Durational Characteristics of Hindi Stop Consonants

Samudravijaya K

Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai, 400005, India
chief@tifr.res.in

Abstract

A study of the durational characteristics of Hindi stop consonants in spoken sentences was carried out. An annotated and time-aligned Hindi speech database was used in the experiment. The influences of aspiration, voicing and gemination on the durations of closure and post-release segments of plosives as well as the duration of the preceding vowel were studied. It was observed that the post-release duration of a plosive changes systematically with manner of articulation. However, due to its large variation in continuous speech, the post-release duration alone is not sufficient to identify the manner of articulation of Hindi stops as hypothesised in earlier studies. A low value of the ratio of the duration of a vowel to the closure duration of the following plosive is a reliable indicator of gemination in Hindi stop consonants in continuous speech.

1. Introduction

Aspiration is a phonemic characteristic of Hindi, a major Indian language belonging to the Indo-European family. Retroflexion and gemination are important features of Indian languages. Consequently, there are many more stop consonants in Hindi than, say, in English language. The study of supra-segmental characteristics of Hindi consonants is, therefore, interesting not only because of theoretical importance, but also because of practical significance. A detailed knowledge of the prosodic characteristics of the language will facilitate the formation of prosodic rules for synthesis of natural sounding speech. This paper deals with a study of durational characteristics of Hindi stop consonants in read speech with an emphasis on aspiration and gemination.

There have been several studies of the durational attributes of Hindi stop consonants in the past [1], [2], [3], [4], [5], [6]. However, these have been controlled studies in the sense that subjects were asked to speak carefully selected words containing consonants of interest. In contrast, the current study deals with the characteristics of Hindi consonants in naturally read sentences. Controlled experiments have an advantage that they permit a systematic study of a particular characteristic of speech sounds while maintaining most other characteristics invariant. In addition, it is easier to collect data sufficient to draw statistically meaningful conclusions. However, the prosodic characteristics of consonants in isolated words tend to differ from those embedded in continuously spoken sentences due to factors such as speaking rate. Thus, a study of characteristics of consonants embedded in sentences is important.

Davis [2] studied voicing characteristics of Hindi plosives in word initial position. She examined a previous hypothesis that a special phonological feature of [murmur] or [breathy voice] is needed to distinguish voiced aspirated stops from their unvoiced counterparts. The study concluded that all 4 Hindi *velar* stops may be distinguished from each other on the basis of voice "lag" by redefining it as the post-release "noise offset time". One of goals of this study is to examine whether this observation holds good for Hindi plosives in general (i.e., other 3 places of articulation as well) in *continuous* speech.

In a study of acoustic properties of Hindi plosives in VCV syllables, Tyagi et al [1] observed that voicing and aspiration tend to compress the closure duration of plosives. Shrotriya et al [4] studied the durational characteristics of Hindi consonant clusters in word medial and word final positions. They observed that (a) duration of the vowel preceding a consonant cluster is reduced, (b) voice onset time of a stop is more in a cluster word as compared to non-cluster word, (c) the closure duration of a cluster depends on the place of articulation. The data used in this study was about 80 cluster words spoken by 5 speakers. Another aim of the current study is to investigate the validity of these observations in the context of continuous speech.

In an acoustic analysis of singleton and geminate fricatives of Italian language [7], a general tendency of shortening the pre-consonant vowel and of lengthening the consonant in a geminate utterance was observed. Consequently, the ratio of the duration of the consonant to that of pre-consonant vowel was significantly related to gemination. In particular, this ratio was more than 1 for most geminate fricatives and less than 1 for singleton fricatives. A Maximum Likelihood classification of fricatives based on this ratio was found to give good results (92% accuracy). In addition, the threshold value of 1 was verified to be significant for Italian stops as well. It should be noted that while fricatives are continuant sounds, stops are not. The authors suggested that "speaker intention in producing a geminate is reflected in the production of a consonant at least longer than the vowel preceding it". The data used for the experiment consisted of 3 repetitions of VCV type bisyllabic words spoken by 6 speakers. It would be interesting to see whether this hypothesis holds good in case of geminated Hindi consonants embedded in continuous speech.

The rest of the paper is organised as follows. A brief description of the speech database used in the experiment is given in section 2. The observed changes in the duration of closure and release segments of Hindi plosives due to aspiration and gemination are presented and discussed in section 3. A summary of the conclusions of the study is given in section 4.

