INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

Hypernasality Severity Analysis in Cleft Lip and Palate Speech Using Vowel Space Area

Nikitha.K[1], *Sishir Kalita*[2], *CM Vikram*[2], *M.Pushpavathi*[1], *S.R.M. Prasanna*[2]

[1]Department of Speech Language Pathology
All India Institute of Speech and Hearing, Mysore-570 006, India
[2]Department of Electronics and Electrical Engineering
Indian Institute of Technology Guwahati, Guwahati-781039, India
nikithakgowda93@gmail.com, (sishir, cmvikram)@iitg.ernet.in, pushpa19@yahoo.co.in, prasanna@iitg.ernet.in

Abstract

Vowel space area (VSA) refers to a two-dimensional area, which is bounded by lines joining $F_1$and F_2 coordinates of vowels. In the speech of individuals with cleft lip and palate (CLP), the effect of hypernasality introduces the pole-zero pairs in the speech spectrum, which will shift the formants of a target sound. As a result, vowel space in hypernasal speech gets affected. In this work, analysis of vowel space area in normal, mild and moderate-severe hypernasality groups is analyzed and compared across the three groups. Also, the effect of hypernasality severity ratings across different phonetic contexts i.e, /*p*/, /*t*/, and /*k*/ is studied. The results revealed that VSA is reduced in CLP children, compared to control participants, across sustained vowels and different phonetic contexts. Compared to normal, the reduction in the vowel space is more for the moderate-severe hypernasality group than that of mild. The CLP group exhibited a trend of having larger VSA for /*p*/, followed by /*t*/, and lastly by /*k*/. The statistical analysis revealed overall significant difference among the three groups ($p < 0.05$).

Index Terms: Cleft Lip and Palate (CLP) speech, vowel space area, formants, hypernasality severity.

1. Introduction

The children with cleft lip and palate exhibit typical speech deviances due to inadequate functioning of velopharyngeal port, structural abnormalities and due to mislearning [1]. Most frequently their speech will always be associated with hypernasality which results from a velopharyngeal dysfunction (VPD). It is a condition where individuals fail to achieve velopharyngeal closure adequately during speech production leading to leakage of air through the nose. Additionally, articulation errors including compensatory productions along with disturbed voice quality affect the speech intelligibility [2], [3], [4]. Among these abnormal speech characteristics, hypernasality is devastating and is a major concern in the assessment and management field. Evaluation of hypernasality and its severity rating has proved to be a challenge for SLPs in the clinical and research field. The issues like the accuracy, reliability of perceptual judgments of hypernasality and the disagreement between subjective and objective evaluation have always interfered with accurate clinical evaluation. There is also the question of whether clinical evaluation reflects the impact of hypernasality in a speakers everyday life. Hence to analyze hypernasality and its effect on speech intelligibility, a number of noninvasive techniques like perceptual, physiological and acoustic methods are used. Among other techniques, spectral analysis is considered as simple, more cost effective and quantitative method to analyze hypernasality [5]. In the literature, acoustic cues such as weakening of formants, decrease in the strength and enhanced bandwidth of F_1 and F_2, rise in the amplitude between F_1 and F_2, introduction of pole/zero pairs in the vicinity of F1 and shifts in the formant frequencies are reported for the analysis of hypernasal speech [5]. Hence, these spectral modifications in the hypernasal speech will have an impact on the articulatory dynamics while producing vowels resulting in vowel centralization and in turn affecting speech intelligibility [6].

