frequencies above 2000 Hz respectively. Nasal consonants involve the velopharyngeal opening that the sound energy could enter the nasal cavity (Kent, & Read, 2002). Therefore, people with or without hypernasality could produce the nasal consonant /n/ with very similar spectra as shown in this study that involves the opening of the velar port.

Consonants /s/ and /n/ are of the same place of articulation, lingual-alveolar, but they are different in manner. For the nasal consonant /n/, the tongue completely occludes the oral cavity at the alveolar place and the velum is lowered. The oral consonant /s/ involves narrow constriction formed by tongue with the alveolar place and the closure of velum. The spectral structure of nasal consonants is formed by the resonance (pole) in the nasal and pharyngeal cavities and by the antiformants (zero) of the oral cavity (Fant, 1970). When a consonant /s/ is nasalized, sound energy also propagates into the nasal cavity with velar port opened. This sound from nose, called “nasal murmur”, is characterized acoustically by the strong amplitude at the region below 500 Hz and weaker amplitude above 500 Hz. Moreover, antiformants presents in the nasal consonant /n/ at around 1500 to 2000 Hz are associated with decrease in amplitude. Antiformants are capable to absorb acoustic and amplitude of the local bands would be decreased (Pickett, 1999). Hypernasal /s/ was characterized by decrease in amplitude at high frequencies (above 2000 Hz) and increase in energy at low (below 500 Hz) and mid (500 to 2000 Hz) frequencies. The spectral characteristics of nasalized /s/ were the consequences of nasalization. It was suggested that increase in amplitude at low and mid frequencies was due to the presence of nasal murmur and decrease in amplitude at high frequencies was due to the presence of antiformants (zero). Resonance and antiformants formed from the nasalization was shown acoustically in the spectra of /s/ produced by the hypernasal speakers.

Generalizability of the one-third-octave analysis on consonants

A range of aetiologies (cleft palate, maxillectomy, and dysarthria) were included in this study with the application of the one-third-octave analysis on the hypernasality in consonants. There were four sub-groups for the hypernasal speakers: 1) all aetiologies included; 2) cleft palate; 3) cleft palate and maxillectomy; and 4) dysarhtria. The results showed that the scores of “High5” and “High-Mid3” of in all the sub-groups had significant differences ($p < 0.01$) with that of the speakers with normal resonance. The generalizability of the spectral measures (“High5” and “High-Mid3”) of consonants was noted. Spectral characteristics of consonant /s/ were consistent across different causes of the hypernasal speech: increase in amplitude at mid frequencies and decrease in amplitude at high frequencies. However, for "Low5", significant differences ($p < 0.05$) were noted in the sub-groups of all aetiologies included, cleft palate and maxillectomy, and dysarthria, except for cleft palate, with the control group. This showed that the increase in amplitude at low frequencies was characterized only in the sub-groups of all aetiologies included, cleft palate and maxillectomy, and dysarthria. In the sub-group of cleft palate, significant differences on the mean amplitude were shown only at high and mid frequencies. It was suggested that the nasal murmur presented in the hypernasal /s/ produced by this sub-group was not prominent at low frequencies below 500 Hz. Rather that, only the lower amplitude peak above 500 Hz characterized in the nasal murmur was prominent in the sub-group of cleft palate.

When comparing with the study by Lee et al. (2003), greater differences on the mean spectral measures between the hypernasal and control group were observed for in the consonant /s/ than for vowels. This showed that hypernasality speech could be visualized and quantified by the 16 bands in the consonant /s/ and /n/. In analyzing /s/ and /n/ samples of each individual, nasalization on /s/ could be measured with reference to its counterpart /n/. However, there is no reference point for vowels in individuals. It was suggested that hypernasality could be more easily noted with these acoustic cues in the spectra of /s/ and /n/ than vowel /i/. In previous studies (Kataoka et al., 1996; Kataoka et al., 2001; Yoshida et al., 2000), one-third-octave analysis applied to the sustained vowel /i/. Vowel /i/ is considered to be the most ideal one to judge hypernasality as there is narrow tongue passage. Therefore, low degrees of nasal coupling would lead to noticeable nasal sound transmission (Fant, 1970). While in this study, spectral analysis applied to consonants in single words and connected speech samples. The results found in this study applied to the connected speech, and the effect of hypernasality on communication is thus explored.

