The syllable structure of Bangla in Optimality Theory and its application to the analysis of verbal inflectional paradigms in Distributed Morphology

von

Somdev Kar

Philosophische Dissertation

angenommen von der Neuphilologischen Fakultät

der Universität Tübingen

am 09. Januar 2009

Tübingen

2009

Gedruckt mit Genehmigung der Neuphilologischen Fakultät
der Universität Tübingen

Hauptberichterstatter	:	Prof. Hubert Truckenbrodt, Ph.D.
Mitberichterstatter	:	PD Dr. Ingo Hertrich
Dekan	:	Prof. Dr. Joachim Knape

To my parents...

ACKNOWLEDGEMENTS

First and foremost, I owe a great debt of gratitude to Prof. Hubert Truckenbrodt who was extremely kind to agree to be my research adviser and to help me to formulate this work. His invaluable guidance, suggestions, feedbacks and above all his robust optimism steered me to come up with this study. Prof. Probal Dasgupta (ISI) and Prof. Gautam Sengupta (HCU) provided insightful comments that have given me a different perspective to various linguistic issues of Bangla. I thank them for their valuable time and kind help to me. I thank Prof. Sengupta, Dr. Niladri Sekhar Dash and CIIL, Mysore for their help, cooperation and support to access the Bangla corpus I used in this work. In this connection I thank Armin Buch (Tübingen) who worked on the extraction of data from the raw files of the corpus used in this study. And, I wish to thank Ronny Medda, who read a draft of this work with much patience and gave me valuable feedbacks.

Many people have helped in different ways. I would like to express my sincere thanks and gratefulness to Prof. Josef Bayer for sending me some important literature, Prof. Meena Dan, Prof. Aditi Lahiri, Prof. Fritz Hamm and Prof. Wolfgang Sternefeld for their kind suggestions on various issues. I am grateful to Prof. Udaya Narayana Singh, Director, CIIL for giving me the opportunity to join CIIL as a researcher after the submission of my thesis and Mrs. Suchita Singh for all her helps during my stay in Mysore. I have also benefited immensely during the preparation of this work from interactions with my co-researchers and friends in Tübingen and in Berlin. Thanks to Elena Bilan, Beatriz Lopez-Jiménez, Panagiotis

Kavassakalis, Ventsislav Zhechev, Sveta Krasokova and others. I am grateful to all of them. I would like to thank Beate Starke and Christl Tierney for their help in many academic and official matters. I specially thank Prof. Arnim von Stechow for providing me a nice and comfortable academic environment to work at the Seminar für Sprachwissenschaft (Universität Tübingen). I thank Prof. Manfred Krifka and Prof. Truckenbrodt for the office accommodation at ZAS, Berlin. I also thank the German Research Foundation (Deutsche Forschungsgemeinschaft – DFG) for the funding I received through SFB441 and SPP1234 during my doctoral research.

I also thank to the wonderful friend circle I had in Tübingen and in Berlin. I must mention about Raja, Lalitha, Avijit, Susmita, Hasan, Bhavani, Soumya, Rajesh, Prakash, Frank, Amaramend and all my flatemates in Fichtenweg 3 (Tübingen). Life was fun and research is never a hard job when one gets friends like them. Thanks to all of you. I am also grateful to Mr. Hakim Arif (Dhaka University) for being such a good friend of mine and for all the linguistic discussions we had during our stay in Berlin.

My greatest debt, of course, is to my parents and my only sister, who are amazing, supportive, and wonderful, who I can never repay and must settle for emulating. Finally, I thank the most special person in my life, my wife, Devitoma, who always encourages me to reach my goal keeping all the problems aside and helps me to share my feelings with her.

TABLE OF CONTENTS

Abstract

In the first part of this thesis, an extensive corpus study is used in order to determine the frequency of occurrences of different consonant clusters. It turns out that the clusters are best described with the help of a division of the Bangla lexicon into three strata: Sanskrit borrowings (SB), Native Bangla words (NB), and the other borrowings (OB). The NB stratum does not allow complex onsets while SB and OB allows complex onsets. Further, SB and NB do not allow complex codas, while these are found in OB. A range of other restrictions are discussed. Special attention is paid to voicing and to aspiration. An agreement analysis of voicing at the word medial position is argued for. A positional faithfulness analysis is presented for syllable final deaspiration. The analysis is presented in Optimality Theory (OT), following the stratification of the lexicon by Ito and Mester (in Japanese).

In the second part of the thesis, a morphological analysis in the Distributed Morphology framework (Halle and Marantz) is provided for standard verbal inflectional paradigms of Bangla. The inflectional categories that are covered by the analysis are ten categories of tense/mood (perfect, conditional etc), three levels of politeness (Formal, Polite and Intimate) in three persons. The Bangla analysis is compared with a similar analysis of the much simpler case of English verbal inflectional morphology. The analysis of Bangla in this second part compares a consonant-final stem with a vowel-final stem for all forms. Differences between the two Bangla cases show the existence of a number of phonologically motivated changes, some of which also relate to syllable structure (diphthongization,

gemination). These are analyzed in Optimality Theory in the third part of the thesis, extending the analysis of the first part. The forms are thus accounted for by the morphological analysis in DM and following phonological changes on them, analyzed in OT.

1 Introduction

1.1 Introduction to this thesis

The syllable structure in Bangla has a unique space of its own in the overall research activities seen in this language. Hence, even in Bangla, it is a well cultivated area in the phonological and morphological research activities. But, most of these works are typically based on segmental and metrical approach. With respect to the relevant developments in phonological works, segments and their features are the micro units of phonology. This could be compared to the molecules and atoms of mass in the physical sciences. Among these two concepts, macro unit scheme of phonology begins at the syllable level and goes up to the higher level of foot and ultimately to the lexical category of word. In Bangla, however, a small amount of the research is done in the macro level. The primary goal of this thesis is an attempt to offer a constraint-based account of the syllable structure (of Bangla) which is considered as the first and foremost macro unit of phonology and then a morphological analysis of certain syllable structures available in verbal paradigms in Bangla. It will be a twofold study consisting a phonological structuring and then a general application on a morphological paradigm. First, we examine different aspects of the syllable structure of Bangla in a stratified model of its lexicon. Various constraints would be deployed to figure out the overall structure of the syllable. This part would be done in the framework of optimality theory (McCarthy & Prince 1993a, 1993b; Prince & Smolensky 1993). Afterwards, there will be an

attempt to construct a verbal inflectional paradigm in Bangla within the framework of the distributed morphology (Halle & Marantz 1993). One could point out that both the theoretical frameworks (phonological and morphological) used in this study are very much contemporary and developed in the early 90's. Among them, optimality theory (OT) is probably more cultivated compared to distributed morphology.

This thesis is organized as follows: in the remainder of this introduction, I will sketch quickly the basic theoretical assumptions of optimality theory focusing on those most relevant to my analysis. Then I will tell about a corpus used in this work, the transcription convention and the language (Bangla) we are going to deal with. Then some basic concepts of Bangla phonology, such as vowels, semivowels, consonants etc followed by an account of syllable structure in Bangla will be discussed. In chapter 2, I will illustrate the basic stratification of Bangla vocabulary and the constraint rankings as per stratum. Chapter 3 deals with the OT analysis of word medial clusters. Then, gemination cases are discussed in the following chapter. Lastly, chapter 4 discusses the word peripheral consonant clusters. In the next level, chapter 5 shows a set of possible verbal inflections of two roots in Bangla in the DM framework (given in tabular form). A section of this chapter shows possible applications of some of the inflected forms in a syntactic tree structure of Bangla. In the following section, first I will illustrate syntactic analysis of some English inflections; then gradually all the inflections of Bangla.

1.2 Introduction to Optimality Theory

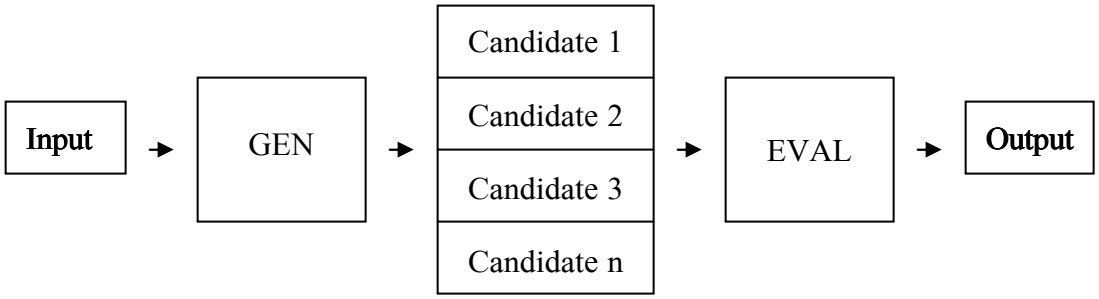
The analysis of syllable structure has received considerable attention in the phonological literature by opening numerous avenues in the sphere of linguistic research. Among different approaches applied in this endeavor, optimality theory is one of the most important and powerful methodology in the recent years after its introduction in 1993 (McCarthy & Prince 1993a, 1993b; Prince & Smolensky 1993). The constraint based approach of this theory has a very strong impact on the linguistic studies in general.

In recent years, there has been a shift in focus in much of the studies on phonological theory, from rule-based system to sets of constraints on well-formedness principles making way to the formation of optimality theory (Bird & Klein 1994; Burzio 1994; Frank & Satta 1997; McCarthy 2001; McCarthy & Prince 1993a, 1993b; Paradis 1988; Prince & Smolensky 1997; Scobbie 1991, *inter alia*). This theory was developed as a response to a "conceptual crisis at the center of phonological thought" (Prince & Smolensky 1993) concerning the role of output constraints. It was also (partly) inspired by the concepts of neural networks, as we can see the ideas of optimization, parallel evaluation, competition, and soft, conflicting constraints are familiar in this framework. Optimality theory (henceforth, OT) is widely adopted by scholars not only in the area of phonology, where OT was initially developed and applied, but also in other areas of linguistic studies, such as in syntax and semantics (*see* Legendre, Grimshaw & Vikner 2001).

In a compact introduction, phonological (rather, grammatical) constraints are ranked and violable by the phonetic forms of their underlying representations in the OT structure. These constraints are minimally violated by a set of candidates (potential surface forms) and the one which incurs the least serious violations wins. The seriousness of a violation is defined in terms of hierarchies of constraints; the violations of higher-ranked constraints are most serious.

More elaborately, OT works in a constraint-based competition system among a possibly infinite set of candidates (at least two). In classical representation, the generation of utterances in the optimality theory involves two very important functions, viz., GEN and EVAL. From an input, GEN returns a set of unique output candidates. Among these candidates at least one could be identical to the input and the rest are somewhat modified in their structure. Then, EVAL functions to choose the optimal candidate that best satisfies a set of specially ranked constraints depending on the violation. That means, in OT the constraints are violable. The ranking process of the constraints is very crucial here, because it is the most important criterion that chooses the optimal candidate as output. EVAL chooses the out from a set of candidate starting from two to an infinite number (n). The figure in (1) bellow illustrates the process to reach an output from the input through the function of GEN and EVAL (Davenport & Hannahs 2005). Note that the following graphic representation is indicative only. The actual output tableaux differ from this representation.

(1) Graphic representation of OT



There are two types of constraints which act as EVAL: markedness and well-formedness constraints. Markedness constraints enforce well-formedness of the output candidate, prohibiting structures that are difficult to produce or comprehend, such as consonant clusters or phrases without overt heads (Arbib 2002; Kager 1999). These constraints usually prohibit some phenomenon or impose restrictions on the occurrence of certain segments. Examples of such markedness constraint are:

(2) Examples of markedness constraints

a. Syllables must not have codas (NoCoda)
b. Syllables must have onsets (Onset)
c. Obstruents at coda position must not be voiced (*VdObs$_{(Coda)}$)
d. Obstruents must be voiced (*VdObs)

On the other hand, faithfulness constraints enforce similarity between input and output. For instance, all morphosyntactic features in the input to be overtly realized in the output. Kager 1999 lists some typical instances of faithfulness constraints that are available in most languages.

(3) Examples of faithfulness constraint

a. The output must present all segments present in the input (DEP-IO)
b. Elements adjacent in the input must be adjacent in the output (CONTIGUITY)
c. Input segments must have counterparts in the output (MAX-IO)
d. The specification for place of articulation of an input segment must be preserved in its output correspondent (IDENT-IO(PLACE))

These constraints, both markedness and faithfulness, are ranked in a language-specific order. In an analysis, different markedness and faithfulness constraints usually do conflict, so the ranking of the constraints decides the right candidate as the output depending on the violation (of constraints) pattern. This ranking of the constraints is not a strict universal ranking; rather it differs from language to language. In other words, different languages have their own constraint ranking which applies to that particular language only. But, for every language, the constraint ranking is very strict. That means a candidate violating a high-ranking constraint can never be a winner by satisfying lower-ranked constraints. Here, the other important issue is the violability of constraints. Violability ensures that the optimal candidate is not required to satisfy all constraints. It may violate a constraint and still win as the optimal candidate, if it satisfies the top-ranked candidate(s). Another way of describing EVAL is that a candidate x is optimal if and only if, for any constraint that prefers another candidate y to x, there is a higher-ranked constraint that prefers x to y (*see* Zuraw 2003). So, the main point in this section is that OT allows the specification of a ranking among the constraints and allows

lower ranked constraints to be violated in order for higher ranked constraints to be satisfied.

The constraint ranking and their interaction among the input candidates is typically showed in tableaux in the OT analysis. In such a tableau, the candidates are listed vertically while the constraints are ranked in the horizontal line. For a hypothetical language X, let us assume that constraint 1 (con_1) is satisfied by candidate 1 and on the other hand, candidate 2 satisfies constraint 2 (con_2). In the following tableau, it is shown how the optimal candidate is being chosen by EVAL through the constraint interaction.

(4) Constraint ranking for language X

input	con_1	con_2
☞ 1. cand 1		*
2. cand 2	*!	

In (4), con_1 is ranked above con_2. Con_1 is violated by the candidate 2. This is a fatal violation for this candidate, because of the high ranking of con_1. The fatal violation is marked by an exclamation (!) symbol. As a result, candidate 1 becomes the optimal candidate, even after violating the low ranked con_2. A hand symbol is used in OT to indicate the optimal candidate in a tableau (see the above tableau).

Optimality theory examines several restrictions available in the phonological processes in a language and relevant constraints are formed to account for those restrictions. Syllable structure is one of the prominent topics in the research

activities in the OT framework. In recent time, Féry & van de Vijver 2003 presented a collection of studies in this topic that opens up new ways of further research on several issues of syllable structure. But, so far OT is comparatively a less preferred methodology among researchers who work in Bangla phonology. There are only a few works (Das 2002, Vijayakrishnan 2003, Kothari 2004 *inter alia*) done in different phonological studies in Bangla using this theoretical framework. So, there is always a deficit of literature for anyone who would study this issue in Bangla. But, on the other hand, OT itself is a very innovative methodology and it's always a challenging matter to deal with. In this thesis, I try to give an account for how consonant clusters work in different positions in a syllable and word. This account is furnished with the corpus study (Bangla) results and then, applied to the morphological analysis of the verbal paradigms in a later section of this study.

1.3 The Focus Language (Bangla)

Bangla (also known as Bengali) is typologically an agglutinative language mainly spoken in the Indian sub-continent. It is the national language of Bangladesh (144 million speakers – 98% of total population – ranked first)[1] and official (and regional official) language of the states of West Bengal, Tripura and Assam of the Republic of India (80million speakers – 8.3% of the total population – ranked second)[2]. Bangla is spoken in many other states of India and a significant number of

[1] Bangladesh Census 2001

[2] India Census 2001 and 1991

populations are in the USA, UK, Singapore, Nepal and several other countries (Gordon 2005). With more than 224 million speakers, Bangla secures the 6th position among the world's languages in terms of together first and second language speakers (Comrie 2005, Katzner 2002, Weber 1997).

There are marked dialectal differences between the spoken variety of the Bengalis living on the western side and the eastern side of the Padma River. During standardization of Bangla in the late 19th and early 20th century, the cultural elite were mostly from West Bengal, especially Kolkata (formerly Calcutta). Hence, the dialect of that area was considered the standard. However, at present, the accepted standard language in West Bengal and (all parts of) Bangladesh are identical, i.e., the West Bengal variety. However, there is another important division. In Bangla, there exists what is known as *Shadhu bhasha* (the elegant language; literally "language of sages"; also called *Shuddho bhasha*) and *Cholti bhasha* (the current, or colloquial, language; literally "the current or running language"; also called *Cholito bhasha* or *Cholit bhasha* in common speech). The major differences between these two are the adherence to traditional grammar (i.e. the archaic forms of Medieval Bengali) and to a heavily Sanskritized vocabulary in Shadhu bhasha. However, Shadhu bhasha is not spoken in commonplace settings but confined to some literary and formal contexts. Here we discuss mainly the standard colloquial Bangla (SCB), i.e., *cholit* bhasa and hereafter the term "Bangla" will denote only the SCB. On the other hand, old literary form, i.e., *sadhu* bhasa will be mentioned as OLB (Old Literary Bangla).

1.4 Transcription Convention

Several different transcriptions are used in the linguistic study of Bangla. I decided to follow the standard IPA convention of transcription. However, many recent works follow another transcription convention developed by Ray, Hai, & Ray 1966. Since, many works of my references use the Ray et al. system; I will give a few examples where the symbols differ from the standard IPA system.

(5) Transcription comparison (IPA and Ray *et al.* 1966)

Characters	**IPA**	**Ray *et al.***
Front mid-high vowel	e	e
Front mid-low vowel	æ	E
Back mid-high vowel	o	o
Back mid-low vowel	ɔ	O
Front high semivowel	e̯	y
Front mid-high semivowel	i̯	Y
Back high semivowel	u̯	w
Back mid-high semivowel	o̯	W
Voiceless unaspirated retroflex stop	ʈ	T
voiced unaspirated retroflex stop	ɖ	D
Retroflex flap	ɽ	R
Paleto-alveolar sibilant	ʃ	S
Velar nasal	ŋ	N
Nasalizes the nucleus (mainly vowel)	-	M

The remaining symbols agree with the IPA conventions; hence they are not listed here again. One can see section 2 below for references.

1.5 The Corpus

A Bangla corpus is used in this work in order to obtain certain data, mostly the frequency of occurrences of specific sound or sound combinations. This corpus was developed under the TDIL program of the erstwhile DoE (Dept of Electronics, Govt. of India, *now* Ministry of Communication and Information Technology). It was developed under the supervision of Prof. D.P. Pattanayak. Subsequently, the corpora were passed on to the Central Institute of Indian Languages (CIIL), Mysore and have been in their custody since then. The original corpus contains approximately a total of 1.6 million words (with repetition of same entries) and is compiled in Bangla script. This is available in Indian Standard Code for Information Interchange or Indian Script Code for Information Interchange (ISCII) format. However, a Romanized version of the same is used here, with certain modifications. This is a thoroughly cleaned Romanized version of the above said corpus. It was produced under the supervision of Prof. Gautam Sengupta (University of Hyderabad) with assistance from Dr. Soma Paul of the International Institute of Information Technology (IIIT, Hyderabad).

For the present study, only the distinct entries are extracted from the whole corpus using a small PERL program and approx. 57,000 entries were taken in a single file. From now onwards, we will refer this file as corpus. Since the original corpus was created using sampling methods and an orthography-based collection, I had to apply certain modifications to that compilation in order to make it useful for phonological studies, too. Still there are some cases where the orthography and

phonemic mapping could not be done. For instance, Bangla অ [ɔ] is frequently pronounced as ও [o] (in many places). But it is not completely done in this modified version that I am going to use in this work. So the frequency of occurrences of these two is shown together. A similar method applies to [e] and [æ] (as in এক /ek/ [æk] 'one') which often interchange their places.

1.6 Bangla Sound Alphabet: Basic Structure

The segmental phones of Bangla are slightly different from those of other Indo-European languages. According to Ray *et al.* 1966, in Bangla there are seven oral vowels, four semi-vowels and thirty consonants. The consonants are subdivided into twenty plosives (ten unaspirated and ten aspirated), three nasals, three liquids and four fricatives.

1.6.1 Vowels

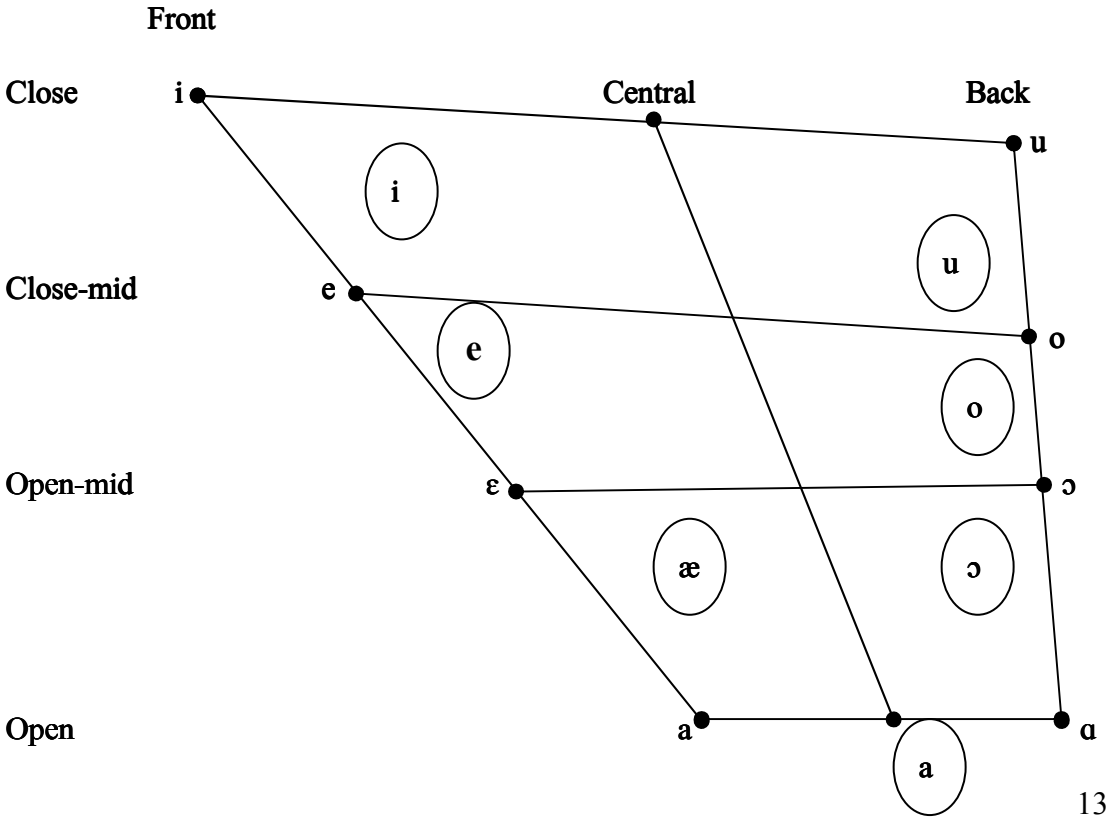
Bangla vowels systems basically follow the structure of cardinal vowels. All seven vowels are opposed to each other according to the height of the tongue (openness) and the articulatory area. The roundness of the lips is another point to be considered.

The vowels in Bangla are generally articulated a little lower in the oral cavity than those of cardinal vowels. The main difference between the two systems is the nasalized counterparts of the vowels in Bengali. Cardinal vowel system does

not represent nasalized vowels, but Bangla does. A comparison between Bangla vowels and the IPA transcription convention shows that the later system has rounded front vowels ([y], [ʏ], [ø] and [œ]) and two middle vowels (the schwa [ə] and the vocalized [ɐ]) that are absent in Bangla (see Appendix – I).

In principle, all the vowels in Bangla can be nasalized. This is a characteristic of Bangla and some other Indian languages. However, each of them may be nasalized only in a word initial syllable (except in some very rare cases) which is not a standard feature of many languages. The following figure illustrates the location of Bangla vowels in the cardinal vowel structure with reference to the IPA symbols. Bangla vowels are shown in circles while Cardinal vowels are bold-lettered. Remaining symbols represent the IPA convention.

(6) Cardinal vowels (position of Bangla)



Bearing in mind the observation by Ray *et al.* 1966 that

> "Each of the above (the Bangla vowels) may, but only in word initial syllable, be nasalized by simultaneous opening of the passage to the nose in the back of the mouth. There is no suggestion of a nasal consonant or semivowel following the nasalized vowel. Also, no nasalized vowel is either contrastive or obligatory or even frequent before or after a nasal consonant."

in this work I shall consider to put a nasalized counterpart for all vowels of Bangla following the IPA convention of transcription. The simple table below shows the Bangla vowels with their nasalized counterparts:

(7) Bengali Vowels with corresponding nasals

	Front	Central	Back
	Unrounded		*Rounded*
Close	**i ĩ**		**u ũ**
Close-mid	**e ẽ**		**o õ**
Open-mid	**æ æ̃**		**ɔ ɔ̃**
Open		**a ã**	

Now, let us consider some examples of pure and nasalized vowels of Bangla. A comparative set of such cases are listed in the following illustration.

(8) Pure and nasalized vowels in Bangla.

1. i -- ĩ ihar (of this) -- ĩhar (of him/her-polite)
2. u -- ũ uki (a hiccup) -- ũki (a peep)
3. e -- ẽ eke (one person-informal) -- ẽke (one person-formal)
4. o -- õ ora (they-informal) -- õra (they-formal)
5. æ -- æ̃ æk (one) -- hæ̃ (yes)
6. ɔ -- ɔ̃ ɔkejo (useless) -- ɔ̃k (a nasal sound)
7. a -- ã baʃi (stale) -- bãʃi (flute)

Here, most of the pairs are minimal pairs (except æ -- æ̃ and ɔ -- ɔ̃). Among all the vowels in Bangla [a] (and [ã] from nasal ones) is the most frequent vowel. This set of pure and nasalized vowels provide the maximum number of minimal pairs in Bangla.

Very prominent nasalized vowels are also found in other major human languages in the world. For instance, French has nasal vowels like Bangla. But, compared to seven nasal vowels in Bangla, French has only four, viz., [ɛ̃], [ɑ̃], [ɔ̃] and [æ̃] (Ghosh 2003).

Bykova 1981 has pointed out two important characteristics of the nasalization of Bangla vowels, namely, the frequency of occurrences of the nasalized vowels and the dependency of nasality.

A. The frequency of the usage of nasalized vowels in Bangla is much lower than that of the non-nasalized ones (i.e., the pure vowels). A calculation by

Ferguson & Chowdhury 1960 (pp. 50-51) states that the total percentage of oral vowels is 44.72 while that of nasal vowels is hardly .87, a ratio of about 50:1. These figures are drawn on the basis of transcribed texts running to 10,130 phonemes and are counts of the relative frequency of individual phonemes in a running text. A similar statistics was found here when I checked for occurrences of these sounds in our present corpus.

(9) Frequency of occurrences of nasalized vowels

	i	u	e + æ	o + ɔ	a
Non-nasal (pure)	28199	12914	23511	36727 *(8730+35857)*	33768
Nasal	67	209	125	141	824

Although the grapheme to phoneme mapping is not so strong in this corpus, we get a fairly strong support of the data provided by Ferguson & Chowdhury 1960. We see the most frequent nasal vowel in Bangla is [ã] and the least frequent one is [ɔ̃]. These observations are almost identical to those found in the data presented by Bykova 1981 where [ã] and [õ] are reported as the most frequent ones.

B. The other aspect of Bangla nasal vowels is the dependency factor. It is essential to distinguish between dependent or positional and independent or phonological vowels in Bangla (Hai 1975). Dependent or positional nasality of a vowel is caused by the influence of the preceding or (partly) following nasal consonant (of neighboring syllable). This is just an echo of the other nasal sound

and not a phonological element of real vowel nasality.

On the other hand, we have pure nasal vowels which occur irrespective of their environment. Hai 1975 termed this type of events as phonological nasality. All the nasalized vowels shown in table 3 belong to this category. They can occur independently even in a single syllable.

1.6.2 Semi-vowels / Glides

By definition, a semi-vowel or glide is a sound that has the quality of one of the high vowels and that functions as a consonant before or after vowels. According to Ray *et al.* 1966 Bangla distinguishes four semi-vowels, viz., [i̯], [u̯], [e̯] and [o̯], which may be described as non-syllabic varieties of [i u e o] respectively. For instances, let us consider the following instances:

(10) Usage of semi-vowels in Bangla

1. খাই [k^hai̯] -- I eat
2. লাউ [lau̯] -- the bottle-ground
3. খায় [k^hae̯] – (he) eats
4. খাও [k^hao̯] -- you eat

The observation by Ferguson & Chowdhury 1960 says the three low vowels [æ a ɔ] have no semivowel form of their own. And their occurrence is restricted solely as pure vowels and never with another vowel. On the other hand, [i u e o]

occur post- and inter-vocalically. A detailed study confirms that [u o] occur infrequently in prevocalic position when preceded by a consonant or juncture. Although [æ a ɔ] do not occur post or inter-vocalically, they may sometimes occur with semivowels or glides.

The positions of the semi-vowels are very close to their vowel counterparts. The following table (11) shows the locations of Bangla semi-vowels in the following vowel chart.

(11) Semi-vowels in Bangla

	Front	Central	Back
	Unrounded		*Rounded*
Close	i ĩ i̯		u ũ u̯
Close-mid	e ẽ e̯		o õ o̯
Open-mid	æ		ɔ
Open		a	

According to Chatterjee 1962 Bangla semivowels are non-syllabic, and predictably non-phonemic. He argues that the all the four distinct vowel phonemes /i u e o/ have non-syllabic allophones in complementary distribution with the syllabic ones. But, Ray *et al.* 1966 gives a different account on the syllabicity of allophones. According to this later work Bangla semivowels are non-syllabic, but phonemic and the non-syllabic phoneme /y w Y W/ (in our work, /i̯ u̯ e̯ o̯/) can be distinguished from their syllabic counterparts /i u e o/. Although most of the linguists have

accepted the view of Ray *et al.* 1966 on this issue, this controversy in the linguistic study of Bangla has still not ceased.

1.6.3 Diphthongs

The most important characteristic of a diphthong claimed to be the "… must necessarily consist of one semivowel" (Jones 1964). Based on this principle Sarkar 1985 establishes a set of 17 diphthongs in Bangla. They are:

i̯i̯ i̯u̯ u̯i̯ ei̯ eu̯ oi̯ ou̯ oe̯ oo̯ æe̯ æo̯
ai̯ au̯ ae̯ ao̯

But this list is not exhaustive. Hai 1975 describes a total of 31diphthongs while Chatterji 1986 gives an account of 25 diphthongs in Bangla. So, it is not clear what the exact number of diphthongs in Bangla is. However, in the study by Chatterji 1986 the number of diphthongs are counted solely based on vowels, and not on vowel-semivowel (glide) combination like Sarkar 1985. And Hai 1975 has taken both vowels and semivowels (glides) into account at the second place of diphthongs.

Actually, Bangla has only two diphthongal letters in the alphabet. They are ঐ [oi] and ঔ [ou]. But the basis of counting the number of diphthongs is phonetic and which results in numbers ranging from 17 to 31. And as a matter of fact, all diphthongs in Bangla are falling in nature, which means, they start with a vowel of relatively higher sonority and end in a vowel with relatively lower sonority.

Lists of diphthongs from all three works are given here for a comparative study. The first is according to the system of Sarkar 1985.

(12) Bangla diphthongs (Sarkar 1985)

Vowels \ Glides	i̯	u̯	e̯	o̯
i	ii̯	iu̯		
u	ui̯			
e	ei̯	eu̯		
o	oi̯	ou̯	oe̯	oo̯
æ			ɔe̯	ɔo̯
ɔ				
a	ai̯	au̯	ae̯	ao̯

Note that all the second members of the diphthongs are semivowels (glides) in this table. But the other two linguists do not accept the idea that the second member is compulsorily a semivowel.

But the table in (12) represents a fairly acceptable set of diphthongs where the second member is a falling vocoid (semivowels or glides). In every case, the second member is weaker than the first one. This is considered as the most recent and widely acceptable inventory on the diphthong issues in Bangla. But, Chatterji

1986 gives another set of diphthongs in Bangla that differs significantly from Sarkar 1985.

(13) Bangla Diphthongs (Chatterji 1986)

2nd member / 1st member	**i**	**u**	**e**	**o**	**æ**	**ɔ**	**a**
i		iu	ie	io			ia
u	ui		ue	uo			ua
e	ei	eu		eo			ea
o	oi	ou	oe				oa
æ			æe	æo			
ɔ				ɔo			ɔa
a	ai	au	ae	ao			

Here, we see a group of 25 diphthongs where no entry contains any semivowel (glide). Only the set of 7 regular vowels are used in both the X and Y-hands of the table to generate the diphthongs. However, this table is not exhaustive. On the other hand, Hai 1975 argued for a total of 31 diphthongs among which 19 are regular and 12 are irregular.

(14) Bangla diphthongs (Ray et al. 1966)

Regular Diphthongs	Irregular Diphthongs
ii, iu, ei, eo, eu, æo, æe̯, ai, ao, au, ae̯, ɔo, ɔe̯, oo, ou, oi, oe̯, ui, uu	ia, ie, io, ea, eo, æa, ɔa, oa, oe, ue, ua, uo

It is clear that Hai includes pure vowels as well as semivowels in his list of diphthongs. In his work, diphthong [e̯] is used as the second member in 4 regular diphthongs viz., [æe̯ ae̯ ɔe̯ oe̯]. The other 27 diphthongs consist of only vowels.

It is evident from the preceding tables and discussion that Bangla doesn't have a uniform and standardized set of diphthongs. None of these approaches is considered beyond any controversy. Rather, they co-exist and are used by the scholars depending on the intention of the very purpose of specific studies. In the present work, we will follow the inventory of diphthongs presented by Sarkar 1985.

1.6.4 Consonants

Bangla has a total of 30 consonants. Twenty of these (thirty) consonants are either plosives or affricates. This two-third portion forms a block of regular consonants in Bangla. They can be divided into the two categories of 'voiced' and 'voiceless' where each category can be further subdivided into unaspirated and aspirated categories.

In terms of the place of articulation, this block has five different classes. They are velar (k-class), Paleto-alveolar (c-class), Retroflex (ʈ-class), dental (t-class) and bilabial (p-class). Each class consists of four sounds: voiceless-unaspirated, voiceless-aspirated, voiced-unaspirated and voiced-aspirated. Other regular consonants, which can be found in most languages, are two fricatives [ʃ s] (considered as allophones), three nasals [ŋ m n], one lateral [l], three flaps [ɽ ʈ^{h} r] (among which the first two are uncommon) and one glottal [h].

Most of the consonants are very common in nature and can be represented in a table following the IPA standard. However, a few of them may need some elucidation for a better understanding. For instance, [ɽ] is a sound unfamiliar to non-native speakers of Bangla. According to Ray *et al.* 1966, this sound "is made by the underside of the tongue tip flapping down just once against the surface above the upper teeth ridge and not completely stopping the flow of breath down through the middle." This sound is more like /rd/ in American /hardy/. [ɽ] has a very limited frequency of occurrence compared to alveolar flap [r]. On the other hand, [ɽ^{h}] is the voiced counterpart of [ɽ]. This sound also has a very rare occurrence in Bangla. Let us have a look to the following table of frequency of occurrences of the three flaps in Bangla.

(15) Frequency of Flaps in Bangla in the corpus used here

r	ɽ	ɽ^{h}
29921	1947	37

A detailed illustration of the place and manner of articulation of the consonants can be obtained from Appendix I.

1.7 Syllables

In phonology, a syllable is a unit of pronunciation typically larger than a single sound and smaller than a word. We find the term 'syllable' in English being used

starting from the time of Chaucer and it is used in linguistic descriptions very frequently. But this is not a comprehensive definition of syllable. Actually, the notion of syllable is more intuitive than linguistic to native speakers. However, there are at least two levels of representation of the notion of syllable, viz., at phonetic and phonological level. We will focus on the phonological level here.

Mainly, the rules of syllabification in Bangla, as they are available in literature, will be illustrated in this section. Till now, Bangla syllabification has received an appropriate attention. According to Dan 1992, not only the phonetics study, but almost all the representative works in the field of Bangla phonology are developed on the basis of so called intuitive syllables. Hence, the core works on syllabification were neglected more or less. Two important works in this field, Chatterji 1986 and Kostić & Das 1972, dealt with the articulatory and acoustic aspects of Bangla phonemes. But, one can still find some amount of emphasis on the syllabification issues.

Later, Sarkar 1986 and Dan 1992 did some notable developments in the study of Bangla syllabification. A fairly candid recapitulation rather than a critical discussion of their work will be stated here.

1.7.1 Bangla Syllabification Rules

The most preferred syllable pattern in Bangla is CV, which is, according to Jakobson 1972, an unmarked syllabic pattern. Actually CV is the not only the most preferred one but the most natural syllabic pattern in Bangla which strongly

conforms the syllabic universals proposed by Jakobson 1972 and others.

Dan 1992 has listed 10 phonetic restrictions involved in syllabifications of Bangla. They are presented as rules in the following paragraphs here:

Rule-1: Except for vowels no other sounds in Bangla bear the feature [+syll][3] which denotes the ability to be the peak of a syllable.

Rule-2: In Bangla, any unitary V + V sequence, i.e., the true diphthongs, never consists of a first member which is higher than the second and interpretable as being [-Syll].

Rule-3: In Bangla only semivowels can occur as the second member of a diphthong.

Some scholars also include vowels as the second member of a diphthong. For example Hai 1975, Chatterji 1986 etc.

Rule-4: Word initial semivowels and glides are not frequent in Bangla and hence a word internal sequence of a diphthong followed by a vowel is often syllabified as diphthong plus syllable boundary plus vowel. For example, অয়ন [ɔ̯on] 'orbit' is syllabified as [ɔ̯-on]. Note that the preceding restriction is also a controversial one and considered as a problem area in Bangla phonology.

[3] [±Syll] stands for the presence or absence of the feature "Syllabic". Hence, [+Syll] designates a sound that is or can be the most sonorant segment of a syllable, as a vowel or a resonant.

Rule-5: The word initial clusters[4] become the onset of the following syllable, the cluster 'tr' is the onset to the peak i in the word তৃণ [tri-no] 'grass'.

Rule-6: Word-final CC sequences[5] form the coda of the preceding syllable; for example, the sequence 'rɖ' is the coda in the word হোমগার্ড [hom-garɖ] 'home-guard'.

Rule-7: Word medial intervocalic CC sequences may be either homorganic or heterorganic. However, all intervocalic CC sequences (regardless of whether they are homorganic or heterorganic), except those with r/l as the second member occurring in *tatsama* words[6], are heterosyllabic, the first member of the sequence being the coda of the syllable to the left and the last member being the onset of the following one.

For example, the sequences 'rb', 'ʃc', 'tt' and 'kk^h' are divided between two syllables in the words পূর্ব [pur-bo] 'east', পশ্চিম [poʃ-cim] 'west', উত্তর [ut-tor] 'north' and দক্ষিন [dok-k^{h}in] 'south'.

Rule-8[7]: In the case of a word medial CC sequence with r/l as the second member, the first member of the sequence is geminated which results in a CCr/l sequence.

[4] Words consisting of CC(C) clusters belong to the borrowed level of Bangla Vocabulary (Sarkar 1986)

[5] Only a few borrowed elements show final CC sequences. On the whole SCB phonology tends to avoid more than one consonant word-finally, even in the case of borrowed items (Sarkar 1986)

[6] Sanskrit words that are more or less unmodified in Bangla

[7] This rule of gemination applies only to the level of tatsama vocabulary as is mentioned in Rule-7. In case of typical Bangla words such a rule does not apply. For example, [ʃãt-ra] 'a surname', [ʃap-la] 'a type of water-weed'

Then the first member of the CCr/l sequence becomes the coda of the preceding peak, and the second and the third members, i.e., Cr/l, together form a cluster which becomes the onset of the following syllable.

For example, the sequences 'ttr' and 'mml' in the words পুত্র [put-tro] 'son', অম্লান [ɔm-mlan] 'untarnished' etc.