The vowel space area (VSA) analysis is one of the efficient measure to study the vowel centralization effect. This metric has helped speech-language pathologists to study the spectral changes, the articulatory precision and found useful in predicting speech intelligibility in various speech disorders like Down syndrome [7], cerebral palsy [8], hearing impairment [9], dysarthria [10]. A recent study on VSA on CLP speakers reported the reduced vowel space in their speech [6]. However, the effect of severity of hypernasality on vowel space was not explicitly studied in the literature. Apart from pathological speech, studies have also been focused on investigating vowel space in the normal population, across age groups, gender, languages and different phonetic contexts. The studies infer that with the increase in age, vowel space reduces [11], the presence of larger vowel space in females as compared to males [12] and also cross-linguistic variations in vowel space. In a study which investigated the vowel space across phonetic contexts, the authors concluded that larger VSA was found in the context of velar /k/, followed by bilabial /p/ and retroflex /t/ at last [13]. Motivated by the importance of vowel space in the analysis of pathological speech, intelligibility, and variation of vowel space across contexts, age, and gender, the current study focused on studying the effect of hypernasality in CLP children. Also in this work, the analysis of vowel space across different severity ratings of hypernasality and across different contexts is carried out. The comparison of vowel space across three severity levels of hypernasality is carried out two different conditions:

1. The vowel space in sustained vowels /a/, /i/, /u/
2. The vowel space across different phonetic contexts (/papa/, /pipi/, /pupu/, /tata/, /titi/, /tutu/, /kaka/, /kiki/, /kuku/) and to look for the trend across the three contexts in CLPs and make comparison with normal children.

Further, the paper is organized as follows: section II describes the data collection and perceptual evaluation. Section III describes the extraction of formants and computation of vowel

http://dx.doi.org/10.21437/Interspeech.2017-1245

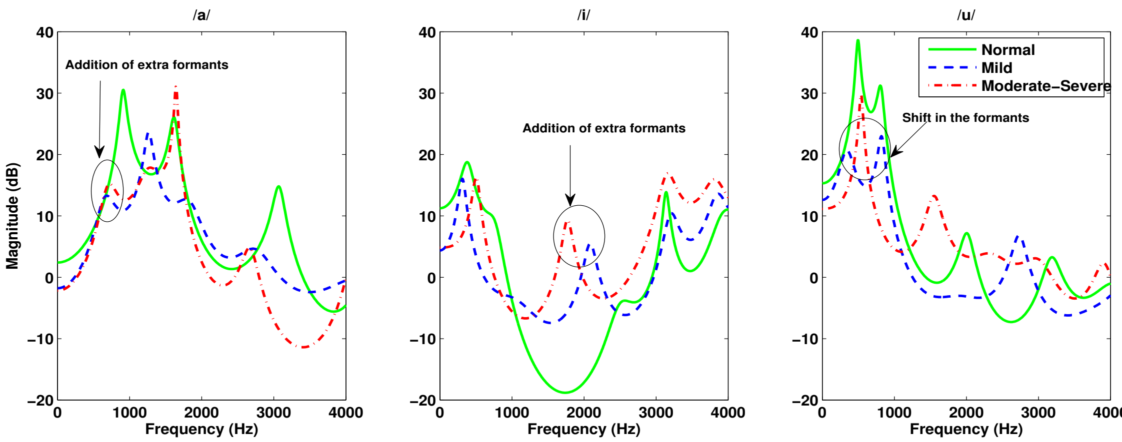

Figure 1: *LP magnitude spectra of phonation /a/, /i/, and /u/ for the normal, mild, and moderate-severe hypernsality cases. Figure shows that introduction of extra nasal formants in the visinity of F_1, increase in the valley amplitude between F_1 and F_2, and shift of formants in hypernasal speech.*

space area. Section IV explains the results and analysis of the current study. Finally, section V provides the conclusion and mentions the future scope.

2. Database and perceptual evaluation

Eighteen children (11 males and 7 females) with cleft lip and palate in the age range of 7-12 years are considered for the study. All the participants had repaired cleft of lip and palate and are native speakers of Kannada. Kannada is a Dravidian language, spoken in Karnataka, a state located in Southern part of India. The participants are chosen based on the inclusionary criteria of not having associated syndromic conditions and other sensory and neurological impairments. The language abilities of all the individuals with cleft palate are age adequate. Fifteen children with normal speech and language characteristics who are age and gender matched served as controls for the study.