Correlations of the spectral measurements with the perceptual ratings on hypernasality

Across the sub-groups of the experimental group, “High5” and “High-Mid3” showed significant correlations with all aetiologies included, cleft palate and maxillectomy, and cleft palate, except for dyarthria. “Low5” showed significant correlations in all aetiologies included, cleft palate and maxillectomy; and no significant correlation was noted in the subgroups of cleft palate and dysarthria. This showed that all the three spectral scores from /s/ and /n/ do not have significant correlations with the perceptual ratings in the sub-group of dysarthria. Poorer relationship between the spectral and perceptual measurements was noted in this sub-group.

By using the same method of adjustment on the amplitude of the one-third-octave bands, this study replicated the studies by Lee et al. (2003) and Lee (2004) with the use of same set of subjects with a range of aetiologies (cleft palate, mexillectomy, and dysarthria) included. Stepwise multiple regression analysis was conducted to compare the scores from consonants in the present study and the octave bands in vowels in previous studies (Lee et al., 2003; & Lee, 2004) with the perceptual scores given to these subjects in the study by Lee (2004). It specified that “High-Mid3”, “Low5”, bands centered at 1000 Hz for vowel /i/ and at 250 Hz for vowel /ɔ/ associated with the perceptual judgments (Direct Magnitude Estimation: Practice-Feedback). A high correlation between the predicted values from these variables and the perceptual ratings was shown (R = 0.80). Lee (2004) found that the three variables (1000 Hz for /i/, 800 and 250 Hz for /ɔ/) best predicted the perceptual ratings with a moderate correlation (R = 0.72) by the stepwise multiple regression analysis. This showed that the spectral scores “High-Mid3 “and “Low5” for consonants further increase the correlation. It was suggested that the acoustic measurements on the vowel /i/ and /ɔ/ and consonant /s/ and /n/ could help to predict the perceptual ratings on hypernasality as shown in the high correlation. The higher the amplitude in the two bands in vowels and the lower the two spectral scores in consonants would lead to the higher perceptual ratings on hypernasality.

Clinical implication

Recent studies on hypernasality with the use of one-third-octave analysis showed promising results. Yoshida et al. (2000) and Kataoka et al. (2001) showed that high correlations were noted between the perceptual ratings and the predicted values from amplitudes of certain one-third-octave bands with $R = 0.84$. In the current study, spectral measurements from the one-third-octave analysis in consonants and vowels were correlated with perceptual judgments; and this suggested that this instrumental means of analysis quantify hypernasality successfully. These measurements are considered to be clinically useful as strong correlation with the subjective measures is evident (Gerratt, Kreiman, Barroso, & Berke, 1993).

In previous studies (Kataoka et al., 1996; Kataoka et al., 2001; & Yoshida et al., 2000), the sustained vowel /i/ was examined. However, in the studies of Lee et al. (2003), Lee (2004) and the current study, vowels segmented from words and consonants segmented from words and oral sentences were examined. The speakers produced connected speech in the communication rather than sustained vowels. Therefore, measuring the nasality level in the connected speech would give more information on how the resonance disorders affecting communication (Lee, 2004). Connected speech materials were then used in these studies instead of sustained vowels. As certain one-third-octave bands in vowels and spectral measurements on consonants showed high correlation with the perceptual ratings, naturalistic connected speech with vowels /i/ and /ɔ/, and consonant /s/, and /n/ could then be reliable to quantify the hypernasality of individuals in clinical setting.

This study showed that the one-third-octave analysis has high intra-judge ($r = 0.97$) and inter-judge reliability ($r = 0.96$). These results are consistent with previous study by Lee