Rule-9: The word medial CCC sequences, as Mallik 1960 observes correctly, have 'r' as their third member[8]. In such sequences the syllable boundary is placed right after the first member. In other words, the first member becomes the coda of the preceding peak, and the second and the third members together form a cluster which becomes the onset of the following peak.

For example, he sequences 'str', 'ntr' and 'ʃpr' in the words অস্ত্র [ɔs-tro] 'weapon', মন্ত্রী [mon-tri] 'minister', নিস্প্রাণ [niʃ-pran] 'lifeless' etc.

Rule-10: According to Mallik 1960, there is only one word medial CCCC sequence, viz., 'ŋʃkr'. The syllable boundary is placed between the second and the third member and the sequence, that is, the first and second member form a closer 'combination than sequence', which becomes the coda to the preceding peak, and the third and fourth member form a cluster which becomes the onset to the following peak. For example, সংস্কৃতি /sangskriti/ [ʃɔŋʃ-kriti] 'culture'.

---

[8] The only exception to his observation is the sequence 'ŋʃk' to which the application of Rule-9 produces quite counterintuitive results. As for example, [ʃɔŋ-ʃkar] 'renovation' is less acceptable than [ʃɔŋʃ-kar].

Based on similar observations, Sarkar 1986 established 16 canonical syllable patterns in Bangla, arranged in descending order of frequency here : **CV, CVC, V, VC, VV, CVV, CCV, CCVC, CVVC, CCVV, CCVVC, CVCC, CCCV, CCCVC, VVC** and **CCCVV.** Out of these sixteen syllable patterns in Bangla, CCCV, CCCVC and CCCVV are not very frequent in nature. According to Dan 1992, "among these 16 patterns CV has the maximum number of exponents, approximately 54% in the whole language."

1.7.2 Representation of the Syllables

The representation of syllabic structures is a matter of continuing debate. Kaye & Lowenstamm 1981a and Prince 1984 assume the syllabic terminals to be pure positional elements and without any possible intrinsic content. Later Hyman 1986 and McCarthy & Prince 1985 proposed the concept of weight units or moras in the skeleton. The concepts of Onset and Rime were recognized by some linguists (Selkirk 1978, Steriade 1982 etc.). Many other distinctions like Margin, Nucleus and Coda were introduced by other scholars (Cairns & Feinstein 1982, etc). Among these perhaps the most useful element for this study is the mora or syllable weight.

In moraic theory (Hyman 1986, McCarthy & Prince 1985, Hayes 1989, Hayes 1995; Hyman 1986 etc), one can notice a prosodic hierarchy as commonly seen in syntactic or morphological structure. The following diagram shows a seven-level prosodic hierarchy of an intonational phrase. But in regular moraic tree structures, we use only the lowest three or four levels.

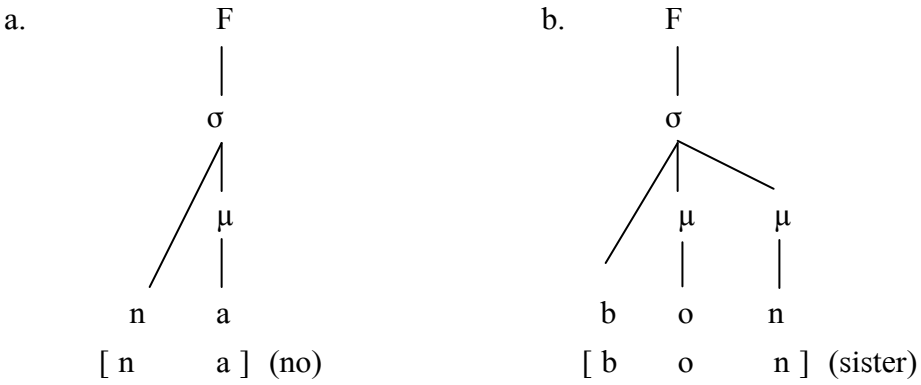
(16) Phosodic hierarchy

Morphological/syntactic hierarchy		Prosodic hierarchy
Sentence		Intonational Phrase
|		|
Phrase		Phonological Phrase
|	⇔	|
Word	mapping	Accentual Phrase
|		|
Morpheme		Prosodic (phonological) Word
(root, affix)		|
		Foot (F)
		|
		Syllable (σ)
		|
		Mora (μ)
		|
		(Segment)

In an influential syllable theory by Itô 1986, a syllable template based on a universal principle of wellformedness is present in all the languages of the world. The preceding prosodic hierarchy shows the elements of this template. Good examples of Moraic structure can be found in the short poems in Japanese called haiku. In general, Japanese is a language famous for its moraic qualities. Most dialects including the standard Japanese use moras as the basis of the sound system rather than syllables. Most linguists believe that haiku in modern Japanese do not follow the pattern 5 syllables/7 syllables/5 syllables, as commonly believed, but rather the pattern 5 morae/7 morae/5 moras. As one example, the Japanese syllable-final n is moraic. According to moraic theory, basic structures of Bangla prosodic

words would be as follows.

(16) Moraic structure of Bangla prosodic word



By definition, (16a) is light in nature (monomoraic) and (16b) is heavy in nature (bimoraic). In any case, onsets do not have any mora. Only the nucleus always has a mora and in Bangla, coda also gets a moraic node. In the following sections we will see more cases analyzed in the moraic structures.

In the next chapter, the basic phonotactic rules of Bangla syllable structure will be introduced in the OT format. Before that, a stratified lexicon of Bangla vocabulary will also be discussed in short.

2 Word initial and word-final sounds and clusters

2.1 Lexical Stratification

The lexicon (vocabulary) of a language consists of the lexemes used to actualize words in that language. These lexemes are formed according to the prominent morphosyntactic rules of every language. According to Aitchison 2003, in a language, the lexicon organizes the mental vocabulary in a speaker's mind by certain principles and then producing more complex words as per certain lexical rules of that language. Bangla has a varied kind of lexicon which contains different types of words. A multi-layer stratification of Bangla lexicon (hereafter, B-lexicon) is possible depending on the origin that the lexical elements.

In general, the whole B-lexicon can be stratified in three major groups, in terms of the origin of the words.

I. **Tadbhaba**

Native Bangla words, rooted in Sanskrit and Prakrit. This will be abbreviated as **NB** (Native Bangla).

E.g., কাঠ /kāṭh/ [kaʈ^{h}] 'wood', ফুল /phul/ [p^{h}ul] 'flower' etc

II. Tatsama

Words directly borrowed from Sanskrit. This will be abbreviated as **SB** (Sanskrit Borrowing).

E.g., গ্রাম /grām/ [gram] 'village', কবি /kabi/ [kobi] 'poet' etc.

III. Deshi o Bideshi

Words borrowed from Indian and foreign languages. This will be abbreviated as **OB** (Other borrowings).

E.g., আনারস /ānāras/ [anarɔʃ] ‘ananas’ (<Portuguese), বুর্জোয়া /burjoyā/ [burɟoa̯] ‘bourgeois’ (<French), বাদুড় /bāduṛ/ [baduɽ] 'bat' (<Austro-Asiatic), হরতাল /haratāl/ [hɔrotal] 'strike' (<Gujarati: [hɔʈtal]) etc.

According to Chatterji 1926a, SB consists of approx. 51.5% of the B-lexicon, when OB fills nearly 4.5%. The rest of the vocabulary, i.e., 44% belongs to the NB category. This is quite an old statistics, though. A Wikipedia[9] database suggests that the SB and NB strata consist of 67% and 28% of the whole lexicon, respectively. The rest (5%) belongs to the OB stratum. However, these figures are not presenting the actual usage of the lexicon, since a lot of the present vocabulary elements are archaic or technical in nature. A recent study on the various strata of Bangla vocabulary usage records a mere 25% share of the NB category words where as SB consists of 67% as registered earlier (Bhowmik 2003a, 2003b). Because of the highly receptive nature of Bangla vocabulary, it takes several words

[9] http://en.wikipedia.org/wiki/Bengali_vocabulary

from other languages in a constant process. It is rational that over time a lot of loan words are added in the Bangla vocabulary. It is also supported in the figures presented in the abode said works, where share of OB stratum is raised to 8% (3% up from 5% of the Wiki data).

Now, the SB stratum could be divided into two categories, viz.,

A) Samoccarito

Written and pronounced as in Sanskrit.

E.g., *San.* গ্রাম [gram] = *Ban.* গ্রাম /grām/ [gram] 'village'

San. সাধু [ʃad^{h}u] = *Ban.* সাধু /sadhu/ [ʃad^{h}u] ‘saint’ etc.

B) Asamoccarito

Written as Sanskrit, but pronounced differently.

E.g., *San.* লক্ষ্মী /lakṣmi/ [lɔksmi] = *Ban.* লক্ষ্মী / lakṣmi / [lok^{h}k^{h}i]

‘goddess of wealth in Hindu mythology’

San. পদ্ম /padma/ [pɔdma] = *Ban.* পদ্ম /padma/ [pɔddo] ‘lotus’ etc.

On the other hand, NB consists of only Bengali pure word forms and some restructured loan word from Sanskrit which came to the B-lexicon through a transitional stage of MIA (Prakrit).

The OB category can again be divided into two categories depending on the origin of the borrowed words, considering whether it is from another Indian language or from a foreign language.

A) Deshi

Borrowed words from Indian languages.

B) Bideshi

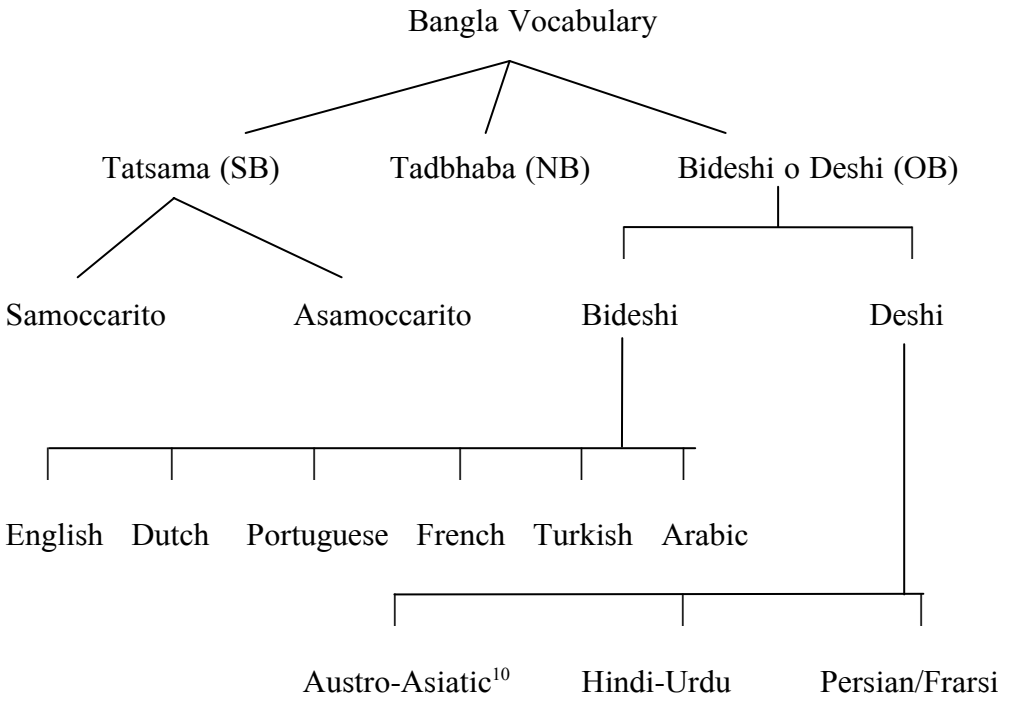
Borrowed words from foreign languages.

These two portions of the B-lexicon contain borrowed words from mainly Persian/Farsi (approx. 3000 words), English (approx. 1500 words), Dutch, Portuguese, French, Turkish, Austro-Asiatic, Hindi and Arabic (Chaki 2001; Chakrabarti 1994). According to Chaki 2001, there are also some minor borrowings from some other Indian and foreign languages, like Gujarati, Tamil, German, Tibetan, Chinese etc.

There are lots of foreign words in Bangla which have been modified from their original form. E.g., *Ban.* বাক্স /bāksa/ [bakʃo] 'box' < *Eng.* box [bŏks], *Ban.* ইস্কুল /iskul/ [iʃkul] < *Eng.* school [skūl] etc. But there are many of them which maintain their original structure and phonological qualities even when they are part of the Bangla vocabulary. *E.g., Ban.* প্যান্ট /pænṭ/ [pænʈ] 'pant' < *Eng.* pant [pănt], *Ban.* ফোন /phon/ [p^{h}on] 'phone' < *Eng.* phone [fōn] etc. We mainly refer this type element that fits to the OB strata in the B-lexicon.

We shall consider mainly the SB and NB words in this study. Other borrowed words (OB: *bideshi* and *deshi*) bear lesser importance; but, will not be totally excluded. The following illustration (13) shows the vocabulary structure of Bangla in terms of the origin of the lexicon.

(1) Bangla vocabulary structure



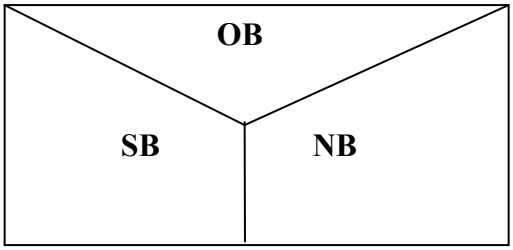
A large number of natural languages show an internal stratification based on native and loan words. The stratification could be done by taking the native words as the core of the lexicon. All the loan words could form peripheral strata around the core. This is called core-periphery structure. Itô & Mester 1995 introduced such stratification for Japanese. In this core-periphery structure of Japanese lexicon, there are three strata, viz., Sino-Japanese, Foreign and Unassimilated foreign words, over the core vocabulary consists of native Japanese words (Yamato). This type of

10 The Austro-Asiatic languages are a large language family of South-east Asia, and also scattered throughout India and Bangladesh. Among these languages, only Vietnamese, Khmer, and Mon have a long recorded history, and only Vietnamese and Khmer have official status (in Vietnam and Cambodia, respectively). The rest of the languages are spoken by minority groups.

stratification of lexicon is not entirely etymological in nature, rather some constrains can play an important role to determine certain sublexica of the whole lexicon.

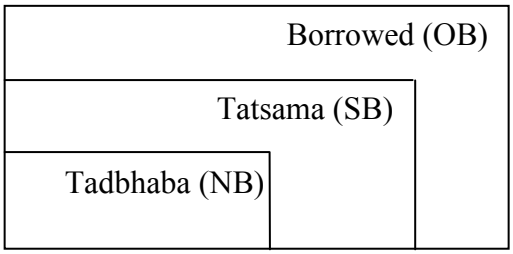
If we divide the B-lexicon in distinct sublexica, it could appear as the following illustration where the sublexica are seen as blocks in the whole lexicon.

(2) Lexicon and sublexica organization



Compared to the stratification illustrated in (2), the core-periphery organization is more like a layered stratification rather blocks of sublexica. In our case, following the core-periphery structure, the stratification of the B-lexicon can be illustrated as in (3).

(3) Core-periphery Structure of B-lexicon



The core-periphery organization is the perfect system of stratification of lexicon for languages like Bangla. We adopt it here in this study.

2.2 Onset

Onset and coda are two important parts of a syllable, besides the nucleus (peak). In this section we will discuss the onset in general and its position in the Bangla syllables. Many typological studies of syllable structures in the earlier century (For example, Greenberg 1978, Kaye & Lowenstamm 1981b, Itô 1986, Blevins 1995) have led to an implicational universal for syllable onset (as mentioned by Kager 1999) which basically says that even if a language does not have onset-less syllables, it still has syllables with onsets. That means, no language entirely exclude onset, though some of the syllable in that language may lack onsets. Onset-less syllables may occur in many natural languages. For example, Japanese, Ponapean, English etc. allow onset-less syllables (Kager 1999). But, universally, syllables prefer to have onsets. On the other hand, some languages, such as Axininca Campa, Arabic, Temier etc. do not allow syllables without any onset (Itô 1989).

In a restrictive definition, the well-formedness of syllable leads to the structures which can be surfaced without phonotactic modifications, such as assimilation, gemination, epenthesis, syllabification and so on. Those who fail to do so, and require certain modifications to coop up, are considered as ill-formed. Bangla has both kinds of structures. But, it likes to have the simple onsets more than the complex ones. In the following table we have some cases with simple onset. They are from both SB and NB categories: examples shown in a-c belong to SB and d-f to NB in (19).

(4) Simple onset (SB and NB)

Stratum	Orthography	Transcription	Meaning
SB	a. কবি	ko.bi	poet
	b. আকাশ	a.kaʃ	sky
	c. দান	dan	bestowal
NB	d. আজ	aj̈	today
	e. পুকুর	pu.kur	pond
	f. মাসি	ma.ʃi	aunt

The words are shown in syllabic structure. A dot (.) separates two syllables in a word. The syllables will be shown in moraic structures in the later discussions. In these examples, the onset contains only one consonant, either at the word-final position or in the word medial position. This is the most favorable onset structure (simple onset) in Bangla. We can refer to the faithfulness constraint described in Kager 1999 which requires that a syllable should begin with a consonant (onset).

(5) ONSET

Syllables must have onsets (Kager 1999)

Unlike some languages like Arabic or Axininca Campa, Bangla is not very strict about having an onset in every syllable. There are both types of syllables present in the B-lexicon: onset-less and with onset. For example, Bangla words like আম /ām/ [am] 'mango', ইলিশ /ilis/ [i.liʃ] 'Hilsa fish' etc. do not have onsets (in word

initial position). But most of the other syllables in Bangla have an onset. There is no restriction as per the strata of the B-lexicon in terms of the presence or absence of onset.

But, the constraint Onset does not say anything about the quantitative and qualitative features of the onset. Generally, onsets could be of two types: simple and complex where a simple onset contains only one consonantal sound and a complex onset consists of more than one consonant. Bangla has a mixed type of vocabulary where different types of words are present with different types of onset.

In a recent work, Yavaş 2003 claimed Bangla does not allow consonant clusters at the onset position. This observation is partially true. Although, the pure and reformed Bangla words (NB) do not allow consonant clusters at the onset and coda position, but, consonant clusters are present at the onset position quite frequently among the SB (Sanskrit) words. Some of the borrowed words (mainly of foreign origin) also contain complex onsets. We should not rule out these cases, because collectively they construct a fairly large portion (approx. 48-53%) of the Bangla vocabulary.

A corpus study (Bangla) suggests that the most possible cases of word initial consonant cluster in the onset position are present in the SB words. A small amount of such clusters are visible among the other borrowed (OB) words. And these borrowed words are mostly of English origin. Some clusters are common in OB and SB words.

The consonant clusters at the onset position, seen among SB words, are /kr/, /kl/, /gr/, /gl/, /tr/, /dr/, /d$^{\text{fi}}$r/, /nr/, /pr/, /pl/, /br/, /b$^{\text{fi}}$r/, /mr/ and /ml/. It is clear that all

these clusters have a distinct combination of whether obstruent and liquid or nasal and liquid sounds. The segmental structure of these occurrences is restricted by a markedness constraint that requires a rising and minimal sonority distance among the sounds of these clusters. The following table represents a set of cases with complex onsets found in the SB.

(6) Word initial clusters in Bangla SB words

Stratum	Type	Cluster	Orthography	Transcription	Meaning
SB	Obstruent + Liquid	kr	ক্রোধ	krod^{h}	anger
		kl	ক্লান্ত	klan.to	tired
		gr	গ্রাম	gram	village
		gl	গ্লানি	gla.ni	mortification
		tr	ত্রান	tran	rescue
		dr	দৃশ্য	driʃ.ʃo	scene
		d^{h}r	ধ্রুব	d^{h}ru.bo	absolute
		pr	প্রাণ	pran	life
		pl	প্লাবন	pla.bon	flood
		br	বৃষ্টি	briʃ.ʈi	rain
		b^{h}r	ভ্রু	b^{h}ru	eye-brow
	Nasal + Liquid	mr	মৃত	mri.to	dead
		ml	ম্লান	mlan	melancholy
		nr	নৃত্য	nrit.to	dance

On the other hand, there are many such clusters found in the OB stratum which represents the same combinations available in SB. For instance, clusters like /kr/, /kl/, /gr/, /gl/, /pr/, /pl/ and /br/ are found in both SB and OB strata. Other than this lot, there are some distinct consonant clusters which could only be available among words of English origin (OB) only. These are /ʈr/, /ɖr/, /t^hr/, /fr/, /fl/ and /bl/.

(7) Word initial clusters in OB

Stratum	Type	Cluster	Orthography	Transcription	Meaning
OB	Obstruent + Liquid	ʈr	ট্রাম	ʈram	tram
		ɖr	ড্রেন	ɖren	drain
		t^hr	থ্রো	t^hro	throw
		fr	ফ্রান্স	frans	france
		fl	ফ্ল্যাট	flæʈ	flat
		bl	ব্লাস্ট	blaʃʈ	blast

All the consonant clusters shown in (6) and (7) follow a particular sound pattern. The second member in these clusters is always of higher sonority than that of the first member. That means they follow the regular sonority sequencing principle (SSP) which will be discussed in the sections below. This principle requires the sonority in a syllable to be rising toward nucleus and falling afterwards (more on SSP in the next section). But, there are cases where this principle is not obeyed.

In languages like English and German, some words have coronal obstruents (fricative) at word initial and word-final consonant clusters. These clusters

systematically violate the sonority sequencing principle. Some of the English words starting with a consonant cluster containing a coronal obstruent [s] are also present in Bangla as borrowed words. Interestingly, Bangla allows these words to preserve their original sound structure, though in native Bangla words not only these patterns, but any sort of consonant clusters are prohibited. Actually, in some non-standard spoken forms of Bangla, a vowel is added before a word initial coronal obstruent. E.g., *Bang. (dia.)* ইস্কুল /iskul/ [iʃ.kul] 'school' < *Eng.* school [skūl], where the consonant cluster is broken into two syllables. In the standard form of Bangla, this borrowed word is represented as স্কুল /skul/ [ʃkul], where the coronal obstruent is part of the complex onset. The following table shows a list of clusters that has a coronal obstruent as the first member. These are borrowed in Bangla from English (OB).

(8) Word initial consonant clusters with coronal obstruents in OB

Stratum	Type	Cluster	Orthography	Transcription	Meaning
OB	Coronal Obstruent + Plosive	skr	স্ক্র্যাচ	skræc	scratch
		st̪	স্টেশন	st̪e.ʃan	station
		st̪r	স্ট্রীট	st̪rit̪	street

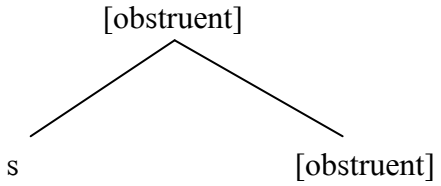
Not only in OB, but word initial appearance of coronal obstruents in consonant clusters is also found in Bangla SB words. A list of such cases is shown in the following table.

(9) Word initial consonant clusters with coronal obstruents in SB

Stratum	Type	Cluster	Orthography	Transcription	Meaning
SB	Coronal Obstruent + Plosive	sk	স্কন্ধ	skɔn.d^{h}o	shoulder
		sk^{h}	স্খলিত	sk^{h}o.li.to	aberrant
		st	স্তম্ভ	stɔm.b^{h}o	pillar
		st^{h}	স্থান	st^{h}an	place
		sp	স্পর্ধা	spɔr.d^{h}a	emulation
		sn	স্নান	snan	bath

The problem with this type of consonant clusters is that they violate the sonority sequencing principle which is not allowed in a systematic syllabic structure. In phonological studies, the word initial consonant clusters with /s-/ as the first element are traditionally accounted for as a special feature of the syllable structure. If the /s/ is followed by an obstruent, then it is claimed that the said s-obstruent cluster cannot be broken up by epenthesis or any other phonological process (Broselow 1999; Gouskova 2002).

(10) s-obstruent structure

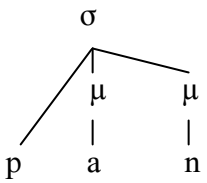


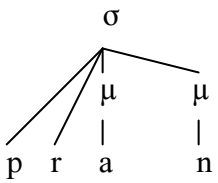
(Structure from Selkirk 1982)

It is widely claimed that all human languages adopt this representation about

s-obstruents. In a recent work, Kenstowicz 1994a proposed the concept of extrasyllabicity that comes after the 'appendix' concept of Fudge 1969 to deal with this type of consonant clusters. The device of extrasyllabicity for word-edge clusters allows certain segments of the said clusters to be invisible to the syllabification rules or algorithm. Let us consider the following moraic structures of 'regular' word initial consonant clusters in Bangla: প্রাণ [pran] 'life' and পান [pan] 'beetle leaf'. These are illustrated in (11) as A and B figures.

(12) Syllable structure without initial /ʃ/

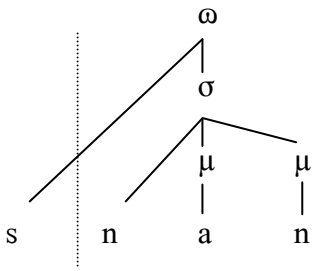
A. 

B. 

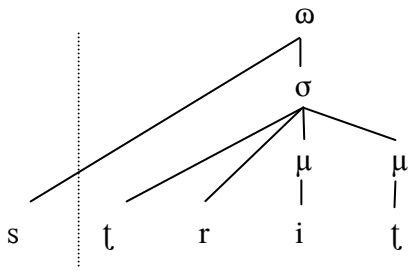
These structures show a fairly rule-binding syllabification in Bangla. In illustration B, the word initial consonant cluster /pr/ is syllabified as a complex onset which follows the sonority sequencing principle (rising sonority). Now, a comparative study can be done by taking a couple of examples of complex onsets with /s-/. Let us consider two cases from SB and OB: স্নান /śnan/ [snan] 'bath' and স্ট্রীট /sṭriṭ/ [sṭriṭ̪] 'street', respectively. These two cases are structured in the extrasyllabicity model in the followed illustration.

(13) Syllable structure with initial /ʃ/

C.



D.



In the above moraic structures, the initial /s-/ segment is considered as an external element and hence not a part of the syllabic structure. Instead, the /s-/ segment is attached directly to the prosodic word level. This concept leads to a simpler analysis of the syllabic structures of these words. Now, they follow SSP that makes it easier to locate them in the regular group of complex onsets.

The extrasyllabicity feature of /s-/ is restricted to its word-peripheral position. Once any sort of affixation moves this /s-/ towards a word medial position, the extrasyllabicity of this segment is lost and then this segment would be a subject to the regular syllabification rules. We have already seen a similar case in স্কুল /skul/ [ʃkul] : [iʃ.kul] 'school' in Bangla, where the extrasyllabicity is lost after the insertion of the vowel /i/ after /s-/.

2.3 Sonority Sequencing Principle (SSP)

In general, the sonority of a segment in a syllable (mainly at the onset position) start rising towards the nucleus (peak) and then falls towards to end of the syllable (coda). The phenomena are formalized as the sonority sequencing principle by G.

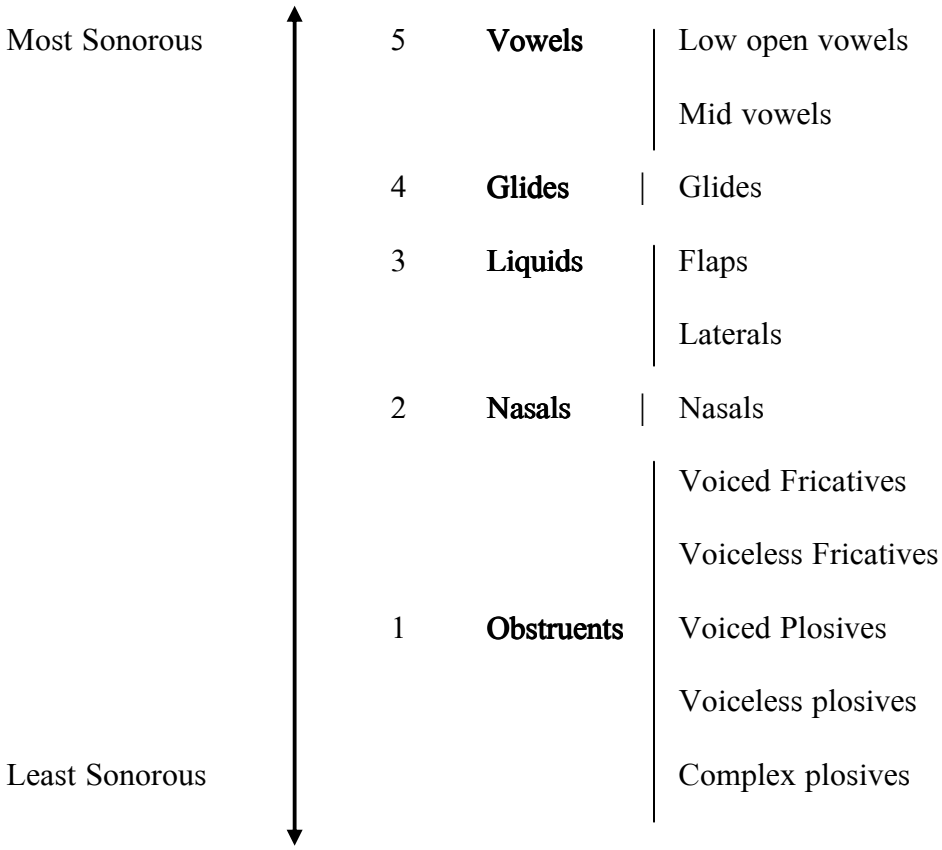
Clements 1990; Dost 2004 and others. The sonority sequencing principle (SSP) says:

> Between any member x of a syllable and the syllable peak p, only sounds of higher sonority rank than x are permitted. (G. Clements 1990, Dost 2004)

In a regular syllable structure, a consonant forms the onset, followed by a vowel at the peak (which is, surely, with a higher sonority value than the onset). Then, in some cases, another consonant lower than the vowel at nucleus (peak), forms the coda. But one should keep in mind that many languages do not allow a coda when onset is a universal occurrence. Following SSP, in a consonant cluster at the onset position, the first consonant must have a lower sonority value than the following consonant. Similarly, at the coda position, the final consonant should be lower than the previous consonant (in terms of sonority), in case of a complex coda.

Now, the rising and falling of sonority could be marked in a sonority scale. This scale is not a physical scale, but sort of a measuring index. In this index, plosives are the least sonorous sounds and vowels are most sonorous. Plosives and fricatives are taken together in one category, viz., obstruents. Then come nasals, liquids, glides and vowels. We can make use of the basic sonority scale as shown in Burquest & Payne 1993, where eleven sound categories are shown in terms of increasing (or decreasing) sonority level. The following chart illustrates a slightly modified version of the sonority chart of Burquest & Payne 1993, with a further step to group some of the sound categories in their respective broader category.

(14) Sonority scale (Bangla)



Many languages in the world have a minimal sonority distance (henceforth, MSD) in between two elements in a syllable. Some of them show a strict MSD while others have a varied MSD. Spanish is a classic example of strict minimal sonority distance. Harris 1983 has shown a strict MSD value of 2 in Spanish, taking the sonority scale values as shown in (15).

(16) Sonority scale (Spanish)

| Obstruents (1) > Nasals (2) > Liquids (3) > Glides (4) > Vowels (5) |

The MSD value is calculated not only for onset-nucleus and nucleus-coda, but also for the distance between two elements in an onset or coda. In Spanish, this distance value between two adjacent elements in a complex onset is always 2. For instance, Spanish words primo [prí.mo] 'cousin', ancla [áŋ.kla] 'anchor' etc. have the MSD value of 2 among the first and second consonants in a consonant cluster at the onset (Shepherd 2003).

Bangla, on the other hand, show different MSD at different parts in a syllable. In a complex onset the MSD is always either 1 or 2. More specifically, the obstruent-liquid clusters have the MSD value of 2 and the nasal-liquid clusters have the value as 1. The MSD between the onset and nucleus is not that specific in Bangla. It could be anything between 2 and 4. Same observation is applicable for the MSD between nucleus and coda, too.

(17) Different sonority distances (Bangla)

Onset-nucleus	MSD value 2	লাল /lāl/ [lal] 'red' (3-5-3)
Onset-nucleus	MSD value 3	নাম /nām/ [nam] 'name' (2-5-2)
Onset-nucleus	MSD value 4	বই /bai/ [boi̯] 'book' (1-5 + 4)

We should note that glides are used as the second member in a diphthong in Bangla, while first member is always a vowel (Sarkar 1986). Hence, the glide goes to the nucleus together with the vowel as one unit.

So, in this sonority scale, we can easily accommodate all the consonant clusters we have found at the onset position in SB words. In the SB cases, the

obstruent-liquid clusters (/kr/, /kl/, /gr/, /gl/, /tr/, /dr/, /d^{h}r/, /pr/, /pl/, /br/ and /b^{h}r/) have a sonority distance of 2, while nasal-liquid cases (/mr/, /nr/ and /ml/) have the value as 1. On the other hand, the borrowed words also have some occurrences of complex onset. Those cases, too, have a sonority distance of 1 or more. For instance, the obstruent-liquid cases (/ʈr/, /ɖr/, /t^{h}r/, /fr/, /fl/ and /bl/) have a sonority distance of 2. So, it is evident that, in Bangla, the minimal sonority distance is never less than 1. This observation leads to the following faithfulness constraint:

(18) MSD-1

From the margins of the syllable to its nucleus, there has to be a sonority rise from one segment to the next with the sonority distance value of at least one.

The minimal sonority distance constraint not only secures a minimum value of sonority distance, but also governs the rise and fall of the sonority. Since onset and coda are the two margins in a syllable, it requires the sonority rise from both onset and coda ends towards the nucleus. That means, in onset position it requires the sonority to be always rising towards the nucleus and then falls towards the coda. In both ends, the minimum sonority distance should be no lesser than 1.

Some languages do not allow any complex onset (i.e., consonant cluster at the onset position). Such a language is Japanese where not only complex onset, but complex coda is also prohibited. Japanese has a strong dislike for complex onset. Almost all the Japanese syllables, including borrowed words even from foreign languages, consist of only simple onsets. For instance, *Eng.* free > *Jap.* furii, to

avoid the consonant cluster at onset by the insertion of a vowel. On the other hand, some languages allow a complex onset even with the sonority distance of 0. Georgian allows a consonant cluster [tb] where two obstruents form a complex onset with MSD value of zero at the onset position. E.g., *Geor.* თბილისი /Tbilisi/ [tbiˈliːsi] 'the name of the capital of Georgia'.

In case of Bangla, it allows a complex onset for certain types of words (SB and OB) and with the MSD value of at least 1 at the onset. Therefore we consider the following well-formedness constraint from Kager 1999.

(19) *COMPLEX^{ONS} (Kager 1999)

Onsets are simple.

This constraint requires the onsets in a word to be simple. *COMPLEX^{ONS} (read as No Complex Onset) would punish any candidate with a complex onset in it. However, this is not a widely acceptable constraint in many languages in the world. Now, the previous constraint MSD-1, depending upon its ranking relative to *COMPLEX^{ONS}, along with other relevant faithfulness constraints (FAITH), can generate a possible factorial typology as follows:

(20) Factorial typology (complex onset)

A. (MSD-1), *COMPLEX^{ONS} ›› FAITH

"No complex onsets allowed"

B. MSD-1 ›› FAITH ›› *COMPLEX^{ONS}

"Complex onsets are allowed with a minimal sonority distance of one"

C. FAITH ›› *COMPLEX^{ONS}, MSD-1

"Complex onsets are allowed with no minimal sonority distance."

In the factorial typology (20A), *COMPLEX^{ONS} is ranked higher than FAITH and hence the concerned stratum of B-lexicon (or language in general) that follows this typology should have only simple onsets. Since the NB words do not allow any complex onset, only the first ranking in the above options caters the syllable structure of NB words. But if we consider SB and OB, the second ranking fits to both of them. But, the constraint ranking shown in (20C) does not work for any stratum of Bangla lexicon, because of the low ranking of the MSD-1 constraint that would make it useless in the minimal sonority requirement issue.

From the above discussion, a preliminary constraint ranking is drawn for each stratum or group of strata where the same ranking is application.

(21) Constraint ranking for complex onset

A. Tadbhaba (NB)

(MSD-1) *COMPLEX^{ONS}

|

FAITH

B. Tadbhaba (SB) and Borrowed (OB)

MSD-1

|

FAITH

|

*COMPLEX^{ONS}

In the factorial typology, we have a constraint group termed as FAITH that includes some relevant faithfulness constraints such as MAX-IO, DEP-IO etc. McCarthy & Prince 1995 argued for the anti-epenthesis constraint Dependency-IO (DEP-IO) which says that nothing should be inserted in the output segments that does not correspond to the input segments. That means there should not be any epenthesis in the output candidates.

(22) DEP-IO (McCarthy & Prince 1995)

Output segments must have input segments.

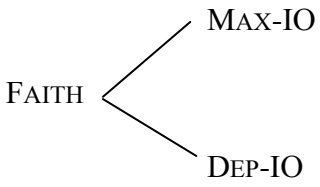
On the other hand, MAX-IO requires all the input segments to be preserved in the output candidates. Hence, no deletion is allowed in the output segments. The following faithfulness constraint could account for such a requirement. The constraint shown in (23) punishes any deletion of sound in the candidates which is not part of the input.

(23) MAX-IO (Kager 1999)

Input segments must have output correspondence.

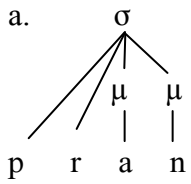
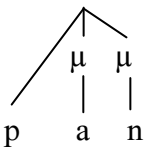
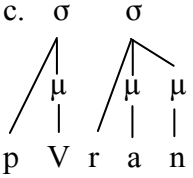
These two faithfulness constraints are used together in the previous constraint rankings as FAITH. Later in this work, they will be shown in the tableaux separately as MAX-IO and DEP-IO.

(24) Faithfulness constrains (Kager 1999)



In the core-periphery organization, one stratum (NB) follows the constraint ranking illustrated in (20A) while the other two strata follow ranking (20B). Let us consider an example from the SB stratum: প্রাণ /pran/ [pran] 'life', to verify the effectiveness of these constraint rankings.

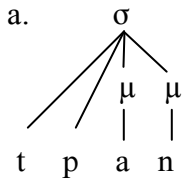
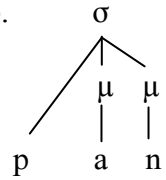
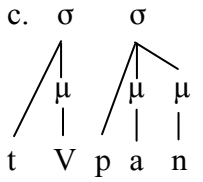
(25) Tableau: [pran] (SB and OB)

/pran/	MSD-1	MAX-IO	DEP-IO	*COMPLEX^{ONS}
☞ a. 				*
b. 		*!		
c. 			*!	*

In the tableau shown in (25), we have one candidate (a) with consonant cluster at the onset and one candidate (b) where a segment is dropped from the input. A third candidate (c) shows an addition of vowel (V) at the word initial consonant cluster. Now, each of them incurs a violation of at least one constraint. The optimal output (a) violates the lowest ranking markedness constrain *COMPLEX^{ONS}, which is the least expensive violation of the constraints. The second candidate (b) violates the faithfulness constraint MAX which is ranked above *COMPLEX^{ONS}. It is a fatal violation and hence ruled out. The third candidate is broken into two syllables after the insertion of a vowel at the word initial position. Here, the DEP-IO constraint is violated which is, again, ranked above *COMPLEX^{ONS} in this tableau.

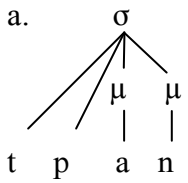
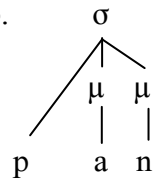
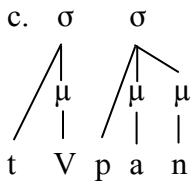
This very ranking allows complex onset in the SB and OB. But the appearance of complex onsets is restricted to the minimal sonority distance (MSD) value of 1 or more. Hence, words containing the MSD value of 0 are ruled out in this ranking. It rather allows a simple onset than a complex onset with MSD-0. The following tableau shows an imaginary word */tpan/ which has a complex onset /tp/.