The speech stimuli consists of CVCV non-words which contain the consonants /k/, /t/, /p/ with corner vowels (/a/, /i/, and /u/) combination (/kaka/, /tata/, /papa/, /pipi/, /titi/, /kiki/, /pupu/, /tutu/, /kuku/). The speech recording has been carried out in All India Institute of Speech and Hearing Mysore, India [14]. All the participants seated comfortably in a sound treated room and the data is recorded using a sound level meter at a distance of 15 cm from each child at a sampling rate of $48,000$ Hz with 16 bits resolution. The three repetitions of CVCV words and a sustained phonation of /a/, /i/, /u/ were taken from all the participants for the purpose of calculating the vowel space. A conversational sample of five to ten minutes and repetition of five oral and oronasal sentences were recorded for the purpose of perceptual analysis to rate the hypernasality level.

To rate the severity of hypernasality and based on which, to group the participants into the mild and moderate-severe category, the perceptual analysis is carried out. The three experienced speech-language pathologists used the conversational sample, oral and oronasal sentences uttered by children with CLP for hypernasality judgment based on a 4-point rating scale given by Henningson, 2008 [15]. Among 18 children with CLP, 10 of them are categorized into a mild hypernasality and other 8 participants are grouped into the moderate-severe category.

3. Vowel space area computation

All the data recorded are transferred to a personal computer. PRAAT software [16] is used to extract the formants of three corner vowels /a/, /i/, /u/ from sustained phonation and from three phonetic contexts to calculate the VSA. By visualizing the vowel region using waveform and spectrogram, the F_1 and F_2 of the target corner vowel are measured at the midpoint of the vowel. Here, for CVCV unit the first vowel is considered for the analysis. The formant frequencies from all the participants are averaged and a single average value representing each group of normal, mild and moderate-severe hypernasality category was obtained for further analysis. A sustained part in the isolated vowel and the first vowel in each word are considered. The samples are analyzed using Praat and F1 and F2 formants are extracted for the analysis. Using the extracted F1 and F2 values for each category, the vowel triangles are computed using MATLAB.

PRAAT calculates the formants by linear prediction (LP) analysis of speech signal, where LP coefficients are computed by Burg's method. The LP magnitude spectra plotted for the sustained phonations of /a/, /i/, and /u/ of the normal, mild, and moderate-severe hypernasal speakers are as shown in Fig. 1. As shown in the Fig. 1, there is an addition of extra formants in the vicinity of the first formant in /a/. There is a shift in the amplitude between the amplitudes of first and second formants /i/, due to the addition of extra poles. Also, a shift in the formant location for the hypernasal /u/ is observed. The LP analysis shows that there is a shift in the formant locations in the hypernasal speech, compared to that of normals.

4. Results and discussion

The aim of the present study is to study and compare VSA across two severity levels of hypernasality in CLP individuals within two different conditions. The vowel triangles and the mean VSA values were obtained for making the comparison between the normal group, CLP group having mild and moderate-severe levels of hypernasality.