2. Speech database

The current work deals with durational characteristics of stop consonants in spoken Hindi sentences. An annotated and time-aligned Hindi speech database was used in the experiments. The relevant features of this database are mentioned below; the details can be found in [8].

The database consists of a total of 1000 sentences spoken

by 100 speakers. There are two sets of sentences. The first set of two sentences (called dialect sentences) contains most Hindi phonemes. The dialect sentences were read by each and every person. The second sets of sentences (one distinct set for each speaker) were designed such that they collectively cover most phonemic contexts. Each plosive was represented as a sequence of 2 segments: closure and release. Special symbols were used to signify the geminate consonants and nasalised vowels.

Retroflexion is a prominent feature of Hindi. There are 4 places of articulation: velar, dental, retroflex and bilabial. Aspiration is a phonemic feature in Hindi. Thus, there are 8 aspirated plosives and 2 aspirated affricates in addition to their unaspirated counterparts.

3. Results and discussion

The means and durations of various units of speech were computed and used to study the durational characteristics of consonants as well as the preceding vowels. The observed changes in the durations due to aspiration and gemination are presented and discussed in this section.

A plosive (geminate or otherwise) is represented as a sequence of 2 segments: closure and release. The release segment starts at the burst of the plosive; it ends when the voicing of the following vowel begins in case of unvoiced plosives, and when the frication/aspiration ceases in case of voiced plosives.

3.1. Aspiration

Systematic change in the durations of the closure as well as the release segment of plosives was observed due to aspiration.

3.1.1. Release duration

In case of unvoiced plosives, the release segment starts at the burst of the plosive and ends when the voicing of the following vowel begins. In case of voiced plosives, the release segment is defined to start at the burst, and end when the post-burst frication/aspiration ceases. In the latter case, the duration of the release segment is equivalent to "lag" or post-release "noise offset time" as defined by Davis [2].

The mean and standard deviation of durations of "release" segments of plosives are shown in Table 1. The number of segments of each type in the database is also shown in the fourth column.

The 5 blocks in the Table corresponds to plosives corresponding to 4 places of articulation and affricates. The symbols D,T,D^h,T^h represent retroflex plosives. It is easy to see that the duration of the "release" segment varies systematically as a function of the manner of articulation. The noise offset time of aspirated plosives are longer than those of unaspirated plosives as expected. Similarly, unvoiced plosives have longer noise offset time than the voiced plosives. However, the large values of standard deviation indicate high variability in the absolute duration of the release segments. Therefore, the release duration alone is not sufficient to identify the manner of articulation of plosives (of a given place of articulation) in continuous speech. This is not in agreement with the observation of Davis [2] that all 4 Hindi velar stops may be distinguished from each other on the basis of voice "lag" alone. However, the differences in the stimuli between her experiment (monosyllable words containing initial velar stops) and the current work (sentences containing all plosives in all contexts) should be kept in mind while comparing the results of the two experiments. The dependence of phoneme duration on speaking rate of different persons is more pronounced in case of continuous speech than in case of words. For example, the duration of the phoneme /b^h/ varies from 40msec to 100msec for different speakers in the Hindi speech database. Secondly, in controlled experiments, subjects tend to articulate words clearly; this may result in enhancement of the distinguishing features of phonemes. It may be noted that the range of the mean noise offset times of 4 velar plosives in experiments of Davis [2] (63msec) is significantly more that that in the present work (29msec). The average noise offset times of g, k, g^h, k^h were $11, 26, 34, 74$ msec respectively in the former case whereas the corresponding values in the present work are $22, 28, 36, 51$ msec respectively (see the topmost cell of column 2 of Table 1). In another study of Hindi stops, the noise offset times of g, k, g^h, k^h were $5, 10, 75, 86$ msec respectively [1](Table 1). Here again, the range (81msec) of the noise offset times of 4 velar plosives is larger than that in case of continuous speech (29msec).

label	mean	SD	ntoken	ratio
g	22	19	453	1.61
k	28	16	1879	1.82
g^h	36	26	129	
k^h	51	26	205	
j	47	28	615	1.29
c	66	37	542	1.32
j^h	60	21	46	
c^h	87	28	98	
D	14	15	147	1.64
T	16	16	213	1.62
D^h	24	32	27	
T^h	27	18	111	
d	16	09	650	2.04
t	23	14	1277	2.17
d^h	33	30	208	
t^h	50	21	244	
b	15	13	571	2.32
p	17	13	843	3.53
b^h	35	33	261	
p^h	61	37	79	

