A SPECTRAL ANALYSIS OF BREATHINESS OF INTERVOCALIC VOICED ASPIRATED PLOSIVES IN DELHI HINDI

Kalika Bali
Dept. of Literature and Language, The University of the South Pacific, Fiji

ABSTRACT

Breathiness is used for phonological contrast in Hindi giving rise to two stop series, viz., Voiced Plosives (VP) and Voiced Aspirated Plosives (VAP). In Delhi Hindi, intervocalic VAP may be produced without a period of voiced aspiration after the release of the stop. A spectral analysis of the vowel following the three categories of stops, namely, VAP, VP and the non-aspirated tokens of VAP (NAVAP), indicates that they differ along several parameters like relative amplitude of the first harmonic, first formant bandwidth, spectral tilt and noise at higher frequencies. The spectral analysis also reveals a difference between male and female speakers. The results indicate that females are overall more breathy suggesting that female and male speakers employ different glottal mechanisms for the production of all three categories of stops.

1. INTRODUCTION

The difference in the voice quality of Voiced Aspirated Plosives (VAP) and Voiced Plosives (VP) in Hindi is reflected in their glottal configurations. The two different types of phonation, modal and breathy, associated with the two types of plosives involve two different states of glottis. The glottal configuration for the production of VAP is very different from that of VP in terms of timing of the opening-closing gesture of the glottis and the degree of opening of the glottis [3, 4]. VP are produced with an approximated glottis through out the closure duration. For a VAP, on the other hand, the glottal opening starts in the middle of the articulatory closure duration and peaks in the middle of the noise duration following the release of the articulatory closure. The glottis closes considerably after the beginning of the following vowel [4]. Thus, it would seem reasonable to assume that the vowels following a VAP would have a larger airflow through glottis than the corresponding vowel following a VP. In other words, the vowel following VAP would be more breathy than the one following VP, hence, it would be expected to show spectral characteristics associated with a breathy vowel.

The Hindi speakers from Delhi present an interesting case in that many of them do not seem to be in the conventional sense. Their speech seems to contain plain VP where a standard Hindi speaker would produce VAP. This is markedly so in the case of intervocalic VAP. Thus, for example, [əbʱi] “now” in std. Hindi would become [əbi] in Delhi Hindi. Although some of the speakers from Delhi do indeed produce non-aspirated tokens of Voiced Aspirated Plosives (NAVAP), these are not phonetically identical to the regular VP [1]. The spectral characteristics of the vowel following NAVAP could assist in evaluating whether an absence of a voiced aspirated period following the stop release suggests a glottal configuration similar to that of a VP.

This paper focuses on the spectral differences between the vowels following intervocalic VAP, VP and NAVAP with the aim of understanding the glottal mechanism involved in the production of these stops. Previous studies [6,7,8,10] on the difference in glottal characteristics of male and female speakers have consistently found females to be more breathy than males. These studies, however, are restricted to languages where breathy voice is not a phonetic exponent of a phonological category. This paper will also consider the spectral differences in male and female speech in view of the contrastive system of stops in Hindi. If there is a physiological basis for breathiness in females then how does it affect the contrast between VAP and VP in female speech ?

2. SPECTRAL MEASUREMENTS FOR GLOTTAL CHARACTERISTICS

The difference in glottal configurations during modal and breathy phonation is reflected by differences in the low and high frequency regions of the speech spectrum. A change in the open quotient is reflected in the lower frequencies, primarily in the amplitude of the first harmonic [2,5,6,7,8,9]. The manner of glottal closure, simultaneous or non-simultaneous, and the rate of closure is reflected in the downward slope of the spectral tilt from lower to higher frequencies [6,7,8]. Any energy loss in the vocal tract affects the first formant bandwidth [6,8]. The energy losses affecting the first formant can be due to resistance offered by the walls of the vocal tract or due to loss at source or glottis [6,8]. Further, an opening at the glottis could also result in noise generation at the glottis and this is reflected as noise or aperiodicity in the spectrum at high frequencies [5,6,7,8,9]. These spectral characteristics could thus, be used as possible measures for breathiness.