(26) Tableau: *[tpan] (SB and OB)

*/tpan/	MSD-1	DEP-IO	MAX-IO	*COMPLEX^{ONS}
a. 	*!			*
☞ b. 			*	
c. 		*!		

In this case, the first candidate violates the highest ranking markedness constraint MSD-1 which is the most expensive violation and hence ruled out. Candidate (b) violates the faithfulness constraint MAX-IO which is not a fatal violation here. The last candidate (c) violates the faithfulness constraint DEP-IO which is ranked higher than MAX-IO. Hence, it is also ruled out for violating a high ranked constraint. Interestingly, the winning candidate /pan/ is a valid grammatical word and could be even from NB stratum. So, the ranking for SB and OB likes to have a complex onset only when it has a minimal sonority distance value of 1, otherwise it leads to a simple onset case irrespective of the lexical stratification.

(27) Tableau: *[tpan] (NB)

*/tpan/	MSD-1	*COMPLEX^{ONS}	DEP-IO	MAX-IO
a. 	*(!)	*(!)		
☞ b. 				*
c. 			*!	

In the above tableau, the second candidate (cand b) violates the lowest ranked faithfulness constraint MAX-IO. But, the first candidate with a complex onset, violates the markedness constraint *COMPLEX^{ONS} which is ranked above the faithfulness constraints. Again, the third candidate violates the faithfulness constraint DEP-IO and that is a fatal violation in this ranking. So, the least expensive violation is in the second case which has a simple onset. The relative ranking of MSD-1 and *COMPLEX^{ONS} are not crucial here. But they must be ranked higher than the faithfulness constraints.

2.4 Coda

Codas are not such an essential part of a syllable (like onsets) in many languages. In general, languages prefer to end in a vowel (nucleus) and not with a consonant (coda) (Kager 1999). But, Bangla allow codas in the syllable. The same phenomenon is seen in some other languages like English, Arabic etc. On the other hand, Blevins 1995 has shown that some languages, such as Mazateco, Fijian and Cayuvava entirely disallow coda in their syllable structure. This type of languages may employ certain phonological operations to avoid codas in a syllable. In Boumaa Fijian, vowel epenthesis takes place in order to avoid a possible coda in a word (Dixon 1988). Interestingly, there are no such languages that strictly require codas in its syllables.

In the B-lexicon, NB and SB elements do not contain any word-final consonant cluster. Although Masica 1991 claims that even borrowed words in Bangla do not allow consonant clusters at the word final position, a few of that sort are indeed present in the B-lexicon. The only cases of such clusters are found in OB cases. For instance, প্যান্ট /pænṭ/ [pænʈ] 'pant' (< English), দোস্ত /dost/ [doʃt] 'friend' (< Persian) etc. In the following table we have a set of word final consonant clusters available in B-lexicon.

(28) Word-final consonant clusters in OB

Type	Cluster	Orthography	Transcription	Meaning
Obstruent + Obstruent	p^ht	লিফট	lip^ht	lift
Nasa + Obstruent	ŋk	ব্যাংক	bæŋk	bank
	nc	লঞ্চ	lɔnc	launch
	nt	প্যান্ট	pænt	pant
	nɖ	পাউন্ড	paunɖ	pound
	mp	ল্যাম্প	læmp	lamp
Liquid + Obstruent	rk	পার্ক	park	park
	rc	টর্চ	tɔrc	torch
	rʄ	সারচার্জ	ʃarcarʄ	surcharge
	rb^h	নার্ভ	$narb^h$	nerve
	rt	শার্ট	sart	shirt
	rɖ	বোর্ড	bɔrɖ	board
Liquid + Nasal	lt	বেল্ট	belt	belt
	rn	হর্ন	hɔrn	horn
	rm	ফর্ম	$p^hɔrm$	form

A set of examples of word-final consonant clusters which have a coronal sound as either first or second member of the cluster, are presented in (29).

(29) Word-final consonant clusters with coronal obstruents in OB

Cluster	Orthography	Transcription	Meaning
rs	নার্স	nars	nurse
rsʈ	ফার্স্ট	p^harsʈ	first
sk	রিস্ক	risk	risk
ks	ট্যাক্স	ʈæks	tax

If complex coda is allowed in a particular stratum of Bangla, then the implicational universal for coda complexity says that Bangla also allows simple codas in its syllable structure. A well-formedness constraint NoCoda is relavent here, which requires that syllables must not end in consonants (coda).

(30) NoCoda (Kager 1999)

Syllables are open.

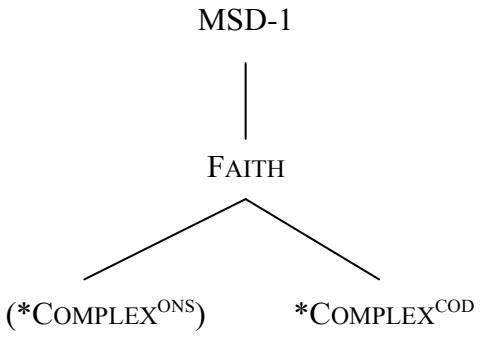
For the other two strata (NB and SB), we need a constraint that prevents any complex coda in a syllable. A similar constraint like *Complex^{ons} is proposed by Kager 1999, that fits the requirement of OB in this respect.

(31) *COMPLEX^{COD} (Kager 1999)

Codas are simple.

This constraint requires that the coda should not be a complex one. We have cases from OB where the original form of the borrowed words are preserved with complex coda in their syllable. The following constraint ranking allows a complex coda in a syllable, which is applicable for OB elements.

(32) Constraint ranking for complex coda



In this ranking, MSD-1 and faithfulness constraints dominate the marginal complexity constraints *COMPLEX^{ONS} and *COMPLEX^{COD}. But, *COMPLEX^{ONS} is not that important in this ranking compared to the other constraints. Let us consider a case from the OB stratum: প্যান্ট /pænt/ [pænt̪] 'pant' where a complex coda is preserved from its original form in English. A full constraint ranking with all the constraints proposed so far are included here in (33).

(33) Tableau: [pænt̬] (OB)

/pænt̬/	MSD-1	MAX-IO	DEP-IO	*COMPLEX^{ONS}	*COMPLEX^{COD}
☞ a. pænt̬					*
b. pæn		*!			
c. prænt̬			*!	*	*
d. pt̬ænt̬	*!			*	*

In the above tableau, each candidate has violated at least one constraint. The first candidate violated the lowest ranking constraint, which is, by definition, least expensive. But, the mutual ranking of the last two candidates are irrelevant here. So is the raking between MAX-IO and DEP-IO. Candidate (33d) violates MSD-1 because of the flat sonority value of zero. And, candidate (33b) violates the no-deletion constraint MAX-IO which is ranked above the marginal complexity constraints.

Now, this constraint ranking leads to a new stratification among other rankings. The two marginal complexity constraints (*COMPLEX^{ONS} and *COMPLEX^{COD}) make a significant difference between NB, SB and OB. Both of these constraints are satisfied in NB. That means, complex onsets and codas are not allowed in NB. Only complex onsets are allowed in SB, but not complex coda. And, in OB both complex onset and coda are allowed. The hierarchical inclusion relations among the said constraints and other faithfulness constraints are illustrated in the following table.

(34) Hierarchical inclusion relations in B-lexicon

	MSD-1	*COMPLEX^{COD}	*COMPLEX^{ONS}
NB	✓	✓	✓
SB	✓	✓	X
OB	✓	X	X

The table in (34) shows the inclusion and exclusion of certain constraints in different domains of B-lexicon. This hierarchy pattern is a sort of the "hierarchy of foreignness" as proposed by Kiparsky 1968 (also see Bobaljik 2006; Itô & Mester 1999). In this table, everything subject to a certain constraint may or may not be a subject to the next or any other constraint. For instance, NB and SB are subject to the constraint *COMPLEX^{ONS}, and both of them are subject to MSD-1 as well. So, it can be said that those strata which satisfy *COMPLEX^{ONS} also satisfy MSD-1, but not vice versa. Because, the words in the OB of B-lexicon satisfy MSD-1, they violate the markedness constraint *COMPLEX^{ONS}. A similar pattern is seen for *COMPLEX^{COD} which is satisfied in NB stratum, but not in the other two strata of the lexicon, viz., SB and OB. This specificity of constraints would help us to account for consonant clusters at different positions in a word of Bangla in the next sections.

2.5 Word initial consonant cluster

A preliminary introduction on the word initial consonant clusters in Bangla is

already given in the previous sections in this chapter. In general, native Bangla words that belong to the NB stratum in the lexicon stratification framework of this study totally restricts any consonant clusters to occur at the word initial position. However, the other two strata (SB and OB) allow such clusters more freely (see T. Bhattacharya 2000; Chatterji 1921, 1926a; Dasgupta 2003; Ferguson & Chowdhury 1960; Fitzpatrick-Cole 1990; Lahiri 2000; Masica 1991; Radice 1994; Sen 1993 a.o). Among these two strata other than NB, the SB stratum contains words directly borrowed from Sanskrit, while the OB stratum consists of foreign and some Indian language borrowings. These words retain their original consonant clusters at the word initial position even after their inclusion in the Bangla vocabulary. Though, there is a tendency to break these clusters in some non-standard variations of Bangla, the SCB typically tries to retain the clusters intact at this position.

A study conducted to search for possible consonant clusters in Bangla has turned out to be an important tool to work on this issue. The following table (35) presents the results from the study our available Bangla corpus. This table is structured in a different manner which does not show the exact number of occurrences of a certain consonant cluster, due to the difficulty of getting the exact frequency of some clusters. Rather a distinct number, ten (10) is taken as standard and every cell in the table shows whether the frequency of occurrence value is either more than ten (denoted by the symbol $>$) or ten or less, but not zero (denoted by $\leq$). The cells having a zero represent cases where the frequency is absolutely zero. The tables below represent the frequency of occurrences of word initial clusters where an obstruent sound is followed by another consonantal sound (both obstruents

and sonorants) in Bangla vocabulary.

(35) Frequency of occurrences of the word initial clusters in Bangla

a. Obstruents followed by another consonant

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	ɽ	l	m	n	ŋ	ʃ	p(h)	t(h)	ʈ(h)	k(h)	b(h)	d(h)	ɖ(h)	g(h)
1	p	>	0	>	0	0	0	0	0	0	0	0	0	0	0	0
2	t	>	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	ʈ	>	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	k	>	0	>	0	0	0	0	0	0	0	0	0	0	0	0
5	b	>	0	>	0	0	0	0	0	0	0	0	0	0	0	0
6	d	>	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	ɖ	>	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	g	>	0	>	0	0	0	0	0	0	0	0	0	0	0	0
9	p(h)	>	0	> OB	0	0	0	0	0	0	0	0	0	0	0	0
10	t(h)	> OB	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	ʈ(h)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	k(h)	>	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	b(h)	>	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	d(h)	>	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	ɖ(h)	>	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	g(h)	>	0	0	0	0	0	0	0	0	0	0	0	0	0	0

In the above table, cells 9a (/p^hl/) and 10a (/t^hr/) represent the frequency of occurrence of the word initial clusters which are available only in the OB stratum of

the B-lexicon. The other cases belong to all three strata.

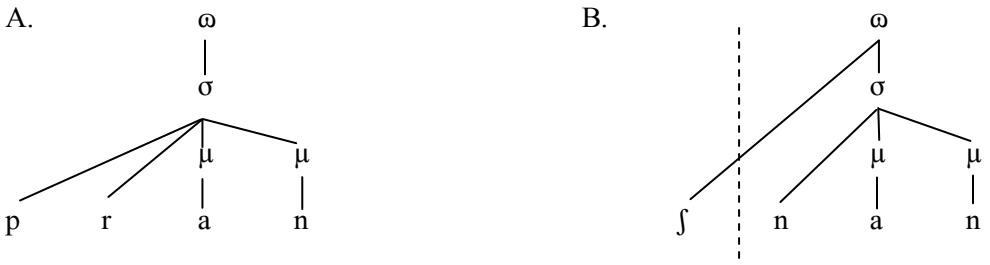
(35b). Sonorants followed by another consonant

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	ɽ	l	m	n	ŋ	ʃ	$p^{(h)}$	$ʈ^{(h)}$	$ʧ^{(h)}$	$k^{(h)}$	$b^{(h)}$	$d^{(h)}$	$ɖ^{(h)}$	$g^{(h)}$
17	r	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	ɽ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	l	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	m	>	0	≤	0	0	0	0	0	0	0	0	0	0	0	0
21	n	≤	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	ŋ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	ʃ	>	0	>	>	>	0	0	>	>	>	>	0	0	0	0

According to Masica 1991, the maximum syllable structure in the NB words is CVC, but the most preferred syllable throughout the language is the very common CV format (Dan 1992; Sarkar 1986). By claiming this, it is actually meant that the NB category words are prone to follow this structure more than the other two strata (SB and OB). But, in any case, NB words do not allow consonant clusters at the word-edge position.

Table (35b) shows more cases of occurrences of word initial clusters in the B-lexicon. In addition to the /mr/ (cell: 20a), /ml/ (cell: 20c) and /nr/ (cell: 21a) cases, there are many instances of s-obstruent and s-sonorant clusters. But, being an extrasybalbic element (Kenstowicz 1994a), word initial or word-final /s/ draws a special attention in the analysis of word-edge consonant clusters. It is not considered inside a syllable; rather a node directly connected to the prosodic word level.

(36) Word initial clusters in Bangla



In the above illustration, it is shown how the word initial (also word-final) clusters with /ʃ/ sounds compared to other obstruents (and sonorants) are taken in the picture. In the underlying level, (B) type structures are not considered as syllabic 'clusters' and hence, the consonant clusters like /ʃr/, /ʃl/, /ʃm/ etc. will not be taken in the analysis of syllabic clusters.

2.5.1 Liquid /r/ as the second member (non-liquide)

In table (35a), the column representing the frequency of occurrences of clusters where a liquid /r/ is preceded by an obstruent (cells: 1a-16a) has positive values in almost all cells (only except 11a: /ʈ^{h}r/). Again, in the following table, only the nasal sounds /n/ and /m/ preceded by /r/ shows more than zero occurrences in this corpus. Hence, it is clear that in Bangla, those word initial clusters are allowed where a liquid /r/ precedes any other consonantal sound (obstruent or sonorant), but not another liquid. That means word initial liquid-liquid clusters are prohibited even in the SB and OB strata in the B-lexicon. The most prominent reason for this

restriction could be the non-maintenance of the requirement of minimum sonority distance of one between two sounds in Bangla.

In Bangla, the minimal sonority distance between two elements in a syllable is never less than 1. This observation leads to the following faithfulness constraint already proposed in the previous sections.

(37) MSD-1

From the margins of the syllable to its nucleus, there has to be a sonority rise from one segment to the next with the sonority distance value of at least one.

This markedness constraint requires that besides the sonority rise and fall, two elements in the syllable must maintain a sonority distance of at least 1. A five point scale of sonority hierarchy in used in this study. In Bangla, the main sound groups in terms of sonority are presented in the following sorority scale (1-5: Low to high).

(38) Sonority hierarchy in Bangla (low to high)

(1) Obstruents > (2) Nasals > (3) Liquids > (4) Glides > (5) Vowels

A cluster like /lr/ at the beginning of a word (complex onset), would record a sonority distance of 0 (zero) since both the elements in this cluster belong to the same level (level 3) in the above mention sonority scale. So, such a cluster would definitely violate the MSD-1 constraint. Additionally, from the previous chapter

(chapter 2: Medial clusters), we would include the IDENT-IO(PLACE) constraint to account for the place issue that might rise in this analysis.

(39) IDENT-IO(PLACE) (Kager 1999)
The specification for place of articulation of an input segment must be preserved in its output correspondent.

This faithfulness constraint preserves the place feature of the input in the output candidates, and hence is ranked lower than the markedness constraints in most cases. We can use this constraint (and MSD-1) along with the two common faithfulness constraints, DEP-IO and MAX-IO. For this analysis, let us consider a hypothetical case with a word initial complex onset: /lran/.

(40) No liquid-liquid cluster word initially

/lran/	DEP-IO	MAX-IO-C	MSD-1	IDENT-IO(PLACE)	MAX-IO-V
1. lran			*!		
2. lman			*!	*	
☞ 3. pran				*	
4. laran	*!				
5. lan		*!			

In the above tableau, for candidate 3 to win, the markedness constraint MSD-1 needs to be ranked above the faithfulness constraint IDENT-IO(PLACE). However, MSD-1 is not ranked crucially relative to the faithfulness constraints DEP-IO and MAX-IO. The crucial ranking of MSD-1 above IDENT-IO(PLACE) eventually

rules out candidate 1 and 2 for fatal violations of MSD-1. In the first candidate, the sonority distance between /l/ and /r/ in the word initial cluster is zero and the second candidate attests a falling sonority towards the nucleus (/l/ > /r/). On the other hand, candidate 4 and five violates faithfulness by addition (DEP-IO) and deletion of sounds (MAX-IO-C), respectively. The minimal candidate (3) is an acceptable word form in Bangla and belongs to the SB stratum.

2.5.2 Liquid /l/ as the second member of a word initial cluster

From the frequency tables and the follow-up discussion, it is clear that Bangla allows word initial consonant clusters only where plosives and nasals are followed by liquids. While the liquid /r/ is preceded by plosives and nasals, the other liquid /l/ (lateral) in Bangla, allows only non-coronal sounds before it at the word initial position. That means at this position, coronal-lateral clusters are forbidden in Bangla vocabulary. One must keep in mind that this restriction is applicable typically to the SB and OB strata of the B-lexicon, because the occurrence of word medial cluster is totally forbidden in the NB stratum. So, there is no point to raise a special restriction applicable to the lexicons in this stratum. A markedness constraint could be employed to account for the coronal-lateral restriction at the word initial position in Bangla.

(41) *[COR-LAT

Word initially no coronal is followed by a lateral sound.

The above constraint would punish any consonant cluster where a coronal is followed by a lateral sound. We take this new constraint into account keeping the last ranking and see the outcome of the conflict. Let us consider a hypothetical example /nlan/ in this case as well.

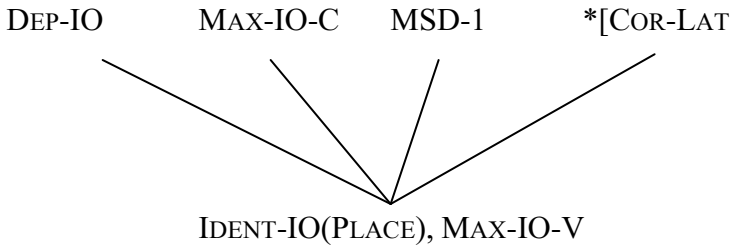
(42) No coronal is followed by a lateral word initially

/nlan/	DEP-IO	MAX-IO-C	MSD-1	*[COR-LAT	IDENT-IO (PLACE)	MAX-IO-V
1. nlan				*!		
☞ 2. mlan					*	
3. nalan	*!					
4. lan		*!				

As shown in the above tableau, MSD-1 is not an active constraint in this section of analysis. It is not violated by any candidate so far. But, the other markedness constraint *[COR-LAT is fatally violated by the input (as candidate 1) for the unwanted cluster of coronal and lateral. Candidates 3 and 4 are punishes by the faithfulness constraints DEP-IO and MAX-IO-C. The other faithfulness constraint IDENT-IO(PLACE) is ranked below the markedness constraint *[COR-LAT, to let candidate 2 be minimal. The winning candidate is a valid word in Bangla (SB stratum): ম্লান /mlan/ [mlan] ‘gloomy’.

From the above discussion, we can figure out the constraints playing in this analysis. The constraint ranking for the whole word initial cluster issue would be as illustrated below.

(43) Constraint ranking for word initial cluster in Bangla (SB and OB)



It should be also noted that this ranking is applicable to the words of the SB and OB strata of B-lexicon. But, Bangla has a set of different dialects that sometimes record deviations in certain words at the spoken level. That means the pronunciation of a set of words in these dialects may differ from the SCB for the same set of words. We will discuss some of such variations in the following section.

We have ruled out the ruled out the word initial consonant clusters starting with /ʃ/ in Bangla, because of the extrasyllabicity issue that suggests the initial (or final) /ʃ/ is not directly connected to onset (or coda). But, there are some instances in the SB and OB strata where such word initial cluster has an internal consonant cluster even after the initial /ʃ/. For example, SB স্ত্রী /stri/ [ʃtri] 'wife', OB স্ট্রীট /sṭriṭ/ [ʃʈriʈ] 'street' etc. These cases could actually be treated as normal consonant cluster just after taking away the /ʃ/ segment. That means, [ʃtri] could be treated as [(ʃ)tri] considering the /ʃ/ segment as extrasyllabic and /tr/ as the effective consonant cluster. Hence, this will fall in the OT analysis we have discussed so far.

2.5.3 Dialectal variations

In many dialects of Bangla (except the Kolkata-Nadia variation, which is considered to be the base of Standard Colloquial Bangla) show distinct sound-changes in the usage of several SB and OB words with initial consonant cluster. A typical change is to break the word initial cluster by inserting a vowel in between the consonants in the cluster. A strong dislike of word-edge clusters in native Bangla words might influence such an anti-cluster operation at the beginning of a word even in the SB and OB strata in these dialects. For instance,

(44) Dialectal variation in the word initial consonant (SB and OB)

Source Language	Bangla (Stratum)	Dialectal Variation	Meaning
a. *Eng.* club [klʌb]	*OB* ক্লাব /klab/ [klab]	কেলাব [kelab]	'club'
b. *Eng.* train [treɪn]	*OB* ট্রেন /ṭren/ [ṭren]	টেরেন [ṭeren]	'train'
c. *Eng.* glass [glas]	*OB* গ্লাস /glas/ [glaʃ]	গেলাস [gelaʃ]	'glass'
d. *San.* গ্রাম /gram/ [gram]	*SB* গ্রাম /gram/ [gram]	গেরাম [geram]	'village'
e. *San.* ম্লান /mlan/ [mlan]	*SB* ম্লান /mlan/ [mlan]	মেলান [melan]	'gloomy'

The examples given in the above set of SB and OB words clearly show an intention to break the word initial consonant cluster by inserting a vowel in between the members of that cluster. So, for these dialectal analyses, the constraint ranking must be restructured to make way for epenthesis. Apparently, a high ranked faithfulness constraint *COMPLEX^{ON} could account for this type of repair process.

But, this constraint is already ranked low throughout the SB and OB stratum. Then we need to deal the issue differently, with some additional constraints. But let us first recall the said faithfulness constraint from the previous sections, which would punish any complex onset in Bangla.

(45) *COMPLEX^{ONS} (Kager 1999)

Onsets are simple.

Typically, this constraint would punish any word having a complex onset irrespective of its stratum identity. Now, let us take a word from (44a) *OB* ক্লাব /klab/ [klab] > কেলাব [kelab] 'club' for an analysis. The split pattern of epenthesis requires *COMPLEX^{ONS} to be ranked above DEP-IO to get rid of the cluster in the loanword.

(46) Epenthesis in word initial cluster

/klab/	*COMPLEX^{ONS}	DEP-IO
1. klab	*!	
☞ 2. ke.lab		*

The high ranked *COMPLEX^{ONS} punishes the word initial cluster in the above tableau. But, after a vowel insertion in the word initial cluster, there could actually be two possible candidates: [kelab] or [eklab]. That means an internal epenthesis is also possible besides the external epenthesis. Sonority must not rise across a syllable boundary. Keeping the *COMPLEX^{ONS} in the account, let us also consider a constraint to account for such possibilities. A relevant constraint would be the SYLCONTACT

(already discussed in the gemination chapter), which punishes the rise of sonority from a coda to the following onset.

(47) SYLCONTACT (Gouskova 2002)

Sonority must not rise across a syllable boundary.

(48) Two types of epenthesis in word initial cluster

/klab/	SYLCONTACT	DEP-IO
1. ek.lab	*!	*
☞ 2. ke.lab		*

So, when a vowel is inserted outside the consonant cluster, it breaks the previous syllable structure and forms a new structure with rising sonority from coda to next syllable's onset. Hence, candidate 2 in the above tableau is punishes by SYLCONTACT.

Additionally, we need a contiguity constraint that requires the input-output mappings to involve contiguous substrings (see Gouskova 2002; Kager 1999; McCarthy & Prince 1999; Silverman 1992).

(49) CONTIGUITY (Gouskova 2002; Kager 1999)

Elements adjacent in the input must be adjacent in the output.

The CONTIGUITY constraint (termed as CONTIGUITY-IO by Kager 1999) is violated by any sort of medial epenthesis or medial deletion as well. Basically, this constraint is motivated by such languages where morpheme-internal epenthesis is

prohibited (e.g., Chukchee; Kenstowicz 1994b; Spencer 1994) and also by language like Diyari, which doesn't allow morpheme-peripheral deletion of any segments (McCarthy & Prince 1994).

(50) Contiguity

/klab/	SYLCONTACT	CONTIGUITY
1. ek.lab	*!	
☞ 2. ke.lab		*

It is clear that Candidate 2 ([kelab]) to win, CONTIGUITY must be crucially ranked below SYLCONTACT, being the other ranking intact. A full scenario of this analysis in presented in the following tableau.

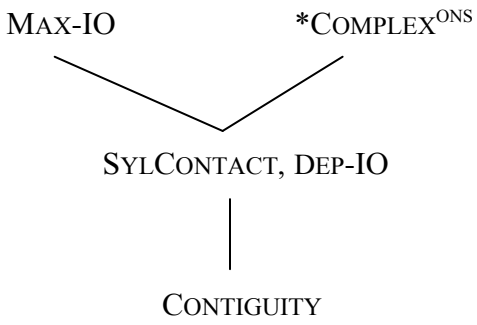
(51) No complex onset in SB (Dialectal variety)

/klab/	MAX-IO	*COMPLEX^{ONS}	SYLCONTACT	DEP-IO	CONTIGUITY
1. klab		*!			
☞ 2. ke.lab				*	*
3. ek.lab			*!	*	
4. lab	*!				

Candidate 2 turns to be the minimal case in this ranking, even after violating two constraints DEP-IO and CONTIGUITY. But both these constraints are ranked lowest in the above tableau. The input candidate (cand 1: [klab]) is ruled out by *COMPLEX^{ONS}, which is actually ranked lower in SB and OB stratum. But, in this analysis it is ranked as of NB stratum ranking. Apparently, such type of epenthesis is cross-linguistically the preferred repair of choice for borrowed words and

second-language acquisition (Davidson, Smolensky, & Jusczyk 2004; Fleischhacker 2000). According to Gouskova 2002, this is more prominent compared to the first-language acquisition where epenthesis is much less frequent than deletion.

(52) Constraint ranking for SB and OB in Dialectal use



Though SYLCONTACT and DEP-IO are not ranked crucially, but together they dominate CONTIGUITY in this tableau. These two constraints, again, ranked bellow MAX-IO and *COMPLEX^{ONS}. One must keep in mind that this ranking is very much specific to the dialectal use of SB and OB category words in the B-lexicon and it never comes to effect in regular use of those words in the Standard Colloquial Bangla (SCB).

2.6 Word final Cluster

Word final clusters are not present in native Bangla words as well as in Sanskrit borrowings. It is only available in the borrowed words (OB stratum) mainly from foreign languages. To be more specific, most of these bowed words with a word-

final consonant cluster are borrowed from English. Some rare cases are recorded as from Persian or some other foreign language. Words like, লিফট /liphṭ/ [lip^{h}ʈ] 'lift/elevator' (< English), প্যান্ট /pænṭ/ [pænʈ] 'pant' (< English), হর্ন /harn/ [hɔrn] 'horn' (< English), দোস্ত /dost/ [doʃt] 'friend' (< Persian) etc words with word-final consonant clusters belong to the OB stratum of B-lexicon. These words simply preserve the cluster settings and rules of their respective source language. There are no particular constraints in Bangla operational for the word-final clusters of this group of words. Only a low ranked markedness constraint *COMPLEX^{COD} would be effective for the OB stratum, and the other two strata (SB and NB) would require this constraint to be ranked high. Let us repeat this constraint from the previous section for our present analysis.

(53) *COMPLEX^{COD} (Kager 1999)
Codas are simple.

This constraint would punish a candidate with a word-final consonant cluster (complex coda) in any stratum. An example (প্যান্ট /pænṭ/ [pænʈ] 'pant') from the OB cluster is taken here to show the effect of this constraint ranking.

(54) Word-final consonant cluster in OB

/pænʈ/	DEP-IO	MAX-IO	*COMPLEX^{COD}
☞ 1. pænʈ			*
2. pænaʈ	*!		
3. pæn		*!	

In this tableau, for the input (candidate 1) to win, the markedness constraint *COMPLEX^{COD} (which prohibits any word-final consonant cluster) is ranked lower than the faithfulness constraint DEP-IO and MAX-IO. This ranking would also rule out any insertion or deletion of sound in the input. In case of SB or NB words, *COMPLEX^{COD} will be ranked comparatively higher, so that a word-final cluster would be punished by it.

WORD MEDIAL CLUSTERS

3.1 Word medial clusters in the B-lexicon

Bangla allows a large range of word medial clusters. These clusters mostly occur in monomorphemic words. Some borrowed words could be treated as monomorphemic in Bangla lexicon, though in their source language, they could be polymorphemic.

3.2 Corpus output

The following tables (1a and b) show the corpus search results for word medial clusters of the specified sound combinations. This is a similarly conducted corpus search as we have seen in the previous chapter. All the three strata of the B-lexicon (NB, SB and OB) are taken into account in this study. However, no compound word is considered in this representation. Also, the derivational and inflectional forms are not counted separately in this corpus study. In the following tables, the vertical X hand (1-23) shows the first member of the cluster and the horizontal Y hand (a-o) shows the send members. The vertical numbers are shown as a continuous numbering in (1a) and (1b).

(1) Frequency of word medial consonant clusters in Bangla

a. Obstruents followed by another consonant

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	ɽ	l	m	n	ŋ	ʃ	p^h	t^h	$ʈ^h$	k^h	b^h	d^h	$ɖ^h$	g^h
1	p	>	≤	>	≤	≤	0	>	G	>	>	0	0	0	0	0
2	t	>	≤	≤	0	≤	0	>	>	G	0	≤	0	0	0	0
3	ʈ	>	0	0	0	≤	0	0	≤	≤	G	≤	0	0	0	0
4	k	>	≤	>	0	≤	0	>	0	>	>	G	0	0	0	0
5	b	>	≤	≤	0	≤	0	≤	0	0	0	0	G	>	0	0
6	d	>	0	0	≤	≤	0	0	0	0	0	0	≤	G	0	≤
7	ɖ	>	≤	≤	≤	0	0	0	0	0	0	0	≤	≤	G	0
8	g	>	≤	≤	≤	>	0	0	0	0	0	0	≤	≤	0	G
9	p^h	≤	0	0	0	0	0	0	G	0	0	0	0	0	0	0
10	t^h	≤	0	0	0	0	0	0	0	G	0	0	0	0	0	0
11	$ʈ^h$	0	0	0	0	0	0	0	0	0	G	0	0	0	0	0
12	k^h	≤	0	0	0	≤	0	0	0	0	0	G	0	0	0	0
13	b^h	>	0	0	0	0	0	0	0	0	0	0	G	0	0	0
14	d^h	≤	0	0	0	0	0	0	0	0	0	0	0	G	0	0
15	$ɖ^h$	≤	0	0	0	0	0	0	0	0	0	0	0	0	G	0
16	g^h	≤	0	0	0	≤	0	0	0	0	0	0	0	0	0	G

The table is broken is two parts (a) and (b), keeping the plosives in one section and non-plosives in another.

b. Sonorants followed by another consonant

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	ɽ	l	m	n	ŋ	ʃ	$p^{(h)}$	$t^{(h)}$	$ʈ^{(h)}$	$k^{(h)}$	$b^{(h)}$	$d^{(h)}$	$ɖ^{(h)}$	$g^{(h)}$
17	r	G	≤	≤	>	>	0	>	>	>	>	>	>	>	>	>
18	ɽ	0	G	0	0	0	0	0	≤	≤	0	≤	0	0	0	0
19	l	0	0	G	>	≤	0	>	>	≤	≤	>	≤	≤	≤	≤
20	m	>	≤	≤	G	≤	0	≤	>	0	0	≤	>	≤	0	0
21	n	≤	0	0	≤	G	0	≤	≤	>	>	≤	0	>	>	0
22	ŋ	0	0	≤	0	0	G	0	0	0	0	≤	≤	0	0	≤
23	ʃ	>	≤	≤	>	>	0	G	>	>	≤	>	≤	0	0	0

Though the results shown in (1a) and (1b) together were collected by certain search commands[11], it still required a great deal of manual processing to gather the required data, because of the non-phonemic nature of the corpus. The regular expression used to search for a word medial cluster XY would be "[A-Za-z] XY [A-Za-z]" (Sound cluster XY is preceded and followed by any upper or lower case character). This regular expression should consider all intervocalic clusters as well as clusters followed and preceded by other consonants.

Interestingly, a straight search for XY cluster does not work perfectly in this corpus, since a large number of words include a silent 'a' in between the members of the cluster in the orthography. They have to be chosen and considered manually which usually changes the overall frequency of occurrence of a cluster. The decisive frequency (≤ 10) required to be hand-counted from the search output in some cases

[11] Commands used: (1) **grep –n [A-Za-z]*XY*[A-Za-z] filename** and, (2)**grep –n [A-Za-z]*XaY*[A-Za-z] filename** where XY is the sound cluster.

with fewer occurrences. A straight search for cluster XY (excluding the XaY instances) give a fairly clear view of an approximate frequency of occurrences for all the clusters shown in (1). It should be noted that this number are always increased when more instances from a XaY search are added to the results.

There are some cases where an acceptable frequency of occurrences is present for the underlying form, but at the surface representation (phonetically) they are different than their lexical forms. Usually, they are assimilated to another form in the surface representation (the form of our present study). For instance, the plosive-plosive cluster /td/ shows the frequency of occurrence as 0. But, there is an example of this very cluster available in the corpus, namely, তড়িদ্বার /taṛidvar/ [tɔɽiddar] 'electrode'. But, this is the lexical representation of the word. When it is pronounced in SCB, the cluster /td/ is assimilated to [dd] and the word becomes /tɔɽiddar/. Since we are studying the surface representation of the sounds, we will take the assimilated form of this word and the /td/ form will not be counted in frequency of occurrences presented in this particular cell. Rather it will belong to the result of [dd] frequency.

3.3 Coronals

In Bangla, word medial consonant clusters are typically divided into coda and onset of two different syllables. There are more than one reasons of this behavior of the word medial clusters. The most frequent words in Bangla belong to NB stratum (Native Bangla words) followed by the SB words. Native Bangla words never allow

a complex onset while SB words do allow them to a certain extent (Masica 1991). But, there are many cases of complex-onset present in OB category of the Bangla lexicon which mostly consists of borrowings from foreign languages like English, Portuguese, French etc. On the other hand, syllable-final consonant clusters are very rare in native Bangla words (NB) with just a few exceptions. For instance, গঞ্জ /ganj/ [gɔnʲ] 'part of a place name' as in নবাবগঞ্জ /nabābganj/ [nɔbabgɔnʲ] 'name of a place'. Similarly, the SB words, too, do not allow a complex coda but OB does.

3.3.1 Non-coronal and coronal

In many instances, coronals occur in certain places where consonants of other places of articulation cannot (see G. N. Clements 1988; Itô 1986; Smolensky 2005; Steriade 1982; Yip 1991). In this connection, Paradis & Prunet 1991 claimed, following Yip 1991, that monomorphemic words in English never contain more than one non-coronal in a cluster. Yip 1991 studied several such cases in English and proposed a cluster condition as follows:

(2) CLUSTER CONDITION (Yip 1991)

Adjacent consonants are limited to at most one place specification.

This account of cluster condition of English is valid only for monomorphemic words. But, Bangla consists of a huge number of words those are not monomorphemic in nature. According to the phonetic inventory of Bangla, there

are 11 coronal sounds present in this language. They are /t/, /t^h/, /d/, /d^h/, /ɖ/ /ɖ^h/, /s/, /ʃ/, /n/, /r/ and /l/. The corpus study suggests a set of clusters which includes at least one coronal in word medial position. The following combinations of cluster are drawn from the above mentioned frequency tables (1a and 1b).

(3) Cluster combinations

a)	stop-stop	pt, pʈ
b)	stop-fricative	ps, ks
c)	stop-nasal	pn, tn, k^hn, gn, g^hn
d)	stop-liquid	pl, p^hl, kl, bl, gl
e)	nasal-stop	mp, mp^h, mb, mb^h
f)	liquid-stop	lp, lk, lb, lg
g)	fricative-fricative	ss
h)	fricative-nasal	sn, sm
i)	nasal-fricative	ns

In case of nasal-stop (3e) or nasal-fricative (3i), the clusters are homorganic in nature. On the other hand, fricative-fricative (3g) clusters represent gemination which could be explained by the double linking. So, it is clear that in Bangla, [-cor] sounds occur only before a [+cor] sound and vice versa, in a word medial cluster. From this discussion, a constraint could be proposed to deal with the place association of coronal and non-coronal sounds in a cluster.

(4) PLACE(COR)

There is only one place association other than coronal in each consonant cluster.

The PLACE(COR) constraint requires a sound which belongs to the same place of articulation in a cluster to be the only non-coronal and if the cluster lacks such a sound, then the rightmost non-liquid be the only non-coronal. This specification is important, because in case of fricative-fricative or nasal-fricative, either a gemination or a homorganic cluster takes place. In the following table (depicted from 1a), the shaded cells show the cases where this constraint applies. These cases have either zero or less than (or equal to) ten occurrences found in the corpus. In Bangla, there are very rare occurrences of this type of clusters. For instance, the /km/ cluster in the cell 4d in (5) occurs only in OB stratum and in an extremely rare case in SB stratum রুক্মিনী (/rukmini/), which is actually a proper noun (name of a woman). But, there is no instance of this cluster available in the NB stratum.

(5) Place agreement (coronals)

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	ʈ	l	m	n	ɳ	ʃ	p(ʰ)	t(ʰ)	ʈ(ʰ)	k(ʰ)	b(ʰ)	d(ʰ)	ɖ(ʰ)	g(ʰ)
1	p	>	≤	>	≤	≤	0	>	G	>	>	0	0	0	0	0
2	t	>	≤	≤	0	≤	0	>	>	G	0	≤	0	0	0	0
3	ʈ	>	0	0	0	≤	0	0	≤	≤	G	≤	0	0	0	0
4	k	>	≤	>	0	≤	0	>	0	>	>	G	0	0	0	0
5	b	>	≤	≤	0	≤	0	≤	0	0	0	0	G	>	0	0
6	d	>	0	0	≤	≤	0	0	0	0	0	0	≤	G	0	≤
7	ɖ	>	≤	≤	≤	0	0	0	0	0	0	0	≤	≤	G	0
8	g	>	≤	≤	≤	>	0	0	0	0	0	0	≤	≤	0	G
9	pʰ	≤	0	0	0	0	0	0	G	0	0	0	0	0	0	0
10	tʰ	≤	0	0	0	0	0	0	0	G	0	0	0	0	0	0
11	ʈʰ	0	0	0	0	0	0	0	0	0	G	0	0	0	0	0
12	kʰ	≤	0	0	0	≤	0	0	0	0	0	G	0	0	0	0
13	bʰ	>	0	0	0	0	0	0	0	0	0	0	G	0	0	0
14	dʰ	≤	0	0	0	0	0	0	0	0	0	0	0	G	0	0
15	ɖʰ	≤	0	0	0	0	0	0	0	0	0	0	0	0	G	0
16	gʰ	≤	0	0	0	≤	0	0	0	0	0	0	0	0	0	G

So, the place of articulation is important here. Kager 1999 introduces a relevant constraint called IDENT-IO(PLACE) that plays an important role in the neutralization of the place-of-articulation feature.

(6) IDENT-IO(PLACE) (Kager 1999)

The specification for place of articulation of an input segment must be preserved in its output correspondent.

This faithfulness constraint preserves the place feature of the input in the output candidates. Usually IDENT-IO(PLACE) is ranked lower than the markedness constraints. Now, let us consider a hypothetical case where a word medial cluster has two non-coronal (and non-liquid) sounds.