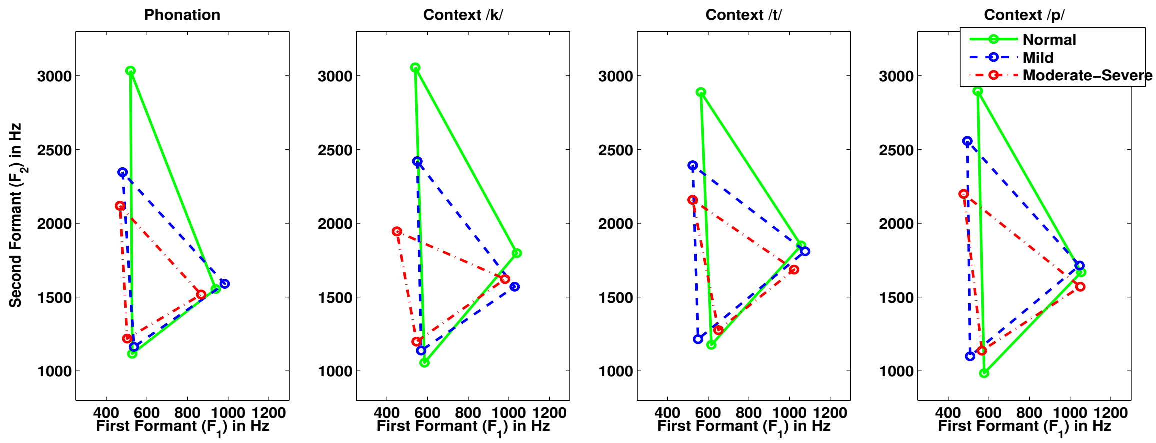

Figure 2: *Vowel triangles of normal, mild and moderate-severe group in phonation and in the contexts of /k/, /t/ and /p/. Figure shows that there is reduction in the vowel space in the gradation from normal, mild to moderate-severe. Also, among different contexts, vowel space is severely affected for the context /k/, and least for that of /p/*

4.1. Comparison of vowel triangles between normal, mild and moderate-severe groups in sustained phonation and across three phonetic contexts of /k/, /t/ and /p/.

The descriptive statistical analysis was performed to obtain the mean values of VSA across each group in two conditions (Phonation and phonetic contexts). Table 1 shows mean (in KHz) of VSA of all the three groups. The obtained mean values of VSA are used to plot vowel triangles as shown in Fig. 2. Figure 3 shows the distribution of vowel space area (VSA) for normal, mild, and moderate-severe hypernasality levels across phonation and contexts (/k/, /t/, and /p/). For all the contexts, the median values are significantly different for each category of hypernasality. However, it can be observed from the distribution for phonation context that the discrimination between hypernasal and normal is highest, whereas, for the /p/ context this discrimination is very less. The discrimination within the CLP categories is very less in phonation context. However, in the /t/, and /k/ contexts this discrimination is more. A further statistical analysis is presented in the next subsection to analyze the statistical significance among them.

It is observed that the larger VSA is found in the normal group, followed by the mild and moderate-severe group in both the conditions of sustained phonation and phonetic contexts. The moderate-severe group shows significantly reduced VSA when compared to the other groups and vowel space was comparatively better in mild hypernasality group. This significant reduction in VSA as a function of severity of hypernasality is attributed to the changes in articulatory dynamics (lack of precise and finer movements of the articulators during the production of vowels) as a result of hypernasality leading to vowel centralization effect. As the severity of hypernasality increases, it leads to unintelligible speech and considerable reduction of VSA is observed. In the phonetic context condition, the typically developing children exhibited a trend of larger vowel space in /k/, followed by /p/ and lastly by /t/. The smaller vowel space among all the three was found in the context of retroflex /t/. It can be explained that for velar stop /k/, the coarticulatory effects are found to be higher and since the point of constriction is moved backward in the oral cavity, there will be increased F1 and reduced F2 for vowels /a/ and /i/. Bilabial /p/ as an active articulator has the highest freedom of movement enabling the optimal

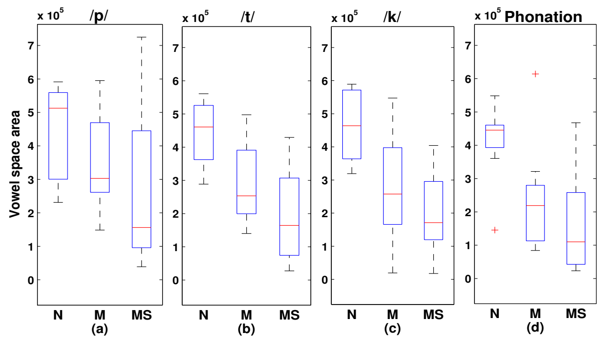

Figure 3: *Distribution of vowel space area (VSA) for normal, mild, and moderate-severe hypernsality levels across phonation, contexts /k/, /t/, and /p/.*

Table 1: *Vowel space area (VSA) for normal, mild, and moderate-severe hypernsality levels across phonation, contexts /k/, /t/, and /p/*

	VSA x10^3 Hz^2		
	Normal	Mild	Moderate-Severe
Phonation	396.18	276.48	169.19
/k/	471.35	299.47	183.22
/t/	395.60	318.45	190.58
/p/	467.55	397.28	276.58

tongue movements required for the production of a target vowel. The production of retroflex /t/ involves complex tongue dynamics which places high demands on tongue movements and compromises the articulatory precision. These results are in consonance with the previous study conducted in [13].

