The Phonetics of Bengali Consonant Clusters

Linda Ösp Heimisdóttir *

1 Introduction

This paper presents an acoustic study of clusters of plosives + sonorants in Standard Colloquial Bengali (henceforth referred to as SCB). One of the goals of this paper is simply to provide an adequate phonetic account of these kinds of clusters in the language since such an account does not seem to be available in the literature on Bengali. Equally important is the fact that descriptive accounts of Bengali phonotactics disagree on various issues related to consonant clusters. It is therefore a second goal of this paper to sort out these differences by collecting phonetic data.

The study focused on two aspects of consonant cluster formation in SCB: (i) which combinations of a stop + sonorant are acceptable clusters to speakers of the language and (ii) which, if any, consonant clusters exhibit gemination of a stop in pre-sonorant position. Results show that all combinations of stop + sonorant are licit clusters in SCB. Furthermore, we conclude that the pre-sonorant gemination of stop consonants is most likely bound to certain lexical strata in the language.

2 Background

2.1 Inventory

Bengali has 28 phonemic consonants as well as 3 allophonic ones. Labial [pʰ] is often pronounced as [f], [z] appears in the pronunciation of some proper names instead of [ʤ]. [s] mostly appears in foreign loanwords and certain consonant clusters and is interchangeable with [ʃ] for most speakers (Kostić and Das 1972).

	Labial	Dental/ Alveolar	Retroflex	Palatoalveolar	Velar	Glottal
Nasal	m	n			ŋ	
Plosive	p pʰ b bʰ	t tʰ d dʰ	ʈ ʈʰ ɖ ɖʰ		k kʰ g gʰ	
Affricate				tʃ tʃʰ ʤ ʤʰ		
Fricative	(f)	(s) (z)		ʃ		h ɦ
Approximant		l				
Rhotic		r	ɽ			

Table 1: Bengali consonants. Allophonic consonants are shown in parentheses.

All Bengali consonants except velar [ŋ], voiceless [h], and [ɾ], occur as geminates. Geminate plosives have, instead of two occlusive periods and two explosions, a prolonged occlusive period and one common explosion, hence they are not pronounced with two separate releases (Kostić and Das 1972).

*I thank Sam Tilsen for guidance and feedback.

2.2 Consonant Clusters

2.2.1 Possible Clusters

Medial consonant clusters are very common in SCB. These clusters are usually heterosyllabic since the occurrence of syllable-initial clusters is very restricted and coda clusters are generally disallowed. There is a lack of agreement in the literature regarding possible consonant clusters in the language. According to Ferguson and Chowdhury (1960), possible clusters are generally equivalent to the sum of final and initial consonant possibilites (although homorganic voiced and voiceless stops rarely form clusters, and in general there is a tendency towards making a cluster either wholly voiced or voiceless). The number of possible clusters is large since limitations on the occurrence of single consonants are few, with [ŋ] and [r] not occurring initially and [h] not occurring finally. Interestingly, these are also the only consonants that don't appear as geminates in the Bengali language.

Other sources are not quite as permissive in their account of possible clusters as Ferguson and Chowdhury (1960) are. Since this paper is solely concerned with clusters of plosives followed by sonorants, the discussion will henceforth be limited to those types of clusters only. There is a general agreement that all (unaspirated) stops may be followed by [r] (Bhattacharya 1988, Morshed 1985, Ferguson and Chowdhury 1960, Kar 2010). As Table 2 shows, sources also agree that [pn], [pl], [tn], [kl] and [gn] are possible clusters in SCB. No such agreement exists for [bn], [tl], [dn], [dl], [kn] and [gl] however. None of the aforementioned sources present phonetic data in support of their claims regarding SCB clusters. Hence, one of the goals of this study will be to determine how native speakers of SCB pronounce clusters deemed illicit in the literature.

	Bhattacharya	Morshed	Ferguson & Chowdhury	Kar
[pn]	✓	✓	✓	✓
[pl]	✓	✓	✓	✓
[bn]		✓	✓	
[bl]	✓	✓	✓	✓
[tn]	✓	✓	✓	✓
[tl]		✓	✓	
[dn]		✓	✓	
[dl]		✓	✓	
[kn]	✓	✓		
[kl]	✓	✓	✓	✓
[gn]	✓	✓	✓	✓
[gl]		✓	✓	✓

Table 2: Possible consonant clusters in SCB according to four sources.

2.2.2 Preconsonantal Gemination

It has been claimed in some of the literature on Bengali (Bhattacharya 1988, Ferguson and Chowdhury 1960, Kar 2010) that stops geminate when directly followed by a liquid consonant, resulting in a triconsonantal cluster consisting of a coda followed by a complex onset.

(1) /p^{h}utro/ [p^{h}ut.tro] 'son'

According to Bhattacharya (1988), this rule applies to all stops followed by [r]; he furthermore claims that /p/ and /k/ geminate before [l] while /t, b, d, g/ remain singletons. Ferguson and

Chowdhury (1960) claim that only labial and dental stops geminate and that gemination is restricted to the position before [r]. Kar (2010), on the other hand, limits the occurrence of gemination not to types of stops or liquids but to types of words, claiming that this kind of gemination only takes place in borrowed words. These include *tatsama* words, i.e. words directly borrowed from Sanskrit, and borrowings from foreign languages. He supports his claim by maintaining that gemination is not possible in native Bengali words (*tadbhava*) because those words neither allow complex codas nor complex onsets and therefore exclude all clusters of three or more consonants. Much like before, none of the sources on Bengali phonotactics present phonetic data to support claims of preconsonantal gemination and therefore one of the goals of the present study is to collect data that will settle this issue.