(2004) on the spectral analysis of vowels. High intra- and inter-judge reliability further showed that one-third-octave analysis was a reliable way to quantify hypernasality. For Practice-Feedback group in rating hypernasality, Pearson product-moment correlation showed a mean intra-judge reliability for rating the male speakers of $r = 0.55$ and for rating female speakers of $r = 0.64$; while Cronbach's alpha showed the inter-judge reliability for rating the male speakers of $\alpha = 0.87$ and for rating female speakers of $\alpha = 0.91$. Therefore, one-third-octave analysis showed much higher intra-judge reliability and slightly higher inter-judge reliability than the perceptual judgments. This showed that acoustic analysis is a more reliable way than perceptual judgments on evaluating hypernasal speech. With basic acoustic knowledge, raters could do the analysis by extracting the segment from the speech samples reliably. One-third-octave analysis is a non-invasive and reproducible that can be performed at real time (Kataoka et al., 2001). Speech samples of connected speech and single words could be easily collected at clinical settings with the microphones. Adjustment on the amplitude of the one-third-octave bands makes the comparison possible even though the speech samples produced are of natural loudness and pitch and variable intensity levels. This one-third-octave spectral analysis is, therefore, reliable and suitable to be used in clinical settings.

Limitations in using one-third-octave analysis in consonants

The use of one-third-octave analysis in consonants was limited to the articulation errors due to hypernasality. Therefore, for those speech samples /s/ from the subjects perceived as hypernasal only were analyzed. The substitutions of /s/ by /ts/ and /t/ were then all discarded in this study. For example, a subject, S19, showed stopping (/s/ → [t]) and deletion errors in all his speech samples; and he was discarded. Individuals with cleft palate and dysarthria are at high risk for disordered articulation, such as hypernasality, deletion and substitution (Duffy, 1995; McWilliams, Morris, & Shelton, 1990). Those who have consistent substitution errors for /s/ might not be eligible for this spectral analysis.

Further research

In this study, the subjects (12 had cleft palate, 6 had dysarthria, and 2 had maxillectomy) of the experimental group were previously described in Lee et al. (2003) and Lee (2004). In general, the spectral characteristics of hypernasal /s/ were found to be consistent across different aetiologies as shown in the scores of “High5” and “High-Mid3”. No significant difference on mean amplitude at low frequencies between the control and the sub-group of cleft palate in the score of “Low5”. The sub-group of dysarthria showed no significant correlation with all the spectral scores. As there was limited number of subjects in each of the sub-group, the results found in this study could not truly reflect the population. Clients who had maxillectomy after oral cancer, cleft palate and dysarthria are high risk groups on resonance problems. Therefore, it was suggested that further study could be on the spectral characteristics on the production of /s/ of each of the aetiologies with a larger number of subjects following the same procedures in the studies of Kataoka et al. (2001), Lee et al. (2003), Lee (2004) and the current study.

Second, the relationship between the spectral measurements on consonants by hypernasal speakers and other instrumental measurements, such as nasalance, could be measured. The relationship between the spectral analysis and other objective measurements could be revealed. This could further investigate the validity of the spectral measures on consonants with the one-third-octave analysis.

Third, the sensitivity and specificity of the spectral scores in each aetiology with a larger number of subjects could be calculated. The discriminate function analysis could then be performed to determine which spectral scores or combination of them best discriminate the hypernasal and normal speakers.

Acknowledgement

I would like to take this opportunity to thank my supervisors, Dr. Valter Ciocca and Dr. Tara Whitehill, for their enormous help and valuable comments on my dissertation project. I would also like to express my hearty thanks to Miss Alice Lee for her permission to use the speech samples and perceptual ratings. Special thanks should be given to my classmates, Miss Karen Tse, Miss Kobe Lam, and Miss Mahomed Ramona, for their support and participation.

References

Boersma, P., & Weenink, D. (2001). Praat (Version 4.0.1) (Amsterdam). Available at: *http://www.praat.org.*

Borden, G. J., Harris, K. S., & Raphael, L. J. (2003). *Speech science primer: Physiology, acoustics, and perception of speech* (4th ed.). New York: Lippincott Williams & Wilkins.

Bradford, L. J., Brooks, A. R., & Shelton, R. L. (1964). Clinical judgment of hypernasality in cleft palate children. *Cleft Palate Journal, 1*, 329-331.

Bzoch, K. R. (1989). Measurement and assessment of categorical aspects of cleft palate language, voice, and speech disorders. In Bzoch, K. R. (Ed.), *Communication disorders related to cleft lip and palate* (pp. 138-146). Boston: Little, Brown.

Counihan, D. T., & Cullinan, V. L. (1970). Reliability and dispersion of nasality ratings. *Cleft Palate Journal, 7*, 261-270.

Duffy, J. R. (1995). *Motor speech disorders: Substrates, differential diagnosis, and management.* St. Louis: Mosby.