2.1 Amplitude of the First Harmonic

Any change in the open quotient affects the spectrum mainly at the lower frequencies. This property of the speech spectrum has been employed for characterising breathiness in languages which use breathy voice to contrast phonologically [2,5,9] as well as gender differences [6,7,8]

Fischer-Jϕrgenson [5] in her study of phonologically breathy vowels in Gujrati found an increase in the level of the fundamental component or H1 by around 3 dB for breathy vowels. Other studies [2,9] have also found the relative amplitude of the first harmonic to be greater for breathy than for modal vowels. Klatt and Klatt [8] found that an increase in the difference in the amplitude of the first and second harmonic [H1-H2] corresponded to an increased breathiness. Hanson[6] too found that [H1-H2] correlated well with the open quotient measured by airflow and fiberscopic methods.

2.2 Spectral Tilt

The spectral tilt or the decrease in the amplitude of harmonics at higher frequencies of the spectrum, is a direct consequence of the nature of closure of the glottis. Closure of vocal folds can be brought about in two ways. In the first case, the vocal folds come together simultaneously causing an abrupt cut-off of airflow through the glottis. In the second case, the vocal folds close non-simultaneously with the closure beginning at the anterior end and gradually proceeding along the length of the folds towards the posterior. This results in a more gradual cessation of airflow and causes an additional decrease in the amplitude of the harmonics in the higher frequencies region resulting in an increase in the spectral tilt [6,8]. This change in spectral tilt is reflected by relative decrease or increase of the third formant peak or A3. Thus, H1-A3 is taken as a reasonable measure of spectral tilt. [6,8]

2.3 First Formant Bandwidth

An incomplete glottal closure causes an increase in the acoustic coupling which increases losses at the glottis which primarily affect the first formant bandwidth.[8]. Thus, the loss of energy at the glottis in case of breathy phonation when there is a glottal chink during the closed phase results in a greater first formant bandwidth. The difference in the amplitude of the first harmonic and that of the first formant peak has been used as an indicator of first formant bandwidth [6,8] . An increase in the first formant bandwidth causes a reduction in the peak amplitude of the first formant, resulting in a less prominent A1, the peak corresponding to first formant.

3. METHOD

The present study is based on the data collected from 21 informants, 9 males and 11 females. All informants were from Delhi and claimed Hindi as their first language. All of them had varieties in which they could produce VAP, VP and NAVAP. Recordings were made of speakers reading out passages with words containing VAP and VP intervocalically. The data was transferred to a computer and analysed with the help of the signal processing software xwaves. A total of 164 tokens were analyse

All spectral measures were made at the beginning and the middle of the vowel for all three categories of stops as determined with the help of sound pressure waveform and spectrogram. The vowel beginning measurements (Vbeg) were made at the point where the first formant becomes visible after plosion. The mid-point was taken as the middle of the steady state of the vowel (Vmid). No spectral measures were made at the end of the vowel as a preliminary examination indicated the end of the vowels tend to be breathy across all stop contexts for all speakers. This is not unexpected as these vowels were mostly at utterance end and there is an increase in airflow at the end of an utterance as glottis prepares for natural breathing [4]. The low first formant frequency of close vowels makes it difficult to determine the location of first formant peak, thus, in this study only tokens containing open vowel [a] were included.

The following measures were used for analysis :

H1-H2 : The difference between the amplitude of the first harmonic, H1, and the second harmonic, H2 in dB was used as an indicator for open quotient.

H1-A1 : The difference between the amplitude of the first harmonic, H1, and the peak corresponding to the first formant, A1, in dB was taken as an indicator for first formant bandwidth.

H1-A3 : The difference between the amplitude of the first harmonic, H1, and the peak corresponding to the third formant, A3, in dB was taken as a measure for spectral tilt.

Measurements for male and female speakers were analysed separately.

4. RESULTS

4.1 H1-H2

The results obtained for the spectral measurements (h1-H2) for female and male speakers are shown in Table 1 (a) and (b). Though there is an overlap between the three categories of stops, overall the (H1-H2) values are of the order VAP>NAVAP>VP for both female and male speakers. The overall difference between the three categories of stops is significant at $p < 0.0001$. This supports the analysis that overall the open quotient is largest for VAP and smallest for VP with NAVAP lying in the middle.

When we take Vbeg and Vmid into consideration, we find that (H1-H2) value is of the order Vbeg>Vmid for VAP but for NAVAP and VP, the order is Vbeg <Vmid. Thus, it is reasonable to assume that the open quotient decreases as we go from the beginning towards the middle of the vowel for VAP. For VP and NAVAP, the reverse is true, however, the open quotient is much greater for NAVAP in both cases than it is for VP.