2.3 The Bengali Writing System

Bengali is written in a so-called *alphasyllabary*, a segmental writing system in which consonant-vowel sequences are written as a unit. Each unit is based on a consonant letter which by itself denotes a certain consonant followed by an inherent vowel [ɔ] (vowel sounds may be modified by adding a secondary vowel letter). Since consonants are automatically associated with a following vowel sound, a special denotation is needed to symbolize consonant clusters. This is done by conjoining two or more consonants into a single symbol, indicating that they should be pronounced as a sequence. However, not all pairs of consonants can be conjoined into a single symbol (although the writing system does offer the option of adding a secondary symbol between two consonants to indicate that they are not separated by a vowel). Interestingly enough, the set of stops + sonorants that may not be conjoined into a single symbol in the writing system roughly correlates to the sequences of consonants that are the subject of disagreement in the literature as to whether they are possible clusters or not (with the exception that [gl], claimed by Bhattacharya (1988) not to be a licit cluster, can in fact be represented as one).

To complicate matters even further, just because a sequence of two consonants is written with separate symbols and not as a single conjoint letter, that doesn't necessarily mean that these consonants should or need to be pronounced with an intervening vowel. In fact, it appears that conjoint consonant symbols are mainly utilized in words of Sanskrit origin whereas clusters in native Bengali words are simply represented as a sequence of two consonants. This writing convention is again reflected in the fact that, in some cases, native Bengali words may be pronounced either with a cluster or a CVC sequence.

This section about the Bengali writing system is included here for two reasons. First of all, since none of the sources referenced here on the subject of Bengali consonant clusters actually present phonetic data in support of their claims about licit and illicit clusters, it is not completely impossible that some of these sources base their claims entirely or in part on writing conventions (hence the aforementioned correlation between the gap in the system of conjoint symbols and clusters assumed to be illicit).

Secondly, it is interesting to note that words of Sanskrit origin, claimed by Kar (2010) to be the only ones in the Bengali language to exhibit gemination of stops before liquids, traditionally contain conjoint letters while native Bengali words are spelled without any conjunction of consonants, as if there were in fact an inherent vowel separating the two consonants. This writing convention may suggest that historically native Bengali words didn't contain any medial consonant clusters which would mean that all consonant clusters in native words are the result of vowel syncope. If that is the case, there might quite possibly be a difference in the pronunciation of 'original' clusters and 'derived' clusters, respectively, with only the former group exhibiting gemination.

3 Experiment

3.1 Research Questions

Following the discussion of the status of consonant clusters in SCB in previous sections, the experiment reported on here was designed to answer the following research questions:

(2) Which sequences of plosive + sonorant are possible clusters in SCB?

(3) Which, if any, plosives geminate and in what environment?

(4) Is gemination limited to words of Sanskrit origin or does it extend to all Bengali words (in other words, is gemination avoided in clusters that are more likely to be broken up by an intervening vowel)?

Regarding (2), it is assumed that if any sequence of a stop and a liquid is not a possible cluster in the Bengali language, speakers will break that sequence up by inserting an ‘inherent’ vowel [ɔ].

3.2 Methodology

In order to answer the research questions posed in the previous section, an acoustic experiment was conducted in a soundproof booth at the Cornell Phonetics Lab and the results were analyzed by hand in Praat.

A list of 51 words was created, containing a mixture of actual Bengali words and nonce words. Each sequence of a stop + sonorant was represented at least twice, as an orthographic cluster (henceforth referred to as a ‘real’ cluster) and as a sequence of two consonant symbols, that could either be pronounced as a cluster or as a CVC sequence (henceforth referred to as a ‘fake’ cluster). Actual words were used where possible but the list was compensated with nonce words where real words could not be found to represent a cluster (possibly because some clusters are simply not possible in SCB).

In addition to clusters with stops followed by [r] and [l], clusters of stop + [n] were included in the word list. This was done for two reasons: First, to answer the research question in (2), i.e. which clusters of stop and sonorant are possible in Bengali. Second, to provide a point of reference for stop duration in pre-sonorant position. None of the resources consulted on stop gemination in Bengali mention gemination before a nasal as a possibility so it may be assumed that all pre-nasal stops will be singletons whereas it is quite possible that all stops preceding either [r] or [l] are geminates.

Finally, the word-list included intervocalic singleton and geminate stops. These were supposed to give an idea of how much durational difference should be expected between singleton and geminate stops in Bengali.

Five native speakers of SCB were recorded, all graduate students at Cornell University. All of the subjects were in their mid- to late twenties and had grown up in the city of Kolkata in West Bengal.

The experiment was conducted using E-Prime. The words appeared in a randomized order on a monitor and subjects were asked to utter them in the carrier phrase *amake* ______ *bolo* ‘say ______ to me’. Each word was repeated 8 times. Subjects were given a few minute break half way through the experiment. Before starting the experiment, subjects were given a few minutes to familiarize themselves with the wordlist.