Fant, G. (1970). *Acoustic theory of speech production.* Netherlands: Mouton, Hague.

Kataoka, R., Michi, K. I., Okabe, K., Miura, T., & Yoshida, H. (1996). Spectral properties and quantitative evaluation of hypernasality in vowels. *Cleft Palate-Craniofacial Journal, 33*, 43-50.

Kataoka, R., Warren, D. W., Zajac, D. J., Mayo, R., & Lutz, R. Z. (2001). The relationship between spectral characteristics and perceived hypernasality in children. *Journal of Acoustical Society of America, 109*, 2181-2189.

Kent, R. D., Liss, J., & Philips, B. J. (1989). Acoustic analysis of velopharyngeal dysfunction in speech. In K. R. Bzoch (Ed.), *Communicative disorders related to cleft lip and palate* (pp. 258-270). Boston: Little-Brown.

Kent R. D. & Kim, Y. J. (2003). Toward an acoustic typology of motor speech disorders. *Clinical Linguistics & Phonetics, 17*, 427-445.

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech* (2nd ed.). Australia: Thomson Learning.

Kurowski, K. M., & Blumstein, S. E. (1993). Acoustic properties for the perception of nasal consonants. In M. K. Huffman and R. A. Krakow (Eds.), *Phonetics and phonology: Nasals, nasalization, and the velum, volume 5* (pp. 147-167). San Diego, CA: Academic Press.

Gerrat, 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*, 14-25.

Lee, S. Y. A. (2004). *Perceptual and Instrumental Analysis of Hypernasality*, unplublished doctoral thesis. Hong Kong: The University of Hong Kong.

Lee, S. Y. A., Ciocca V., & Whitehill, T. L. (2003). Acoustic correlates of hypernasality. *Clinical Linguistic & Phonetics, 17*, 259-264.

Maeda, S. (1993). Acoustic of vowel nasalization and articulatory shifts in French nasal vowels. In M. K. Huffman and R. A. Krakow (Eds.), *Phonetics and phonology: Nasals, nasalization, and the velum, volume 5* (pp. 147-167). San Diego, CA: Academic Press.

Pickett, J. (1999). *Acoustics of speech communication.* Boston: Allyn & Bacon.

So, K. H. L., & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language, 22*, 473-495.

Whitehill, T. L., Lee, S. Y. A., & Chun C. J. (2002). Direct magnitude estimation and interval scaling of hypernasality. *Journal of Speech, Language, and Hearing Research, 45*, 80-88.

Yoshida, H., Furuya, Y., Shimodaira, K., Kanazawa, T., Kataoka, R., & Takahashi, K. (2000). Spectral characteristics of hypernasality in maxillectomy patients. *Journal of Oral Rehabilitation, 27*, 723-730.

Zarick, R. I., & Liss, J. M. (2000). A comparison of the equal-appearing interval scaling and direct magnitude estimation of nasal voice quality. *Journal of Speech, Language and Hearing Research, 43*, 979-988.

Appendix A

Nine oral sentences with the phonetic transcriptions in IPA system and English translation

1. 哩個係一個好舊嘅筆盒。

   /lei$_{55}$ kɔ$_{33}$ hɐi$_{22}$ jɐt$_{5}$ kɔ$_{33}$ hou$_{25}$ kɐu$_{22}$ ke$_{33}$ pɐt$_{5}$ hɐp$_{2}$/
   (This is a very old pencil case.)

2. 話畀佢知哩隻狗叫羅拔。

   /wa$_{22}$ pei$_{25}$ k^{h}œy$_{23}$ tsi$_{55}$ lei$_{55}$ tsek$_{3}$ kɐu$_{25}$ kiu$_{33}$ lɔ$_{55}$ pɐt$_{2}$/
   (Tell him/her hat this dog is called Robert.)

3. 哩對波鞋嘅設計好特別。

   /lei$_{55}$ tœy$_{33}$ pɔ$_{55}$ hai $_{21}$ ke$_{33}$ ts^{h}it$_{3}$ kɐi$_{33}$ hou$_{25}$ tɐk$_{2}$ pit$_{2}$/
   (The design of this pair of sports shoes is very special.)