(7) Interaction of coronals with the place feature

/rɔpko/	PLACE(COR)	DEP-IO	MAX-IO	IDENT-IO(PLACE)	NOCODA
1. rɔp.ko	*!				*
☞ 2. rɔp.to				*	*
3. rɔ.pɔ.to		*!			
4. rɔ.ko			*!		

The constraint PLACE(COR) is fatally violated by the first candidate in the above tableau. This candidate (*/rɔpko/) contains a cluster of two non-coronal stops /p/ and /k/. On the other hand, candidate 2 has a coronal stop /t/ with the non-coronal /p/. This combination results a violation of the constraint IDENT(PLACE), but satisfies the PLACE(COR) constraint which is ranked higher than the violated constraint. Candidates 3 and 4 have violated DEP-IO and MAX-IO, respectively, as results of addition and deletion of sounds in the input. So, candidate 2 is the winner

for the least fatal violence of the constraints.

3.4 Obstruents

3.4.1 Obstruent (Voice)

There are some languages (such as Maori, Ainu) which do not allow any voiced obstruent to occur inside the syllable at all (Maddieson 1984). On the other hand, some other languages like Kannada, Georgian, Tulu etc. allow unrestricted voicing in obstruents (Cho 1990). Bangla does not fit in either type. It allows voicing, but in a restricted manner.

Obstruent-obstruent clusters could be of four types depending on voicing. A possible list of these cases available in Bangla is illustrated below in (44).

(8) Potential obstruent-obstruent clusters in Bangla (in terms of [voice])

a.	Voiced-voiceless	:	bp, bt, dk, dʈ etc.
b.	Voiceless-voiced	:	pb, tg, ʈḍ, kd etc.
c.	Voiced-voiced	:	bd, ḍg, dḍ, gb etc.
d.	Voiceless-voiceless	:	pt, tk, ʈt, kp etc.

Now, except the geminate cases, the above listed combinations occur in a certain pattern. Among these, voiceless-voiced and voiced-voiceless combination including their aspirated variants, show a unique feature of obstruent-obstruent cluster restriction. The corpus study suggests that neither a voiced obstruent is ever

followed by a voiceless obstruent, nor is a voiceless obstruent followed by a voiced one. Lombardi 1999 noted that in Sanskrit (and also in Polish, Dutch, Catalan etc), voicing assimilation takes place in obstruent clusters with word-final neutralization, whereas some other languages, such as Yiddish, Serbo-Croatian, Romanian etc. show voicing assimilation (as in the earlier languages), but with word-final voicing contrast. Being a closely related language, Bangla shows a similar behavior as in Sanskrit, in case of obstruent clusters. In the following table, the highlighted areas are representing the frequency of occurrences of such obstruent clusters.

From the table below, it is evident that the heterorganic obstruent clusters do not occur in Bangla. This could be a case of positional neutralization where a contrast is maintained only in a prominent position[12] in a word. Outside of that position, the inventory is a less-marked subset of the full inventory attested in positions of privilege; the contrast in question is neutralized in favor of an unmarked value (Beckman 1998). Such phenomenon typically discussed in earlier studies under the heading of positional neutralization documented in several languages for many different positions of privilege. (see Bach 1968; Beckman 1998; Goldsmith 1985, 1989, 1990; Itô 1986, 1989; Kingston 1985, 1990; Kiparsky 1981; Ringen 1975; Steriade 1979, 1982, 1995; Trubetzkoy 1939). This observation could be a case of voicing assimilation in obstruent clusters.

[12] Privileged positions are those positions which enjoy some perceptual advantage in the processing system, via either psycholinguistic or phonetic prominence (Beckman 1998).

(8) Obstruent-obstruent clusters (in terms of [voice])

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	ɽ	l	m	n	ŋ	ʃ	p(ʰ)	t(ʰ)	ʈ(ʰ)	k(ʰ)	b(ʰ)	d(ʰ)	ɖ(ʰ)	g(ʰ)
1	p	>	≤	>	≤	≤	0	>	G	>	>	0	0	0	0	0
2	t	>	≤	≤	0	≤	0	>	>	G	0	≤	0	0	0	0
3	ʈ	>	0	0	0	≤	0	0	≤	≤	G	≤	0	0	0	0
4	k	>	≤	>	0	≤	0	>	0	>	>	G	0	0	0	0
5	b	>	≤	≤	0	≤	0	≤	0	0	0	0	G	>	0	0
6	d	>	0	0	≤	≤	0	0	0	0	0	0	≤	G	0	≤
7	ɖ	>	≤	≤	≤	0	0	0	0	0	0	0	≤	≤	G	0
8	g	>	≤	≤	≤	>	0	0	0	0	0	0	≤	≤	0	G
9	pʰ	≤	0	0	0	0	0	0	G	0	0	0	0	0	0	0
10	tʰ	≤	0	0	0	0	0	0	0	G	0	0	0	0	0	0
11	ʈʰ	0	0	0	0	0	0	0	0	0	G	0	0	0	0	0
12	kʰ	≤	0	0	0	≤	0	0	0	0	0	G	0	0	0	0
13	bʰ	>	0	0	0	0	0	0	0	0	0	0	G	0	0	0
14	dʰ	≤	0	0	0	0	0	0	0	0	0	0	0	G	0	0
15	ɖʰ	≤	0	0	0	0	0	0	0	0	0	0	0	0	G	0
16	gʰ	≤	0	0	0	≤	0	0	0	0	0	0	0	0	0	G

There is a constraint punishing any consonant cluster other than voiced-voiceless ones (in either order) which was originally proposed by Lombardi 1999 in the optimality theoretic study on laryngeal sounds.

(9) AGREE(VOICE) (Lombardi 1999)

Two adjacent obstruents must agree in voicing.

The above constraint requires that both the elements in an obstruent-obstruent cluster at word medial position must be either voiced or voiceless. Here, the place of articulation is less important than the manner of articulation. What matters here is the [voice] feature of the concerned sound. For such cases, the positional faithfulness is relevant which decomposes a faithfulness constraint according to the position and nature of the segment (see Beckman 1997, 1998; Féry 2003). Beckman's approach to the preservation of [voice] in the surface form is adopted in the constraint IDENT-IO(VOICE).

(10) IDENT-IO(VOICE) (Beckman 1998)

Correspondent segments must agree in voicing.

The constraint IDENT-IO(VOICE) punishes any change in the [voice] feature in any segment of the output form. This constraint is equally applicable to both onset and coda. An additional constraint (markedness) could also be introduced here which dislikes voiced obstruents. It is named VOICED OBSTRUENT PROHIBITION, and abbreviated as VOP (Itô & Mester 1998; Kager 1999). The VOP constraint states that the unmarked value of the feature [voice] in obstruent [-son] sounds is [-voice] sounds.

(11) VOP (Itô & Mester 1998; Kager 1999)

*[+voi, -son]

No obstruent must be voiced.

Hence, the context-free markedness constraint VOP punishes any occurrence of voiced obstruents in the candidates. In the following tableau, we see five possible candidates contesting to win where these new constraints are applied. As NOCODA is not playing a decisive role in this section, this tableau does not include that constraint. We will take a hypothetical case [ʃubto] for the analysis shown in (12).

(12) Voice agreement among obstruent clusters

/ʃub.to/	AGREE (VOICE)	DEP-IO	MAX-IO	IDENT-IO (VOICE)	VOP
1. ʃub.to	*!				*
2. ʃu.ba.to		*!			*
3. ʃub.do				*!	**
4. ʃu.to			*!		
☞ 5. ʃup.to				*	

In the above tableau the input (first candidate) is ruled out as it violates the AGREE(VOICE) constraint by allowing a voiceless-voiced cluster at the word medial position. A vowel insertion and a deletion of sound are also punished by the constraints DEP-IO and MAX-IO, respectively (candidate 2 and 4). Eventually, candidate 5 ([ʃupto] 'inactive') wins for the least fatal violation of constraint IDENT-

IO(VOICE), which is ranked lower than the other three constraints, but dominates VOP. Though the winning candidate (cand 5) and candidate 3 violate the same constraint IDENT-IO(VOICE) one time each, the latter additionally violates constraint VOP twice for the presence of two voiced obstruents in it and hence, ruled out. That violation is fatal for this candidate, because of the comparatively higher ranking of the IDENT-IO(VOICE) constraint. The winner (cand 5: [ʃupto]) is an acceptable word in Bangla. This word belongs to the SB stratum which contains borrowed from Sanskrit.

The ranking of AGREE(VOICE) above IDENT-IO(VOICE) and VOP allows the occurrence of voiced-voiced clusters, too. For instance, let us consider an OB stratum word জব্দ /jabda/ [ɟɔbdo] 'punished'.

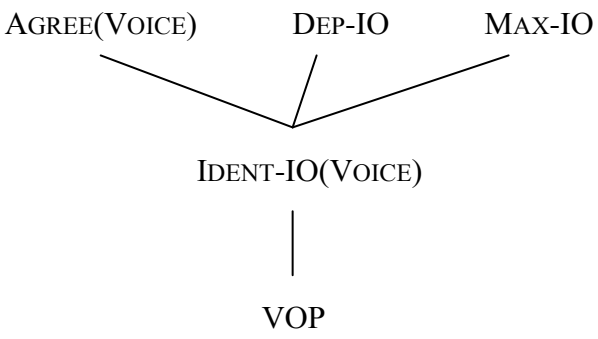
(13) Voice agreement among obstruent clusters

[ɟɔb.do]	AGREE (VOICE)	DEP-IO	MAX-IO	IDENT-IO (VOICE)	VOP
☞ 1. ɟɔb.do					**
2. ɟɔ.ba.do		*!			*
3. ɟɔp.to				**!	**
4. ɟɔb.to	*!			*	
5. ɟɔ.to			*!		

In the above tableau, the lowest ranking of the VOP constraint leads the input candidate to be optimal one, in spite a double violation of the very constraint. Two cases of devoicing (candidate 3 and 4) violate the faithfulness candidate IDENT-

IO(VOICE). Candidate 3 punished for the double violation of this constraint., but the latter is ruled out by another fatal violation of the voice agreement constraint. The rest of the candidates are punished by DEP-IO and MAX-IO. Below is the constraint ranking for these two tableaus.

(14) Constraint ranking favoring voice agreement in obstruent clusters



In this tableau, the first three constraints are not crucially ranked, but they dominate the last two constraints which are ranked crucially against them.

3.4.2 Final Devoicing

There could be an alternative approach to the voicing issue in Bangla. Since voiced obstruents are never followed by a voiceless obstruent in this language (as depicted from table 6), it could be assumed that a phonological neutralization might taken place in such cases. For instance, there is no word medial /bt/ cluster available in this corpus, but /bd/ is available. This type of neutralization takes place in many languages in just the opposite direction. It is well established that several West

Germanic languages like, German, Dutch, Bulgarian, Czech, Russian etc as well as Afrikaans, Türkish etc. have a common feature called *final devoicing* where an underlyingly voiced obstruent devoices when it occurs word-finally or at the syllable final position at the surface level (see Brockhaus 1995; Charles-Luce 1985; Féry 2003; Grijzenhout & Krämer 2000; Kopkalli 1993; Lombardi 1991; Oostendorp 2007; Rubach 2004; Vennemann & Murray 1972; Wiese 1996). Syllable final devoicing in German is a classical example of neutralization. For instance, Féry 2003 lists a set of German words where final devoicing takes place when the voiced sound is not followed by another sound (that is, word-finally).

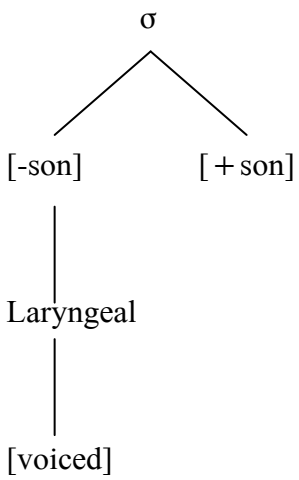
(15) Final devoicing in German

a. Hände [hɛndə] 'hands'	Hand [hant] 'hand'
b. Brave [bʁɑ:və] 'good, inflected'	Brav [bʁɑ:f] 'good, inflected'

In the earlier literature, two different approaches are already discussed. Neutralizing coda is one of them, which is defended by many scholars (see Brockhaus 1995; Hall 1992; Rubach 1989; Vennemann & Murray 1972; Wiese 1996 a. o.). In this approach, the voiced obstruent sound at the coda position is responsible for the neutralization process. The other FD approach is the onset-based process formulated by Lombardi 1991, 1995 and later Féry 2003. In this process FD works as a filter restricting the occurrence of voiced obstruent sounds to the onset position of syllable. As Lombardi's filter (see below) suggests, obstruent sounds become voiced only when followed by tautosyllabic sonorants, and in German,

according to Féry 2003, this amounts to restricting the occurrence of voiced obstruents to the onset position.

(16) Voicing restriction in German (Féry 2003)



In the present study, the first approach is more relevant, because of the good presence of voiced obstruent at the coda position in many cases. In this connection, Féry 2003 introduces the following constraint for German.

(17) *VDOBS$_{(CODA)}$ (Féry 2003)

Obstruents in the syllable coda are voiceless.

This constraint prohibits any occurrence of a voiced obstruent at the coda position in a syllable. In the word medial cluster, usually the first member is the coda and the latter is the onset of the next syllable. So, many clusters will violate

this constraint while some will satisfy it. For instance, let us consider the same example from the last tableau: /ʃubto/.

(18) Final devoicing

/ʃub.to/	*VDOBS$_{(CODA)}$	DEP-IO	MAX-IO	IDENT-IO (VOICE)	VOP
1. ʃub.to	*!				*
2. ʃu.ba.to		*!			*
3. ʃub.do	*!			*	**
4. ʃu.to			*!		
☞ 5. ʃup.to				*	

In (18), candidates 2 and 4 are ruled out by the violation of the two high ranking constraints DEP-IO and MAX-IO in the above tableau. The faithfulness constraint IDENT-IO(VOICE) is violated by candidate 3 and 5 for the change of the [voice] feature inside the word, but not punished for the low ranking of this constraint. But, candidate 3 is ruled out for another violation of the higher ranked constraint *VDOBS$_{(CODA)}$ for having a voiced obstruent at the coda position of the first syllable. This constraint is directly void by Candidate 1 (input) as well. So, Candidate 5 (/ʃupto/) is the least violated option in the above tableau which has a devoiced sound at the coda position (first syllable).

However, this promising new candidate *VDOBS$_{(CODA)}$ works with the word medial coda cases, but it would actually fail in the word final cases. Let us keep the same ranking of the previous tableau to analyze the word form সুদ /sud/ [ʃud]

'interest' from the SB stratum.

(19) Word-final devoicing

/ʃud/	*VDOBS$_{(CODA)}$	DEP-IO	MAX-IO	IDENT-IO (VOICE)	VOP
1. ʃud	*!				*
2. ʃu.do		*!			*
☹ 3. ʃut				*!	
4. ʃu			*!		

This ranking of the constraints let candidate 3 win over the other candidates for the least violation of the faithfulness constraint (only IDENT-IO(VOICE), which is ranked lowest). The absence of a voiced coda in candidate 3 made it the most favourite among all others. But, the loosing candidate 1 (input /sud/ [ʃud]), which retains the word-final voicing, is part of the lexicon (OB stratum).

From the above discussion it is clear that the final devoicing hypothesis actually does not lead to a credible outcome in this issue. Bangla perhaps does not allow an unconditional final devoicing in its lexicon. Only, word medial cases are devoiced in some instances. So far the first approach (voicing agreement) would be taken more accurate to analyze the voicing issue in Bangla.

3.5 Aspirated obstruents

3.5.1 Agreement in aspiration

In general, Bangla does not allow any other consonantal sound (with a few exceptions) after an aspirated obstruent at the word medial position. In the B-lexicon, there are very few cases of such exceptional clusters. These are depicted from the in the following table (20).

(20) Aspiration exceptions

i. Geminates, which is always treated differently (marked as G)

ii. Aspirated obstruents followed by liquid /r/ (9a-16a)

iii. Aspirated velars followed by /n/ (12e and 16e)

Among these three types of exceptions, (i) and (ii) cases are taken care of in a later stage, but cases noted in (iii) are discussed in the present section itself. Let us consider the shadowed portion of the following table, which represents the word medial consonant clusters where aspirated obstruents are the first members. In (21), almost all aspirated sounds followed by /r/ have more than zero frequency of occurrence and particularly /b^{h}r/ has more than ten occurrences registered in the B-lexicon. Apparently, this observation may contradict the previous claim that Bangla does not allow any consonant (here, /r/) preceded by an aspirated obstruent at word medial position. But, any obstruent followed by /r/ or /l/ word medially acts in a different manner compared to any other consonant. Dan 1992 has mentioned this

characteristic of such sound sequences as a gemination process.

(21) Aspirated Obstruent which is followed by another consonantal sound

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	ʈ	l	m	n	ŋ	ʃ	p(ʰ)	t(ʰ)	ʈ(ʰ)	k(ʰ)	b(ʰ)	d(ʰ)	ɖ(ʰ)	g(ʰ)
1	p	>	≤	>	≤	≤	0	>	G	>	>	0	0	0	0	0
2	t	>	≤	≤	0	≤	0	>	>	G	0	≤	0	0	0	0
3	ʈ	>	0	0	0	≤	0	0	≤	≤	G	≤	0	0	0	0
4	k	>	≤	>	0	≤	0	>	0	>	>	G	0	0	0	0
5	b	>	≤	≤	0	≤	0	≤	0	0	0	0	G	>	0	0
6	d	>	0	0	≤	≤	0	0	0	0	0	0	≤	G	0	≤
7	ɖ	>	≤	≤	≤	0	0	0	0	0	0	0	≤	≤	G	0
8	g	>	≤	≤	≤	>	0	0	0	0	0	0	≤	≤	0	G
9	pʰ	≤	0	0	0	0	0	0	G	0	0	0	0	0	0	0
10	tʰ	≤	0	0	0	0	0	0	0	G	0	0	0	0	0	0
11	ʈʰ	0	0	0	0	0	0	0	0	0	G	0	0	0	0	0
12	kʰ	≤	0	0	0	≤	0	0	0	0	0	G	0	0	0	0
13	bʰ	>	0	0	0	0	0	0	0	0	0	0	G	0	0	0
14	dʰ	≤	0	0	0	0	0	0	0	0	0	0	0	G	0	0
15	ɖʰ	≤	0	0	0	0	0	0	0	0	0	0	0	0	G	0
16	gʰ	≤	0	0	0	≤	0	0	0	0	0	0	0	0	0	G

In a word medial C_1C_2 sequence with /r/ (or /l/) as the second member (C_2), the first member of the sequence (C_1) is geminated which results in a /C_1C_1r/ (or C_1C_1l) sequence. So, basically the consonant clusters we are dealing with in (9a-16a) have the geminate of the C_1. In this case, C_1 will form a complex onset with a liquid in the following syllable. But, complex onset is not allowed in the NB stratum of B-

lexicon. The corpus result supports this observation, as only SB and OB words show such kind of geminated clusters at the surface representation. So, this gemination process is applicable only to the SB and OB strata, but not to the NB stratum. For instance, গোখরো /gok^{h}aro/ [gok^{h}.ro] 'cobra' (NB) does not allow a gemination at /k^{h}r/, but অভ্র [ɔb^{h}b^{h}ro] (</ɔb^{h}ro/) 'mica' does (from SB). This particular type of gemination will be discussed in a later section. But, in some other cases, geminates occur as aspirated-aspirated sound sequence as well. That means $C^{h}_{1}C^{h}_{2}$ becomes $C^{h}_{1}C^{h}_{1}$ or one of them is dropped, where C^{h} denotes an aspirated consonant. For instance, Sanskrit word সাধ্বী /sad^{h}vi/ [ʃad^{h}ʊi] 'faithful to husband' becomes সাধ্বী [sad^{h}d^{h}i] in Bangla. Here, C^{h}_{2} (/v/) merges with C_{1} (or deleted) and forms a geminate $C^{h}_{1}C^{h}_{1}$ ([d^{h}d^{h}]). This issue will be discussed in the next chapters.

Now, apart from gemination cases (particularly, with /r/ as second member), the shaded area in the above table shows a complete emptiness (except a couple of cases) in terms of frequency of occurrences. Almost all word medial consonant clusters where an aspirated consonant is followed by another consonant (not /r/) have no occurrences in B-lexicon, which leads to a very strong generalization that Bangla does not allow such clusters at all. An agreement constraint is introduced below in order to account for this situation with the same example given above. It should be noted that this analysis is a tentative one and might be modified in a later part of this chapter.

(22) AGREE(ASP)

Two adjacent obstruents must agree in aspiration.

The constraint AGREE(ASP) requires an agreement in the aspiration feature of two adjacent obstruents. So, both the members of a consonant cluster must be either aspirated or unaspirated, but not a combination of both types. In some cases a change in [aspiration] is possible, which could be focused by the following IDENT-IO constraint.

(23) IDENT-IO(ASP)

Correspondent segments must agree in aspiration.

This constraint punishes any deviation in the candidates from the input form in terms of aspiration. Now, let us consider an imaginary word-form /bod^{h}.bi/ for the present analysis.

(24) No aspirated obstruent followed by another obstruent

/bod^{h}.b^{h}i/	DEP-IO	MAX-IO	AGREE (ASP)	IDENT-IO (ASP)	NOCODA
☞ 1. bod^{h}.b^{h}i					*
2. bo.d^{h}a.b^{h}i	*!				
3. bo.d^{h}i		*!			
4. bo.di		*!		*	
5. bod^{h}.bi			*!	*	*
6. bod.bi				**!	*

In the above tableau, the NOCODA constraint is reintroduced for the clarity of the tableau. Here, candidate 2 violates the DEP-IO constraint, for the addition of segments in the input form. Candidate 5 is punished by the AGREE(ASP) which is ranked comparatively lower. Candidate 4 violates the IDENT-IO(ASP) constraint twice for the deviation (deaspiration) both at the coda and onset positions. This constraint is violated also by candidates 4 and 5. But, these two candidates are already punished by higher ranked constraints MAX-IO and AGREE(ASP). MAX-IO also punishes candidate 3 for the deletion of a segment. The input candidate 1 wins for the least fatal violation of the lowest ranked constraint (NOCODA). But, the winning candidate [bod^{h}b^{h}i] is a not valid word form in Bangla.

Now, an aspirated velar sound followed by /n/ is not totally absent in B-lexicon. It is already mentioned as an expectation case. A few cases are found in the corpus where these two clusters appear. Let us introduce a markedness constraint to account for the /k^{h}n/ and /g^{h}n/ cases available in Bangla.

(25) IDENT-ASP(vel, -n)

Input and output of a velar-nasal cluster must agree in aspiration, if the velar precedes a nasal /n/.

This context specific constraint works in a very restricted environment. It punishes any change in aspiration where an aspirated velar consonant precedes a nasal /n/. Let us take an example /big^{h}bo/, where such a cluster will be shown in action.

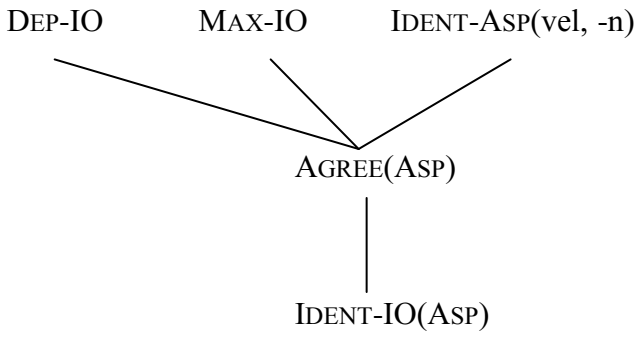
(26) Aspirated velar obstruent followed by /n/

/big^{h}.no/	DEP-IO	MAX-IO	IDENT-ASP (vel, -n)	AGREE (ASP)	IDENT-IO (ASP)
☞ 1. big^{h}.no				*	
2. bi.g^{h}a.no	*!				
3. bi.g^{h}o		*!			
4. big.no			*!		*

In (26), candidate 4 with unaspirated /g/ before /n/ is punished by the context sensitive constraint IDENT-ASP(vel, -n) as well as by IDENT-IO(ASP) for the change in aspiration. The input candidate (cand 1) could not fulfil agreement requirement for [aspiration] but turns out to be minimal, because of the low ranking of AGREE (ASP). Candidates 2 and 3 are ruled out by the faithfulness candidates, DEP-IO and MAX-IO, respectively. It is a valid word form in Bangla: বিঘ্ন [big^{h}no] 'disturbance'.

The constraint ranking remains unchanged from the previous tableau, even after the introduction of the new context specific faithfulness constraint IDENT-ASP(vel, -n).

(27) Constraint ranking for agreement in aspiration



The above illustration shows the ranking of the constraints used in this analysis. The last two constraints are ranked crucially, but both of them are dominated by the first three faithfulness constraints.

But, the agreement in aspiration cannot account for the nonoccurrence of the word medial cluster where aspirated obstruents are followed by another consonant. The aspirated-aspirated clusters (other than geminates) turn to be the optimal candidates in this approach. Alternatively, there are two distinct but related approaches which could be introduced here. These are the positional neutralization and positional faithfulness approach.

3.5.2 Final deaspiration

As in the voicing issue discussed in the previous section, aspiration, too, has room for an alternative way of analysis. A syllable-final deaspiration may play an important role here. Some Indian languages register a final deaspiration, for example, Sanskrit (Selkirk 1980), Gujarati (Cardona & Suthar 2003) to name a few. Chatterji 1926a gives a specific account for Bangla where he mentions that the final aspirates are deaspirated while the paleto-alveolar (voiceless) sounds tend to preserve the aspiration only word initially, but not very strongly. The focus of Chatterji's study was on the word-final coda, though. As we have seen in final devoicing issue in the previous section, deaspiration could also be analysed in two different ways, viz., positional neutralization and positional faithfulness.

3.5.2.1 Positional Neutralization (Aspiration)

Markedness constraints along with the context-free faithfulness constraint IDENT-IO(ASP) could be used here in the positional neutralization process. A similar approach is proposed in the earlier works of positional neutralization analysis of Dutch final devoicing (see Flemming 1995; Jun 1995, 2001; Kager 1999; Kirchner 1998; Steriade 1995; Zoll 1996, 1998 a.o). We can take the Dutch data from Kager 1999:

(28) Final devoicing in Dutch (Kager 1999)

a. /bɛd/ > [bɛt] 'bed'

b. /bɛd-ən/ > [bɛdn] 'beds'

The following markedness constraints related to a faithfulness constraint ([voice]) can account for the final devoicing in Dutch (Jun 2001).

(29) a. *VDOBS : Obstruents must not be voiced.

b. *$\text{VDOBS}_{(\text{CODA})}$: Obstruents at the syllable coda are voiceless.

The above mentioned markedness constraints are ranked in relation to the said context-free faithfulness constraint for voice feature as follows:

DEP-IO, *$\text{VDOBS}_{(\text{CODA})}$ > > IDENT-IO(ASP) > > *VDOBS

Then, the actual output can come out as an optimal output of the analysis, as illustrated in the following tableau.

(30) Final devoicing in Dutch (from Kager 1999)

/bɛd/	DEP-IO	*VDOBS$_{(CODA)}$	IDENT-IO(VOICE)	*VDOBS
1. bɛd		*!		**
2. bɛ.da	*!			**
☞ 3. bɛt			*	*

In the above tableau, the top ranked marked constraint *VDPOBS$_{(CODA)}$ rejects the voiced coda obstruent in candidate 1. Hence, candidate 3 becomes the optimal one after the fatal violation of DEP-IO of Candidate 2. A similar framework could be adopted in the account of positional neutralization of [aspiration] in our present study.

A constraint prohibiting any aspirated obstruent in the syllable would be a marked constraint which may interact with another and more marked constraint obstructing the occurrence of any aspirated obstruent at the coda obstruent.

(31) *ASPOBS

Obstruents must not be aspirated.

In a word medial cluster, the first member of the cluster is usually a coda which precedes the onset of the following syllable. A constraint opposing the

occurrence of an aspirated obstruent at the coda position will affect such a cluster by restricting its first member only to be an unaspirated sound.

(32) *ASPOBS$_{(CODA)}$

Obstruents at the syllable coda are unaspirated.

Now, let us take a hypothetical example /ud^{h}gɔm/, to see how this ranking works for the aspiration issue in Bangla.

(33) Final deaspiration (positional neutralization): case 1

/ud^h.gɔm/	DEP-IO	MAX-IO	*ASPOBS$_{(CODA)}$	IDENT-IO(ASP)	*ASPOBS
1. ud^h.gɔm			*!		*
2. u.d^{h}u.gɔm	*!				*
3. u.d^{h}ɔm		*!			*
☞ 4. ud.gɔm				*	
5. ud^h.g^{h}ɔm			*!	*	**

In the (33), candidates 1 and 5 violate *ASPOBS$_{(CODA)}$ constraint because of the presence of an aspirated obstruent in the coda position in the first syllable. In both cases, the violation is fatal enough to rule out the candidates. The faithfulness constraint DEP-IO punishes candidate 2 for an addition of sound in it whereas the faithfulness constraint MAX-IO punishes candidate 3 for a deletion of sound in the output. On the other hand candidate 4 fails to satisfy the constraint IDENT-IO(ASP),

but still wins for the low ranking of the violated constraint. The winning candidate represents a valid word-form in Bangla lexicon.

In general, Bangla also allows obstruent-obstruent (both unaspirated) instances at the word medial position, only when it is not part of a geminate. For instance, let us take a compound word in Bangla, বুধবার /bud^{h}-bar/ 'Wednesday' which consists of two morphemes: বুধ /bud^{h}/ and বার /bar/. Here, the word final coda /d^{h}/ actually becomes deaspirated in colloquial Bangla.

(34) Final deaspiration (positional neutralization): case 2

/bud^{h}-bar/	DEP-IO	MAX-IO	*ASPOBS$_{(CODA)}$	IDENT-IO(ASP)	*ASPOBS
1. bud^{h}.-bar			*!		*
2. bu.d^{h}a.-bar	*!				*
3. bud^{h}.ar		*!	*		*
4. bu.d^{h}ar				*!	*
☞ 5. bud.-bar				*	

Here, the ranking order of the constraints remains unchanged from the previous tableau in (34). The syllabic structure of the candidates and the violations of the constraints are also similar to the previous one. Eventually, candidate 5 emerges as the optimal candidate for its least amount of viobility. And, this is a valid word in spoken Bangla (SCB) as well.

3.5.2.2 Positional Faithfulness (Aspiration)

The very character of the phonological inventory of any language depends on the relationship of markedness constraints to faithfulness constraints. The relevant faithfulness constraint here is that which controls the mapping between input [aspiration] and output [aspiration]. According to Beckman 1998, faithfulness constraints reflect the intuition that phonological alternations are costly, occurring only under duress (that means, under compulsion by a higher-ranking constraint). In the current framework of optimality theoretic analysis, Beckman 1998 (cf. Féry 2003, 2008) introduces the positional faithfulness concept in the voicing feature of word segments. A similar structure could be followed here to analyze the aspiration properties of various segments in Bangla lexical entities. It is different from the positional neutralization approach. To analyze the same fact as stated in the previous section, the positional faithfulness approach adopts an opposite way in the invention of position-specific constraints. Beckman 1998 proposes the positional faithfulness constraint in terms of faithfulness to [voice]:

(35) IDENT-IO(VOICE)

Correspondent segments must agree in voicing.

(36) $\text{IDENT-IO(VOICE)}_{\text{ONSET}}$

Onset segments and their input correspondents must agree in voicing.

We have already introduced a constraint in this chapter to deal with the faithfulness of [aspiration], namely, IDENT-IO(VOICE). Now the positional (here, onset) faithfulness constrain influences another similar constraint which could be used for [aspiration] at the onset position. We will call it IDENT-IO(ASP)$_{ONSET}$.

(37) IDENT-IO(ASP)$_{ONSET}$

Onset segments and their input correspondents must agree in aspiration.

The above constraint requires the onset elements in the potential candidates to be faithful to their input form in terms of [aspiration]. Any change in the [aspiration] feature at the onset position would be punishable.

Let us analyze the same fact illustrated in the positional neutralization approach in the previous section. The input is /ud^{h}gɔm/ in the following tableau preserves the ranking of (33) above.

(38) Final deaspiration (positional faithfulness)

/ud^h.gɔm/	DEP-IO	MAX-IO	IDENT-IO (ASP)$_{ONSET}$	*ASPOBS	IDENT-IO (ASP)
1. ud^h.gɔm				*!	
2. u.d^{h}u.gɔm	*!			*	
3. u.d^{h}ɔm		*!	*	*	
☞ 4. ud.gɔm					*
5. ud^h.g^{h}ɔm			*!	**	*

In this constraint ranking, *ASPOBS prohibits voiced obstruents everywhere and hence the deaspirated cases are often required to obey it. That rules out candidate 1 (input) in this tableau. The high-ranked positional faithfulness constraint IDENT-IO(ASP)$_{ONSET}$ requires no featural change in the onset position. Candidate 5 violates that requirement. Thus candidate 4 becomes optimal, because it has no change in the onset position. But, the same constraint punishes candidate 3 for the aspirated onset, due to the structural change in it. The other two highest ranked faithfulness constraint DEP-IO and MAX-IO rule out candidate 2 and 3, respectively for the violation of epenthesis prohibition and no-deletion requirements.

Positional faithfulness disintegrates a faithfulness constraint into multiple others depending on the position of the certain segment in the syllable. Here the faithfulness of [aspiration] of obstruents at the onset position is more important to account for, compared to the general faithfulness of [aspiration] of the obstruents.

Both these approaches (positional neutralization and positional faithfulness) are working in the analysis of final deaspiration so far. But, a deeper study opens the possibility for one of them to perform better when the gemination cases are taken in the account. In the present framework, positional neutralization would rule out any gemination candidate. This approach would neutralize the aspiration present in either element of a cluster. But, Bangla shows a lot of gemination cases which are not preferred in this approach of de-accentuation process. On the other hand, the positional faithfulness ranking would allow geminates to be the optimal candidate in the OT analysis. So, finally the latter should be the preferred approach in the [aspiration] issue in this study.

3.6 Nasals

Nasals in Bangla have some restrictions of occurrences in general and particularly in word medial consonant clusters. The two main nasal sounds in Bangla /m/ and /n/ usually occur freely with other nasals and liquids. But, when they occur with plosives in a word medial-cluster, certain restrictions are imposed.

3.6.1 Plosive and nasal /m/

From the frequency of occurrences of word medial clusters presented in the frequency table (83), it is evident that the non-coronal nasal /m/ occurs restrictively after a plosive. This observation is presented by the selected area in the following table, where the frequency of occurrences of such clusters is always below ten (even zero). Here, only the unaspirated plosive cases followed my /m/ is shown as shaded cells (1d-8d). The as pirated counterparts of these plosives are already discussed in the previous sections as final deaspiration issue.

The shaded column in the following table represents the frequency of occurrences of the clusters where a plosive precedes /m/. Now, some of the cells record a few occurrences of the relevant clusters, that are more than zero but less than 10 cases each. A more close investigation in the data collected from the corpus reveals that most of these cases have more than zero instances belong to a particular stratum of B-lexicon, viz., OB. So, NB and SB class words following this restriction more faithfully, but OB does not.

(39) Plosives followed by nasal /m/

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	ʈ	l	m	n	ŋ	ʃ	p^h	t^h	$ʈ^h$	k^h	b^h	d^h	$ɖ^h$	g^h
1	p	>	≤	>	≤	≤	0	>	G	>	>	0	0	0	0	0
2	t	>	≤	≤	0	≤	0	>	>	G	0	≤	0	0	0	0
3	ʈ	>	0	0	0	≤	0	0	≤	≤	G	≤	0	0	0	0
4	k	>	≤	>	≤	≤	0	>	0	>	>	G	0	0	0	0
5	b	>	≤	≤	0	≤	0	≤	0	0	0	0	G	>	0	0
6	d	>	0	0	≤	≤	0	0	0	0	0	0	≤	G	0	≤
7	ɖ	>	≤	≤	≤	0	0	0	0	0	0	0	≤	≤	G	0
8	g	>	≤	≤	≤	>	0	0	0	0	0	0	≤	≤	0	G
9	p^h	≤	0	0	0	0	0	0	G	0	0	0	0	0	0	0
10	t^h	≤	0	0	0	0	0	0	0	G	0	0	0	0	0	0
11	$ʈ^h$	0	0	0	0	0	0	0	0	0	G	0	0	0	0	0
12	k^h	≤	0	0	0	≤	0	0	0	0	0	G	0	0	0	0
13	b^h	>	0	0	0	0	0	0	0	0	0	0	G	0	0	0
14	d^h	≤	0	0	0	0	0	0	0	0	0	0	0	G	0	0
15	$ɖ^h$	≤	0	0	0	0	0	0	0	0	0	0	0	0	G	0
16	g^h	≤	0	0	0	≤	0	0	0	0	0	0	0	0	0	G

Let us have a look in the following illustration of such clusters where the frequency of the clusters is shown according to their respective stratum. Below, the X mark denotes that the frequencies of the corresponding clusters are zero and the tick (✓) mark is for frequencies between one and ten.

(40) Frequency of plosive-/m/ clusters

Cell	Cluster	Frequency			Cell	Cluster	Frequency		
		NB	SB	OB			NB	SB	OB
1d	/pm/	X	X	✓	9d	/p^{h}m/	X	X	X
2d	/tm/	X	X	X	10d	/t^{h}m/	X	X	X
3d	/ʈm/	X	X	X	11d	/ʈ^{h}m/	X	X	X
4d	/km/	X	✓	X	12d	/k^{h}m/	X	X	X
5d	/bm/	X	X	X	13d	/b^{h}m/	X	X	X
6d	/dm/	X	X	✓	14d	/d^{h}m/	X	X	X
7d	/ɖm/	X	X	✓	15d	/ɖ^{h}m/	X	X	X
8d	/gm/	X	✓	✓	16d	/g^{h}m/	X	X	X

So, it is clear from the above illustration, that plosives are not followed by /m/ for maximum number of entries in the NB and SB strata. But, two SB strata do record a nominal number of such cases, which will be taken into the account later in this section. We can draw the general plosive-/m/ restriction in terms of a constraint prohibiting any plosive to occur before /m/ in a word medial cluster in Bangla, which will be applied only to the NB and SB stratum.

(41) *PLOSIVE-/m/

No plosive is followed by nasal /m/.

In Sanskrit, there are numerous instances where a plosive precedes /m/ at the

word medial position. In general, many such words in Bangla are borrowed from Sanskrit (hence, belong to the SB stratum). These words project an interesting behavior when integrated in the B-lexicon. They always form a geminate of the first element of the cluster, by either merging or dropping the second element. That means, C_1C_2 becomes C_1C_1 where C_1 is the plosive and C_2 is /m/. For instance,

(42) Gemination: Consonant followed by /m/ (SB)

a. *San.* पद्म /padma/ > *Bang.* পদ্ম /padda/ [pɔd.do] 'lotus'

b. *San.* छद्म /c^{h}adma/ > *Bang.* ছদ্ম /c^{h}adma/ [c^{h}ɔd.do] 'disguise'

c. *San.* आत्मा /atma/ > *Bang.* আত্মা /atma/ [at.ta] 'soul'

Now, to analyze such changes, we need a constraint that accounts for gemination in general. The faithfulness constraint UNIFORMITY (see McCarthy & Prince 1995, cf. Keer 1999; Wheeler 2005) could be used in order to dispose of certain codas by coalescence with a following consonant (violation of UNIFORMITY). Since this constraint blocks any merger in the syllable or cross-syllabically, it must be ranked lower than markedness for gemination effect to take place.

(43) UNIFORMITY (McCarthy & Prince 1995)

No element of the output has multiple correspondents in the input.

(No coalescence)

The constraint UNIFORMITY punishes the mapping of a pair geminate onto a

single segment. This constraint punishes any candidate where two elements are merged (fusion) into one from the input form and form a geminate. A more detailed discussion on this and on gemination processes in general, will be discussed in the following chapter.