When comparing the duration of a stop in a specific cluster between the ‘real’ and the ‘fake’ condition, all instances of ‘fake’ clusters with an epenthetic vowel intervening between the stop and the following sonorant were excluded. This was done to ensure that only the duration of preconsonantal stops would be taken into consideration.

4 Results

4.1 Possible Consonant Clusters

As discussed in Section 2.2.1, sources vary somewhat on what are considered possible consonant clusters in SCB. For reference, Table 2, which shows possible consonant clusters in Bengali according to various sources, is repeated here as Table 3 with the addition of the present study’s results.

The results of this study suggest that there is no combination of a plain stop followed by a sonorant, [r], [l] or [n], that cannot be pronounced as a cluster without an intervening vowel in Bengali. The only ‘fake’ cluster that was consistently broken up by an epenthesized [ɔ] was /tn/ in

	Bhattacharya	Morshed	Ferguson & Chowdhury	Kar	Heimisdóttir
[pn]	✓	✓	✓	✓	✓
[pl]	✓	✓	✓	✓	✓
[bn]		✓	✓		✓
[bl]	✓	✓	✓	✓	✓
[tn]	✓	✓	✓	✓	✓
[tl]		✓	✓		✓
[dn]		✓	✓		✓
[dl]		✓	✓		✓
[kn]	✓	✓			✓
[kl]	✓	✓	✓	✓	✓
[gn]	✓	✓	✓	✓	✓
[gl]		✓	✓	✓	✓

Table 3: Possible consonant clusters in SCB according to four sources.

the word *tʃetna* ‘thought’. However, the ‘real’ counterpart of that cluster (in the word *rotno* ‘gem’) was pronounced without any epenthesis. Table 4 shows all instances of [ɔ]-epenthesis that were found among the five speakers.

Word	Cluster	Tokens with epenthesis	Subjects with epenthesis
tʃetna	/tn/	100%	5/5
bhabna	/bn/	77.5%	5/5
tʃopla	/pl/	75.0%	4/5
dapna	/pn/	46.2%	3/5
tʃhotri	/tr/	37.5%	3/5
nadna	/dn/	28.2%	4/5
mogra	/gr/	28.2%	2/5
khabra	/br/	23.1%	2/5
bokna	/kn/	7.7%	2/5
tʃhapra	/pr/	2.6%	1/5
ʃãtra	/tr/	2.6%	1/5
gadna	/dn/	2.5%	1/5

Table 4: Words that have occurrences of [ɔ]-epenthesis.

As Table 4 indicates, the occurrence of [ɔ]-epenthesis varies somewhat between subjects. Table 5 gives a subject-specific overview of where epenthesis occurred (either in all or some tokens containing each cluster).

According to the data in Table 3, six consonant clusters should be illicit in SCB according to one or more sources on the language’s phonotactics. These are /bn/, /tl/, /dn/, /dl/, /kn/ and /gl/. Out of these six clusters, three were never broken up by an epenthetic vowel in our experiment, i.e. [tl], [dl] and [gl]. It may, therefore, be concluded that these are in fact licit clusters in SCB despite claims to the contrary.

Out of the three remaining clusters, /dn/ and /kn/ were pronounced without an epenthetic vowel in a vast majority of tokens. While /kn/ was not consistently broken up by any individual subject, the fake cluster /dn/ was broken up in all instances by one subject (three other subjects exhibited epenthesis in this cluster but only in one out of eight tokens each). Interestingly, one instance of [ɔ]-epenthesis took place between [d] and [n] in the nonce word *gadna*, which was presented to the subjects with a ‘real’ cluster.

	AD	KS	OP	SG	SC
/pn/	✓	✓	✓		
/pl/	✓	✓	✓		✓
/pr/		✓			
/bn/	✓	✓	✓	✓	✓
/bl/					
/br/		✓			
/tn/	✓	✓	✓	✓	✓
/tl/					
/tr/	✓	✓	✓		
/dn/	✓		✓	✓	✓
/dl/					
/dr/					
/kn/		✓		✓	
/kl/					
/kr/					
/gn/				✓	
/gl/					
/gr/	✓		✓		

Table 5: Occurrences of [o]-epenthesis by subject and cluster.

This leaves us with the cluster /bn/ which was broken up by an epenthetic vowel in a majority of tokens, as Table 4 shows. However, only one subject displayed epenthesis in all eight tokens of the word *bhabna*. Other subjects varied between forms with and without epenthesis. Note, that none of the subjects had any problems pronouncing the 'real' counterpart of this cluster (in the nonce word *ʃabna*).

To conclude, the data show that none of the consonant clusters recorded in this experiment can be considered illicit in SCB. In articulatory terms, the data indicate that, while some pairs of consonants are more resistant to overlap than others and have a tendency to be pronounced with an audible vowel intervening, these occurrences of epenthesis are not robust enough to be considered a regular pattern in the language. These results contradict all the sources consulted on Bengali phonotactics, as discussed in Section 2.2.1.

4.2 Comparison of Intervocalic Singletons and Geminates

The five subjects were recorded uttering words that contained singleton and geminate stops in intervocalic position. Results show that geminate stop closure tends to be at least twice the duration of singleton stop closure. They also suggest that the durational difference between singletons and geminates is greater in voiced stops than voiceless ones. It should be noted, however, that all of the subjects had a tendency to spirantize the voiced singleton stops to some degree intervocalically which might explain why the voiced singletons are shorter and why voiced stops show a greater durational difference between singletons and geminates than voiceless ones do. Results are shown in Figure 1 and Table 6.