4. 食嘢嗰時至扑開個椰子。

   /sik$_{2}$ je$_{23}$ kɔ$_{25}$ si$_{21}$ tsi$_{33}$ pɔk$_{5}$ hɔi$_{55}$ kɔ$_{33}$ je$_{21}$ tsi$_{25}$/
   (Break open the coconut when (you) eat.)

5. 四條灰色褲都係好薄嘅。

   /sei$_{33}$ t^{h}iu$_{21}$ fui$_{55}$ sik$_{5}$ fu$_{33}$ tou$_{55}$ hɐi$_{22}$ hou$_{25}$ pɔk$_{2}$ ke$_{33}$/
   (The four gray trousers are very thin.)

6. 會撥四億畀佢去起大學。

   /wui$_{23}$ put$_{2}$ sei$_{33}$ jik$_{5}$ pei$_{25}$ k^{h}œy$_{23}$ hœy$_{23}$ hei$_{25}$ tai$_{22}$ hɔk$_{2}$/
   (Will allocate him/her forty billion (dollars) to build a university.)

7. 哩嗖都有蝴蝶飛嚟飛去。

   /lei$_{55}$ tou$_{22}$ tou$_{55}$ jɐu$_{23}$ wu$_{21}$ tip$_{2}$ fei$_{55}$ lɐi$_{21}$ fei$_{55}$ hœy$_{33}$/
   (There are butterflies flying here.)

8. 下晝嚟度吓露台有幾大。

   /ha$_{22}$ tsɐu$_{33}$ lɐi$_{21}$ tɔk$_{2}$ ha$_{23}$ lou$_{22}$ t^{h}ɔi $_{21}$ jau$_{23}$ kei$_{25}$ tai$_{22}$/
   (Come this afternoon to measure how big the balcony is.)

9. 佢要吃高腳至攞倒啲書。

   /k^{h}œy$_{23}$ jiu$_{33}$ kɐk$_{2}$ kou$_{55}$ kœk$_{3}$ tsi$_{33}$ lɔ$_{25}$ tou$_{25}$ ti$_{55}$ sy$_{55}$/
   (He/She has to stand on his/her toes to get the books.)

Appendix B

Articulation of /s/ in the oral sentences in the experimental group

		Sentence 4		Sentence 5		Sentence 6	Sentence 7
Subject	Etiologies	/sik$_2$/	/si$_{21}$/	/sei$_{33}$/	/sik$_5$/	/sei$_{33}$/	/sy$_{55}$/
S1	dysarthria	1	1	missed	missed	1	1
S2	dysarthria	deleted	2	2*	2*	2*	2
S3	dysarthria	3*	3	3	1	3*	3*
S4	dysarthria	1	1	1	1	1	1
S5	dysarthria	1	1	1	1	1	1
S6	dysarthria	1	1	1	1	1	1
S7	maxillectomy	1	1	1	1	1	1
S8	maxillectomy	1	1	1*	1	1	1
S9	cleft palate	1	1	1	1	1	1
S10	cleft palate	1	1	1	1	1	1

1-perceived as /s/, 2-perceived as /n/, 3-perceived as /ts/, 4-perceived as /t/, deleted-deletion of /s/, missed-the oral sentence was not recorded

* durations of its extracted segment /s/ is less than 50ms

continued

Appendix B, continued

		Sentence 4		Sentence 5		Sentence 6	Sentence 7
Subject	Etiologies	/sik$_2$/	/si$_{21}$/	/sei$_{33}$/	/sik$_5$/	/sei$_{33}$/	/sy$_{55}$/
S11	cleft palate	1	1	1	1	1	1
S12	cleft palate	deleted	2	deleted	1	4*	1
S13	cleft palate	1	1	1	1	1	1
S14	cleft palate	1	1	1	1	4	1
S15	cleft palate	1	1	1	1	1	1
S16	cleft palate	1*	1	1	1	1	1
S17	cleft palate	1	1	1	1	1	1
S18	cleft palate	1	1	1	1	1	1
S19	cleft palate	deleted	4*	deleted	4*	4*	4*
S20	cleft palate	1	1	1	1	1	1

1-perceived as /s/, 2-perceived as /n/, 3-perceived as /ts/, 4-perceived as /t/, deleted-deletion of /s/, missed-the oral sentence was not recorded

*duration of its extracted segment /s/ is less than 50 ms