In CLP group, in all the contexts of consonants, results revealed reduced VSA when compared to typically developing children and found in the order of $VSA/p/ > VSA/t/ > VSA/k/$. VSA is highest in the context of bilabial /p/ and least in the context of velar /k/. This reduction in VSA in all the three phonetic contexts in CLP in comparison with the normal group is most likely because of the presence of hypernasality and compensatory articulation errors in their speech. The compensatory articulation errors are seen as a learned pattern to compensate for physiological constraints and are likely to persist even after the surgical correction of the structural anomaly in CLP children. The most common error pattern observed in repaired cleft palate speakers is the posterior placement of oral targets (Glottal stops, pharyngeal and laryngeal stops) [17]. These errors are thought to occur due to an attempt to achieve a valve at a point posterior to the oral cavity in an effort to build up pressure to produce plosives and even to prevent nasal air escape [17]. These compensatory errors cause a shift in the place of articulation where the normal tongue posture may not be established resulting in deviant tongue movements. The altered tongue movements can also occur due to the presence of oronasal fistulae, dental anomalies, and inadequate velopharyngeal closure. As a consequence of the deviant articulatory dynamics involved in the production of consonants, the vowels associated with it get affected due to the coarticulatory effect leading to reduced vowel space area.

Table 2: *MANOVA Test of Between Subjects Effects*

	Phonation	/k/	/t/	/p/
F(2,25)	8.52	11.56	3.14	6.93
Significance (p-value)	p <0.05	p <0.05	p <0.05	p >0.05

4.2. Statistical analysis

To check the normality, Shapiro test of normality was performed. The normal and moderate-severe groups showed normal distribution, whereas two outliers were found in mild hypernasality group. To achieve normality, the outliers from mild hypernasality group are removed from the data. As the results showed the normal distribution in all groups ($p < 0.05$), the parametric tests were performed to carry out different comparisons across three groups. MANOVA test is performed to find the differences in mean VSA values between the three groups. It is revealed that overall there is a significant difference in VSA between the three groups [$F(8, 44) = 3.419, p < 0.05$]. Subsequent MANOVA is done to compare the groups (Normal, mild and moderate-severe hypernasality) within phonation and phonetic contexts. The results are shown in the Table 2. From the table it can be observed that, there was a highly significant difference between the groups for sustained phonation [$F(2, 25) = 8.52, p < 0.05$], context /k/ [$F(2, 25) = 11.56, p < 0.05$] and context /t/ [$F(2, 25) = 6.93, p < 0.05$]. There is no significant difference found across the groups for context /p/ at $p < 0.05$ level of significance. The post hoc comparison using Tukey test revealed that normal group showed significant difference with other two groups in phonation and context /k/ at $p < 0.05$ level of significance. For context /t/, normal and moderate-severe groups were significantly different at $p < 0.05$ level of significance. The context /p/ did not show significant difference across the groups at $p > 0.05$ level of significance.

5. Conclusion

In this work, VSA analysis and comparison is carried out for different degrees of hypernasality. Statistically, VSA acts as an overall significant feature to discriminate the different hypernasality groups, however, within CLP, mild and moderate-severe groups are not discriminable. Vowel space area obtained for the context /k/ for the moderate-severe group is highly reduced compared to mild and normal groups. The reduction in the vowel space area is found to be least for context /p/, whereas /t/ showed an intermediate result. Further, the study can be extended to develop an automatic tool for the hypernasality level classification using formant estimation techniques. Also, the vowel space area can be used as an acoustic evidence, in addition to existing spectral based features for the better discrimination of hypernasality grading.