4.3 Comparison of Presonorant Stops in 'Real' and 'Fake' Clusters

4.3.1 Before [n]

As discussed in Section 3.2, stops are not reported to geminate when followed by a nasal. Therefore it was expected that 'real' and 'fake' consonant clusters would not differ with respect to the duration

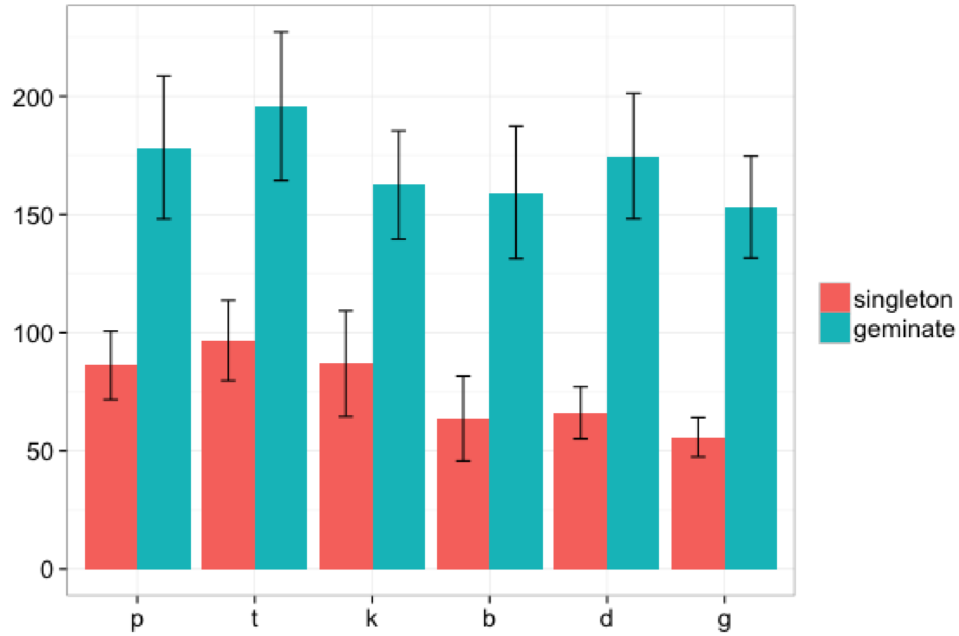

Figure 1: Duration of intervocalic singleton and geminate stops. Error bars represent the sample standard deviation

Stop	Singleton	Geminate	Difference
/p/	86.15	178.31	107%
/t/	96.67	195.83	103%
/k/	86.87	162.44	87%
/b/	63.72	159.33	150%
/d/	65.97	174.72	165%
/g/	55.83	153.13	174%

Table 6: Mean duration (ms.) of intervocalic singleton and geminate stops across all speakers.

of a stop preceding [n]. Results are shown in Figure 2. Note that, as mentioned in Section 3.2, all instances of ‘fake’ clusters with an epenthetic vowel intervening between the stop and the sonorant were excluded from the comparison. The ‘fake’ cluster /tn/ (in the word /tʃetna/) was the only cluster tested that was always broken up by an epenthetic vowel. Therefore, no data are available for the duration of /t/ before a nasal in a ‘fake’ cluster.

A *t*-test was performed, comparing each of the ‘real’ clusters to a corresponding ‘fake’ one, to determine if there was a statistically significant difference in mean duration of stops between the ‘real’ and ‘fake’ conditions. As Table 7 shows, the results were as expected; none of the stops were significantly longer in duration under the ‘real’ condition.

These results will give us a useful point of reference for expected duration of a singleton stop in pre-sonorant position in the following sections. As Figure 3 shows, singleton stops are considerably longer before a consonant than they are in intervocalic position (overall mean duration of intervocalic singleton stops was 76 ms compared to 98 ms for stops preceding [n] in both cluster types combined). It is also expected (see e.g. Kraehenmann (2001)) that geminates will be shorter pre-consonantally than intervocalically. Therefore, the durational difference between singleton and geminate stops before [r] and [l], respectively, will likely be considerably smaller than in intervocalic position.