Now, let us take a similar word: আত্মা /ātmā/ [ãtta] 'soul', which is originally from Sanskrit, and now belongs to the SB stratum of B-lexicon. It is presented in the format of C_1C_2 (that is, /t_1m_2/) in the following tableau.

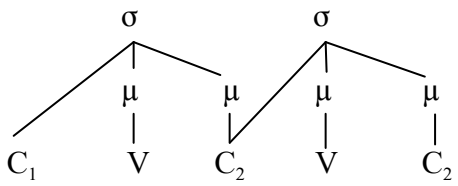
(44) No plosive + /m/ cluster (SB case)

/at_1m_2a/	DEP-IO	MAX-IO	*PLOSIVE-/m/	UNIFORMITY
1. at_1m_2a			*!	
☞ 2. $a\mu t_{1/2}a$				*
3. at_1a-m_2a	*!			
4. $a\mu t_1a$		*!		
5. at_1a		*!		

In the above tableau, the markedness constraint *PLOSIVE-/m/ rules out candidate 1 for the existence of an unwanted cluster /tm/ in the candidate. Candidates 3, 4 and 5 fatally violate faithfulness constraints either by adding sounds or deleting any part of the input. The lowest ranked constraint UNIFORMITY is violated by candidate 2. But, this violation is the weakest among all the candidates. So, this geminated form wins.

The μ symbol in (44) represents the place sharing feature of the geminate structure as illustrated in the following figure.

(45) Geminate structure (moraic)



Now, in the presentation of the frequency of occurrences, very few cases of velar-nasal /m/ (/km/ and /gm/) clusters occur in the SB stratum. For instance, যুগ্ম /jugma/ [ɟugmo] 'joint', বাগ্মী /bagmi/ [bagmi] 'speaker', রুক্মিনী /rukmini/ [rukmini] 'a proper noun' etc do not allow any change in the medial cluster. A context sensitive markedness constraint requiring the nasal /m/ to preserve its nasality after a velar sound could be introduced here to account for such cases in Bangla.

(46) IDENT-NAS(vel, -m)

Input and output of a velar-nasal cluster must agree in nasality, if the velar is followed by a nasal /m/.

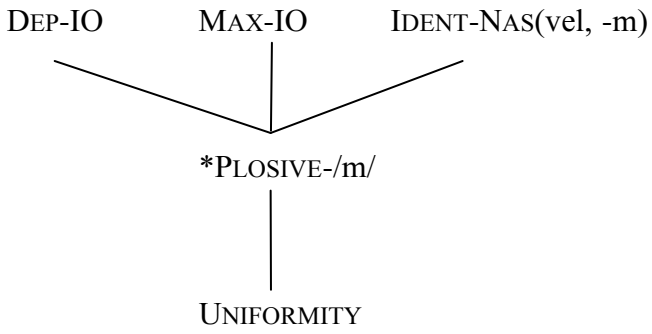
This constraint punishes any deviation in the [nasal] feature in the output candidate, when the input contains a velar-/m/ cluster. Let us consider the above mention SB case: বাগ্মী /bagmi/ [bagmi] 'speaker'

(47) Velar (plosive) + /m/ cluster (SB case)

/bag$_1$m$_2$i/	DEP-IO	MAX-IO	IDENT-NAS (vel, -m)	*PLOSIVE-/m/	UNIFORMITY
☞ 1. bag$_1$m$_2$i				*	
2. ba$_{\mu}$g$_{1/2}$a			*!		*
3. bag$_1$a-m$_2$i	*!				
4. ba$_{\mu}$g$_1$a		*!			
5. bag$_1$a		*!			

In the above tableau, candidates 3, 4 and 5 violate the highest ranking faithfulness constraints DEP-IO and MAX-IO. For the input candidate (cand 1) to win, the context sensitive constraint IDENT-NAS(vel, -m) needs to be ranked above *PLOSIVE-/m/ and UNIFORMITY. This ranking would rule out the gemination case (candidate 2) for the violation of the nasality ([gm] > *[gg]) in this velar-nasal cluster. Hence, the input candidate wins for the least amount of violation (only *PLOSIVE-/m/). There would be a strict constraint ranking for the above mentioned criteria from NB and SB strata of the B-lexicon, keeping the exceptional cases of SB (velar-nasal).

(48) Constraint ranking for plosive-nasal restriction: NB and SB strata



The three top-ranked faithfulness constraints are not crucially ranked relative to the other two constraints in this analysis. These constraints dominate the last two markedness constraints. But, this ranking is valid only for NB and SB strata. If the same constraints are applied in case of an instance from the OB stratum, the ranking would be changed in the latter case.

Let us consider a word from the OB stratum: তকমা /takma/ [tɔkma] 'marking'. This is a borrowed word from Persian. The OT analysis for this word in illustrated in the following tableau. The ranking of the markedness constraints, UNIFORMITY and *PLOSIVE-/m/ is changed here to respect the occurrence of /km/ cluster in the OB stratum.

(49) No plosive + /m/ cluster (OB case)

/tɔk_1m_2a/	DEP-IO	MAX-IO	IDENT-NAS (vel, -m)	*PLOSIVE-/m/	UNIFORMITY
☞ 1. tɔk_1m_2a				*	
2. tɔ$\mu k_{1/2}$a			*!		*
3. tɔk_1a-m_2a	*!				
4. tɔμk_1a		*!			
5. tɔk_1a		*!			

Since, OB stratum does not have any restriction in the occurrence of /m/ after plosives; the relevant constraint prohibiting such occurrences is ranked lowest in the above tableau. Hence, the previous winner (cand 2) is rules out in this tableau by a fatal violation of the anti-merger (no-gemination) constraint UNIFORMITY,

which is ranked comparatively higher. The input candidate (cand 1) wins, because of the low ranking of *PLOSIVE-/m/.

So, the conclusion of this section is that the plosives are followed by /m/ in the OB stratum, but prohibited in NB and SB stratum in Bangla. But, they need different constraint rankings (of the same set of constraint) for the respective stratum.

3.7 Other clusters

There iare some other interesting phenomena that could be found in the word medial clusters. But, their occurrences are not regular enough to give an OT account. Since, optimality theoretic analysis requires solid evidence in favor of a claim; it is hard to give an OT analysis for these cases. Let us discuss these issues in a general patters in this section.

3.7.1 Nasal and plosives

In word medial position, coronal nasal sounds /n/ and /ŋ/ typically avoids any occurrence before a non-coronal plosive. These two sounds usually occur with any other coronal plosives. On the other hand, non-coronal nasal m could be followed only by non-coronal plosives. But, their occurrences are not totally accountable, because of the irregular frequency of occurrence pattern. The following table shows the frequency of occurrences of this sort of clusters.

(50) Nasal sounds followed by plosives

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	**ɽ**	**l**	**m**	**n**	**ŋ**	**ʃ**	**$p^{(h)}$**	**$t^{(h)}$**	**$ʈ^{(h)}$**	**$k^{(h)}$**	**$b^{(h)}$**	**$d^{(h)}$**	**$ɖ^{(h)}$**	**$g^{(h)}$**
17	**r**	G	≤	≤	>	>	0	>	>	>	>	>	>	>	>	>
18	**ɽ**	0	G	*0	*0	*0	0	0	≤	≤	0	≤	0	0	0	0*
19	**l**	0	0	G	>	≤	0	>	>	≤	≤	>	≤	≤	≤	≤
20	**m**	>	≤	≤	G	≤	0	≤	>	*0	0	≤	>	≤	0	*0
21	**n**	≤	0	*0	≤	G	0	≤	≤	>	>	≤	0	>	>	0
22	**ŋ**	0	0	≤	0	0	G	0	0	0	0	≤	≤	0	0	≤
23	**ʃ**	>	≤	≤	>	>	0	G	>	>	≤	>	≤	0	0	0

In (50), the shaded area represents the clusters where nasal sounds are followed by plosives. One can draw two different patterns of the frequency of occurrence among these cases from (1). Below, we illustrate those cases according to their value depicted in the corresponding cells. Here, since we are not giving an OT account of this phenomenon, let us make a broad division in the number of occurrences. Instead of three frequency types, now we will consider only two.

i. Zero (0) ii. Ten or less (≤) iii. More than ten (>)	➔	i. Ten or less (includes zero) (≤) ii. More than ten (>)

In the following illustration, clusters with less than or equal to ten occurrences are shown with an **X** mark and instances where more than ten cases are

available are shown with a tick (✓) mark.

(51) Place agreement of nasal and plosive

Labial + X			Coronal + X			Velar + X		
(20h)	/mp(ʰ)/	✓	(21h)	/np(ʰ)/	**X**	(22h)	/ŋp(ʰ)/	**X**
(20l)	/mb(ʰ)/	✓	(21l)	/n(ʰ)/	**X**	(22l)	/ŋb(ʰ)/	**X**
(20i)	/mt(ʰ)/	**X**	(21i)	/nt(ʰ)/	✓	(22i)	/ŋ(ʰ)/	**X**
(20j)	/mʈ(ʰ)/	**X**	(21j)	/nʈ(ʰ)/	✓	(22j)	/ŋ(ʰ)/	**X**
(20m)	/md(ʰ)/	**X**	(21m)	/nd(ʰ)/	✓	(22m)	/ŋd(ʰ)/	**X**
(20n)	/mɖ(ʰ)/	**X**	(21n)	/nɖ(ʰ)/	✓	(22n)	/ŋɖ(ʰ)/	**X**
(20k)	/k(ʰ)/	**X**	(21k)	/k(ʰ)/	**X**	(22k)	/k(ʰ)/	**X**
(20o)	/g(ʰ)/	**X**	(21o)	/g(ʰ)/	**X**	(22o)	/g(ʰ)/	**X**

The tick marked cases in (51) show that a coronal plosive will go with another coronal (nasal) /n/ and a labial plosive will go with another labial nasal) /m/ in a word medial consonant clusters in Bangla. The velar nasal /ŋ/ has a very low amount of occurrences in general. Here, it does not make any significant difference in the outcome. So, from the first two columns of (51), it is evident that a nasal-plosive cluster usually agrees on the place of articulation. An example of such phenomenon would be, সম্পদ /sampad/ [ʃɔmpɔd] ‘property, asset’. This word belongs to the SB stratum of B-lexicon.

3.7.2 Lateral and coronal

The cluster combinations presented in the following tables, the shaded cells show the frequency of occurrences of liquid-stop clusters in Bangla. These liquids include lateral (alveolar) /l/ and unaspirated flap /r/ in Bangla. The following table shows the cases where /l/ and a coronal together form a cluster.

(52) a. Coronal-lateral interaction

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	**ɽ**	**l**	**m**	**n**	**ŋ**	**ʃ**	**p^{h})**	**t^{h})**	**ʈ^{h})**	**k^{h})**	**b^{h})**	**d^{h})**	**ɖ^{h})**	**g^{h})**
1	**p**	>	≤	>	≤	≤	0	>	G	>	>	0	0	0	0	0
2	**t**	>	≤	≤	0	≤	0	>	>	G	0	≤	0	0	0	0
3	**ʈ**	>	0	0	0	≤	0	0	≤	≤	G	≤	0	0	0	0
4	**k**	>	≤	>	0	≤	0	>	0	>	>	G	0	0	0	0
5	**b**	>	≤	≤	0	≤	0	≤	0	0	0	0	G	>	0	0
6	**d**	>	0	0	≤	≤	0	0	0	0	0	0	≤	G	0	≤
7	**ɖ**	>	≤	≤	≤	0	0	0	0	0	0	0	≤	≤	G	0
8	**g**	>	≤	≤	≤	>	0	0	0	0	0	0	≤	≤	0	G
9	**p^{h}**	≤	0	0	0	0	0	0	G	0	0	0	0	0	0	0
10	**t^{h}**	≤	0	0	0	0	0	0	0	G	0	0	0	0	0	0
11	**ʈ^{h}**	0	0	0	0	0	0	0	0	0	G	0	0	0	0	0
12	**k^{h}**	≤	0	0	0	≤	0	0	0	0	0	G	0	0	0	0
13	**b^{h}**	>	0	0	0	0	0	0	0	0	0	0	G	0	0	0
14	**d^{h}**	≤	0	0	0	0	0	0	0	0	0	0	0	G	0	0
15	**ɖ^{h}**	≤	0	0	0	0	0	0	0	0	0	0	0	0	G	0
16	**g^{h}**	≤	0	0	0	≤	0	0	0	0	0	0	0	0	0	G

b. Lateral-coronal interaction

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	**ʈ**	**l**	**m**	**n**	**ŋ**	**ʃ**	**p^{h}**	**t^{h}**	**ʈ^{h}**	**k^{h}**	**b^{h}**	**d^{h}**	**ɖ^{h}**	**g^{h}**
17	**r**	G	≤	≤	>	>	0	>	>	>	>	>	>	>	>	>
18	**ʈ**	0	G	0	0	0	0	0	≤	≤	0	≤	0	0	0	0
19	**l**	0	0	G	>	≤	0	>	>	≤	≤	>	≤	≤	≤	≤
20	**m**	>	≤	≤	G	≤	0	≤	>	0	0	≤	>	≤	0	0
21	**n**	≤	0	0	≤	G	0	≤	≤	>	>	≤	0	>	>	0
22	**ŋ**	0	0	≤	0	0	G	0	0	0	0	≤	≤	0	0	≤
23	**ʃ**	>	≤	≤	>	>	0	G	>	>	≤	>	≤	0	0	0

A very interesting observation could be made in the above frequency tables, particularly, in the lateral /l/ columns on both X and Y hands. It is evident from the data presented in the table that the lateral /l/ typically disagrees to go with any coronal in Bangla. But, there are a very few exceptions as well as one can notice that not all the cells show null frequencies in the above tables.

Let's look at the frequency of occurrences if /l/ would have created any cluster with a coronal. In (10a), the first member of the cluster is /l/ and in (10b), it is at the final position. Unlike the frequency marking presented in the previous section, here we will take the three layer counting. In the following illustration, the tick (✓) marked clusters denote that the corresponding cases have ten or less (but more than zero) occurrences whereas the **X** marked cases represent a cluster with zero occurrences in the corpus. An exceptional case (19g) representing more than ten cases is shown as ✓* in this illustration.

(53)

a.	l+Coronal		b.	Coronal+l	
Cell	Cluster	Frq	Cell	Cluster	Frq
19a	/lr/	**X**	2c	/rl/	✓
19b	/lɽ/	**X**	3c	/ɽl/	**X**
19e	/ln/	✓	6c	/nl/	✓
19g	/lʃ/	✓*	7c	/ʃl/	✓
19i	/lt(^{h})/	✓	10c	/t(^{h})l/	✓**(X)**
19j	/lʈ(^{h})/	✓	11c	/ʈ(^{h})l/	**X (X)**
19m	/ld(^{h})/	✓	14c	/d(^{h})l/	**X (X)**
19n	***/ld(^{h})/***	✓	***15c***	***/ɖ(^{h})l/***	✓ *(X)*

The above list shows that in Bangla, lateral-coronal or coronal-lateral clusters at word medial position is very restricted as they never count for more than ten cases (actually very few in number). There is, however, one exception (19g) where it is more than ten. From the data presented in (10), it is evident that /lʃ/ cluster (cell: 19g) occurs more frequently in this position. This exceptional phenomenon is explained afterwards.

Word medial consonant clusters are the most important and practically a very complex set of phenomena in the syllable structure of Bangla. In the next chapter, we will discus about the gemination issue followed by word initial and final clusters along with some other phonological issues.

4 Gemination

A large number of Sanskrit words are integrated in Bangla lexicon over a long period of time. These words typically belong to the SB stratum, as termed in the present work. The syllabic pattern of many of these words is changed after their integration in the Bangla vocabulary. There are different types of sound change that are available throughout the process of borrowing from other languages including Sanskrit. One of such changes is the gemination process, where a consonantal sound reduplicates inside a word and sounds audibly longer period of time compared to the original form. Interestingly, these words retain their original orthographic mapping in the target language even after the phonological changes in the spoken form.

There are many instances of geminates in the corpus study presented in this work. Earlier studies show that only a few languages allow all types of geminates, but actually there is a significant amount of study still needs to be done in this area. In this regard, some significant work is done in the case of voiced obstruent geminates which reveals that voiced obstruent geminates are cross-linguistically disfavored, because with long closure, it is difficult to maintain a transglottal air pressure drop sufficient to produce voicing (Hayes & Steriade 2004; Jaeger 1978; Kawahara 2005; Ohala 1983; Taylor 1985). For instance, Luganda[13] has no nongeminate coda (McCarthy 2003).

[13] Luganda belongs to the Bantu branch of the Niger-Congo language family. It is spoken by more than three million people in Uganda (Africa).

In a C_1C_2 structure, the loss (or merger) of C_2 in the geminated form is subsidized as the former consonant (C_1) reduplicates to compensate the loss (or merger) of the following sound. In other cases, the structure could be C_1V_2 as well (V stands for a vowel here), where V_2 denotes the presence of a semi-vowel, assuming that this cluster is located word medially and hence V_1 is taken as the peak of the first syllable. There is a third type of gemination found in Bangla lexicon where no sound is dropped or merged, but the first consonant is geminated keeping the following sound intact. This type of gemination is seen when an obstruent is followed by a liquid (/r/ or /l/) sound in the SB stratum words. Though, all the above said combinations are found in Bangla, but depending on the sound preceded by the consonant C_1, the geminates available in Bangla could be divided into three categories as shown in (1). All the geminate cases are available in the word medial position, where, in Sanskrit, the following conditions prevail.

(1) Gemination process in Bangla (SB and OB)

i. Obstruents followed by an semi-vowel

($C_1C_2 > C_1C_1$, where $C_1 =$ obstruent and $C_2 =$ Semi-vowel)

ii. Obstruents followed by a liquid

($C_1C_2 > C_1C_1C_2$, where $C_1 =$ obstruent and $C_2 =$ liquid)

iii. Plosives followed by a nasal /m/

($C_1C_2 > C_1C_1$, where $C_1 =$ Plosive and $C_2 =$ /m/)

Among these three distinctive types of gemination in Bangla, the third one is already discussed in the previous chapter. A markedness constraint *Plosive-/m/, along with some faithfulness constraints could explain the behavior of this phonological process.

4.1 Gemination with semi-vowels

The very first type in (1) represents a good number of SB words which lost their post-consonantal semi-vowels in the spoken form of Bangla (SCB, in particular) when borrowed from Sanskrit. Note that, such changes are not visible in some other Indian languages such as Hindi, Gujarati, Marathi etc., which are like Bangla, also closely related to Sanskrit. These languages typically try retaining the original phonological forms of Sanskrit even after the bowing process is over. Sanskrit has two distinct semi-vowels (also called glides in some earlier literature): labio-dental semi-vowel [ʋ] and palatal semi-vowel [j]. Bellow is a set of examples where type (i) gemination process takes place with labio-dental [ʋ]. This observation is also supported by several literary works (see Chatterji 1926a, 1926b, 1988; Dey 1979; Singh 1980 a.o).

(2) Gemination: Consonant followed by a labiodental semi-vowel in Bangla (SB)

a. *San.* साध्वी /sad^{h}.ʋi/ > *Ban.* সাধ্বী [sad^{h}.d^{h}i] ‘faithful wife’

b. *San.* पृथ्वी /pr̩t^{h}.ʋi/ > *Ban.* পৃথ্বী [prit^{h}.t^{h}i] ‘earth’ (as in *‘prithviraj’*)

c. *San.* विश्वास /bis.ʋaś/ > *Ban.* বিশ্বাস [biʃ.ʃaʃ] ‘faith/belief’

(3) Gemination: Consonant followed by a palatal semi-vowel in Bangla (SB)

a. *San.* सत्य /satya/ > *Ban.* সত্য [ʃot.to] ‘truth’

b. *San.* बाल्यकाल /bālyakāl/ > *Ban.* বাল্যকাল [bal.lo.kal] ‘childhood’

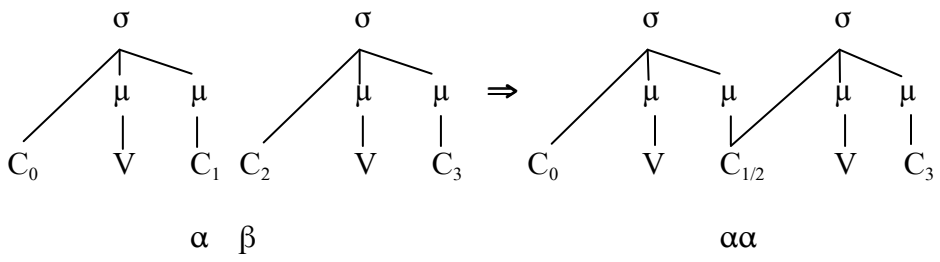
c. *San.* मृत्यु /mṛtyu/ > *Ban.* মৃত্যু [mrit.tu] ‘death’

d. *San.* बश्यता /bashyatā/ > *Ban.* বশ্যতা [bɔʃ.ʃo.ta] ‘domesticize’

e. *San.* पुण्य /punya/ > *Ban.* পুণ্য [pun.no] ‘virtue’

There are two possible gemination processes for this type of words that could be explained in the present section. These are either a merger of C_1 and C_2 into C_1 or the deletion of C_2. It is generally assumed that the deletion process is not favored in many languages. We will consider both possibilities in the present study and see what be the outcome. The above mentioned gemination (semi-vowels) cases could be structurally illustrated as follows. First, the more favored merger is shown in (4) below.

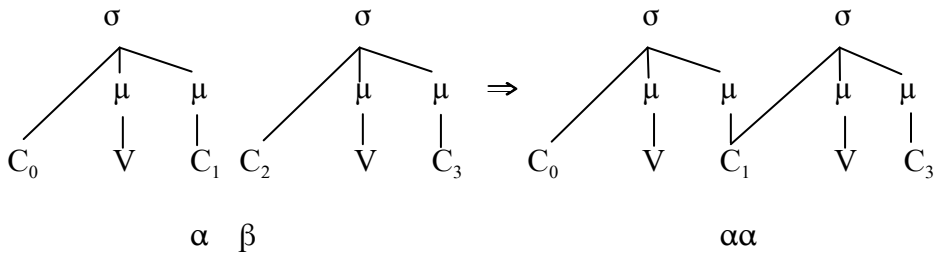
(4) Geminate formation 1: Merger



In this gemination process, C_1 is merged with C_2 and constitutes the geminated form C_1C_1 which is shown as $C_{1/2}$ in the above illustration. That means the merged place feature is shared by the coda of the first syllable and the onset of

the following syllable. This formation would give a [C_1C_1] geminate in the surface structure. Another source of gemination is illustrated bellow where the coda of a syllable shares its position with the onset of the following syllable after the deletion of the original value of that very onset. This is not a case of merger, rather an issue where a segment of the syllable is deleted to make way for a geminate.

(5) Geminate formation 2: Deletion



As we can see in (5), the onset (C_2) of the second syllable is dropped and the coda (C_1) is shared at that position (onset). In general this type of gemination is not favored in many languages. In the OT framework, there could be a constraint favoring this type of gemination as it shows coalescence in the syllable structure. The faithfulness constraint UNIFORMITY (see McCarthy & Prince 1995, cf. Causley 1997; Gnanadesikan 1995, 1997; Keer 1999; Lamontagne & Rice 1995; McCarthy 2000; Pater 1999) could be used in order to dispose of certain codas by coalescence with a following consonant (violation of UNIFORMITY). Since this constraint blocks merger (but allows deletion), it must be ranked lower than markedness constraints for gemination effect to take place.

(6) UNIFORMITY (McCarthy & Prince 1995)

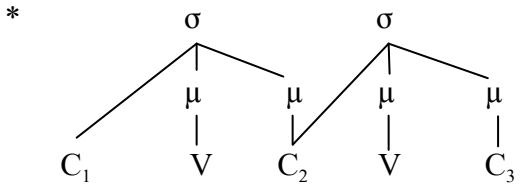
No element of the output has multiple correspondents in the input.

(No coalescence)

The constraint UNIFORMITY punishes the mapping of a pair geminate onto a single segment. This constraint punishes any candidate where two elements are merged (fusion) into one from the input form and form a geminate. In the following figure, this type of fusion occurred where the nearby consonants C_1 and C_2 are merged as $C_{1/2}$ in the later stage. UNIFORMITY does doesn't like this type of merger though.

A constraint prohibiting any kind of gemination in syllable is used in some previous studies (see Al-Ahmadi Al-Harbi 2003; Baković 2005; Hall 2003; Ham 1998; Murray & Vennemann 1983; Shinohara 2002). In many literary works it is called NoGem; we will use the same name in the present study as well. This constraint requires every root node to be linked either to a mora or to an onset. This concept is illustrated below.

(7) No gemination structure

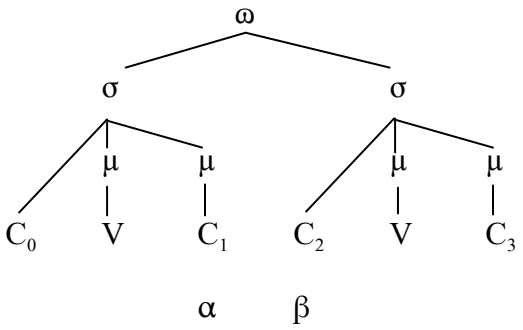


(8) NOGEM (Hall 2003)

No multiple links from a root node to a higher tier.

The well known syllable contact law is a relevant issue in the current section of this work, as it maintains the sonority of the syllable edge intact across the syllable boundary (Davis 1998; Gouskova 2002; Hooper 1976; Murray & Vennemann 1983; Rose 2000; Vennemann 1988). In a disyllabic word, if the coda of the first syllable is α and the onset of the second syllable is β, then the sonority of β must not be greater than α. This law is illustrated in the following moraic structure.

(9) Syllable contact principle



then, $son_{\alpha} \geq son_{\beta}$

The syllable contact law could be represented as a constraint to account for a possible sonority change (rise) from coda to the following onset.

(10) SYLCONTACT (Gouskova 2002)

Sonority must not rise across a syllable boundary.

In order to ommit a consonant or vowel which is not favored in a syllable structure, two faithfulness constraints could be introduced here. Both of these constraints belong to the same family of MAX.

(11) MAX-IO-C

[+con][14] sounds in the input must have output correspondence.

(12) MAX-IO-V

[-con] sounds in the input must have output correspondence.

The faithfulness constraint MAX-IO-C requires every consonant of the input form to be preserved in the output. On the other hand, MAX-IO-V wants the same for every vowel in the input. These MAX constraints will replace the general MAX-IO constraint which is used in this study so far. Additionally, we will not consider DEP-IO in this section (and thereafter), since it is not so crucial in the improvement of the analysis. Instead, we would recall the complex onset related markedness constraint *COMPLEX^{ONS}. This constraint does not prefer consonant clusters at the onset position.

(13) *COMPLEX^{ONS} (Kager 1999)

Onsets are simple.

Now, let us consider (2c) /biʃʊaʃ/ from the example set to account for an OT analysis based on the above mentioned gemination structures. Let us mark the consonant-semivowel cluster /$ʃ_1ʊ_2$/ as /$ʃ_\alpha ʊ_\beta$/ for the simplicity of explanation.

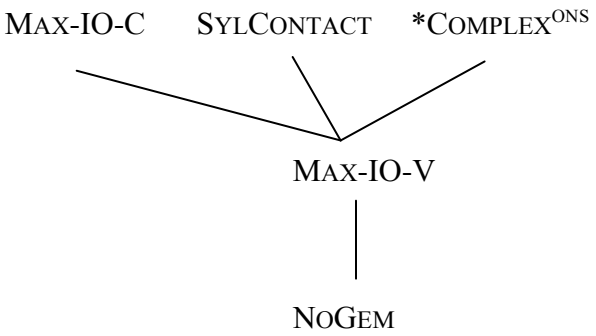
(14) Gemination (SB): with a semi-vowel as C_2

[14] [+/-con] denotes the presence or absence of a consonantal sound.

/biʃ$_{\alpha}$ʊ$_{\beta}$aʃ/	MAX-IO-C	SYL CONTACT	*COMPLEX^{ONS}	MAX-IO-V	NOGEM
1. biʃ$_{\alpha}$.ʊ$_{\beta}$aʃ		*!			
☞ 2. biμʃ$_{\alpha}$aʃ				*	*
3. bi.ʊ$_{\beta}$aʃ	*!				
4. biμʃ$_{\alpha}$ʊ$_{\beta}$ʃ			*!		*

In this analysis, there are two different geminates contesting to win. The above tableau has candidates 2 and 4 with geminated forms which violate the NOGEM constraint once each. But, none of them is ruled out by this violation, because of the low ranking of NOGEM. But, candidate 3 is actually ruled out by the violation of the faithfulness constraint MAX-IO-C, for the deletion of the consonant [ʃ]. On the other hand, candidate 2 violates the other faithfulness constraint MAX-IO-V for a vowel deletion, but still became the minimal candidate for the comperatively low ranking of MAX-IO-V. The input form (candidate 1) has a fatal violation of constraint SYLCONTACT for the rising sonority of the β place from α. Candidate 6 is punished by *COMPLEX^{ONS} which is actually ranked low in SB stratum words. The winning candidate /biμʃ$_{\alpha}$aʃ/ ([biʃʃaʃ]) is a geminated form of the corresponding non-geminated Sanskrit word and an acceptable phonological word-form in Bangla vocabulary. The constraint ranking of this OT account would be as follows.

(15) Constraint ranking for gemination (obstruent-semivowel)



The first three of the five conflicting constraints are not ranked crucially, but they dominate the faithfulness constraint MAX-IO-V. In turn the low ranking anti-gemination constraints NOGEM is dominated by this faithfulness constraint.

4.2 Gemination with liquids

As we have seen in the word medial cluster combinations, voiced plosives are geminate before a liquid (/r/ or /l/), keeping the latter sound intact. That means the adjacent consonants C_1 and C_2 forms a geminate as $C_1C_1C_2$. But, this rule actually extends throughout the obstruent column, and is not limited only to voiced obstruents.

(16) Obstruent followed by liquid

		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
		r	ʈ	l	m	n	ɳ	ʃ	p(h)	t(h)	ʈ(h)	k(h)	b(h)	d(h)	ɖ(h)	g(h)
1	p	>	≤	>	≤	≤	0	>	G	>	>	0	0	0	0	0
2	t	>	≤	≤	0	≤	0	>	>	G	0	≤	0	0	0	0
3	ʈ	>	0	0*	0*	≤	0	0*	≤	≤	G	≤	0	0	0	0
4	k	>	≤	>	0*	≤	0	>	0	>	>	G	0	0	0	0
5	b	>	≤	≤	0	≤	0	≤	0	0	0	0	G	>	*0	0
6	d	>	0	*0	≤	≤	0	*0	0	*0	0	*0	≤	G	0	≤
7	ɖ	>	≤	≤	≤	*0	0	0	0	0	0	0	≤	≤	G	0
8	g	>	≤	≤	≤	>	0	*0	0	0	0	0	≤	≤	0	G
9	p(h)	≤	0	*0	0	0	0	*0	G	*0	*0	0	0	0	0	*0
10	t(h)	≤	0	*0	0	0	0	0	0	G	0	0	0	0	0	0
11	ʈ(h)	0	0	0	0	0	0	0	0	0	G	0	0	0	0	0
12	k(h)	≤	0	0	0	≤	0	0	0	0	0	G	0	0	0	0
13	b(h)	>	0	0	0	0	0	0	0	0	0	0	G	0	0	0
14	d(h)	≤	0	0	0	0	0	0	0	0	0	0	0	G	0	0
15	ɖ(h)	≤	0	0	0	0	0	0	0	0	0	0	0	0	G	0
16	g(h)	≤	0	0	0	≤	0	0	0	0	0	0	0	0	0	G

It is clear that, only except one case (*cell* 11a: /ʈ^{h}r/), all the clusters in this column have more than zero occurrences in the B-lexicon. That means the obstruent-liquid (/r/) clusters occur in geminated form in all possible cases in this category. On the other hand, the (c) column records instances only for unaspirated obstruents followed by /l/. As mentioned in the previous chapter, such sound combinations where $C_1C_1C_2$ type gemination occurs are found only in the SB and

OB stratum of the B-lexicon. The NB words do not allow any gemination of this sort. Actually, there are a very few words present in the NB stratum that have such sound combinations and when the gemination issue is relevant, they refuse to change their structure.

Orthographically, all these combinations are represented in the written form of Bangla using "phala", a distinct feature of Bangla orthography. While used as r-phala, the obstruent and the following /r/ sound creates a ligature so that the regular form of the /r/ sound in Bangla is replaced by a special r-phala symbol. For instance, para = পর, but, pra = প্র. Sarkar 2006 includes also "l-phala" in this type of gemination, though formally there is no evidence of "l-phala" in Bangla orthography. Let us list down some examples of this type of cluster, both from SB and OB strata.

(17) Gemination: Obstruent followed by liquid in Bangla

Sanskrit Borrowings (SB)

a. *Ban.* পুত্র /putra/ [put.tro] 'son' (< Sanskrit)

b. *Ban.* শুভ্র /sub^{h}ra/ [ʃub^{h}.b^{h}ro] 'white/bright' (< Sanskrit)

c. *Ban.* আপ্লুত /apluta/ [ap.pluto] 'inundated' (< Sanskrit)

Other Borrowings (OB)

d. *Ban.* সুপ্রিম /suprim/ [ʃup.prim] 'supreme' (< English)

e. *Ban.* সাপ্লাই /saplai/ [ʃap.plai̯] 'suply' (< English)

f. *Ban.* মাদ্রাসা /madrasa/ [mad.dra.ʃa] '(Islamic) school' (< Arabic)

It should be noted that this type of gemination occurs only when the adjacent consonants (αβ) are preceded by a vowel. Thus, 'Cαβ' clusters do not show such a gemination in Bangla. For instance, ইলেক্ট্রিক /ilekṭrik/ [ilek.ʈrik] 'electric' does not register such gemination at the /ʈr/ cluster, even though it belongs to the OB stratum. On the other hand, it is already mentioned that words from NB category strictly do not follow this gemination pattern at all.

(18) Word medial obstruent-liquid clusters belonging to NB stratum

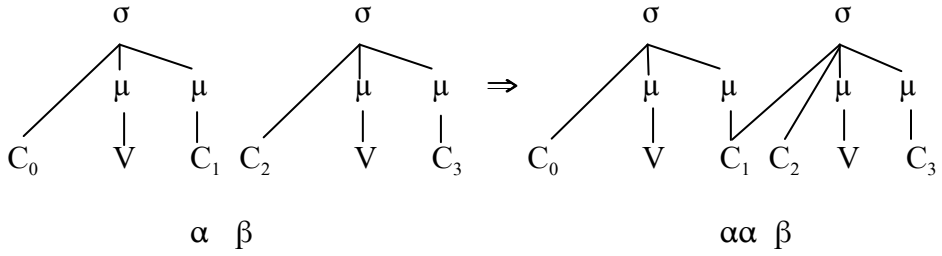
a. *Ban.* সাঁতরা /sãtra/ [ʃãt.ra] 'a Bengali surname'

b. *Ban.* বাবরি /babri/ [bab.ri] 'long curing hair-style'

c. *Ban.* শাপলা /sapla/ [ʃap.la] 'water lily'

In the following figure (19), at the gemination stage, $C_2(\beta)$ remains untouched, but $C_1(\alpha)$ is being shared by the coda of the first syllable and the first segment of the complex onset of the following syllable. SB and OB class words fall in this structure, but NB words do not allow any complex coda.

(19) Geminate formation 3: WGG



This type of gemination has a close resemblance with a very well known process of language change in Germanic language family called West Germanic Gemination (WGG). As claimed in Hall 2003, that a post-short vowel consonant is a geminate in West Germanic languages before the palatal semi-vowel (glide) /j/. The said gemination process is attested in some West Germanic languages, such as Old English (OE), Old High German (OHG), Old Saxon (OS), whereas it is absent in North Germanic (= Old Norse (ON)) or in East Germanic (= Gothic (Got.)). A set of data showing WGG (taken from Hall 2003 and Simmler 1974) is illustrated below.

(20) West Germanic Gemination (WGG) data (from Simmler 1974)

East/North Germanic	West Germanic	Gloss	Geminate
a. *Got.* Skapjan	*OS* skeppian, *OE* scieppan	'to create'	[pp]
b. *Got.* hafjan	*OHG* heffan, *OE* hebban	'to lift'	[ff]
c. *ON* framja	*OHG* fremmen, *OE* fremman	'to carry out'	[mm]
d. *Got.* halja	*OS* hellia, *OHG* hella	'hell'	[ll]

The contrast between the nongeminate forms in the East or North Germanic languages (first column) and the corresponding cases with geminates in the West Germanic laguages (second column) is claimed to be a historical development. It is postulated that in Early West Germanic languages, all of the above listed cases (and many more) were originally VCjV (where j is the palatal glide) and that a historical

process of the form VCjV > VCCjV occurred before the West Germanic daughter languages emerged from this family.

But, one notable difference is that, in WGG, mostly voiceless obstruents are geminated when preceded by a liquid (Gaeta 2001), while Bangla typically allows such geminates for all obstruents irrespective of their voicing. This type of sound change could be explained by the so called Preference Laws for Syllable structure (Murray & Vennemann 1983, cf. Vennemann 1988). The idea of this type of sound change was determined by the influence of marked syllable structure as a consequence of the Germanic preference of the bimoraic stems as asserted by Murray and Vennemann. Bimoraic stem structure is also preferred by Bangla syllables. Hence, a similar mechanism may also be applied to analyze the said gemination process in the present

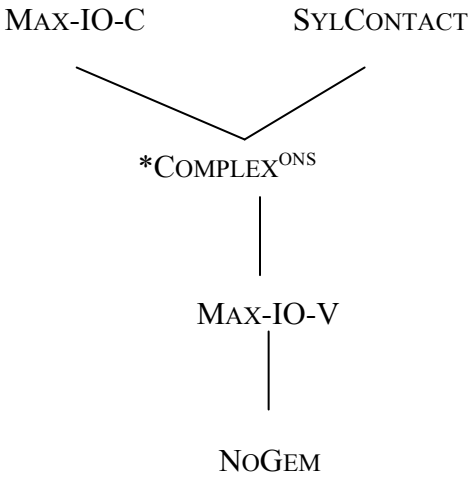
Let us take a word from (17a) *Ban.* পুত্র /putra/ [put.tro] 'son', which belongs to the SB stratum of Bangla vocabulary. The same constraints of the previous section are also used in the following tableau, but the mutual ranking of UNIFORMITY and NoGem is ommited here.

(21) Gemination (SB): with a liquid as C_2

	$[put_1r_2o]$	MAX-IO-C	SYL CONTACT	*COMPLEX^{ONS}	MAX-IO-V	NOGEM
	1. $put_1.r_2o$		*!			
☞	2. $pu_\mu t_1r_2o$			*		*
	3. $pu_\mu t_1o$	*!				*
	4. put_1o	*!				

In (21) the input form (cand 1) violated the SYLCONTACT constraint by allowing a word medial cluster [tr] which registers a rising sonority pattern. This constraint punishes such phonological feature across syllable boundaries. The winning candidate (cand 3: [$pu_{\mu}t_1r_2o$]) satisfies top ranked constraints MAX-IO-C and SYLCONTACT, but fails to satisfy *COMPLEX^{ONS} and the anti-gemination constraint NoGem. But, the lower ranking of these constraints could not rule out candidate 3. The last two candidates simply violate the faithfulness constraint MAX-IO-C while candidate 3 violates another faithfulness constraint NoGem. This explains the double linking issue with a liquid sound in the cluster. Now, the modified consonant ranking for gemination in Bangla would be as followed.

(22) Constraint ranking for gemination (obstruent-liquid)



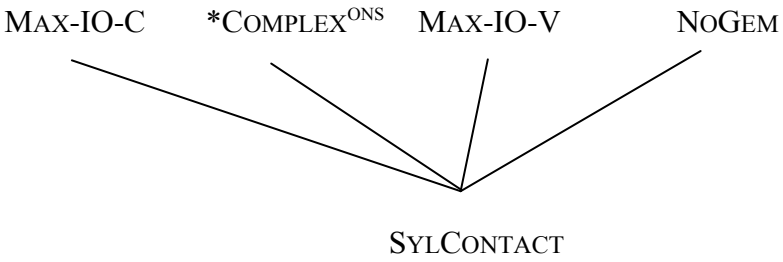
This constraint ranking, however, can not account for NB category words. Hence, the constraint ranking should be changed for this stratum. Let us take a word of such a cluster from the NB stratum: (18c) শাপলা /sapla/ [ʃap.la] 'water lily'.