6. Acknowledgement

This work is part of the ongoing project titled "NASOSPEECH: Development of Diagnostic system for Severity Assessment of the Disordered Speech" funded by the Department of Biotechnology (DBT), Govt. of India and project titled "ARTICULATE +: A system for automated assessment and rehabilitation of persons with articulation diorders" funded by the Ministry of Human Resource Development (MHRD), Department of Higher Education (DHE), Shastri Bhawan, New Delhi.

7. References

[1] D. Sell, A. Harding, and P. Grunwell, "A screening assessment of cleft palate speech (great ormond street speech assessment)," *International Journal of Language & Communication Disorders*, vol. 29, no. 1, pp. 1–15, 1994.

[2] S. J. Peterson-Falzone, M. A. Hardin-Jones, and M. P. Karnell, *Cleft palate speech*. Mosby St. Louis, 2001.

[3] A. Kummer, *Cleft palate & craniofacial anomalies: Effects on speech and resonance*. Nelson Education, 2013.

[4] K. Bzoch, "Introduction to the study of communicative disorders in cleft palate and related craniofacial anomalies," *Communicative disorders related to cleft lip and palate. 5th ed. Austin: pro-ed*, pp. 3–66, 2004.

[5] K. Bettens, F. L. Wuyts, and K. M. Van Lierde, "Instrumental assessment of velopharyngeal function and resonance: A review," *Journal of communication disorders*, vol. 52, pp. 170–183, 2014.

[6] S. Haque, M. H. Ali, and A. F. Haque, "Variability of acoustic features of hypernasality and its assessment," *International Journal of Advanced Computer Science & Applications*, vol. 1, no. 7, pp. 195–201, 2016.

[7] K. Bunton and M. Leddy, "An evaluation of articulatory working space area in vowel production of adults with down syndrome," *Clinical linguistics & phonetics*, vol. 25, no. 4, pp. 321–334, 2011.

[8] H.-M. Liu, F.-M. Tsao, and P. K. Kuhl, "The effect of reduced vowel working space on speech intelligibility in mandarin-speaking young adults with cerebral palsy," *The Journal of the Acoustical Society of America*, vol. 117, no. 6, pp. 3879–3889, 2005.

[9] E. A. Wieland, E. B. Burnham, M. Kondaurova, T. R. Bergeson, and L. C. Dilley, "Vowel space characteristics of speech directed to children with and without hearing loss," *Journal of Speech, Language, and Hearing Research*, vol. 58, no. 2, pp. 254–267, 2015.

[10] H. Kim, M. Hasegawa-Johnson, and A. Perlman, "Vowel contrast and speech intelligibility in dysarthria," *Folia Phoniatrica et Logopaedica*, vol. 63, no. 4, pp. 187–194, 2010.

[11] H. K. Vorperian and R. D. Kent, "Vowel acoustic space development in children: A synthesis of acoustic and anatomic data," *Journal of Speech, Language, and Hearing Research*, vol. 50, no. 6, pp. 1510–1545, 2007.

[12] Y. Krishna and B. Rajashekhar, "Vowel space areas across age, gender and dialects in telugu," *LANGUAGE IN INDIA*, vol. 12, no. 1, pp. 357–369, 2012.

[13] Jyotsna, "Comparison of vowel space area across age groups in malayalam," Ph.D. dissertation, University of Mysore, Mysore, 2015.

[14] "All india institute of speech and hearing, mysore." [Online]. Available: https://www.http://aiishmysore.in/

[15] G. Henningsson, D. P. Kuehn, D. Sell, T. Sweeney, J. E. Trost-Cardamone, and T. L. Whitehill, "Universal parameters for reporting speech outcomes in individuals with cleft palate," *The Cleft Palate-Craniofacial Journal*, vol. 45, no. 1, pp. 1–17, 2008.

[16] P. P. G. Boersma *et al.*, "Praat, a system for doing phonetics by computer," *Glot international*, vol. 5, 2002.

[17] T. L. Whitehill, A. L. Francis, and C. K. Ching, "Perception of place of articulation by children with cleft palate and posterior placement," *Journal of speech, language, and hearing research*, vol. 46, no. 2, pp. 451–461, 2003.