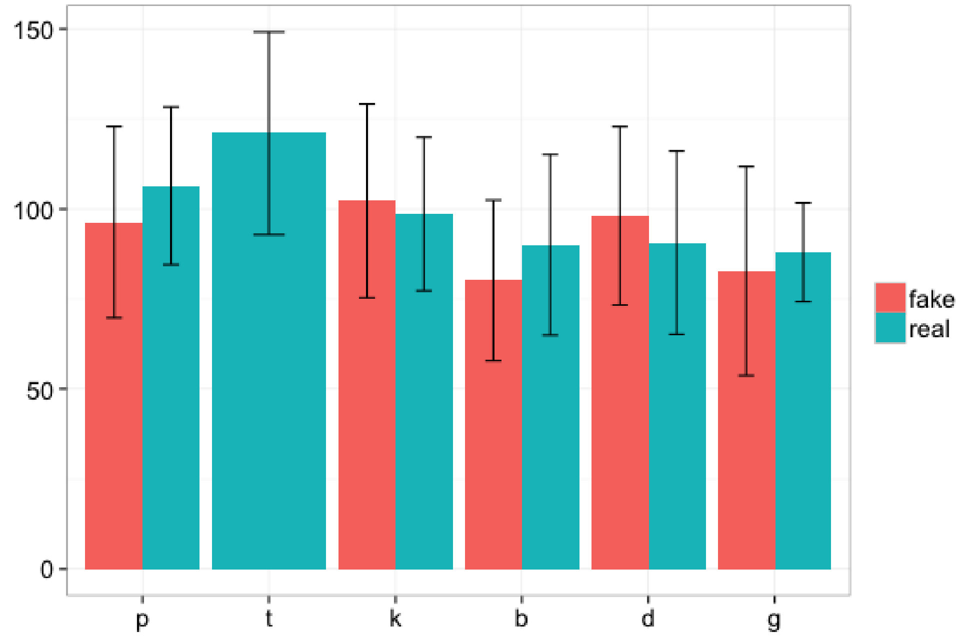

Figure 2: Duration of stops preceding [n] in 'real' and 'fake' clusters. Error bars represent the sample standard deviation.

Cluster	*t*	df	Confidence Int.		Real		Fake		p
			95% L	95% R	M	SD	M	SD	
/pn/	1.489	35.45	-3.68	23.98	106.43	21.86	96.29	26.54	0.073
/kn/	-0.632	66.49	-15.02	7.80	98.58	21.27	102.19	26.89	0.74
/bn/	1.164	14.10	-8.44	28.51	90.15	25.50	80.11	22.29	0.13
/dn/	0.527	48.41	-10.96	18.76	101.04	26.87	97.14	26.24	0.30
/gn/	1.189	29.66	-5.40	20.45	90.27	12.89	82.75	28.97	0.12
*p <0.05 **p <0.01 ***p <0.001									

Table 7: Results of a *t*-test comparing mean duration of stops preceding [n] in real and fake clusters.

4.3.2 Before [r] and [l]

Duration of stops before [r] in 'real' and 'fake' clusters is shown in Figure 4 and durations for [l] are found in Figure 5. In order to determine the relationship between duration of stops and types of clusters ('real' or 'fake'), a mixed effects linear regression with subject as a random factor was conducted. Models were compared using a χ-square test. Comparison of models showed that the best fit was obtained by including an interaction of cluster type with both sonorant and stop type. This model returned an R^2 value of 0.74, with the random variable accounting for 38% of the total variance explained. The mixed model's assumptions of homoscedasticity and normality of residuals was satisfied by using the log values, rather than the absolute values, of the dependent variable (stop duration). P-values for each of the interactions were computed using Satterthwaite's approximation to degrees of freedom by means of the *lmerTest* package in R.

Results showed that interaction of cluster type and stop/sonorant was statistically significant for the following clusters: /pr/, /gr/, /pl/ ($p < 0.001$), /kl/, /tr/ ($p < 0.01$), and /kr/, /dr/ ($p < 0.05$). In all of these cases, the stop duration was longer in 'real' clusters than 'fake' clusters. For convenience, the mean durations for all clusters in both conditions are given in Table 8.

The consonant clusters shown to differ significantly by cluster type correspond to the clusters

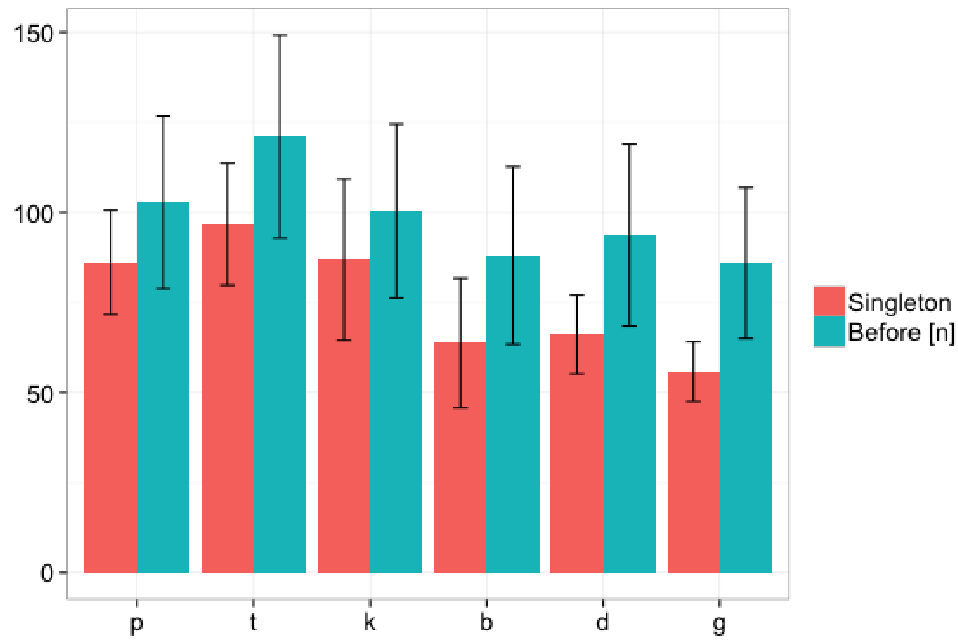

Figure 3: Duration of intervocalic singleton stops compared to the duration of stops preceding [n]. Error bars represent the sample standard deviation.