(23) No gemination (NB): with a liquid as C_2

	[ʃap$_1$l$_2$a]	MAX-IO-C	*COMPLEX^{ONS}	MAX-IO-V	NOGEM	SYLCONTACT
☞	1. ʃap$_1$.l$_2$a					*!
	2. ʃaμp$_1$l$_2$a		*!		*	
	3. ʃaμp$_1$a	*!			*	
	4. ʃap$_1$a	*				

The input is the optimal candidate in the above tableau for its least amount of violation (only the lowest ranked SYLCONTACT). The first four constraints MAX-IO-C, *COMPLEX^{ONS}, MAX-IO-V and NOGEM, eventually rule out all other candidates for fatal violations. The major change in this tableau from the last one is in the ranking of the constraints. The crucially ranked *COMPLEX^{ONS} and NoGem in the previous tableau no more maintain that ranking. It is now ranked with the other Faith (Max-IO-C, MAX-IO-V) constraints. On the other hand, SYLCONTACT is now ranked crucially with the other four constraints, which was mutually ranked with Faith in the earlier tableau.

(24) Constraint ranking for gemination (obstruent-liquid): NB



The OT analysis for the gemination process given in this chapter covers

almost all possible gemination cases in the Bangla vocabulary. One must be careful about the stratification in Bangla lexicon (here, termed as B-lexicon), because, gemination may or may not occur in certain clusters, depending on their stratum. From the above discussion, we can draw a factorial typology with the relevant constraints, taking DEP-IO and MAX-IO in one group: FAITH.

(124) Rankings for strata:

a. FAITH, SYLCONTACT » *COMPLEX^{ONS}, NOGEM

'Gemination in Obstruent-semivowel cluster (SB, OB)'

b. FAITH, SYLCONTACT » *COMPLEX^{ONS} » NOGEM

'Gemination in Obstruent-liquid cluster (SB, OB)'

c. FAITH, *COMPLEX^{ONS}, NOGEM » SYLCONTACT

'No gemination in Obstruent-liquid cluster (NB)'

5 Morphology

5.1 Introduction to Distributed Morphology

The study of linguistics in general, is always very much diversified by different schools of thought. All the branches of linguistic studies have a little theoretical agreement on many issues. Morphology, being one of the core branches of linguistics is no exception and actually has a number of different approaches for every element of it. One of them is the place of the affixes in morphology. Lieber 1992 endorses the traditional lexicalist approach that affixes are lexical (or vocabulary) items, which take important role in creating words and operate in the syntax. In contrast, Anderson 1992 states the *a*-morphous or affixless theory where the word-formation rules (which are by nature morphophonological rules) are the basic processes underlying morphology whereas affixes are considered as mere epiphenomena. But Halle & Marantz 1993 defend a third theory of morphology, called Distributed Morphology, which actually combines some features of the above mentioned lexicalist and a-morphous alternatives.

The original inspiration of this work was the notion and morphological events in a particular language can be analyzed from many different points of view that could allow us to choose between various theoretical frameworks. If one theory could reduce the burden of explanation placed on general principles such as analogy or reanalysis by providing specific motivation for certain types of morphological events, this fact would suggest, all else being equal, that it is a better theory than its

competitors (within a given theory as well as those relying on a different theoretical framework). In particular, I believe Distributed Morphology is quite a strong theory in the realm of the study of linguistics where one can analyze most of the significant morphological events in languages in a well-structured manner.

The key motivation of this work is to analyze certain aspects of Bangla verbal morphology in Distributed Morphology (henceforth DM) framework. But the DM structure is not strictly followed here to show some of the phonological operations. In the DM framework, the scope of the analysis of phonological operations is restricted, as the phonological elements are introduced to the morphological structure at the final stage of the whole syntactic operation. After the syntactic operations are over, a thorough analysis of the phonological events are presented here using the moraic structures.

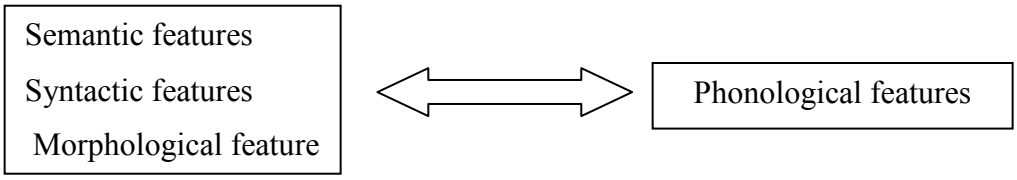
This work has two interrelated goals. First, I hope to contribute to the theory of DM by providing complete analyses of the verbal inflections of two specific types of forms present in Bangla and examining the certain theoretical issues in this context. Second, I will explore interactions between the relationship of some of the morphological forms and their phonological characteristics. I have taken a particular portion of Bangla verbal inflection to show how the components of inflected forms act in the area of verb morphology. The components of each inflected form can be separated in different "slots" and according to their position and relation with neighboring slots; they are termed as Base, Tense AGR etc. It is the fundamental structure of DM that there should be more than one candidate to be inserted at every slot. I have grouped the candidates according to the possible slots where they are

eligible to be inserted. Every element is shown along with their feature sets so that their occurrence in the particular slot could be well-defined. This system works as a computer program where all the elements are structured in a particular manner what makes the computer to run the program perfectly. In a similar manner, the elements in the above said groups are structured in such a manner that they should act without any exception.

5.1.1 The theory of Distributed Morphology

The distributed morphology theory of Halle & Marantz 1993 will be the basis of this study. In Distributed Morphology (DM), the basic unit is the Vocabulary Item as shown below.

(1) DM structure



The terminal nodes in DM framework contain three units, namely, semantic, syntactic and morphological features; but no phonological feature. In other morphological theories, these vocabulary items have morphemes or lexical entries as their counterparts (Halle & Marantz 1993). According to Noyer & Harley 1999, "Vocabulary items provide the set of phonological signals available in a language

for the expression of abstract morphemes. The phonological content of a vocabulary item may be any phonological string, including zero or null-Ø". Now, there are three major properties of the vocabulary items collectively, which distinguish the theory of DM from other approaches of morphology. These are:

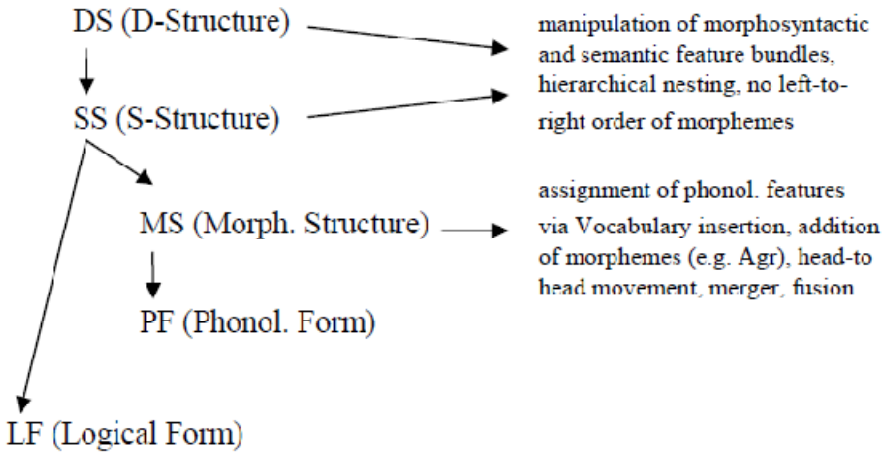
i) *Late Insertion*: Phonological features are supplied by the insertion of vocabulary items into the terminal nodes and only after the syntactic operations. As the name denotes, late insertion contrasts with "Early Insertion" which takes place in some other theories, where the features of the words are formed by the contribution of the features of the lexical entries (such as Lexicons). This operation happens at the syntax level itself.

ii) *Underspecification*: Among several vocabulary items available for insertion, only the highest specified one whose identifying features are a subset of the features of the terminal node is eligible for insertion, i.e., it wins the competition to be eligible for insertion. In most cases, several features are available to be chosen for a specific terminal node.

iii) *Syntactic hierarchical structure all the way down:* In DM, elements within syntax and within morphology enter into the same types of constituent structures and can be diagrammed through binary branching trees. The following diagram from Glück & Pfau 1999 illustrates the five-level conception of grammar adopted by Halle & Marantz 1993 and the sequence of operations to take place.

One can understand that the late insertion has become mandatory by underspecification because the required features that are needed in the syntax and Logical form (LF) may be left unspecified in a specific vocabulary item (Halle & Marantz 1994).

(2) D-structure and S-structure



These features jointly fabricate an important theory in morphology by which we can now analyze the structure of our focus language *Bengali* from a non-traditional point of view.

The DM framework approbates an interesting feature called late insertion, where the terminal nodes get the phonological contents only after the required syntactic operations are complete. This phenomenon happens at the level of morphological structure that represents the surface structure and the phonological form (PF). However, this approach does not follow the standard assumption that the morphemes are not separate entities, but are combined in the lexicon and inserted into the terminal nodes prior to the syntactic elements. Since syntax is the input to

morphology, the DM approach makes a highly economical assumption that morphology simply preserves the structures provided by the syntax unless processes specific to Morphological structure alter these ones (Embick & Noyer 2001). Hence, DM accounts for Anderson's (Anderson 1992:70) dissent that, in some cases, "the one-to-one relation between components of meaning and components of form which is essential to the classical morpheme is violated", which, as a matter of fact, does not cast aside the proposal of Baker 1985 that morphology "mirrors" syntax.

5.2 Structure of Bangla verbal system

5.2.1 Personal Pronouns

A high morpheme-to-word ratio is found where each affix typically represents one unit of meaning (such as number, tense, person etc.). According to Klaiman 1987, morphology in modern Bengali is non-existent for adjectives, minimal for nouns and very productive for verbs. The most interesting area of Bengali morphology is the derivation of inflecting stems from verbal bases. There are mainly two types of verbal bases in Bengali: monosyllabic (e.g., /ken-/ 'buy') and disyllabic (e.g., /kāmɽaā-/ 'bite') with the majority of the former. Here two monosyllabic verbal bases are discussed (one vowel-ending and another consonant-ending).

According to Dasgupta 2003 Bangla has no case or number agreement as well as no grammatical gender phenomena at all. The sole parameters for Bangla subject-verb agreement are persons (1st, 2nd and 3rd) and the status of the person in terms of the degree of formality (applicable for the second and third person). The

formality level is shown as F (Formal), P (Polite) and I (Intimate) for the second person and only F and P for the third person. So the personal pronouns are: /apni/ —You (F), /tumi/ —You (P), /tui/ —You (I), /tini/ — He/She (F) and /ʃe/ —He/She (P). The third person shows only one form for both polite and intimate forms.

T. Bhattacharya 2000 and Dasgupta 2003 termed /tumi/ —You as 'neutral' instead of 'polite'. But it should not be so controversial to use the term 'polite' in this case. Actually, there is no standard terminology for the levels of formality in Bangla. Even the third level of formality, i.e., 'Intimate', is not out of the debate. Some linguists term it as 'Informal' while some others call it 'Intimate'. Anyways, the central point of this discussion is to term the level of formality in Bangla and I think the terms mentioned in this study are clear enough to represent the forms.

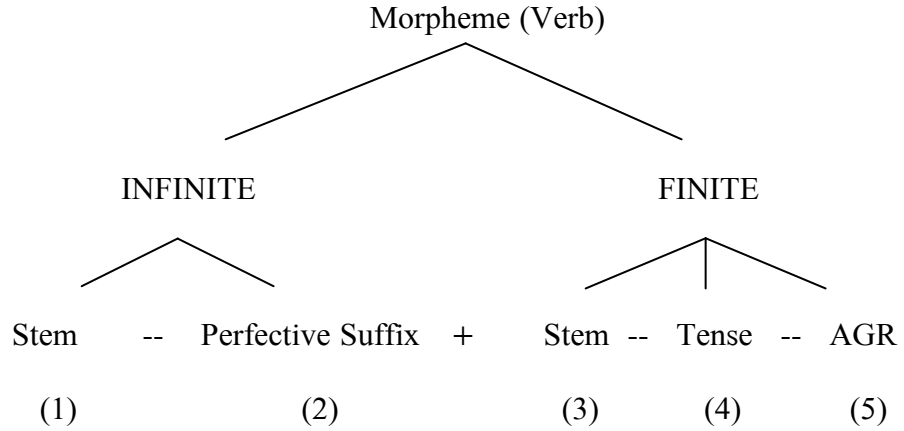
We can compare these forms of Bangla with German *"Sie"* (You -- formal) and *"Du"* (You – informal). However, there are only two levels of politeness in German while Bengali possesses three. English does not exhibit any such divisions in the morphology. On the other hand, most of the Indian languages, such as, Gujarati, Marathi, Assamese, Hindi, Punjabi etc. have this kind of formality levels.

5.2.2 Morpheme (Verb) Structure

In Bengali, each personal pronominal form has its own set of verbal inflection. But all the verbal inflections have a general structure. It contains two parts, viz., infinite and finite, and then the basic elements of any inflected form, such as, stem, tense and AGR take place under them. The finite part has all the three elements in a

particular order (stem – tense – AGR), while the infinite part lacks the tense slot. This structure is not an absolutely unquestionable one, but most of the inflected forms of the verb fit in here. It is illustrated as a tree diagram below.

(3) Bengali morpheme (Verb) structure



The right-hand side of 'Morpheme (Verb)' is the base structure present in all verbal forms. However in some cases the left-hand side structure appears as an additional element. Usually the two can be distinguished as the left becomes the infinite verb and the right as the finite verb. Bengali has both finite and infinite verb forms; but sometimes they appear together. The infinite form is the main verb and finite form is an auxiliary verb there.

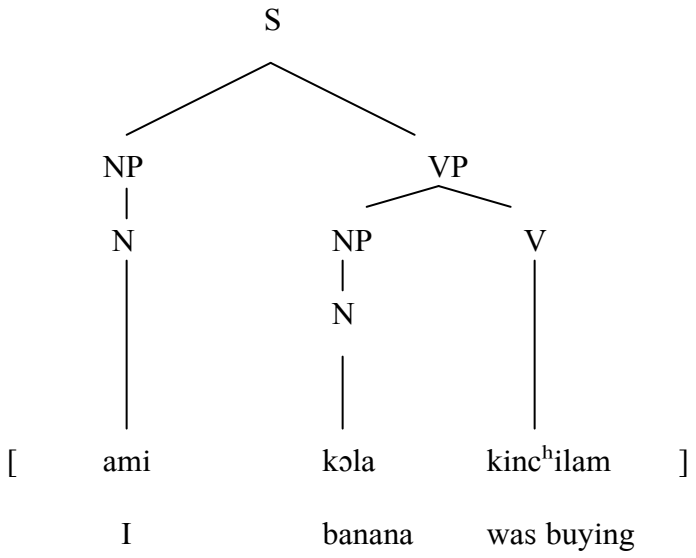
K. Bhattacharya 1993 describes the only auxiliary verb in Bangla /-$c^h(c^h)$-/ as a derived form of the main verb /ac^h-/ (to be) with the loss of the initial stress. For instance we may consider the past perfect (PP) forms in *table 9.* The third, fourth and fifth (i.e., last three) parts of all the forms in that column show a finite form of / ac^h-/ (such as [-c^hilam] in the first person). But these are finite forms only

when they are taken separately. In that case, /kin-/ and /k^{h}a-/ are infinite forms as we see in case of /kin-/ (on the left hand side) in (8). But together they constitute the finite form [kinec^{h}ilam] (have bought) and so on. There are more complicated issues regarding the finite and infinite forms of Bangla verbs which will be brought forward according to the requirements of this discussion.

In this connection, a very brief syntactic structure of Bangla is illustrated here. Bangla projects the syntactic components like V(Verb), N(Noun), A(Adverb) etc. in head-final word order (Fitzpatrick-Cole 1996; Hayes & Lahiri 1991; Truckenbrodt 2002; Vijayakrishnan 2003 a.o.). Let us consider the verb form [kinc^{h}ilam] in a sentence and its syntactic structure:

(4) Syntactic structure (Bangla)

[ami kɔla kinc^{h}ilam] -- *'I was buying banana'*



The structure shows the head-final VP of a common sentence in Bangla. However, our main aim is not the sentence structure, but the syntactic structure of

word-forms. In the next section, we will see different word-forms of Bangla and their syntactic structure in the DM manner.

5.3 Data

5.3.1 Verbal Inflection

Campbell 1995 says that, in Bangla, number is not marked, i.e., the same verbal form is used for both singular and plural and the aspect or tense markers are added to the stem before the personal inflexions. In the following tables these stems, markers and other affixes are separated from each other by "-" (hyphen) marks.

Every column in the following tables shows a different tense form. In order to fit all the columns in these tables, only abbreviations are used here: SPR (Simple Present), PRI (Present Imperative), PRC (Present Continuous), PRP (Present Perfect), PC (Past Continuous), PP (Past Perfect), PHC (Past Habitual Conditional), SP (Simple Past), SF (Simple Future) and FI (Future Imperative). Every cell in (5) and (6) contains two forms. The first (upper) one is the inflected form of /k^{h}a-/ 'eat'. The later (lower) one is for /ken-/ 'buy' and given for an instant comparison. Note that the imperative cannot be applied to the first person; hence, those cells remain empty.

(5) Bangla Verbal Inflection for the base /kha-/ and /ken-/ (Present)

Tense / *Number*		SPR	PRI	PRC	PRP
1P		kʰa-Ø-i̯ kin-Ø-i	----	kʰa-cʰ-cʰ-Ø-i kin-Ø-cʰ-Ø-i	kʰe-e̯e-cʰ-Ø-i kin-e-cʰ-Ø-i
2P	F	kʰa-Ø-n ken-Ø-en	kʰa-Ø-n kin-Ø-un	kʰa-cʰ-cʰ-Ø-en kin-Ø-cʰ-Ø-en	kʰe-e̯e-cʰ-Ø-en kin-e-cʰ-Ø-en
	P	kʰa-Ø-o ken-Ø-o	kʰa-Ø-o ken-Ø-o	kʰa-cʰ-cʰ-Ø-o kin-Ø-cʰ-Ø-o	kʰe-e̯e-cʰ-Ø-o kin-e-cʰ-Ø-o
	I	kʰa-Ø-Ø ken-Ø-Ø	kʰa-Ø-Ø ken-Ø-Ø	kʰa-cʰ-cʰ-Ø-iʃ kin-Ø-cʰ-Ø-iʃ	kʰe-e̯e-cʰ-Ø-iʃ kin-e-cʰ-Ø-iʃ
3P	F	kʰa-Ø-n ken-Ø-en	kʰa-Ø-n kin-Ø-un	kʰa-cʰ-cʰ-Ø-en kin-Ø-cʰ-Ø-en	kʰe-e̯e-cʰ-Ø-en kin-e-cʰ-Ø-en
	P	kʰa-Ø-e̯ ken-Ø-e	kʰa-Ø-k kin-Ø-uk	kʰa-cʰ-cʰ-Ø-e kin-Ø-cʰ-Ø-e	kʰe-e̯e-cʰ-Ø-e kin-e-cʰ-Ø-e

(6) Bangla Verbal Inflection for the base /kʰa-/ and /ken-/ (Past and Future)

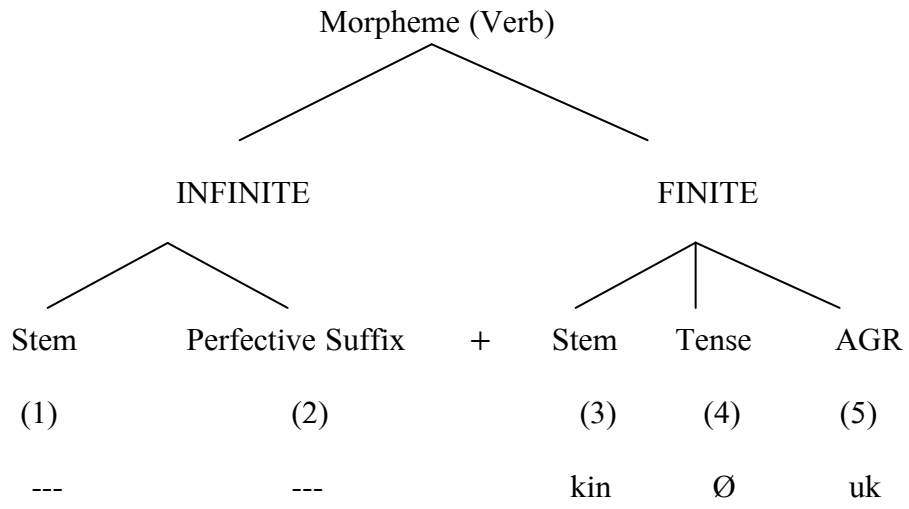
Tense / *Number*		PC	PP	PHC	SP	SF	FI
1P		kʰa-cʰ-cʰi-l-am kin-Ø-cʰi-l-am	kʰe-e̯e-cʰi-l-am kin-e-cʰi-l-am	kʰe-t-am kin-t-am	kʰe-l-am kin-l-am	kʰa-b-o kin-b-o	---
2P	F	kʰa-cʰ-cʰi-l-en kin-Ø-cʰi-l-en	kʰe-e̯e-cʰi-l-en kin-e-cʰi-l-en	kʰe-t-en kin-t-en	kʰe-l-en kin-l-en	kʰa-b-en kin-b-en	kʰa-b-en kin-b-en
	P	kʰa-cʰ-cʰi-l-e kin-Ø-cʰi-l-e	kʰe-e̯e-cʰi-l-e kin-e-cʰi-l-e	kʰe-t-e kin-t-e	kʰe-l-e kin-l-e	kʰa-b-e kin-b-e	kʰe-Ø-e̯o kin-Ø-o
	I	kʰa-cʰ-cʰi-l-i kin-Ø-cʰi-l-i	kʰe-e̯e-cʰi-l-i kin-e-cʰi-l-i	kʰe-t-iʃ kin-t-iʃ	kʰe-l-i kin-l-i	kʰa-b-i kin-b-i	kʰa-Ø̃- kin-Ø-iʃ
3P	F	kʰa-cʰ-cʰi-l-en kin-Ø-cʰi-l-en	kʰe-e̯e-cʰi-l-en kin-e-cʰi-l-en	kʰe-t-en kin-t-en	kʰe-l-en kin-l-en	kʰa-b-en kin-b-en	kʰa-b-en kin-b-en
	P	kʰa-cʰ-cʰi-l-o kin-Ø-cʰi-l-o	kʰe-e̯e-cʰi-l-o kin-e-cʰi-l-o	kʰe-t-o kin-t-o	kʰe-l-o kin-l-o	kʰa-b-e kin-b-e	kʰa-b-e kin-b-e

Note that the forms in the tables are phonological forms, and they differ at many points with their orthographic counterparts.

5.4 Application in the structure

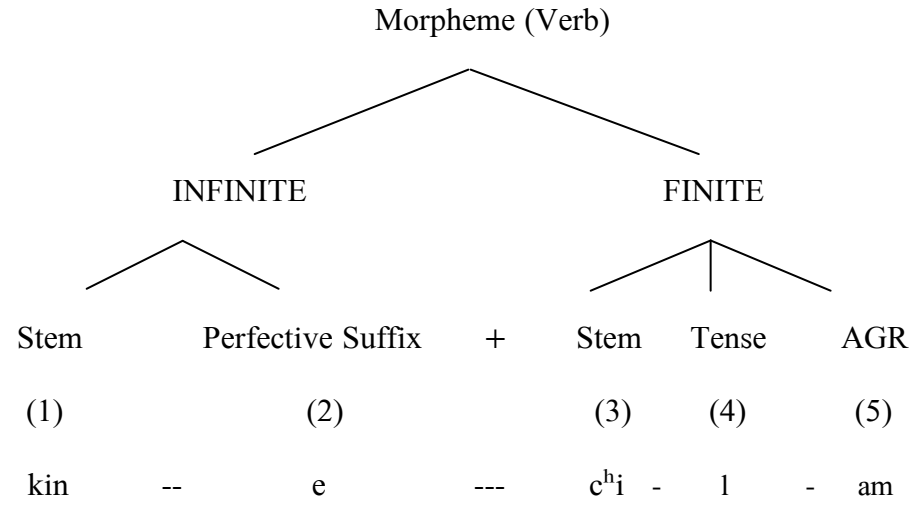
In connection with the earlier discussion, let us see how these inflected forms fit in the structure of (3). We will take one case from each of the tables showing the verb paradigm in Bangla as illustrated in (5) and (6) and check them with said morpheme (verb) structure. The first case is from (5): **/kin-Ø-uk/** (present imperative 3rd person informal form of /kin-/).

(7) [kin-Ø-uk]



In this case, we see that the structure is working for the finite part and there is no element in the infinite part. Now let us take another example to check the infinite part of this morpheme structure as well. From (6), we will take the example: *past perfect first person* **/kin-e-c^{h}i-l-am/**.

(8) [kin-e-c^{h}i-l-am]



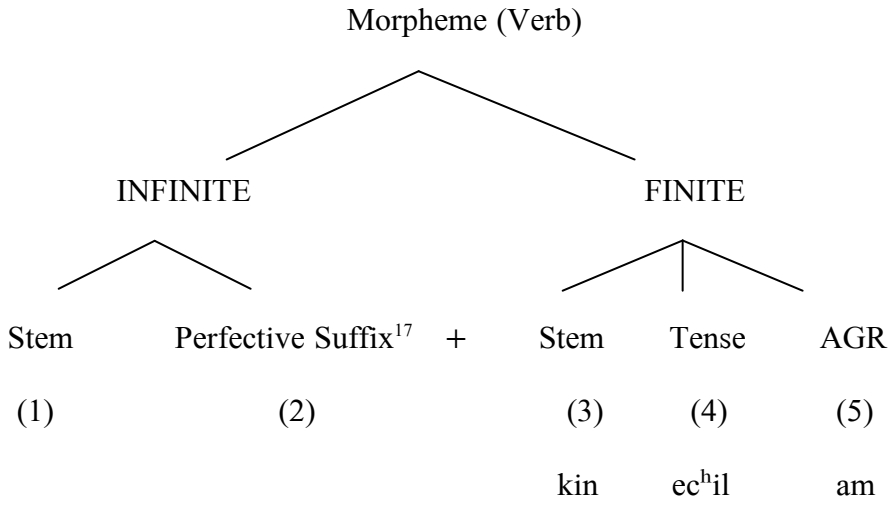
In this analysis, we have morphemes for both infinite and finite parts of the morpheme structure. The infinite part contains the conjunctive participle /**kin-e/**[15] (having bought) and the rest goes to the finite part. Now the finite part breaks into a new structure showing /**c^{h}i-/** as the stem, **/-l-/** as the tense marker and /**–am/** as the agreement (AGR)[16]. The finite part **/c^{h}i-l-am/** denotes a different meaning *'(I) was',* which does not go with the sense of the whole word **/kinec^{h}ilam/** '(I) have bought'. Visibly, the finite part of (8) represents a *be-verb* while the initial word represents a *do-verb*.

But, it is also possible to fit this form only in the finite part of the structure by changing the morph order, such as, **/kin-ec^{h}il-am/**, where the affix **/ec^{h}il/** is taken as a single morpheme.

[15] Conjunctive participles remain same for all persons and numbers. i.e., one is used for all.

[16] Agreement (AGR) is a form of cross-reference between different parts of a sentence or phrase. Agreement happens when one word changes in form depending on which other words it is being related to.

(9) [kin-ec^{h}il-am]



So, it is evident that both of them are eligible to be considered as the "acceptable" form of the morpheme-structure. But, a contrastive study of (8) and (9) reveals the possibility of analyzing the same morphological structure in two different ways. But now we will proceed with the analysis of (8) as the basic structure of Bangla verbal inflection, leaving the third one to be discussed at a later point. The reason for choosing this structure will be justified in the following discussion.

All the possible verb-inflection paradigms in Bangla have already been shown (5) and (6). There are clearly five different morphosyntactic slots present which illustrate the terminal nodes in the structure. These slots are:

(10) Morphosyntactic slots in Bangla

Base1 (Stem) -- Perfective Suffix -- Base2 (Stem) -- Tense -- AGR

[17] A suffix used in describing action that has been completed. Perfective suffix applies to Continuous and perfect tenses.

Some of the inflections make only make use of the last three slots (*Base2, Tense* and *AGR*). And, only those inflected forms containing either the feature [+Cont] or [+Perfect] (what results in a must */-c^{h}-/ or /-c^{h}i-/* in the finite part of the structure) use all the five slots. This category includes *Base1* and *Perfective Suffix* in the infinite part. In the following sections, we will see how they are analyzed syntactically and how the terminal nodes are inserted in the designated slots. But, before that let us see the syntactic analysis of English structures and then we will imply them in the Bangla verbal inflections.

5.5 MORPHOSYNTACTIC ANALYSIS

5.5.1 Morphosyntactic Analysis: English

In English, the main verbal forms of past and present tense cases can be illustrated syntactically within the framework of distributed morphology. Here we have six distinct rules for the third person Present and past tense cases and they work for both singular and plural forms.

1. Present -> Verb[-Past, AGR]
2. Present perfect -> HAVE[-Past, AGR] Verb[+Past_Participle]
3. Present continuous -> [-Past, AGR]BE Verb[+Present_Participle]
4. Past -> Verb[+Past, AGR]
5. Past perfect -> HAVE[+Past, AGR] Verb[+Past_Participle]
6. Past continuous -> [+Past, AGR]BE Verb[+Present_Participle]

Now, following the DM structure, we can draw a table showing the participating suffixes of the verbal inflection of English and the features for which they are inserted at the terminal nodes.

(11) Suffixes and features for the terminal nodes

Affix	**Features**
-ɪŋ (‘-ing’)	[+ Present_Participle]
-(ɪ)d (‘-(e)d’)	[+ Past]
-(ɪ)d (‘-(e)d’)	[+ Past_Participle]
-(ɪ)n (‘-(e)n’)	[+ Past_Participle] / [+ “en”] ___ [18]
-Ø (- Ø)	[+ Past_Participle] / [+ “Ø”] ___ [19]
-(ɪ)z (‘-(e)s’)	[-Past, -S, -A]

In (11), the phonological forms of the affixes are used, in which certain elements are bracketed indicating that they may or may not be there, depending on phonological rules. For example, [-(ɪ)d] means that it is either [–ɪd] or [–d], depending on phonological rules. The second parts of the affix column that are put in brackets and quotation marks, show the orthographical rendition of the suffixes.

In the next step, we take the examples from the *3rd person singular number* cases from the English verb paradigm below. Here, the base form is /*like*/ and the

[18] Some English verbs use -en in the past participle, such as ‘seen’ or ‘eaten’, for which this entry takes effect. Since it is more specific, it will block the preceding lexical entry from taking effect. The assumption with the rule here is that the stems that take ‘-en’ are marked with the feature [+ “en”], and that this rule is then sensitive to this feature. Thus, ‘see’ is marked [+ “en”]. On the other hand, the verbs ‘like’ or ‘heat’ take ‘-ed’. The assumption is that they are not marked for [+”en”], so that the preceding entry then adds ‘-ed’

[19] There are some other verbs in English which do not use either -en or -(e)d in the past participle. For instance, put remains unchanged in past participle (and even in past). so this type of English verbs get no suffix and marked as [+ “Ø”] (null ‘Ø’). They behave in a similar manner as the verbs with [+ “en”].

structure of the sentence would be "he/she [VERB_FORM]". Though we have two forms of each affix (phonetic and orthographic), only the phonetic form will be illustrated in the syntactic structures. However, for the ease of understanding, the orthographic renditions of all the relevant examples are given in brackets in (12).

(12) English verb paradigm for 3rd person singular

	Simple	**Perfect**	**Continuous**
Present	lʌv-z (laɪk-s)	hæz lʌv-d (hæ-z laɪk-d)	ɪ-z lʌv-ɪŋ (ɪ-z laɪk-ɪŋ)
Past	lʌv-d (laɪk-d)	hæd lʌv-d (hæ-d laɪk-d)	wəz lʌv-ɪŋ (wa-z laɪk-ɪŋ)

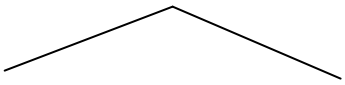
Now, if we take the arrow to point from a mother-node to its daughter-node(s), these rules define morphosyntactic structures as shown illustrated below for two cases from the above examples, 3rd person singular present and present perfect:

(13) 3rd person singular present and present perfect (English): 1

1. Present

Verb[-Past, -S, -A]

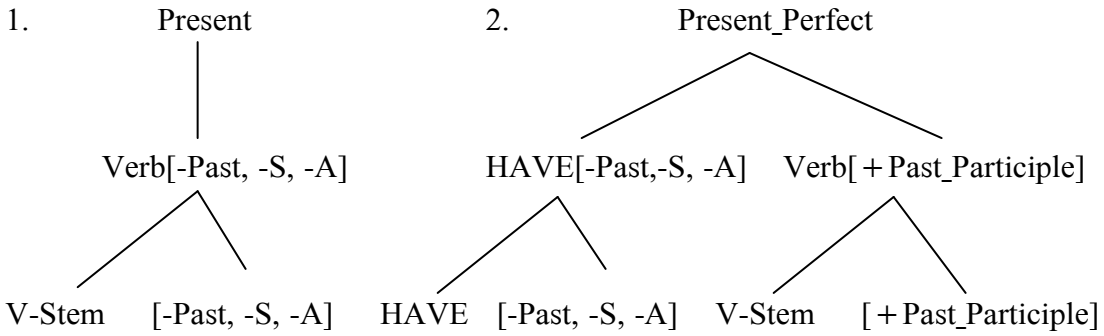
2. Present_Perfect



HAVE[-Past, -S, -A] Verb[+ Past_Participle]

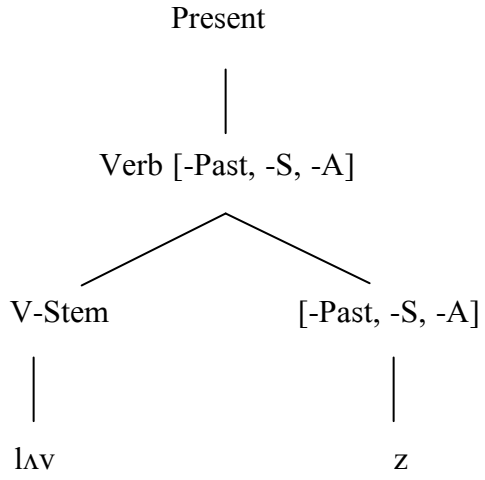
Then we can go further by analyzing Verb[feature(s)] into two nodes, such as (let's take the same two cases as above):

(14) 3rd person singular present and present perfect (English): 2

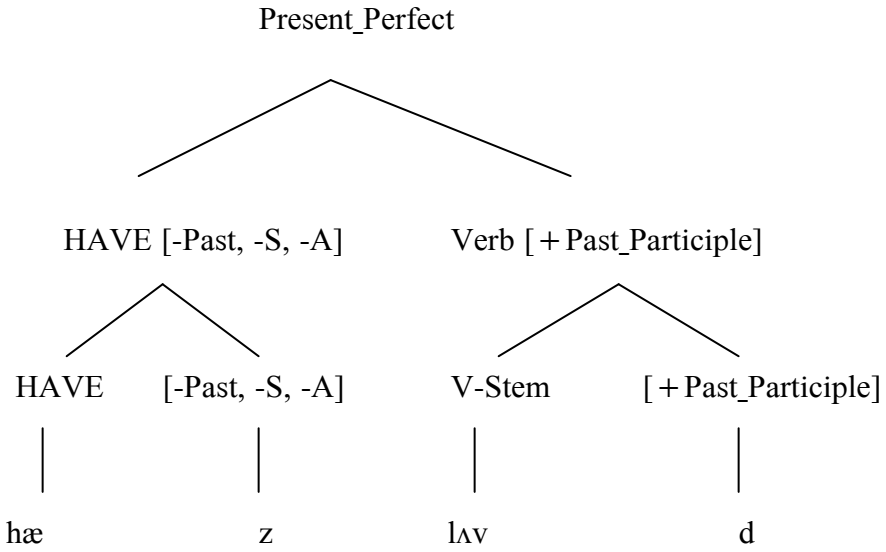


Now, we have a structure where the terminal nodes could be inserted. Following this pattern, we can draw syntactic tree structures for all the above mentioned six ceases. Only two cases have an alternative structure each: Present Continuous and Past Continuous (shown in III and IV below).

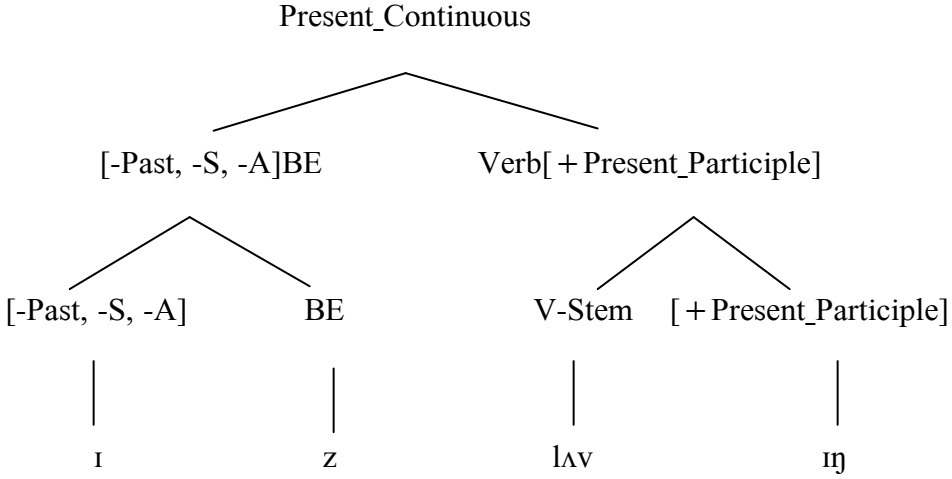
I. Present -> Verb[-Past, AGR]



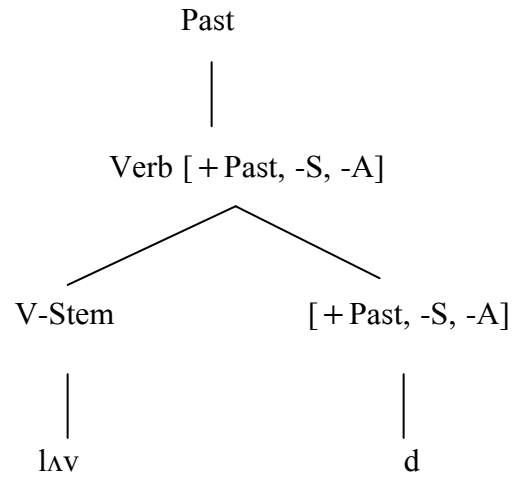
II. **Present perfect -> HAVE[-Past, AGR] Verb[+ Past, + Participle]**



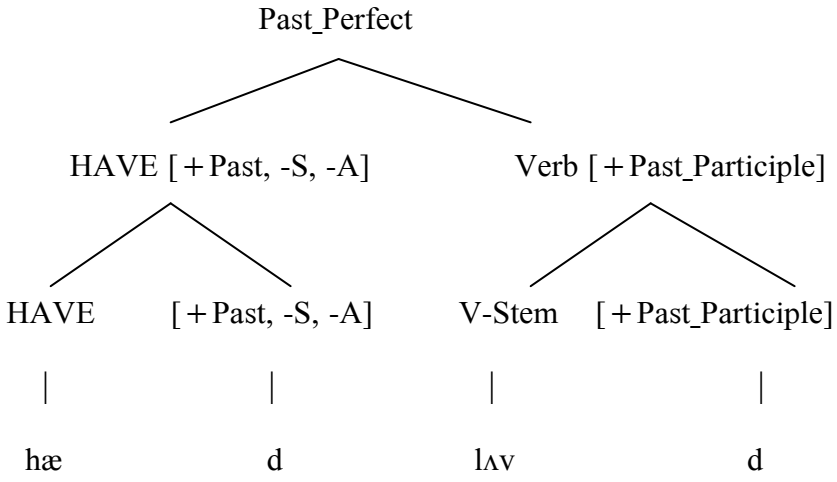
III. **Present continuous -> [-Past, AGR]BE Verb[+ Present, + Participle]**



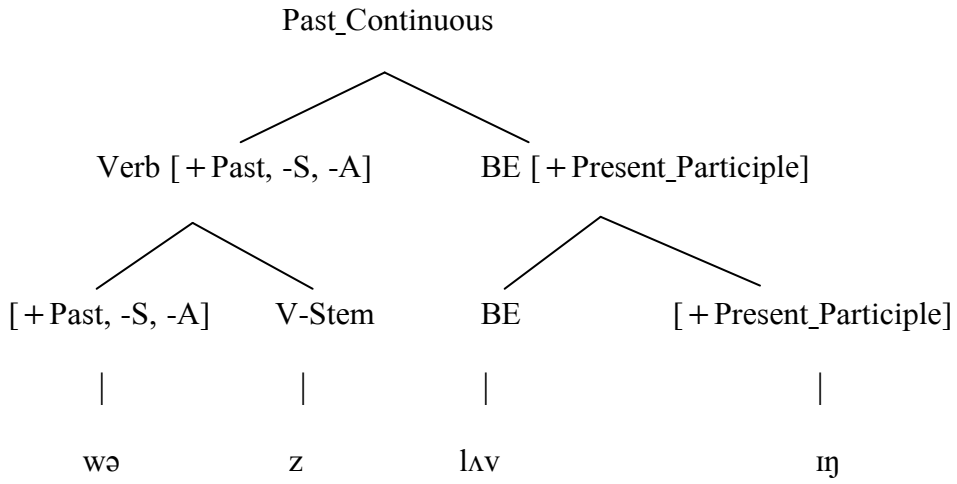
IV. Past -> Verb[+Past, AGR]



V. Past perfect -> HAVE[+Past, AGR] Verb[+Past, +Participle]



VI. Past continuous -> [+Past, AGR]BE Verb[+Present, +Participle]



These are syntactic trees of the main verbal inflections of English. In the next section the Bangla verbal inflections will be analyzed by these structures and with some modifications meant for Bangla structures.