Table 1: *The mean and standard deviation (SD) of durations (in msec) of "release" segments of plosives. The symbols D,T,D^h,T^h represent retroflex plosives. The mean duration changes systematically as a function of manner of articulation. However, due to large values of standard deviation, the duration alone is not sufficient to distinguish plosives of different manner of articulation. The last column shows the ratio of noise offset time of aspirated stop to that of the corresponding unaspirated stop.*

The last column of the Table 1 shows the ratio of noise offset time of aspirated stop to that of the corresponding unaspirated stop. The ratio is smallest in case of affricates due to intrinsically long duration of unaspirated affricates. The ratio varies from 1.6 to 3.5 in case of plosives. The ratio is highest for bilabial stops due to intrinsically short duration of unaspirated bilabial plosives.

3.1.2. Closure duration

It was also observed that the closure duration of aspirated plosives are slightly *shorter* than those of the unaspirated counter-

parts (detailed information in [9]). This trend differs from that observed by Lisker and Abramson [10] in case of word initial Hindi stops; the average closure durations were 63 and 75 msec for /g/ and /g^h/ respectively. On the other hand, the trend is in agreement with those observed by Tyagi et al [1](Table 1) in case of intervocalic stops. The average closure durations for /g/ and /g^h/ were 100 and 84 msec respectively. Such a trend was observed by them in case of other voiced and unvoiced stops as well. In a study of plosives of Marathi language, the measured average closure durations were 116 and 89 msec for /g/ and /g^h/ respectively [10]. Nevertheless, authors of [10] observed that the total (closure + release) duration of an aspirated plosive is longer than that of its unaspirated counterpart. This interesting behavior of decrease in duration of a segment with an increase in the duration of the following acoustic segment is reminiscent of a general trend observed in vowel-consonant pairs; vowels are shorter before unvoiced consonants.

3.2. Gemination

Changes were observed in the durations of the release segment of plosives as well as the duration of the preceding vowel due to gemination.

3.2.1. Duration of the preceding vowel

Shrotriya et al [4] found that the duration of Hindi vowels preceding cluster of stop consonants decreased by about 10 to 15msec as compared to the case when there is no cluster. In case of Italian geminate stops, the decrease in the duration of vowels was about 43msec (26% decrease) [11]. This was true in case of geminate fricatives as well; the duration of the preceding vowel decreased by about 28% [7]. However, we observed that such a decrease is true in 28 out of 44 instances of vowel-geminate stop segments in our Hindi speech database. In fact, the durations of long vowels (a:, e: etc.) *increased* due to gemination. Presumably, the speakers slow down to clearly articulate long vowel followed by a geminate stop.

The duration of vowel preceding voiced stops (singleton or geminate) was found to be longer than that in case of unvoiced stops. This observation is in agreement with previous works including the one on Hindi consonants [4].

3.2.2. Closure duration

The average closure durations of plosives are shown in the form of a bar-chart in Figure 1. The bars occur in triplets. The first, second and third bars of a triplet represent the closure durations of voiced singleton, unvoiced singleton and unvoiced geminate plosives respectively. The corresponding data for voiced geminate plosives are not shown due to insufficient data. The Figure shows 5 bar triplets corresponding to the 5 places of articulation.

From Figure 1, it is clear that the closure duration is shorter if the stop is voiced. The mean closure duration of a Hindi plosive increases due to gemination as expected. The percentage increase in duration varies from 48% in case of bilabial unvoiced stops to 128% in case of retroflex stops. Similar behavior is found in case of voiced stops, albeit to a lesser degree. This is in agreement with studies of Hindi stop consonants [4] and Italian stop consonants [11]. However, the average closure duration of geminate unvoiced stops of Hindi as reported by Shrotriya et al [4] (199msec) is much larger than that observed here (135msec). Thus, there is a difference between the closure durations of geminate plosives found in isolated Hindi words