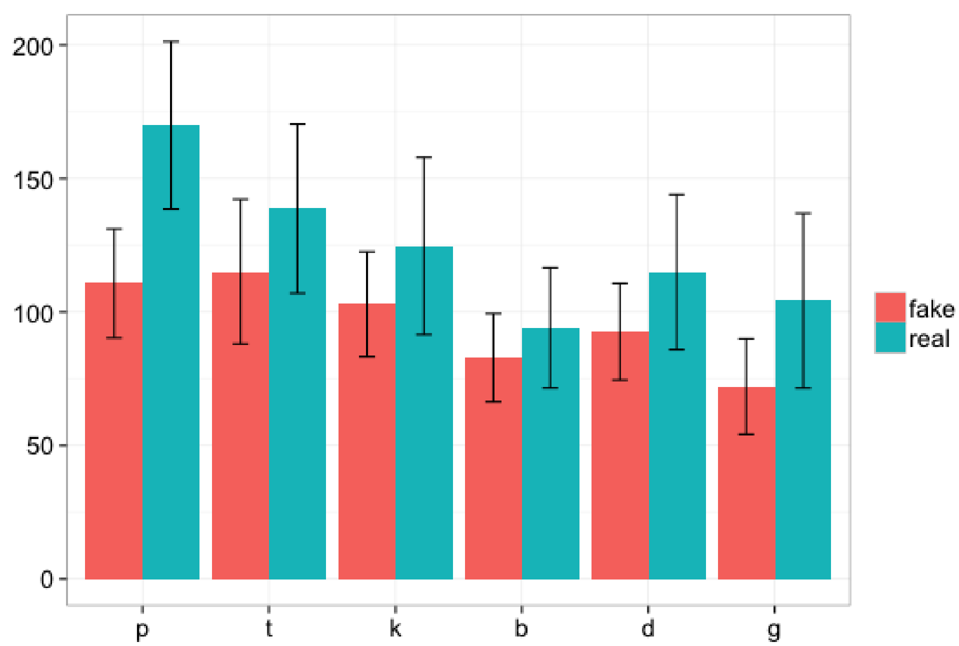

Figure 4: Duration of stops preceding [ɾ] in 'real' and 'fake' clusters. Error bars represent the sample standard deviation.

that were represented by actual words in the experiment. None of the 'real' clusters presented in nonce words exhibited a significant difference between the two conditions. As discussed in Section

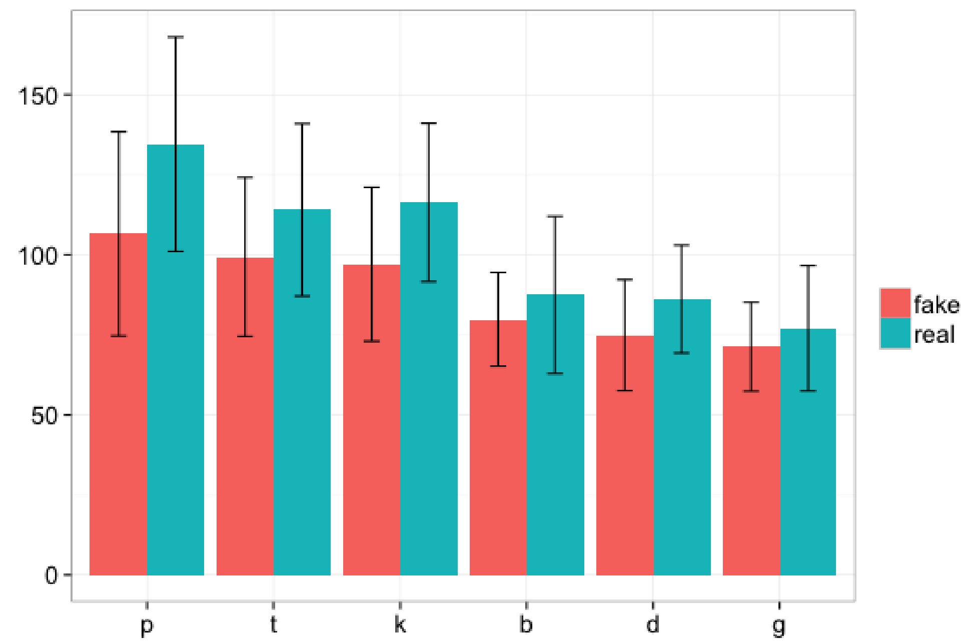

Figure 5: Mean duration of stops preceding [l] in ‘real’ and ‘fake’ clusters. Error bars represent the sample standard deviation.

3.2, nonce words were included in the experiment where no actual words were found containing the cluster in question. Due to imbalanced numbers of actual and nonce words in the experiment, including word type as an independent variable in the model was not possible. However, a separate model was fitted to examine how the interaction between cluster type and word type affects the dependent variable. As before, subject was included as a random variable. Results showed word type to be highly significant ($p < 0.001$) for duration in ‘real’ clusters, i.e. clusters contained in nonce words had on average considerably shorter stop closures than clusters in actual words. However, no significant difference between actual and nonce words was found for ‘fake’ clusters.

Stop	**Duration in ‘real’ clusters**	**Duration in ‘fake’ clusters**	**Difference**
/pr/	169.93	110.66	54%
/gr/	104.23	72.07	45%
/pl/	134.55	106.50	26%
/kl/	116.36	97.08	20%
/tr/	138.70	115.06	21%
/kr/	124.71	102.88	21%
/dr/	114.91	92.59	24%
/tl/	114.05	99.35	15%
/dl/	86.15	74.90	15%
/br/	94.08	82.83	14%
/bl/	87.49	79.87	10%
/gl/	77.08	71.32	8%

Table 8: Mean duration (ms.) of clusters across all speakers.