5.5.2 Morphosyntactic Analysis: Bangla

5.5.2.1 Third Person

As described earlier, the third person has three distinguished forms in Bangla. They are the formal, polite and intimate forms in a decreasing level of formality. Here we will start with the formal cases and gradually move into the other forms.

5.5.2.1.1 Third Person Formal

Now, the following tables represent the 3rd person formal cases of present and past

tense for the Bangla verbal base /*kin-*/ and /*k^{h}a-*/. Since Bangla has no distinction between singular and plural at the verbal level, all the forms illustrated in (15) and (16) will automatically be valid for both 3rd person formal singular and 3rd person formal plural cases. The following tables contain all the ten cases of Bangla third person polite cases for two different verbal bases in Bangla taken from (5) and (6). The following table in (15) shows the inflected forms of the consonant-ending base /*kin-*/ and (16) shows the forms of vowel-ending base /*k^{h}a-*/.

(15) 3PF /kin-/ (third person formal for SPR, PRP, PRC, SP, PP and PC)

	Simple	**Perfect**	**Continuous**
Present	ken-Ø-en	kin-e-c^{h}-Ø-en	kin-Ø-c^{h}-Ø-en
Past	kin-l-en	kin-e-c^{h}i-l-en	kin-Ø-c^{h}i-l-en

(16) 3PF /k^{h}a-/ (third person formal for SPR, PRP, PRC, SP, PP and PC)

	Simple	**Perfect**	**Continuous**
Present	k^{h}a-Ø-n	k^{h}e-e̯e-c^{h}-Ø-en	k^{h}a-c^{h}-c^{h}-Ø-en
Past	k^{h}e-l-en	k^{h}e-e̯e-c^{h}i-l-en	k^{h}a-c^{h}-c^{h}i-l-en

The forms shown in the columns of perfect and continuous cases have five slots while other cases have three slots. The fifth slot, i.e., *Base2* contains the stem of the be-form of Bangla (as discussed in section 6.4.2). But there are phonological variations in this slot. When the present cases have the form /c^{h}-/, past cases have

/c^hi-/. We will take the /c^hi-/ as default suffix for this slot and /c^h-/ as a phonological variation of it. This issue will be addressed in the phonology section of this work.

As in the analysis of English verbal forms, here also, we take the same features: **±BE, ±Past, ±Past Participle, ±Present Participle, ±Speaker (±S)** and **±Addressee (±A).** From the above data, we can proceed further with the rules of Bangla syntactic structure. Although, in the basic structure there are some similarities to the English rules, Bangla rules are not identical to them.

1. Present -> Verb[-Past, AGR]
2. Present Perfect -> Verb[+Past_Participle] BE[-Past, AGR]
3. Present Continuous -> Verb[+Present_Participle] BE[-Past, AGR]
4. Past -> Verb[+Past, AGR]
5. Past Perfect -> Verb[+Past_Participle] BE[+Past, AGR]
6. Past Continuous -> Verb[+Present_Participle] BE[+Past, AGR]

Unlike English, Bangla has a BE-verb part in all perfect and continuous cases. But the feature set [±Past,-S, -A] is always attached to the BE-verb only except the simple present and past cases, where they are attached to the base verb.

Now a table similar to English affix and feature combination similar to (11) is drawn here, showing the affixes and the features related to them in Bangla. As discussed earlier for the English case, here too, the affix in the bracket implies the optional occurrence depending on the phonological rules. However, unlike English, the phonetic versions of the Bangla suffixes follow their orthographic rendition in

most cases. Nevertheless, in some cases Bangla pronunciation differs from the orthography. But, the whole analysis of Bangla is done according to the phonological representation in this study. That is why in *table 11* and henceforth no orthographic rendition is provided separately.

(17) Suffixes and features for the terminal nodes (SPR, PRP, PRC, SP, PP and PC)

Affixes for tense and AGR	Features
-Ø(c^hi)	[+ Present_Participle]
-l	[+ Past]
-(e̯)e	[+ Past_Participle]
-(e)n	[-S, -A][20]

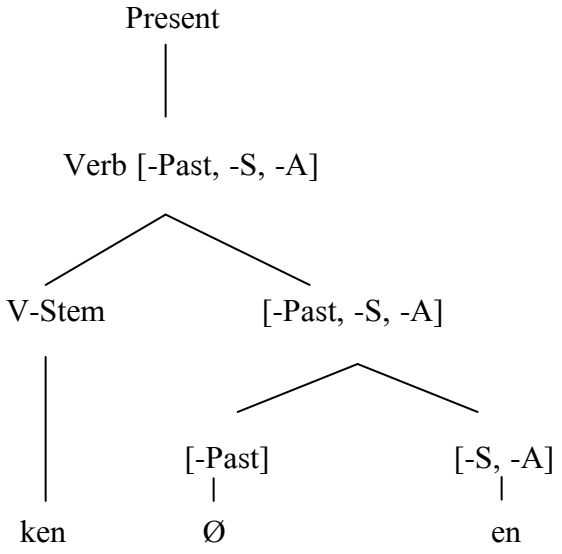
This table shows four affixes and their feature sets. First affix /-Ø(c^hi)/ has something optional in it. Only /-Ø/ or /-c^hi/ may occur at the present participle cases (because the combination of /Ø/ and /c^hi/ will produce only /c^hi/). Hence, it covers the column 'Continuous' for past and present cases in both *tables 10A* and *10B*. Similarly, /-(e̯)e/ accounts for the Past participle cases and covers the 'Perfect' column in both the tables. Affix /-l/ represents the row for 'Past' in both the tables shown in (15) and (16). And /-(e)n/ covers the AGR slots of all the columns and rows of these tables.

Note that (17), illustrates the affix-feature relations only for the above-mentioned tense cases, which means, for SPR, PRP, PRC, SP, PP and PC. Now, the similar process, as seen in English structures, applies here to form the syntactic tree

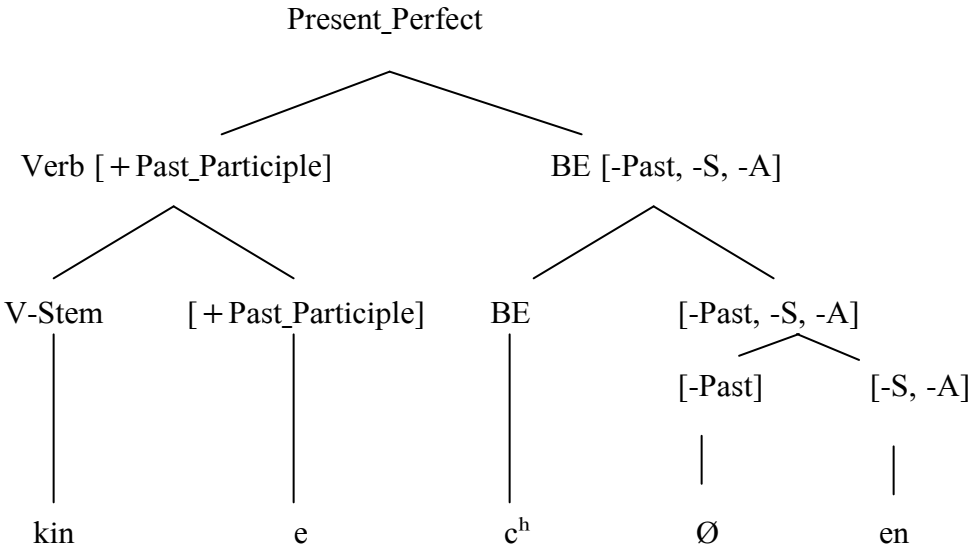
[20] [±S] and [±A] denotes Speaker and Addressee, respectively

and the insertion of the terminal nodes. Again, unlike English, the feature node [±Past, AGR] is further divided into two daughter nodes, namely, tense [±Past] and AGR [±S, ±A]. Let us illustrate the six cases and see how this works.

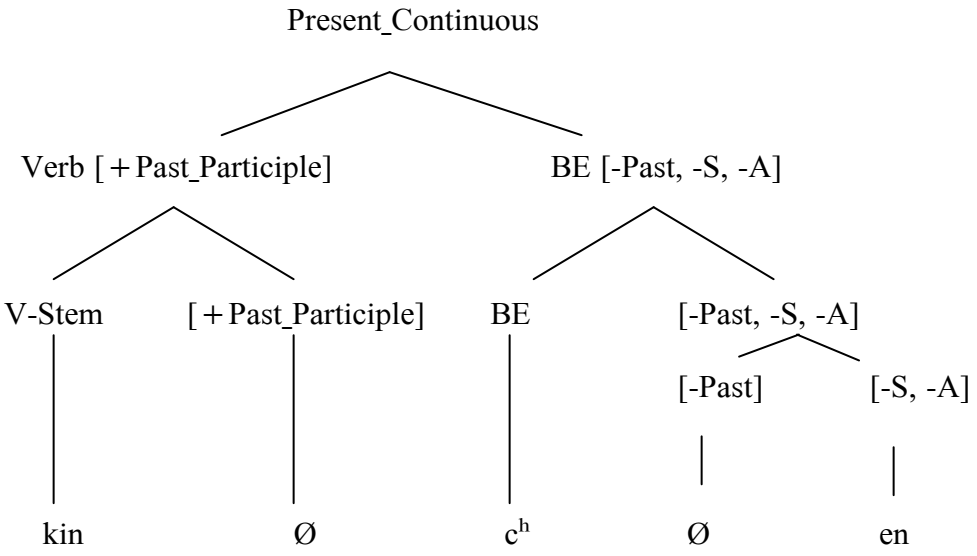
1. **Present -> Verb[-Past, AGR]**



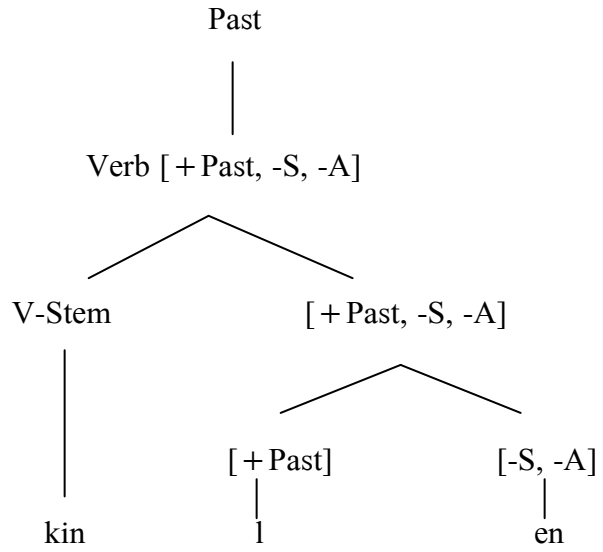
2. **Present Perfect -> Verb[+Past_Participle] BE[-Past, AGR]**



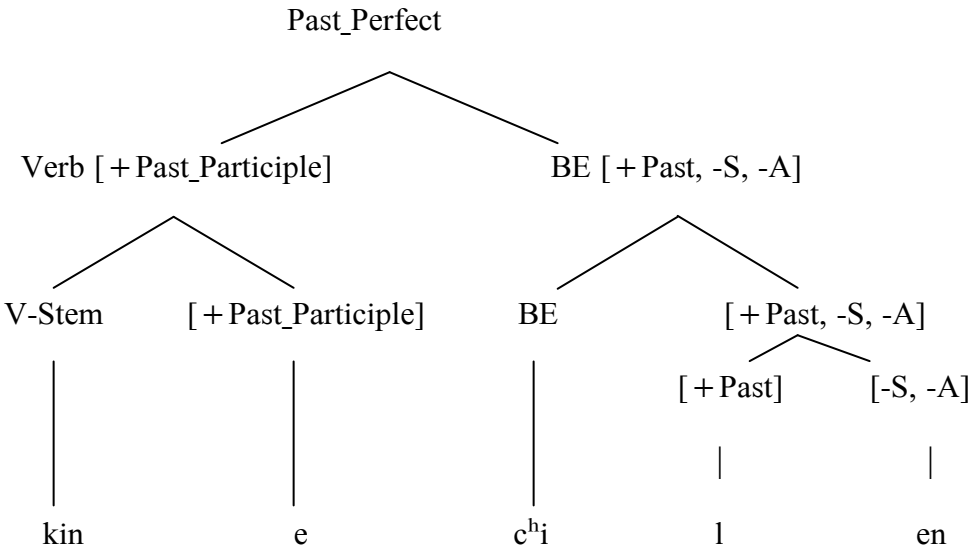
3. **Present Continuous -> Verb[+Present_Participle] BE[-Past, AGR]**



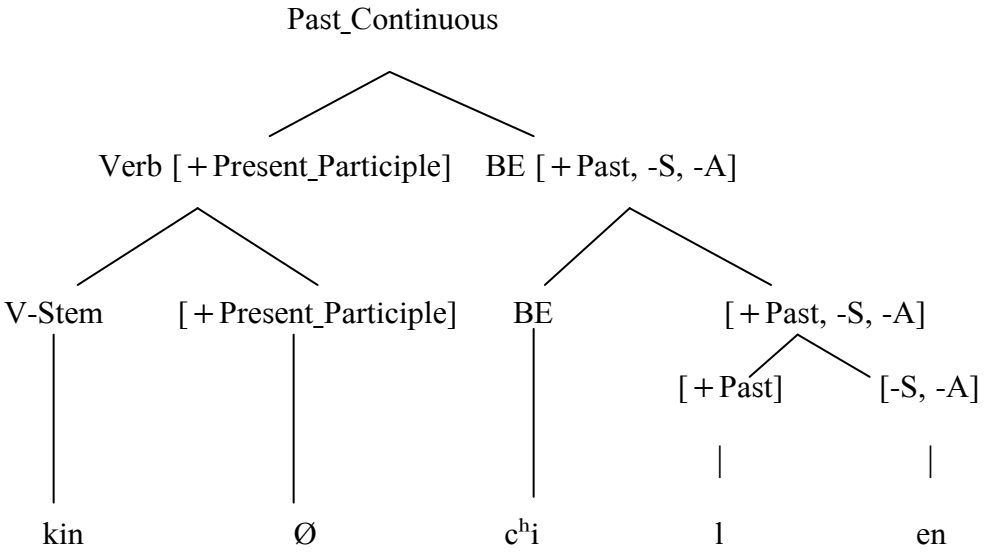
4. **Past -> Verb[+Past, AGR]**



5. **Past Perfect -> Verb[+Past_Participle] BE[+Past, AGR]**



6. **Past Continuous -> Verb[+Present_Participle] BE[+Past, AGR]**



These syntactic tree structures demonstrate the structure of third person formal cases in Bangla. Now the other cases of the Bangla verbal inflections would be dealt with.

5.5.2.1.2 Additional Tenses

There are still four different tense-forms of third person formal, viz., Present Imperative (PRI), Simple Past (SP), Past Habitual Conditional (PHC), Simple Future (SF) and Future Imperative (FI) to be illustrated. We will proceed with all these tenses but only with the 3rd person formal agreement. Later the whole analysis will be extended to all other number and personal agreements as well.

Now, the continuations of the rules from the previous parts are illustrated here. Let us formulate the rules for these new set of tenses. And they are just considered as additions to the previous rules.

7. Present Imperative -> Verb[-Past, +Imperative, AGR]
8. Past Habitual Conditional -> Verb[+Past, +Conditional, AGR]
9. Future -> Verb[+Future, AGR]
10. Future Imperative -> Verb[+Future, +Imperative, AGR]

The following tables contain the remaining four of the tense-forms of Bangla verbal inflection paradigm. The tables in (18) and (19) show the third person formal inflections of consonant-ending base /*kin*-/ and vowel-ending /k^{h}a-/, respectively.

(18) 3PF /kin-/ third person formal for PRI, SP, PHC, SF and FI

	Simple	**Imperative**	**Habitual Conditional**
Present	--	kin-Ø-un	--
Past	--	--	kin-t-en
Future	kin-b-en	kin-b-en	--

(19) 3PF /k^{h}a-/ third person formal for PRI, SP, PHC, SF and FI

	Simple	**Imperative**	**Habitual Conditional**
Present	--	k^{h}a-Ø-n	--
Past	--	--	k^{h}e-t-en
Future	k^{h}a-b-en	k^{h}a-b-en	--

The bold-lettered cells in next table show the affix-feature combinations for the verbal forms of the elements of the tables illustrated in (18) and (19). The old affixes from (17) are shown in italics in (20).

(20) Suffixes and features for the terminal nodes 3rd person Formal (3PF)

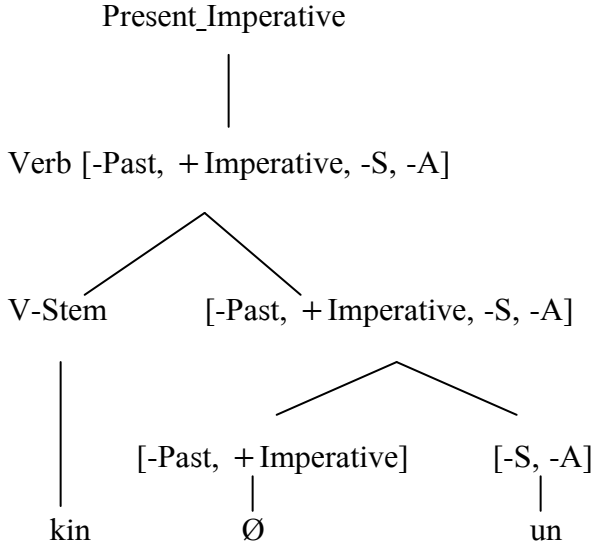
Suffixes for tense and AGR	**Features**
-Ø(c^{h}i)	*[+Present_Participle]*
-l	*[+Past]*
-(ę)e	*[+Past_Participle]*
-(e)n	*[-S, -A]*
-(u)n	**[-Future, +Imperative]**
-t	**[+Conditional]**
-b	**[+Future]**

So, there are actually three new affixes in this table. We will not provide separate tables in each section, but will modify the table from earlier section with new entries. Therefore, at the end there will be only one table for each terminal node (only for those terminal nodes required to be presented in a tabular form). In

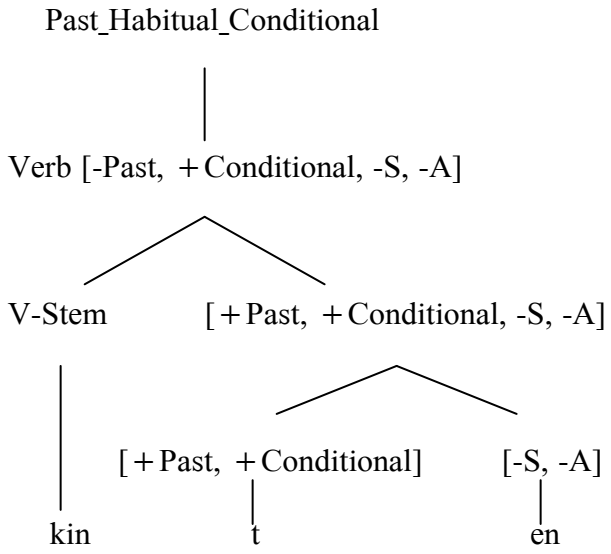
the intermediate stage, all the tables are incomplete and should be considered as 'steps' towards the final and the complete form of it.

Now, we have three new entries in (20). Affix /-(u)n/ accounts for the present imperative cases and hence represents the combination of row 'present' and column 'imperative' in both the tables in (18) and (19). It does not cover the future imperative cases as /-b/ accounts for all the cases of the row for 'Future', which includes this. And, /-t/ covers the tense-affixes in the column for 'Habitual Conditional' in these tables.

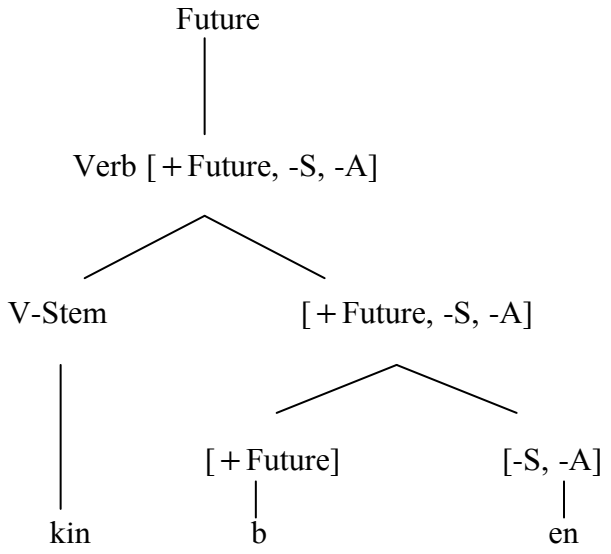
7. Present Imperative -> Verb[-Past, +Imperative, AGR]



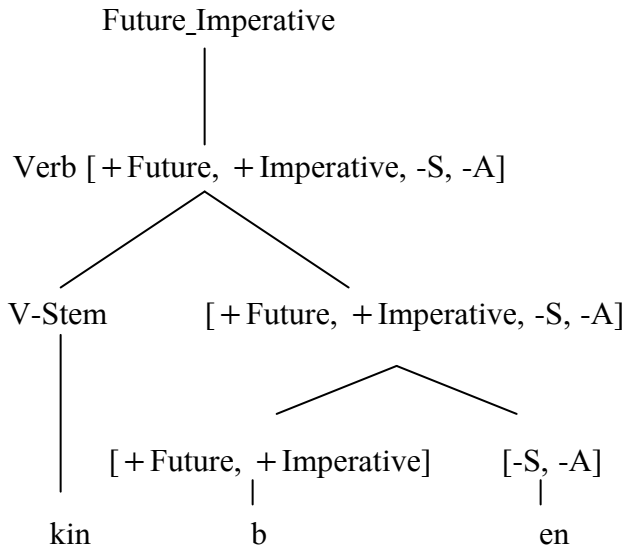
8. **Past Habitual Conditional -> Verb[+Past, +Conditional, AGR]**



9. **Future -> Verb[+Future, AGR]**



10. **Future Imperative -> Verb[+Future, +Imperative, AGR]**



So, this way we can analyze other numbers and tenses in Bangla. The third person formal agreement is already analyzed now. In the next sections, the remaining tenses and agreements are analyzed.

5.5.2.1.3 Third Person Polite

So far, we have considered only 3rd person formal cases. In this section, we will analyze the remaining part of the third person, namely, the polite part. The rules for the previous tree structures carry over to the structures discussed in this section (and hereafter) and no new tree structure illustrations are given here. The following tables contain all the ten cases of Bangla third person polite/informal cases for two different verbal bases in Bangla (from tables in (11) and (12)). Following (21) and

(22) show the inflected forms of the consonant-ending base *ken-* and *table 13B* shows the forms of vowel-ending base /*k^{h}a-/*.

(21) 3PP/ken-/ (Third person polite)

	Simple	**Perfect**	**Continuous**	**Imperative**	**Habitual Conditional**
Present	ken-Ø-e	kin-e-c^{h}-Ø-e	kin-Ø-c^{h}-Ø-e	kin-Ø-uk	--
Past	kin-l-o	kin-e-c^{h}i-l-o	kin-Ø-c^{h}i-l-o	--	kin-t-o
Future	kin-b-e	--	--	kin-b-e	--

(22) 3PP/k^{h}a-/ (Third person polite)

	Simple	**Perfect**	**Continuous**	**Imperative**	**Habitual Conditional**
Present	k^{h}a-Ø-e̯	k^{h}e-e̯e-c^{h}-Ø-e	k^{h}a-c^{h}-c^{h}-Ø-e	k^{h}a-Ø-k	--
Past	k^{h}e-l-o	k^{h}e-e̯e-c^{h}i-l-o	k^{h}a-c^{h}-c^{h}i-l-o	--	k^{h}e-t-o
Future	k^{h}a-b-e	--	--	k^{h}a-b-e	--

Now, instead of a single table for the affix-feature combinations for the verbal forms of the elements of (21) and (22), we will have two different tables. One is for the tense slot and the other is for the AGR slot. These tables specifically show the affixes and their feature sets required for third person formal and polite/informal cases (listed in previous tables).

Let us first consider the tense slot. The following table represents the tense suffixes we have come across so far:

(23) Affixes and features for the tense slot

Affix for Tense	Feature
-Ø(c^hi)-	[+ Present_Participle]
-l-	[+ Past]
-(ẹ)e-	[+ Past_Participle]
-t-	[+ Conditional]
-b-	[+ Future]
Ø	**ELSEWHERE**

First five affixes of (23) have already been discussed in earlier sections. Only the null /Ø/ is a new member in this table. Technically, it stands for the ELSEWHERE position, but in the tables above, it is actually applied in the row for 'Present'.

Another important characteristic feature of (23) is the inclusion of the perfective suffixes. Both /-(ẹ)e-/ and /-Ø(c^h)-/ are not exactly regular tense markers, but here they play an important role in the tense field.

(24) Suffixes and features for the AGR slot

Affix for AGR	Feature
-(e/u)n	**[+ Formal, -S]**
-(u)k	**[-Formal, -S] / [-Future, + Imperative] __**
-o	**[-Formal, -S, -A] / [+ Past] __**
-e/ẹ	**ELSWHERE**

The table illustrated in (24) has a modified version of (24) where /-e(n)/ is inserted only for the feature set [+Formal, -S]. It covers the formal cases of both second person and third person forms. In the next section, we will see how this condition applies in the second person cases.

Earlier we had only /-(u)n/ for present imperative cases. Now, another form for present imperative appears in the tables, viz., /-(u)k/. It is required to specify the position and role of each of them. Suffix /-(u)k/ in *table 15* accounts for the AGR-suffixes in the cell present imperative polite/informal cases in both the tables in (18) and (19).

But the suffix /-(u)n/ is not in a separate row anymore. It is combined with other 'formal' cases of [-S]. Actually, the difference between [-en] and [-un] is merely phonological. So, their occurrence is not considered as morphological in this study. A phonological account of this sound change will be provided in a later section after discussion of the morphological operations.

Suffix /-o / accounts for the AGR-suffixes in the entire row for the 'Past' cases. Present and future cases are taken care of by the ELSEWHERE position's suffix /-e/ which has an alternative form /-e̯/. These two suffixes are in complementary distribution and their occurrence depends upon the phonological environment in each particular case.

5.5.2.2 Second Person

5.5.2.2.1 Second Person Formal

(25) Bangla verb paradigm for second person formal (2PF)

	Simple	Perfect	Continuous	Imperative	Habitual Conditional
Present	ken-Ø-en	kin-e-c^{h}-Ø-en	kin-Ø-c^{h}-Ø-en	kin-Ø-un	--
Past	kin-l-en	kin-e-c^{h}i-l-en	kin-Ø-c^{h}i-l-en	--	kin-t-en
Future	kin-b-en	--	--	kin-b-en	--

(26) Bangla verb paradigm for second person formal (2PF)

	Simple	Perfect	Continuous	Imperative	Habitual Conditional
Present	k^{h}a-Ø-n	k^{h}e-e̯e-c^{h}-Ø-en	k^{h}a-c^{h}-c^{h}-Ø-en	k^{h}a-Ø-n	--
Past	k^{h}e-l-en	k^{h}e-e̯e-c^{h}i-l-en	k^{h}a-c^{h}-c^{h}i-l-en	--	k^{h}e-t-en
Future	k^{h}a-b-en	--	--	k^{h}a-b-en	--

A thorough observation reveals that the tense slot has nothing new to add or modify in (17) at this level. Furthermore, the AGR slot, too, has nothing unique to contribute to the AGR table. All the formal cases consist of the suffix /-(e)n-/ which is featured as [+Formal, -S] and has already been modified in *table 15*. So, there will be no modification of the set of affixes for tense or AGR slots in this section.

5.5.2.2.2 Second Person Polite

The next couple of tables represent the second person polite forms of the inflection for the verbal bases /*kin-*/ and /*k^{h}a-*/.

(27) Bangla verb paradigm for second person polite (2PP)

	Simple	**Perfect**	**Continuous**	**Imperative**	**Habitual Conditional**
Present	ken-Ø-o	kin-e-c^{h}-Ø-o	kin-Ø-c^{h}-Ø-o	ken-Ø-o	--
Past	kin-l-e	kin-e-c^{h}i-l-e	kin-Ø-c^{h}i-l-e	--	kin-t-e
Future	kin-b-e	--	--	kin-Ø-o	--

(28) Bangla verb paradigm for second person polite (2PP)

	Simple	**Perfect**	**Continuous**	**Imperative**	**Habitual Conditional**
Present	k^{h}a-Ø-o	k^{h}e-e̯e-c^{h}-Ø-o	k^{h}a-c^{h}-c^{h}-Ø-o	k^{h}a-Ø-o	--
Past	k^{h}e-l-e	k^{h}e-e̯e-c^{h}i-l-e	k^{h}a-c^{h}-c^{h}i-l-e	--	k^{h}e-t-e
Future	k^{h}a-b-e	--	--	k^{h}e-Ø-e̯o	--

There is one additional element in the tense slot. So far, the future tense was denoted by the affix /-b-/ in all forms. But in (27) and (28), it is seen that the future imperative cases have a different affix, namely, /Ø/ for the tense slot. This affix occurs in the 'present_participle' and the ELSEWHERE position, too. But there is no common feature set or environment present in these cases what could reduce the appearance of the affix /Ø/ in different rows of the tense table.

(29) Suffixes and features for the tense slot (modified) (2PP)

Affix for Tense	**Feature**
-Ø(c^h)-	*[+Present_Participle]*
-l-	*[+Past]*
-(ẹ)e-	*[+Past_Participle]*
-t-	*[+Conditional]*
-b-	*[+Future]*
Ø	**[+Future, +Imperative] / ___ [-Formal, -S, +A]**
Ø	ELSEWHERE

The affixes (and corresponding feature sets) from the previous table are shown in *italics* and the newly introduced ones are given in **bold** letters. The AGR slot is likely to be modified with some unique elements introduced in (24).

(30) Suffixes and features for the AGR slot (modified) (2PP)

Affixes for AGR	**Feature**
-(e/u)n	*[+Formal, -S]*
-(u)k	*[-Formal, -S] / [-Future, +Imperative] __*
-o	*[-Formal, -S, -A] / [+Past] __*
-(ẹ)o	**[-Formal, -S] / [+Future, +Imperative] __**
-o	**[-Formal, -S, +A] / [-Past, -Future] __**
-e/ẹ	*ELSEWHERE*

In (30), a couple of new affixes are introduced to the AGR slot in addition to the previous four affixes. Sometimes the same suffix (like, -o-) has occurred in different but distinct environments and hence received two distinct feature sets.

Affix /-o/ occurs not only for the third person polite of past tense cases, but also for the second person polite of present tense cases. But both of them cannot be merged in a single feature set. So an additional /-o/ is introduced in *table 18.*

Another suffix /-(e̯)e/, which does not belong to the ELSWHERE position, is interpolated in this table. This suffix stands for second person polite of present and future imperative cases. However, the other /-o/ represents the second person present imperative cases, too.

5.5.2.2.3 Second Person Intimate

The following tables in (31) and (32) show the third and final level of formality of the second person agreements in Bangla.

(31) Bangla verb paradigm for second person informal (2PI)

	Simple	**Perfect**	**Continuous**	**Imperative**	**Habitual Conditional**
Present	ken-Ø-Ø	kin-e-c^{h}-Ø-iʃ	kin-Ø-c^{h}-Ø-iʃ	ken-Ø-Ø	--
Past	kin-l-i	kin-e-c^{h}i-l-i	kin-Ø-c^{h}i-l-i	--	kin-t-iʃ
Future	kin-b-i	--	--	kin-Ø-iʃ	--

(32) Bangla verb paradigm for second person informal (2PI)

	Simple	**Perfect**	**Continuous**	**Imperative**	**Habitual Conditional**
Present	k^ha-Ø-Ø	k^he-e̯e-c^h-Ø-iʃ	k^ha-c^h-c^h-Ø-iʃ	k^ha-Ø-Ø	--
Past	k^he-l-i	k^he-e̯e-c^hi-l-i	k^ha-c-c^hi-l-i	--	k^he-t-iʃ
Future	k^ha-b-i	--	--	k^ha-Ø-ʃ	--

The tense slot remains unchanged at this level. But there is a significant number of new elements in the AGR slot. Let us put them down in the following table, which is actually an extended version of (24).

Once again, the *italics* forms are representing the affixes already discussed in previous tables and the **bold** one are the new entries at this level.

(33) Suffixes and features for the AGR slot (modified) (2PI)

Affix	**Feature**
-(e/u)n	*[+Formal, -S]*
-(u)k	*[-Formal, -S] / [-Future, +Imperative] __*
-o	*[-Formal, -S, -A] / [+Past] __*
-(e̯)o	*[-Formal, -S] / [+Future, +Imperative] __*
-o	*[-Formal, -S, +A] / [-Past]__*
-i	**[+Informal, -S, -A] / [+Past, -Conditional] __**
-i	**[+Informal, -S, -A] / [+Future, -Imperative] __**
-Ø	**[+Informal, -S, -A] / V-Stem [-Past, -Future] __**
-(i)ʃ	**[+Informal, -S, -A]**
-e/e̯	ELSEWHERE

In (33), the total number of affixes and their feature sets has been increased from 6 to 10 compared to the previous version. A pair of identical affixes (/i/) is introduced in this table. Since Past and Future cases do not belong to the same natural class, the affix /i/ has to be shown as different affixes. The null suffix (/Ø/) represents the AGR slots of simple and imperative cases. The forms of both these cases are three-slot forms consisting only V-Stem (*Base1*). The remaining forms are covered by the suffix /-(i)ʃ/ where /-i/ is optional. A particular phonological operation determines the occurrence of /-i/ at this position.

But this table of the AGR slot is not the final one. There is another personal division to be analyzed in order to get the complete set of affixes and their features namely the first person.

5.5.2.3 First Person

The first person in Bangla has a pretty simple set of verbal inflections. It consists only a single set of inflections and the imperative mood (present and future) is not present here.

(34) Bangla verb paradigm for first person (1P)

	Simple	**Perfect**	**Continuous**	**Imperative**	**Habitual Conditional**
Present	kin-Ø-i	kin-e-c^h-Ø-i	kin-Ø-c^h-Ø-i	--	--
Past	kin-l-am	kin-e-c^hi-l-am	kin-Ø-c^hi-l-am	--	kin-t-am
Future	kin-b-o	--	--	--	--

(35) Bangla verb paradigm for first person (1P)

	Simple	**Perfect**	**Continuous**	**Imperative**	**Habitual Conditional**
Present	k^ha-Ø-i̯	k^he-e̯e-c^h-Ø-i	k^ha-c^h-c^h-Ø-i	--	--
Past	k^he-l-am	k^he-e̯e-c^hi-l-am	k^ha-c^h-c^hi-l-am	--	k^he-t-am
Future	k^ha-b-o	--	--	--	--

For the first person, again, there is no change in the tense slot. However, the AGR slot has some additional elements which are added in the following table in **bold** face.

In this table we see three more new affixes, viz., [-i/i̯], [-am] and [-o] for the first person cases of present, past and future tense, respectively. It is impossible to merge the [-o] in any existing one, since the natural class settings do not allow for such a grouping.

Once again, there are two different phonological forms [-i] and [-i̯] occur at the same place. It will be analyzed, like similar cases, in the phonological discussion.

(36) Suffixes and features for the AGR (modified)

Affix	**Feature**
-i/i̯	**[+S] / [-Past, -Future] __**
-am	**[+S] / [+Past] __**
-o	**[+S] / [+Future] __**
-(e/u)n	*[+Formal, -S]*
-(u)k	*[-Formal, -Intimate, -S] / [-Future, +Imperative] __*
-(e̯)o	*[-Formal, -Intimate, -S] / [+Future, +Imperative] __*
-o	*[-Formal, -Intimate, -S, +A] / [-Past]__*
-o	*[-Formal,-Intimate, -S, -A] / [+Past] __*
-i	*[+Informal, -S, -A] / [+Past, -Conditional] __*
-i	*[+Informal, -S, -A] / [+Future, -Imperative] __*
-Ø	*[+Informal, -S, -A] / V-Stem [-Past, -Future] __*
-(i)ʃ	*[+Informal, -S, -A]*
-e/e̯	*ELSEWHERE*

The table in (36) contains a total of 10 suffixes and their corresponding feature sets. And the other table (17) for the tense slot is intact since we discussed the third person.

5.6 Summary

Till now, we have discussed the morphological forms of Bangla verbal inflections in the framework of Distributed Morphology. The results of the whole analysis can be summed up in the complete tables illustrated below:

(37) Complete set of suffixes for the tense slot

Affix for Tense	**Feature**
-l-	[+ Past]
-t-	[+ Past, + Conditional]
-(e̤)e-	[+ Past_Participle]
-Ø(c^hi)-	[+ Present_Participle]
-b-	[+ Future]
Ø	[+ Future, + Imperative] / [-Formal, -S, + A] ___
Ø	ELSEWHERE

There is another complete set of affixes for the AGR slot. These there listed down in the following table (38).

(38) Complete set of suffixes for the AGR slot

Affix	Feature
-i/i̯	[+S] / [-Past, -Future] __
-am	[+S] / [+Past] __
-o	[+S] / [+Future] __
-(e/u)n	[+Formal, -S]
-(u)k	[-Formal, -Intimate, -S] / [-Future, +Imperative] __
-(e̯)o	[-Formal, -Intimate, -S] / [+Future, +Imperative] __
-o	[-Formal, -Intimate, -S, +A] / [-Past]__
-o	[-Formal,-Intimate, -S, -A] / [+Past] __
-i	[+Informal, -S, -A] / [+Past, -Conditional] __
-i	[+Informal, -S, -A] / [+Future, -Imperative] __
-Ø	[+Informal, -S, -A] / V-Stem [-Past, -Future] __
-(i)ʃ	[+Informal, -S, -A]
-e/e̯	ELSEWHERE

These two tables present a considerable set of affixes and features of Bangla verbal inflections for both vowel and consonant ending roots in the distributed morphology framework. The vocabulary items listed in the above tables are in competition. The most specific one among each list wins for a particular position and gets inserted there. For each slot, only one vocabulary item wins the competition. There are some cases where an empty morphere (Ø) occurs as a vocabulary item. In those cases, empty morpheme has no vocalic or consonantal

content. But, even this vocabulary item has an empty mora as its phonological content and certain morphosyntactic features that determine its insertion.