		Mean	s.d.	Max.	Min.
VAP	Total	11.21	8.30	29.0	-11.8
	Vbeg	14.63	6.59	29.0	0.9
	Vmid	7.79	8.50	28.2	-11.8
NAVAP	Total	8.87	6.39	23.7	-5.4
	Vbeg	8.14	6.89	21.6	-5.4
	Vmid	9.60	5.85	23.7	-0.8
VP	Total	3.80	6.26	14.5	-9.3
	Vbeg	3.06	6.80	14.5	-9.3
	Vmid	4.60	5.65	14.2	-6.0

Table 1 (a) : Mean (H1-H2) values at Vbeg and Vmid for female speakers.

		Mean	s.d.	Max.	Min.
VAP	Total	5.92	5.58	16.8	-9.8
	Vbeg	6.70	5.26	16.8	-7.6
	Vmid	5.14	5.87	13.3	-9.8
NAVAP	Total	2.44	4.96	12.6	-11.9
	Vbeg	2.08	4.84	8.4	-11.9
	Vmid	2.80	5.13	12.6	-9.8
VP	Total	-2.28	4.69	7.3	-11.6
	Vbeg	-3.32	3.65	4.5	-9.8
	Vmid	-1.25	5.40	7.3	-11.6

Table 1(b) Mean (H1-H2) values at Vbeg and Vmid for male speakers.

The overall (H1-H2) values for male speakers are considerably lower than those for female speakers in each category of stop. The standard deviation (s.d.) for females is larger than males for all categories and contexts. The females thus, have a wider range as compared to males. The higher values of (H1-H2) for female speakers can be taken to indicate a larger open quotient. However, there is a considerable variation within each individual's token and within the group. Thus, we can argue that females can have open quotients from very small to very large. This variation is much less in males. The values for VP present an interesting case. VP are produced with an abducted glottis and should have a small open quotient as suggested by the low values for male speakers. The high values for VP for female speakers indicate a large open quotient even for an abducted glottis. This suggests the presence of a glottal chink as predicted in earlier studies [3,4,5].

4.2 H1-A1

Table 2 (a) and (b) show the results for the spectral measure (H1-A1) for VAP, NAVAP and VP for female and male speakers. Again there is a considerable overlap between the three categories of stops but the overall order is VAP>NAVAP>VP. The overall difference between VAP and VP, and VP and NAVAP are significant at p<0.0001 for all speakers. The difference between VAP and NAVAP are slightly less significant at p-value of 0.01 for male speakers and 0.002 for female speakers. Thus, overall the first formant bandwidth is widest for VAP indicating greatest energy loss at glottis, followed by NAVAP and VAP respectively.

		Mean	s.d.	Max.	Min.
VAP	Total	20.80	15.29	53.6	-3.9
	Vbeg	24.90	14.49	53.6	3.7
	Vmid	16.70	15.13	45.8	-3.9
NAVAP	Total	17.80	14.56	49.9	-6.5
	Vbeg	15.50	14.36	43.7	-6.5
	Vmid	20.09	14.55	49.9	-3.1
VP	Total	12.84	14.83	45.5	-17.2
	Vbeg	12.71	14.67	45.5	-8.5
	Vmid	12.97	15.16	42.8	-17.2

Table 2 (a) Mean (H1-A1) values at Vbeg and Vmid for female speakers.

		Mean	s.d.	Max.	Min.
VAP	Total	7.29	8.76	28.9	-16.2
	Vbeg	13.00	5.92	28.9	2.5
	Vmid	1.58	7.31	15.5	-16.2
NAVAP	Total	3.87	6.67	19.4	-19.7
	Vbeg	0.30	6.09	11.2	-19.7
	Vmid	7.45	5.18	19.4	-0.4
VP	Total	-3.90	4.93	7.6	-19.9
	Vbeg	-3.55	4.48	5.3	-12.4
	Vmid	-4.25	5.39	7.6	-19.9

Table 2 (b) Mean (H1-A1) values at Vbeg and Vmid for male speakers.

Again, Vbeg is greater than Vmid in VAP. For NAVAP , the reverse is true . No significant difference is observed w.r.t. Vbeg and Vmid in VP. Thus, the first formant bandwidth decreases from Vbeg to Vmid for VAP and increases from Vbeg to Vmid for NAVAP.