4.4 Discussion

At this point, it is important to issue certain caveats. Our understanding of the results reported here relies on our ability to distinguish between a singleton and a geminate with some confidence. In other words, the data can only be interpreted if we know where to place a boundary between the duration of a singleton consonant and a geminate one. This is not always an easy task. As Figures 6 and 7 below show, comparison of singleton and geminate stops in intervocalic position does not provide much insight into expected durations in pre-sonorant position because singletons appear to be longer and geminates shorter pre-consonantally than intervocalically.

Similarly, while statistical tests can determine which groups of consonants display a significant difference in duration in a given environment, it is open to interpretation whether a statistically significant difference necessarily reflects an actual difference. To clarify this point, one might ask if it is enough for one stop to be *significantly* longer than another stop in a certain environment for us to conclude that the first stop is a geminate. The answer is: probably not. Gemination is not just a matter of statistically significant difference; *how much* that significant difference is is also important. But a geminate cannot be expected to be twice the length of a singleton in all environments either for reasons mentioned above. For these reasons, some of the results reported here may perhaps not be considered conclusive.

The results discussed in Section 4.3.2 suggest that the following 'real' clusters exhibit gemination of the stop: /pr/, /tr/, /kr/, /dr/, /gr/, /pl/ and /kl/. The following 'real' clusters do not show gemination of the stop: /br/, /tl/, /bl/, /dl/ and /gl/. Mean durations of stops in various environments can be seen in Figures 6 and 7 and Table 9.

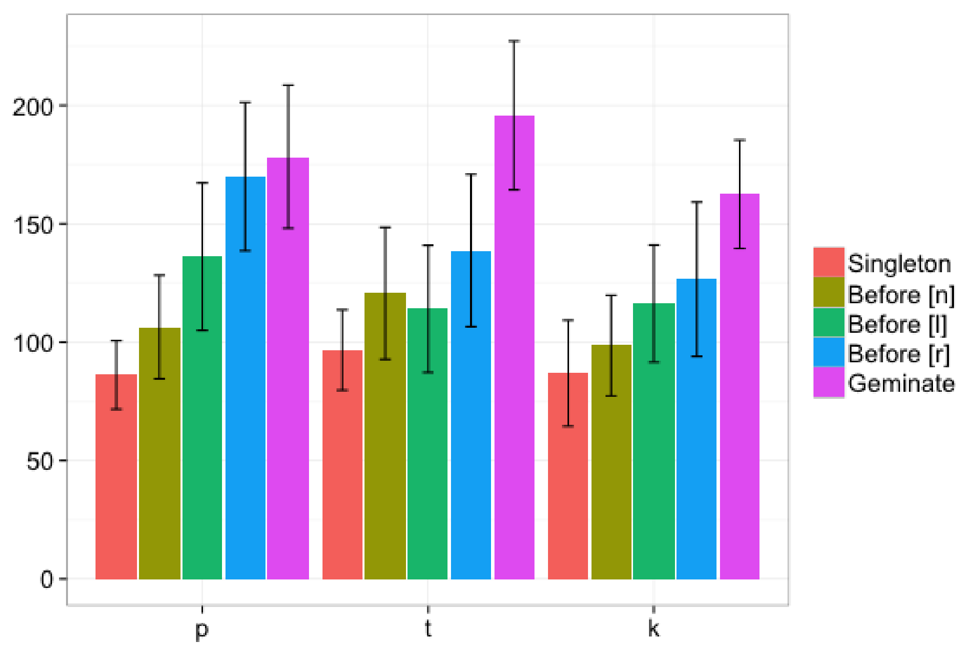

Figure 6: Mean duration of voiceless stops in different environments. Numbers for pre-consonantal stops represent durations in 'real' clusters. Error bars represent the sample standard deviation

We discussed in Section 2.2.2 that various claims are made in the literature on Bengali regarding stop gemination in pre-sonorant position. Those claims are reiterated here:

(i) Bhattacharya (1988): all stops geminate before [ɾ]. Furthermore, /p/ and /k/ geminate before [l] while other stops remain singletons in that position.