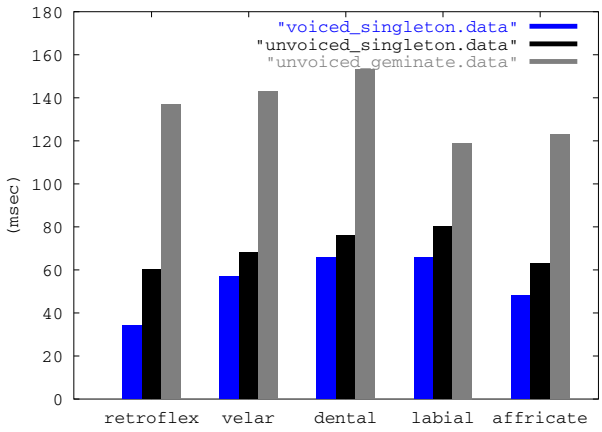

Figure 1: *The average closure durations of singleton and geminate Hindi plosives of 5 place of articulation. The percentage increase in duration due to gemination is dependent on place of articulation, and varies from 48% to 128%.*

and those in continuously spoken sentences.

While it is clear from Figure 1 that the *mean* closure duration increases significantly on gemination, the current work also showed that there is considerable variation in the closure durations of different tokens of a given plosive in natural speech. The mean and standard deviations of various durational attributes of Hindi stops are tabulated in [9]. For example, the mean and standard deviations closure durations of singleton /k/ are 67 and 46msec respectively. The corresponding quantities for the geminate counterpart are 141 and 38msec respectively. Thus, there is considerable overlap between the closure duration distributions of singleton and geminate stops. In other words, the closure duration alone is not sufficient to distinguish singleton and geminated plosives. This is in agreement with conclusions of a study [5] of perceptual cues for gemination in Hindi voiceless unaspirated stops at the bilabial and velar places of articulation, wherein the authors concluded that while closure duration is the primary cue for the geminate vs. singleton distinction in Hindi stops, secondary cues based on preceding and following vowels are also used by listeners.

3.2.3. Vowel consonant duration ratio

It has been reported that the duration of a geminate fricative consonant is almost always more than the duration of the preceding vowel [7]. In fact, the ratio of the duration of the vowel to that of the following consonant was mostly less (greater) than 1 for geminate (singleton) consonants; this was observed to be true for both fricatives and stops in bisyllabic Italian words [7]. It was hypothesised that the speaker intention in producing a geminate is reflected in the production of a consonant at least longer than the vowel preceding it. In the following, we will examine whether this hypothesis holds good in case of Hindi plosives embedded in continuous speech.

Let C and T denote the closure duration and the total duration (including the release duration) of a plosive respectively. Let V represent the duration of the vowel preceding the plosive. We use the subscript g and s to indicate the cases of geminate and singleton stop consonants respectively. Since the database consists of natural sentences, the number of geminate conso-

nants is small. Therefore, only those vowel-geminate stop combinations which occur atleast twice in the database are considered here.

vowel	stop	V_g/T_g	V_s/T_s	V_g/C_g	V_s/C_s
a	k	0.455	0.976	0.579	1.537
e:	k	0.800	1.302	0.932	1.890
a	Th	0.491	0.998	0.610	1.554
a	t	0.392	0.769	0.436	1.063
u	t	0.373	0.888	0.416	1.328
a:	p	1.041	1.257	1.138	1.594
a	dh	0.376	1.192	0.441	1.852
a	b	0.647	1.203	0.705	1.892

Table 2: The ratio of the duration of the vowel to that of the following stop consonant for various combinations of vowel and geminate stops. The ratio (V/C) of the duration of the vowel (V) to the *closure* duration of the following plosive (C) can be used to distinguish geminate and singleton plosives in continuous speech. This is a better measure than (V/T) where T is the *total* duration of the plosive.