The pattern with respect to female-male differences is similar to that observed for (H1-H2). Female speakers have much higher values across all stop categories and contexts than males. Again, the difference is more marked in VP. The relatively higher values for VP in females indicates high energy losses at glottis leading to a wider first formant bandwidth even when complete glottal closure is expected.

4.3 H1-A3

(H1-A3) values for VAP, NAVAP and VP are given in Table 3 for (a) female and (b) male speakers. Though there is a difference of only 2-3 dB between NAVAP and VAP, overall (H1-A3) values are of the order VAP>NAVAP>VP. The difference between NAVAP and VP and VAP and VP is highly significant (p<0.0001). The difference between NAVAP and VAP is less significant (p= 0.07 for females and 0.01 for males). Thus, the spectral tilt observed is steepest for VAP followed by NAVAP and VP respectively.

		Mean	s.d.	Max.	Min.
VAP	Total	36.43	10.86	54.1	10.5
	Vbeg	40.86	10.35	54.1	13.8
	Vmid	32.01	9.56	47.3	10.5
NAVAP	Total	34.63	9.16	49.9	6.1
	Vbeg	33.79	8.48	47.3	12.3
	Vmid	35.47	9.83	49.9	6.1
VP	Total	29.79	8.29	45.5	7.4
	Vbeg	30.34	8.91	45.5	9.7
	Vmid	29.24	7.68	42.8	7.4

Table 3 (a): Mean (H1-A3) values at Vbeg and Vmid for female speakers.

		Mean	s.d.	Max.	Min.
VAP	Total	34.09	6.81	49.4	14.7
	Vbeg	37.91	5.41	49.4	14.7
	Vmid	30.27	5.90	46.7	19.2
NAVAP	Total	31.54	6.35	44.5	21.9
	Vbeg	30.83	6.40	42.3	26.8
	Vmid	32.26	6.31	44.5	21.9
VP	Total	24.04	5.07	34.5	9.5
	Vbeg	23.95	4.40	34.5	11.9
	Vmid	24.13	5.73	33.6	9.5

Table 3 (b) : Mean (H1-A3) values at Vbeg and Vmid for male speakers.

When Vbeg and Vmid are taken into account, we find that (H1-A3) values decrease from Vbeg to Vmid for VAP and increase from Vbeg to Vmid for NAVAP. For VP, there is no significant difference. The difference between Vbeg and Vmid is not very high (approx. 2 dB) for NAVAP but for VAP there is a considerable difference (approx. 8 dB).

The difference in the two sexes follow the same pattern as the previous two measures. Thus, females have higher (H1-A3) values than males in all categories and sub-categories of stop. The distinction is less prominent in absolute terms , however, it is statistically significant (p<0.0001). The range for females is much larger as compared to other categories with the maximum in each category being higher and the minimum being lower than males. Thus, in terms of spectral tilt, females show a large variation, however, overall the spectral tilt for females is larger than for males.

5. DISCUSSION

The results obtained along different spectral parameters can be related to glottal parameters to help give us a better understanding of the glottal configurations involved in the production of NAVAP, VAP and VP. The results show that the values of (H1-H2), (H1-A1) and (H1-A3) are all of the order VAP >NAVAP> VP. Thus, not only are the VAP produced with a larger open quotient but also with a larger area of glottal opening that results in a wider first formant bandwidth, and a steeper spectral tilt. VP show the reverse glottal configuration, that is the open quotient and the area of glottal opening are both much smaller. This is as predicted by previous studies [2,3,4]. NAVAP are interesting in that their values are in between those observed for VAP and VP.

The results show that for NAVAP the open quotient, first formant bandwidth and spectral tilt increases from vowel initial to mid position. Whereas for VAP the reverse is true. VP vary little with vowel position. The decrease in the values of these parameters in the case of VAP is not at all surprising. In the production of VAP, the glottal opening starts in the middle of the articulatory closure, peaks in the middle of the noise duration following plosion and terminates in the middle of the following vowel. Thus, at the beginning of the following vowel, the glottal opening would be larger than it would be at the middle of the vowel. At the middle of the following vowel, the opening would either have terminated or be close to termination. Thus, the results for VAP are entirely as expected. The production of VP involves an abducted glottis and hence, does not involve any glottal opening before or after plosion. Thus, it is to be expected that there would be little difference with position in the reported values of the spectral measures for VP. The slight increase sometimes observed vowel medially could be attributed to the opening of the glottis in anticipation of an utterance end.