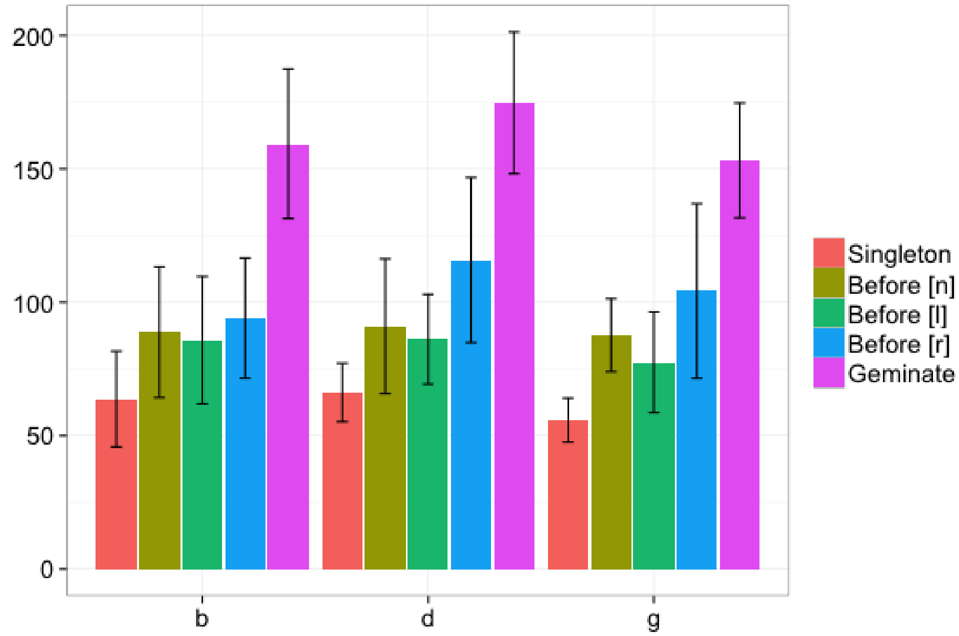

Figure 7: Mean duration of voiced stops in different environments. Numbers for pre-consonantal stops represent durations in 'real' clusters. Error bars represent the sample standard deviation

Stop	Before [n]	Before [l]	Before [r]	Geminate
/p/	124%	*158%	*197%	207%
/t/	125%	118%	*143%	203%
/k/	113%	*134%	*146%	187%
/b/	141%	135%	148%	250%
/d/	153%	131%	*173%	265%
/g/	162%	139%	*187%	274%

Table 9: Mean durations in various environments as a percentage of the duration of an intervocalic singleton. Star-marked values (*) are those concluded to be geminates given the results of statistical tests.

(ii) Ferguson and Chowdhury (1960): only labial and dental stops geminate, gemination only takes place before [r].

(iii) Kar (2010): all stops can geminate before [r] and [l], respectively, but gemination only takes place in borrowed vocabulary (from Sanskrit or foreign languages).

The results of this study indicate that neither Ferguson and Chowdhury (1960) nor Bhattacharya (1988)'s assessment of gemination in SCB is accurate. While Bhattacharya (1988) correctly states that gemination before [l] is limited to voiceless labial and velar stops, he overgeneralizes the gemination effect before [r]. Ferguson and Chowdhury (1960), on the other hand, are too conservative in their estimation of stop gemination, excluding the clusters /kr/, /gr/, /pl/ and /kl/ which all exhibit gemination in the present study (while including /br/ which shows no signs of gemination). This leaves us with Kar (2010)'s theory of the effect of different lexical strata on stop gemination, which turns out to have some merit.

If we compare the group of clusters that show stop gemination to the group of clusters that do not, we discover that there is nothing about the components' place or manner of articulation that sets these groups apart. The only difference between the two groups is that all the clusters in the latter group were produced in nonce words (due to the fact that these clusters were not found in any actual words) while all the clusters in the first group belonged to actual words. These findings support Kar (2010)'s theory since all the actual words that contained a 'real' cluster are words of Sanskrit origin (see discussion in Section 2.3). We, therefore, conclude that gemination of stops in pre-sonorant position in SCB is restricted to a finite number of vocabulary items and is either underlying in those words or the result of a separate phonology that only applies to words of a particular lexical stratum.

5 Conclusion

In this paper, we have presented the results of an acoustic study on stop-sonorant clusters in Standard Colloquial Bengali. The goal was to collect phonetic data that would answer questions regarding two aspects of cluster formation in the language: (i) which clusters are possible and (ii) which clusters, if any, exhibit gemination of a stop in pre-sonorant position.

We found that, despite claims to the contrary, there is no combination of stop + sonorant that cannot be produced as a cluster by a native speaker of SCB. The only 'fake' cluster (i.e. a sequence of two consonants that can either be pronounced as a consonant cluster or as two consonants separated by an epenthetic [ɔ]) that was consistently broken up by all speakers was /tn/ but the 'real' counterpart of that cluster exists in a number of actual Bengali words and was produced by all speakers without problems. It is likely that the occurrence of an epenthetic vowel in 'fake' clusters is somewhat word-specific (perhaps correlated with word frequency) rather than being a reflection of which clusters are acceptable to speakers of SCB.

Regarding stop gemination, we found evidence to support claims that stops only geminate in specific vocabulary items in SCB (words of Sanskrit origin). The results of the acoustic study showed that stops were produced as geminates in all 'real' clusters belonging to actual words while no gemination occurred in nonce words.

References

Bhattacharya, Krishna. 1988. *Bengali Phonetic Reader*. Central Institute of Indian Languages.

Ferguson, Charles A., and Munier Chowdhury. 1960. The phonemes of Bengali. *Language* 36:22–59.

Kar, Somdev. 2010. *Syllable Structure of Bangla. An Optimality-Theoretic Approach.* Cambridge Scholars Publishing.

Kostić, Djordje, and Rhea S. Das. 1972. *A Short Outline of Bengali Phonetics*. Statistical Publishing Society, Calcutta.

Kraehenmann, Astrid. 2001. Swiss German Stops: Geminates All over the Word. *Phonology* 18:109–145.

Morshed, A.K.M. 1985. *A Study of Standard Bengali and the Noakhali Dialect*. Bangla Academy, Dhaka.

Linda Ösp Heimisdóttir
Department of Linguistics
Cornell University
Ithaca, NY 14853
loh5@cornell.edu