The ratio of the duration of the vowel to that of the following stop consonant is shown in Table 2 for various combinations of vowel and geminate stops. The third and the fourth columns of the Table show the ratio of duration of the vowel (V) to the total duration of the stop (T) for geminate and singleton stops respectively. The ratio (V/T) is less than 1 for most geminate plosives. However, the ratio is greater than 1 in only about half of the singleton stops (column 4). This is at variance with the observations of [7] where the ratio was greater than 1 for most singleton consonants. On the other hand, if we consider the duration of the closure of the stop consonant (C) instead of its total duration (T), the ratio (V/C) is greater than 1 in all cases of singleton stops as shown in the last column of the Table. This is not surprising as the major acoustic correlate of gemination is the lengthening of the closure duration of stops [12], [11], [5]. For example, the closure duration, on an average, doubled on gemination in case of Italian stops [11]. In the same study [11], it was also observed that when the duration of the preceding vowel is elongated, the closure duration must be longer for perceiving a geminate consonant. Thus, employment of a ratio of durations rather than the absolute value of closure duration is preferable; it has the advantage that the ratio compensates for variations in the speaking rates of different speakers. We can conclude that the ratio of the duration of the preceding vowel to the closure duration of the plosive can be used to distinguish geminate and singleton Hindi plosives in continuous speech. However, a single (vowel independent) threshold for such a distinction is not be possible due to intrinsically large durations of long vowels of Hindi.

4. Conclusions

The durational characteristics of Hindi stop consonants in read sentences were studied with an emphasis on aspiration and gemination. The following are the main conclusions of this study.

- The post-release duration of a plosive changes systematically with manner of articulation. However, due to its large variation in continuous speech, the post-release duration alone is not sufficient to identify the manner of articulation of the stop as observed earlier [2].
- In many cases, the closure duration of aspirated plosives are slightly shorter than those of the unaspirated counterparts.
- The closure duration of a plosive increases due to gemination; the degree of increase is a function of the place of articulation. However, the closure duration alone is not sufficient to distinguish singleton and geminated plosives.
- A low value of the ratio of the duration of the vowel to the closure duration of the following plosive is a cue for gemination in Hindi stop consonants in continuous speech.

5. References

[1] S.K.Tyagi, S.S.Agrawal, K.D.Pavate, "Acoustic properties of (Hindi) plosives spoken in VCV syllables", J. Ac. Soc. India, **XVII**:313–318, 1989.

[2] Katharine Davis, "Stop Voicing in Hindi", J. Phonetics, **22**:177–193, 1994.

[3] Shrotriya, N., Sharma, A.S.S., Verma, R. and Agrawal, S.S., "Acoustical and Perceptual Characteristics of geminated Hindi stop consonants", Proc. ICPhS-95, **4**:132–135, 1995.

[4] Shrotriya, N., Verma, R., Gupta, S.K. and Agrawal, S.S., "Durational Characteristics of Hindi Consonant Clusters", Proc. ICSLP96, pp. 2427–2430, 1996.

[5] Ohala, M. and J. J. Ohala, "The perception of geminates in Hindi", Seventh Conference on Laboratory Phonology, 2000 (http://www.let.kun.nl/labphon7/abstracts/p47.shtml).

[6] Rajesh Verma and Puneet Chawla, "Comparative analysis of Hindi retroflex and dental CV syllables and their synthesis", Proc. WSLP2003, Eds. K Samudravijaya, Aniruddha Sen, Nandini Bondale, pp. 163-167 (http://speech.tifr.res.in/wslp/proceedings/p55.pdf).

[7] Maurizio Giovanardi, M. and Maria-Gabriella Di Benedetto "Acoustic Analysis of Singleton and Geminate Fricatives in Italian", The European Student Journal of Language and Speech, pp.1-13, June, 1998, available at http://wrangler.essex.ac.uk/web-sls

[8] Samudravijaya K, Rao, P.V.S. and Agrawal S.S. "Hindi Speech Database", Proc. ICSLP00, Beijing, China, October 2000, CDROM paper: 00192.pdf. (http://speech.tifr.res.in/č̃hief/publ/icslp2000.ps.gz)

[9] Samudravijaya K, http://speech.tifr.res.in/č̃hief/publ/03hindiPhonemeDur.pdf

[10] Lisker, L. and Abramson, A.S., "A cross-language study of voicing in initial stops: acoustical measurements", Word, **20**:384-422, 1964.

[11] Esposito, A. and Di Benedetto, M.G., "Acoustical and perceptual study of gemination in Italian stops", J. Acoust. Soc. Am., **106**(4):2051-2062, 1999.

[12] Lahiri, Aditi & Marslen-Wilson, William, "Lexical processing and phonological representation", in Gerard J. Docherty and D. Robert Ladd (eds) Papers in Laboratory Phonology II, Gesture, Segment, Prosody. Cambridge: Cambridge University Press, 1992.