The production of NAVAP along the same parameters shows a mechanism different from that of VAP and VP. For NAVAP, vowel initially, open quotient, first formant bandwidth and spectral tilt is much lower than vowel medially. This suggests that for NAVAP, the glottal opening peaks much later than it does for VAP. To explain this characteristic of NAVAP, we would suggest that the glottal opening in NAVAP starts later, at the end of the articulatory closure or immediately after articulatory release, and peaks well into the middle portion of the following vowel. This fits in with the earlier observation that NAVAP are marked by an absence of a clear breathy or voiced aspirated period after the articulatory release and before the beginning of the vowel formants. NAVAP and VAP have been consistently reported as being significantly different for all measured parameters vowel initially. The start of the glottal opening near articulatory release would explain the difference between NAVAP and VAP at that position. At that particular point where the articulators come apart, the glottis in the case of VAP is proceeding towards its peak opening. This would result in a larger open quotient and a greater rate of air-flow through the glottis at that point. In the case of NAVAP, the glottal opening would have just begun at that point and the open quotient and the rate of airflow through the glottis would be much smaller. The overall higher values reported for VAP as compared to NAVAP also suggests that the glottal opening and the rate of airflow overall are much larger for VAP than for NAVAP. Thus, not only a difference in timing but also a difference in the size of the glottal opening result in NAVAP being less breathy than VAP.

In terms of male and female differences, females had a higher value for all spectral measures as compared to males. The difference is most visible in the case of VP where the spectral measurements clearly indicate that female speakers do not produce VP with a completely abducted glottis. Female speakers have high positive H1-H2 averages as compared to the largely negative values for VPs in male speakers. This difference in VP can be seen in (H1-A1) values and the noise level. Thus, female speakers have large open quotients, higher noise levels as well as wider F1 bandwidth. However, the difference between VP and the other two categories of stops, viz., NAVAP and VAP are maintained by female speakers. That is, even though the spectral analysis of female speech indicates that their VP are "breathy" as compared to those of male speakers, within the female speech the overall order of breathiness, VAP > NAVAP > VP, is maintained.

In terms of spectral tilt, the females have an extremely wide range with their minima being lower and their maxima being higher than those for male speakers. Recall that spectral tilt is influenced by not only the area of glottal opening but also the manner of glottal closure during vocal fold vibration. A large spectral tilt would thus, indicate non-simultaneous vocal fold closures where there is a time lag between the closure at the posterior and anterior ends of the glottis as the vocal folds come together. Thus, the low tilt values in some of the female tokens would indicate a more abrupt cut-off of airflow through the glottis due to a more simultaneous closure of the vocal folds.

ACKNOWLEDGEMENTS

This study was carried out as a part of research for D.Phil. at The University of York, UK.

REFERENCES

[1] Bali, K. 1997. Temporal Aspects of Voiced Aspirated Plosives in Delhi Hindi. Paper presented at the 18th South Asian Language Analysis Roundtable, New Delhi, India.

[2] Bickley, C. 1982. Acoustic analysis and perception of breathy vowels. *Speech Comm. Group Work. Papers I*, Research Lab. of Electronics, MIT, 71-82.

[3] Davis, K. 1994. Stop voicing in Hindi. *Journal of Phonetics,* 22, 177-193.

[4] Dixit, R. P. 1989. Glottal gestures in Hindi plosives. Journal of Phonetics, 17, 213-237.

[5] Fisher-Jørgensen, E. 1967. Phonetic analysis of breathy (murmured) vowels in Gujrati. *Indian Linguistics,* 28, 71-139.

[6] Hanson, H. M. 1995. *Glottal Characteristics of Female Speakers*. PhD Thesis, Harvard Univ.

[7] Holmberg, E. B., Hillman, R. E., Perkell, J. S. 1988. Glottal airflow and transglottal air-pressure measurements for male and female speakers in soft, normal and loud voice. *Journal of the Acous. Soc. of America*, 84, 511-529.

[8] Klatt , D. H. and Klatt, L. C. 1990. Analysis, synthesis and perception of voice quality variations among female and male talkers. Journal of the Acous. Soc. of America, 60, 190-197.

[9] Ladefoged, P. and Antonanzas-Barroso, N. 1985. Computer measures of breathy phonation. *Work. Papers in Phonetics 61*, UCLA, 79-86

[10] Södersten, M. and Lindestad, P-A. 1990. Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech and Hear. Res.*, 33, 601-611.