These tables should work like a computer program which applies the rules to a given input in order to get the right output. But, there are some points which need attention. Firstly, one can find suffixes used repeatedly in different positions. This is unavoidable because the groupings do not always follow the same natural class. As a result, the same suffix occurs more than once with different features and environment. For instance, the suffix /-o/ occurred four times in the AGR slot. Another such case is /Ø/ which is noticeable in the tense slot.

Secondly, a small number of cases have more than one phonological form which could not be analyzed in this sphere of DM framework. Since DM does not allow any such operation at the syntactic level, those cases are addressed in the next section of this study using OT inventory.

6 Application of OT to DM output

The morphological analysis of Bangla verbal inflections results in a further complex set of phenomena where basically phonological issues are to be dealt with. The previous analysis was done with the help of Distributed Morphology, which is not capable of solving the phonological issues. Hence, the outcome of the discussion from previous chapters on optimality theory is used here to address these phonological issues.

From the morphological inventory presented in chapter 6, we have three special cases where optimality theory could be used to analyze the phenomena. They cases are:

(1) DM output > Actual Pronunciation

i. Diphthongization : /k^{h}ae/ > [k^{h}ae̯]

ii. Vowel deletion : /k^{h}aek/ > [k^{h}ak]

iii. Gemination : /k^{h}ac^{h}i/ > [k^{h}ac^{h}c^{h}i]

We would use the constraints already introduced in the OT related chapters in this study. Additionally, there will be some more constraints specially introduced in this section, according to the requirement of the very cases.

6.1 Diphthongization

The most common and useful which are considered very basic in an OT analysis is already introduced in the precious chapters. A few of them are also used in this section along with some of the new constraints that might come in use. To account for the diphthongization issue, it is important to analyze the syllable structure of an input in moraic representation (Clark, Yallop, & Fletcher 2007; Hayes 1989 *inter allia*) which is partially introduced in some earlier sections. Such an important structure-related constraint Branch-μ is relevant here in the analysis of diphthongs (Paradis 1990; Rosenthall 1994).

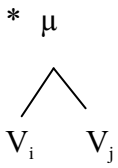
(2) BRANCH-μ

A mora can not be branched into two vowels.

'Diphthongs can not be monomoraic.'

This markedness constraint requires that diphthongs cannot be monomoraic. In other words, every mora in a syllable can dock only one element and if there are more than one vowel (vocoids) in the syllable, then each one of then could be attached to a mora. The following illustration shows this structure clearly. Note that V denotes both vowels as well as semi-vowels in this study. Since Bangla does not allow two pure vowels to occur together, diphthongs are always made of a vowel (first element) and a semi-vowel (final element) in a row in this language.

(3) No branching of mora



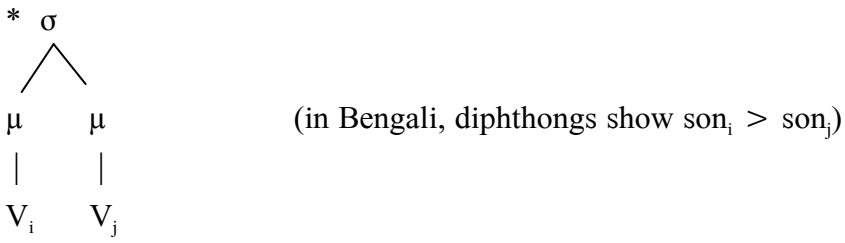
The BRANCH-μ constraint has a very close link to another markedness constraint of similar nature. This is a constraint that punishes any moraic structure where two vowels are linked to two moras. We will call it NODIPH.

(4) NODIPH

Two moras in the same syllable can not be linked to two vowels.

An illustration of this constraint feature is given in the following figure. The sonority of a vowel is always higher than the sonority of a semi-vowel. Hence, structures shown as in (5) always denote a diphthong in moraic theory.

(5) Structure ruled out by NODIPH:



A set of other markedness constraints obstructing the occurrence of certain sounds at some specific positions are also useful in this analysis. Let us consider the following two constraints from Prince & Smolensky 1993.

(6) *M/a

A syllable does not have /a/ in the margin.

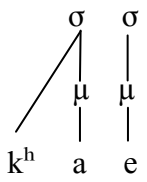
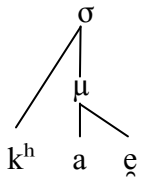
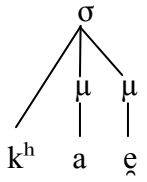
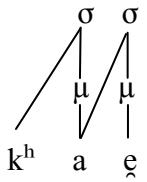
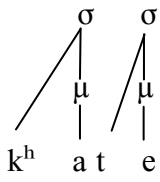
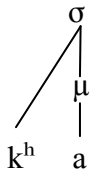
(7) *M/e

A syllable does not have /e/ in the margin.

The above constraints form a group of margin-constraints that required non-occurrence of certain sound at the onset and coda position. Along with all these constraints, we could also use another well-known markedness constraint ONSET which requires that every syllable must have an onset. Two other general constraints which are, eventually faithfulness constraints are also reintroduced in this analysis. These two, namely, MAX-IO and DEP-IO punish any deletion or epenthesis, respectively, in the input.

Now, in this section, we will be discussing the diphthongization of a particular inflectional pattern in Bangla. This is an outcome of the verbal inflectional paradigms analyzed in the distributed morphology framework (see chapter 6).

(8) Diphthongization: /k^{h}ae/ → [k^{h}ae̯]

/k^{h}ae/	ONSET	MAX-IO	DEP-IO	BRANCH-μ	*M/a	NODIPH
1. 	*!					
2. 				*!		
3. ☞ 						*
4. 					*!	
5. 			*!			
6. 		*!				

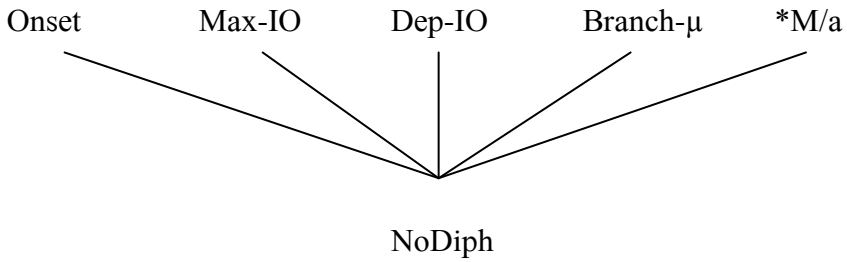
The third person present form of /k^{h}a-/ results in an ill-formed output /k^{h}ae/ (1a) in the DM analysis. Since Bangla would allow only a diphthong instead of a

pure VV sequence, the */-ae/ structure in this word must be analyzed for an acceptable account in the OT framework.

The tableau in (8) shows the constrain hierarchy for the diphthongization process for such cases in Bangla. We have six candidates (1-6) in this tableau, of which candidate 3 turns out to be the optimal one. The first five constraints (ONSET, MAX-IO, DEP-IO, BRANCH-μ and *M/a) are not crucially ranked. And, the final constraint (NoDiph) is dominated by these five.

In this tableau, candidate 1 violates two constraints (ONSET and *M/a) which are not crucially ranked. So do candidates 5 and 6 (Dep-IO, *M/a, and Max-IO, *M/a, respectively). Candidates 2 and 4 also violate one constraint each (BRANCH-μ and *M/a, respectively) which are not crucially ranked. But candidate 3 violates the only crucially ranked constraint (NoDiph) which is dominated by the other constraints, but none else. That makes candidate 3 win over the other candidates. The constraint ranking of these constraints is quite simple.

(8) Constraint ranking for diphthongization



6.2 Vowel deletion

Normally the underlying form /k^{h}aek/ never occurs in Bangla. We have arrived to this form through the DM mechanism. But, even this odd structure could be analyzed to reach the valid word-form in Bangla using a set of constraints in a systematic ranking. Here, some new constraints are required to be introduced for this particular case study.

Sometimes the rime (peak + coda) in a syllable contains more than the regular two moras. This happens when there is a vowel cluster at the peak and and additional consonant is left out to be directly linked to the syllable node. A constraint requiring two or less mora in the rime would punish this tendency.

(9) 2μ-rime

The rime consists of no more than 2 moras (μ).

The above markedness constraint is not crucially ranked with our earlier NODIPH constraint which also punishes such vowel clusters. Next, we would consider a pair of markedness constraints those do not like certain sound at the nucleus (or peak) position. This is also connected to a similar family of constraints (*M).

(10) *P/e

A syllable does not have /e/ in the nucleus (peak).

(11) *P/a

A syllable does not have /a/ in the nucleus (peak).

While analyzing words like /k^{h}aek/, one must rank these two constraints low enough to be above 2μ-rime and NODIPH at least. Internally, *P/e is better placed above *P/a, in order to let candidate 3 win. Again, the same constraint-set replaces the general MAX-IO constraint by a modified version of the same faithfulness constraint. These constraints are already introduced in the gemination chapter in a previous section.

(12) MAX-IO-C

Consonants in the input must have output correspondence.

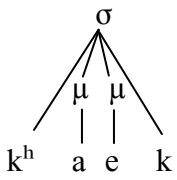
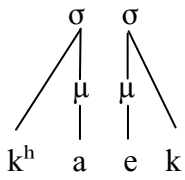
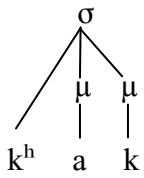
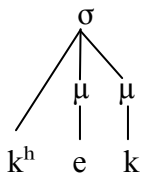
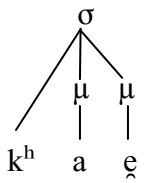
(13) MAX-IO-V

Vowels in the input must have output correspondence.

In these pair of tableaux, MAX-IO-C requires every consonant in the input be preserved in the out, but it doesn't care about any loss of vowel from input. On the other hand, MAX-IO-C behaves exactly similarly, but here the focus is on vowels. This constraint ignores any deletion of consonants from the input.

Taking these constraints in the account, let us analyze the input form resulting from the DM output.

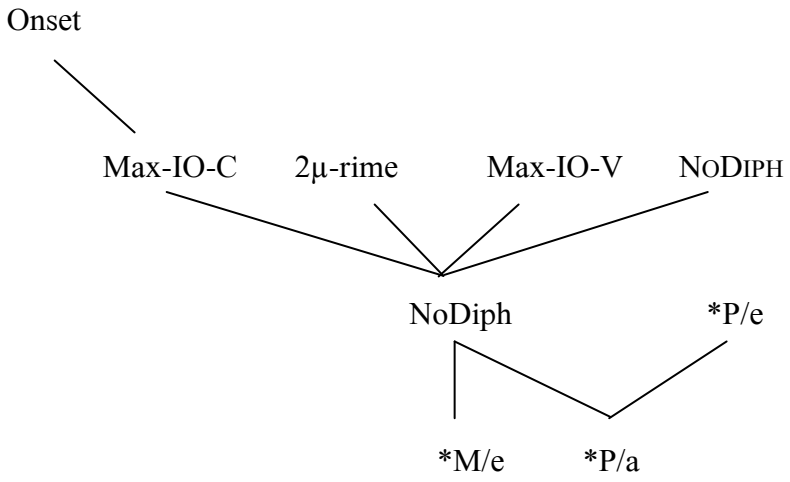
(12) /k^{h}aek/ → [k^{h}ak] (Vowel deletion)

/k^{h}aek/	Onset	Max-IO-C	2μ-rime	Max-IO-V	NoDiph	*M/e	*P/e	*P/a
1. 			*!		*		*	*
2. 	*!					*	*	*
3. ☞ 				*				*
4. 				*			*!	
5. 		*(!)		*	*(!)	*		*

The modified constraint ranking from the previous tableau has included the new constraints we discussed now. But the most important changes are made in the constraints hierarchy. Onset and 2-μ-rime dominate MAX-IO-V which was, as only MAX-IO, an independent constraint along with Onset before. All the high ranked constraints are dominated by *M/e which is an independent constraint with *P/e.

Again, *P/a is dominated by *P/e. Hence, the constraint ranking for this analysis would be as follows.

(13) Constraint ranking for vowel deletion



Keeping this new constraint ranking in mind, let us check another possibility of phonological change. Instead of /k^{h}aek/, let us consider the Sadhu (SB) form of the same root /k^{h}a-/, i.e., /k^{h}auk/. This is used in the classical Bangla for quite a long time.

(14) vowel-deletion 2

a. /k^{h}auk/ → [k^{h}ak]

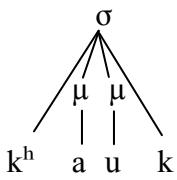
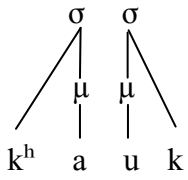
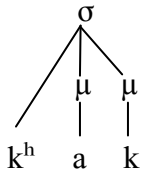
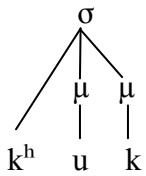
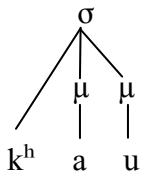
b. /kinuk/ → [kinuk]

c. */k^{h}aek/ → [k^{h}ak]

In the following tableau the same constraint ranking is maintained (from k^{h}aek/ → [k^{h}ak]). Technically, it should produce the same result as the previous

tableau did.

(15) Vowel-deletion: /k^{h}auk/ → [k^{h}ak]

/k^{h}auk/	Onset	2-μ-rime	Max-IO-V	NoDiph	*M/e	*P/u	*P/a
1. 		*!		*		*	*
2. 	*!				*	*	*
3. ☞ 			*				*
4. 			*			*!	
5. 			*	*	*!		*

We can see that this analysis also give the same output. That leads to a conclusion that the DM analysis leaves us an indifferent syllable structure in Bangla. Let us also consider a gemination case in the next section.

6.3 Gemination

In this section, a distinct type of gemination case will be discussed. The DM analsys left us with this morphological structure where in the surface level, a gemination phenomenon takes place.

(16) k^ha-**μ**-c^hi > $k^ha_\mu c^hi$

This gemination input has a morphologically stratified syllable structure. One can clearly see that there is a moraic slot left empty in between two syllables. Let us assumes a new faithfulness constraint here which would like to preserve moras (μ) in candidates from the input.

(17) MAX-IO-μ

Moras (μ) in the input must have output correspondence.

Additionally, we bring the anti-gemination constraint NoCoda from chapter 4 (Gemination) of this work.

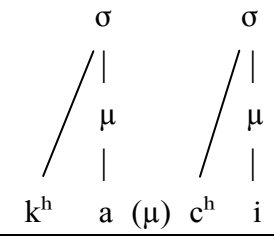
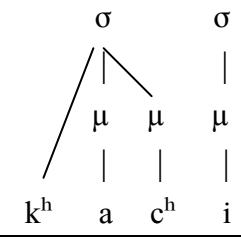
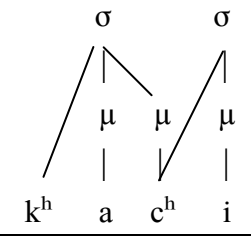
(18) NOGEM (Hall 2003)

No multiple links from a root node to a higher tier.

Keeping the constraint ranking optimal, we will consider only those

constraints and candidates which act most effectively.

(19) Gemination (morphological strcture)

/k^ha-**μ**-c^hi/	Onset	Max-IO-μ	*M/a	NoGem
1. 		*!	*	
2. ☞ 	*!			
3. 				*!

In this gemination tableau, the first candidate has an unsorted μ which claims a fatal violation of the faithfulness constraint Max-IO-μ. Candidate 2 is punished by the violation of Onset constraint. Hence, candidate 3 becomes the minimal candidate with the only violation of the lowest ranked NoGem constraint.

So, it is evident that the morphologically derived syllable structures are also accountable in the optimality theoretic analysis. Bangla shows a very flexible interconnection in between these two systems.

7 Conclusion

An important motivation for this study has been to give a fairly exhaustive account of the syllable structure of Bangla. Many works are already done in this area, but a very few of them have taken optimality theory as their preferred methodology. But, it is widely believed that OT can give a fairly strong and stable account for a structural analysis in a language. In this work, I tried to focus on the gemination patterns in Bangla and then analyze them in terms of optimality theory. The outcome of the study is quite neat. But there are still some areas that need more attention in future.

The data used in this study on syllable structure of Bangla was mainly collected by a corpus study. But, the corpus was primarily not meant for the purpose of phonological research. It is an orthography based corpus, and hence some of the data are not perfectly fetched up. But, it still gives a major input in determining the frequency of occurrences of certain sounds or clusters in Bangla vocabulary.

Many of the constraints used in the OT analysis part are already used by scholars for different languages. Some other constraints, both markedness and faithfulness, are introduced here to account for specific issues in the syllable structure analysis of Bangla. For example, the IDENT-IO(vel, -n) is a context specific faithfulness constraint that is introduced here solely for a specific purpose of this study. This might not be available in any other study. But, the most common constraints are followed here for a consistency in the OT stidies in different languages.

In general, I tried to account for the most typical cases in a syllable structure, such as consonant cluster at word medial position and other positions. This is an issue of high interest for many scholars, but most of the works do not go beyond the micro level. Here, the OT analysis is proposed to deal with those issues in Bangla phonology. I tried to give a stratified lexicon model for the OT analysis, which might not fit in the classical concept of OT analysis. But, a highly stratified language like Bangla deserves such an approach to be properly understood in the sphere of phonological research in optimality theory. But, the important point in this work is not only the OT account, but also a morphological analysis of verbal inflections in Bangla. This analysis is executed in the framework of Distributed Morphology. This part of the study has followed the theoretical account of distributed morphology and later it has been expanded by implementing the constraint based analysis after the syntactic operations are over and completes the analysis within the said framework. In general, both the morphology and the phonology sections are dependent on each other. I hope this study has shown that the surface facts of a language can be reduced to a rule-based framework that interacts with the syntax and phonology in well-defined ways. Moreover, the analysis of this work has suggested some theoretically interesting results for further studies.

In the morphology section, several tables are used to show the different slots of the inflected forms of verbs. One may find the tabular representation more technical. But I hope this type of tabular representation makes the whole discussion well-understandable. In the relevant tables, specific features are assigned to their

corresponding affixes. All the slots consist of their own set of affixes and theoretically they should not clash with other. That means, one affix can take only the specific slot assigned to it and restricted by the corresponding feature set.

There might be some minor phenomena left unaddressed which I would like to discuss in future. Despite the broad scope of a study related to DM and Moraic Theory, I have tried to narrow the discussion to only one issue: verbal inflection. It has been maintained throughout this study that all the forms are given in a particular standardized version of Bangla. Still there are some issues which are observed closely but not addressed in this study. The most prominent one is illustrated here for a closer look. Let us consider the following table where all the affixes of the AGR slot for both /k^{h}a-/ and /ken-/ are shown:

(1) Affixes of the AGR slot (DM)

Tense / *Number*		SPR	PRI	PRC	PRP	PC	PP	PHC	SP	SF	FI
1st		-i/i̯	---	-i	-i	-am	-am	-am	-am	-o	---
2nd	F	-(e)n	-(u)n	-en	-en	-en	-en	-en	-en	-en	-en
	P	-o	-o	-o	-o	-e	-e	-e	-e	-e	-(e̯)o
	I	-Ø	-Ø	-iʃ	-iʃ	-i	-i	-iʃ	-i	-i	-(i)ʃ
3rd	F	-(e)n	-(u)n	-en	-en	-en	-en	-en	-en	-en	-en
	P	-e	-(u)k	-e	-e	-o	-o	o	-o	-e	-e

In (1), two particular rows, viz., the second and third person polite forms, are shaded. One can easily find some kind of relationship present in between the elements of these couple of rows, only except the PRI and SF columns. If we compare the elements of each column, it will be clearly seen that where /-o/ occurs

in the 2P-P row, /-e/ occurs in the corresponding 3P-P row. This could be a shift from either /-e/ to /-o/ (as in PC, PHC and SP) or from /-o/ to /-e/ (as in SPR, PRI, PRC, PRP and FI). In the SP column there is no change (both are /-e/) and in the PRI column this relationship does not exist.

Appendix – I: Consonants in Bangla

	Plosive / Affricate				Fricative (Voiceless)	nasals	lateral	Tap/Flap		Glottal
	Voiceless		Voiced							
	Unaspirated	Aspirated	Unaspirated	Aspirated				Unaspirated	Aspirated	
Velar (k-class)	k	k^h	g	g^h		ŋ				
Paleto-alveolar (c-class)	c	c^h	ɟ	$ɟ^h$	ʃ					
Retroflex (T-class)	ʈ	$ʈ^h$	ɖ	$ɖ^h$				ɽ	$ɽ^h$	
Dental (t-class)	t	t^h	d	d^h	s					
Bilabial (p-class)	p	p^h	b	b^h		m				
Glottal										h
Alveolar						n	l	r		

Appendix- II

Corpus correction Protocol

The Bangla corpus is modified according to the pronunciation pattern. The following changes are made to create a possible phonological corpus (from the orthographic one).

,r.	> >	ri	.y yA	> >	E	(æ)
,r	> >	ri	.yyu	> >	Eu	(æu)
.r	> >	R	~u	> >	uM	
.n	> >	n	\	> >	*DELETED*	
~n	> >	n	~e	> >	eM	
.m	> >	N				
.d	> >	D				
.t	> >	T				
.s	> >	s				
aa	> >	A				
uu	> >	u				
ii	> >	i				
tv	> >	tt				
k.s	> >	kkh				
~a	> >	aM				
~A	> >	AM				

BIBLIOGRAPHY

Aitchison, J. (2003). Words in the mind: an introduction to the mental lexicon. Malden, Massachusetts: Blackwell Publishing.

Al-Ahmadi Al-Harbi, A. (2003). Acehnese coda condition: an optimality-theoretic account. *Umm Al-Qura University Journal of Educational and Social Sciences and Humanities 15*(1), 9-28.

Anderson, S. (1992). A-morphous morphology. Cambridge: Cambridge University Press.

Arbib, M. A. (Ed.). (2002). The Handbook of Brain Theory and Neural Networks (2 ed.). Cambridge, Massacusetts: MIT Press.

Bach, E. (1968). Two proposals concerning the simplicity metric in phonology. *Glossa 2*(2), 128-149.

Baker, M. (1985). The mirror principle and morphosyntactic explanation. *Linguistic Inquiry 16*, 373-416.

Baković, E. (2005). Antigemination, assimilation and the determination of identity. *Phonology 22*, 279-315.

Beckman, J. N. (1997). Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology 14*(1), 1-46.

Beckman, J. N. (1998). Positional Faihfulness. University of Massachusetts Amherst, Amherst. *PhD dissertation.*

Bhattacharya, K. (1993). Bengali-Oriya verb morphology. Calcutta: Dasgupta & Co. Pvt. Ltd.

Bhattacharya, T. (2000). Bangla : Bengali. In J. Gary & C. Rubino (Eds.), *Encyclopedia of World's Languages: Past and Present* New York: WWWilson.

Bhowmik, D. (2003a). Tatbhava. *Banglapedia, the national encyclopedia of Bangladesh*. Dhaka, Asiatic Society of Bangladesh. Access: http://banglapedia.search.com.bd/HT/T_0010.htm.

Bhowmik, D. (2003b). Tatsama. *Banglapedia, the national encyclopedia of Bangladesh*. Dhaka, Asiatic Society of Bangladesh. Access: http://banglapedia.search.com.bd/HT/T_0087.htm.

Bird, S. & E. Klein. (1994). Phonological analysis in typed feature systems. *Computational Linguistics 20*(3), 455-491.

Blevins, J. (1995). The syllable in phonological theory. In J. Goldsmith (Ed.), *Handbook of Phonological Theory* (pp. 106-244). London: Basil Blackwell.

Bobaljik, J. D. (2006). Itelmen reduplication: edge-in association and lexical stratification. *Journal of Linguistics 42*, 1-23.

Brockhaus, W. (1995). Final obstruent devoicing in the phonology of German. Tübingen: Niemeyer.

Broselow, E. (1999). Loanwords and learnability. Handout of a talk given at the University of Maryland, December 3, 1999.

Burquest, D. & D. Payne. (1993). Phonological analysis: a functional approach. Dallas: Summer Institute of Linguistics.

Burzio, L. (1994). Principles of English stress. Cambridge: Cambridge University Press.

Bykova, E. M. (1981). The Bengali language. Moscow: Central department of Oriental Literature, Nauka Publishing House.

Cairns, C. E. & M. H. Feinstein. (1982). Markedness and the theory of syllable structure. *Linguistic Inquiry 13*, 193-226.

Campbell, G. L. (1995). Bengali. In G. L. Campbell (Ed.), *Concise Compendium of the World's Languages*. London & New York: Routledge.

Cardona, G. & B. Suthar. (2003). Gujarati. In G. Cardona & D. Jain (Eds.), *The Indo-Aryan Language*. London: Routledge.

Causley, T. (1997). Identity and featural correspondence: the Athapaskan case. Paper presented at the *NRLS27, Amherst, Massachusetts.* GLSA. 93-105.

Chaki, J. (2001). Bangla bhasar vyakaran (*2nd ed.*). Kolkata: Ananda Publishers.

Chakrabarti, B. (1994). A comparative study of Santali and Bengali. Kolkata: K.P. Bagchi & Co.

Charles-Luce, J. (1985). Word-final devoicing in German and the effects of phonetic and sentential contexts. *The Journal of Acoustic Society of America 77*(S1).

Chatterjee, S. (1962). A study of relationship between written and colloquial Bengali. The Hartford Seminary Foundation, Hartford. *PhD thesis dissertation.*

Chatterji, S. K. (1921). Bengali phonetics. *Bulletin of the School of Oriental Studies 2*(1), 1-25.

Chatterji, S. K. (1926a). The origin and development of the Bengali language (*Vol. 1*). Calcutta: University of Calcutta.

Chatterji, S. K. (1926b). The origin and development of the Bengali language (*Vol. 2*). Calcutta: University of Calcutta.

Chatterji, S. K. (1986). A Bengali phonetic reader. Calcutta: Rupa.

Chatterji, S. K. (1988). Bhasa prakash bangla vyakaran. New Delhi: Rupa.
First published by the University of Calcutta in 1939 (Bengali 1346)

Cho, Y.-M. Y. (1990). Parameters of consonantal assimilation. University of Massachusetts, Amherst, Amherst. *PhD dissertation.*

Clark, J., C. Yallop & J. Fletcher. (2007). Introduction to phonetics and phonology. Oxford: Blackwell.

Clements, G. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (pp. 283-333). Cambridge: Cambridge University Press.

Clements, G. N. (1988). The role of sonority cycle in core syllabification. *Working Papers of the Cornell Phonetic Laboratory 2*, 1-68.

Comrie, B. (2005). Language. *Microsoft Encarta Online Encyclopedia*, 2005. Microsoft Corporation.

Dan, M. (1992). Some issues in metrical phonology : the indigenous research tradition. Deccan College, Pune. *PhD dissertation.*
Published as e-book from Central Institute of Indian Langisges (CIIL), Mysore.
Access: http://www.ciil-ebooks.net/html/mp/coverpage.html.

Das, S. (2002). Some aspects of the phonology of Tripura Bangla and Tripura Bangla English. Central Institute of English and Foreign Languages (CIEFL), Hyderabad. *PhD dissertation.*

Dasgupta, P. (2003). Bangla. In G. Cardona & D. Jain (Eds.), *The Indo-Aryan Languages* (pp. 351-390). London: Routledge.

Davenport, M. & S. J. Hannahs. (2005). Introducing phonetics and phonology (*2 ed.*). London: Hodder Arnold.

Davidson, L., P. Smolensky & P. W. Jusczyk. (2004). The initial and final states: theoretical implications and experimental explorations of richness of the base. In R. Kager, J. Pater & W. Zonneveld (Eds.), *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press. (Reprinted in Smolensky, P. & Legendre, G. 2006. The Harmonic mind: from neural computation to optimality-theoretic grammar. (Vol. 1: Cognitive Architecture; vol. 2: Linguistic and Philosophical Implications). Massachusetts: MIT Press.)

Davis, S. (1998). Syllable contact in optimality theory. *Journal of Korean Linguistics 23*, 181-211.

Dey, P. (1979). On rule ordering in Bengali phonology. *Indian Linguistics 40*(1), 24-34.

Dixon, R. M. W. (1988). A grammar of Boumaa Fijian. Chicago: Chicago University Press.

Dost, A. (2004). Sonority Distance Constraints in Syllable Evaluation. Qualifying paper: University of California, Santa Cruz

Embick, D. & R. Noyer. (2001). Movement operations after syntax. *Linguistic Inquiry 32*(4), 555-595.

Ferguson, C. A. & M. Chowdhury. (1960). The phonemes of Bengali. *Language 36*(1), 22-59.

Féry, C. (2003). Final devoicing and the stratification of the lexicon in German. In J. M. van de Weijer, V. van Heuven & H. van der Hulst (Eds.), *The Phonological Spectrum: Segmental Structure* (pp. 145-170). Amsterdam: John Benjamins.

Féry, C. (2008). *Phonologie des Deutschen: eine optimalitätstheoretische analyse*. Ms., Potsdam.

Féry, C. & R. van de Vijver (Eds.). (2003). The syllable in optimality theory. Cambridge: Cambridge University Press.

Fitzpatrick-Cole, J. (1990). The minimal word in Bengali. In A. L. Halpern (Ed.), *The Proceedings of the Ninth West Coast Conference on Formal Linguistics* (pp. 157-170). Stanford: Center for the Study of Language and Informantion. Published for the Stanford Linguistics association.

Fitzpatrick-Cole, J. (1996). Reduplication meets the phonological phrase in Bengali. *The Linguistic Review 13*, 305-356.

Fleischhacker, H. (2000). *Cluster dependent epenthesis asymmetries*. Ms., UCLA.

Flemming, E. (1995). Auditory features in phonology. UCLA, LosAngeles, California. *PhD dissertation.*

Frank, R. & G. Satta. (1997). Optimality theory and the generative complexity of contraint violability. *Computational Linguistics 24*(2), 307-315.

Fudge, E. (1969). Syllables. *Journal of Linguistics 5*, 253-286.

Gaeta, L. (2001). Striving for optimality. In K. Dziubalska-Kołaczyk (Ed.), *Constraints and Preferences* (pp. 101-142). Berlin: Walter de Gruyter.

Ghosh, A. (2003). The sounds of Bengali and French. Kolkata: The Asiatic Society.

Glück, S. & R. Pfau. (1999). A distributed morphology account of verbal inflection in German sign language. In T. Cambier-Langefeld, A. Lipták, M. Redford & E. J. v. d. Torre (Eds.), *Proceedings of ConSOLE 7* (pp. 65-80). Leiden: SOLE.

Gnanadesikan, A. (1995). *Markedness and faithfulness constraints in child phonology*. Ms., University of Massachusetts.

Gnanadesikan, A. (1997). Phonology with ternary scale. University of Massachusetts, Amherst. *PhD dissertation.*

Goldsmith, J. (1985). Vowel harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian. *Phonology Yearbook 2*, 253-275.

Goldsmith, J. (1989). Licensing, inalterability, and harmonic rule application. In C. Wiltshire (Ed.), *Papers from the 25th Meeting of the Chicago Linguistics Society* (pp. 145-156). Chicago: Chicago Linguistic Society.

Goldsmith, J. (1990). Autosegmental and metrical phonology. Oxford: Basil Blackwell.

Gordon, R. G., Jr. (Ed.). (2005). Ethnologue: languages of the World (Vol. 15). Dallas, Texas: SIL.

Gouskova, M. (2002). Falling sonority onsets, loanwords, and Syllable Contact. In M. Andronis, C. Ball, H. Elston & S. Neuvel (Eds.), *CLS 37: the main session. Papers from the 37th meeting of the Chicago Linguistic Society* (pp. 175-186). Chicago: The Chicago Linguistics Society.

Greenberg, J. H. (1978). Some generalizations concerning initial and final consonant clusters. In J. H. Greenberg (Ed.), *Universals of Human Language, Volume 2: Phonology* (pp. 243-280). Stanford: Stanford University Press.

Grijzenhout, J. & M. Krämer. (2000). Final devoicing and voicing assimilation in Dutch derivation and cliticization. In B. Bierwisch & S. Dieter Wunderlich (Eds.), *Lexicon in Focus* (pp. 55-82). Berlin: Akademie Verlag.

Hai, M. A. (1975). Dhonibiggan o Bangla dhonitatto. Dhaka: Barnamichil.

Hall, T. A. (1992). Syllable structure and syllable-related processes in German. Tübingen: Niemeyer. (Linguistische Arbeiten 276)

Hall, T. A. (2003). On the nongemination of /r/ in West Germanic twenty years later. Paper presented at the *Old World Conference in Phonology I, Leiden.*

Halle, M. & A. Marantz. (1993). Distributed morphology and the pieces of inflection. In K. Halle & S. J. Keyser (Eds.), *In The View from Building 20* (pp. 111-176). Cambridge: MIT Press.

Halle, M. & A. Marantz. (1994). Some key features of distributed morphology. *MIT Working Papers in Linguistics 21*, 275-288.

Ham, W. (1998). A new approach to an old problem: gemination and constraint reranking in West Germanic. *The Journal of Comparative Germanic Linguistics 1*, 225-262.

Harris, J. W. (1983). Syllable structure and stress in Spanish: a nonlinear analysis. In *inguistic Inquiry Monograph 8*. Cambridge, Massachusetts: MIT Press.

Hayes, B. (1989). Compensatory lenthening in moraic phonology. *Linguistic Inquiry 20*, 253-306.

Hayes, B. (1995). Metrical stress theory : principles and case studies. Chicago: The University of Chicago Press.

Hayes, B. & A. Lahiri. (1991). Bengali intonational phonology. *Natural Language and Linguistic Theory 9*(1), 47-96.

Hayes, B. & D. Steriade. (2004). Introduction: the phonetic bases of phonological markedness. In B. Hayes, R. Kirchner & D. Steriade (Eds.), *Phonetically-Based Phonology* (pp. 1-33). Cambridge: Cambridge University Press.

Hooper, J. (1976). An introduction to natural generative phonology. New York: Academic Press.

Hyman, L. M. (1986). A theory of phonological weight. Dordrecht: Foris Publication.

Itô, J. (1986). Syllable theory in prosodic phonology. University of Massachusetts Amherst, Amherst. *PhD dissertation.*

Itô, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory 7*, 217-260.

Itô, J. & A. Mester. (1995). The core-periphery structure of the lexicon and constraints on reranking. In J. N. Beckman, L. Walsh Dickey & S. Urbanczyk (Eds.), *University of Massachusetts occational papers 18* (pp. 181-209). Amherst: GISA.

Itô, J. & A. Mester. (1998). *Markedness and word structure: OCP effects in Japanese*. Ms., University of California, Santa Cruz.

Itô, J. & A. Mester. (1999). The phonological lexicon. In N. Tsujimura (Ed.), *A Handbook of Japanese Linguistics* (pp. 62-100). Oxford: Blackwell.

Jaeger, J. (1978). Speech aerodynamics and phonological universals. Paper presented at the *Proceedings of the Chicago Linguistics Society, Chicago.* 311-329.

Jakobson, R. (1972). Child language aphasia and phonological universals. The Hague: Mouton.

Jones, D. (1964). An outline of English phonetics (*9 ed.*). Cambridge: Heffer.

Jun, J. (1995). Perceptual and articulatory factors in place assimilation: an optimality theoretic approach. UCLA, Los Angeles, California. *PhD dissertation.*

Jun, J. (2001). Positional markedness vs. positional faithfulness. *Eoneohag 29*, 57-84.

Kager, R. (1999). Optimality theory. Cambridge: Cambridge University Press.

Katzner, K. (2002). The languages of the World. London & New York: Routledge.

Kawahara, S. (2005). Sonorancy and geminacy. *MS.*

Kaye, J. & J. Lowenstamm. (1981a). De la syllabicité. Unpublished MS Dissertation. Université du Québec á Montréal and University of Texas/Austin.

Kaye, J. & J. Lowenstamm. (1981b). Syllable structure and markedness theory. In A. Belletti, L. Brandi & L. Rizzi (Eds.), *Theory of markedness in generative grammar: procedings of the 1979 GLOW conference* (pp. 559-604). Pisa: Scuola Normale Superiore di Pisa.

Keer, E. (1999). Geminates, the OCP and the nature of CON. State University of New Jersey, New Brunswick. *PhD dissertation.*

Kenstowicz, M. (1994a). Phonology in generative grammar. Cambridge: Blackwell.

Kenstowicz, M. (1994b). Syllabification in Chukchee: a constraints-based analysis. In A. Davison, N. Maier, G. Silva & W. S. Yan (Eds.), *Proceedings of the Formal Linguistics Society of the Midwest 4* (pp. 160-181). Iwoa City: Department of Linguistics, University of Iwoa.

Kingston, J. (1985). The phonetics and phonology of the timing of oral and glottal events. University of California, Berkeley. *PhD dissertation.*

Kingston, J. (1990). Articulatory binding. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (pp. 406-434). Cambridge: Cambridge University Press.

Kiparsky, P. (1968). How abstract is phonology? In O. Fujimura (Ed.), *Three Dimensions of Linguistic Theory*. Tokyo: TEC.

Kiparsky, P. (1981). *Vowel harmony*. Ms., MIT.

Kirchner, R. (1998). An effort-based approach to consonant lenition. UCLA, Los Angeles, California. *PhD dissertation.*

Klaiman, M. H. (1987). Bengali. In B. Comrie (Ed.), *The World's major languages* (pp. 490-513). London & Sydney: Croom Helm.

Kopkalli, H. (1993). A phonetic and phonological analysis of final devoicing in Turkish. University of Michigan, Michigan. *PhD dissertation.*

Kostić, D. & R. S. Das. (1972). A Short outline of Bengali phonetics. Calcutta: Statistical Publishing Society.

Kothari, A. (2004). Hiatus resolution in Bengali: a non-cyclic account. Paper presented at the *24th Annual Conference of South Asian Languages Analysis (SALA), Stony Brook, New York.*

Lahiri, A. (2000). Hierarchical restructuring in the creation of verbal morphology in Bengali and Germanic: evidence from phonology. In A. Lahiri (Ed.), *Analogy, Levelling, Markedness: Principles of Change in Phonology and Morphology* (pp. 71-124). Berlin: Mouton de Gruyter.

Lamontagne, G. & K. Rice. (1995). A correspondence account of coalescence. *University of Massachusetts Occasional Papers in Linguistics 18*, 211-224.

Legendre, G., J. Grimshaw & S. Vikner (Eds.). (2001). Optimality-theoretic syntax. Cambridge, Massachusetts: MIT Press.

Lieber, R. (1992). Deconstructing morphology. Chicago: University of Chicago Press.

Lombardi, L. (1991). Laryngeal features and laryngeal neutralization. University of Massachusetts, Amherst, Amherst. *PhD dissertation.*

Lombardi, L. (1995). Laryngeal neutralization and syllable wellformedness. *Natural Language & Linguistic Theory 13*(1), 39-74.

Lombardi, L. (1999). Positional Faithfulness and Voicing Assimilation in Optimality Theory. *Natural Language & Linguistic Theory 17*(2), 267-302.

Maddieson, I. (1984). Patterns of sounds. Cambridge: Cambridge University Press.

Mallik, B. P. (1960). Phonemic analysis of the consonant clusters in standard colloquial Bengali. *Bulletin of the Philological Society of Calcuta 1*(2), 37-46.

Masica, C. P. (1991). The Indo-Aryan languages. Cambridge: Cambridge University Press.

McCarthy, J. J. (2000). The prosody of phase in Rotuman. *Natural Language and Linguistic Theory 18*(1), 147-197.

McCarthy, J. J. (2001). A Thematic guide to optimality theory. Cambridge: Cambridge University Press.

McCarthy, J. J. (2003). Sympathy, cumulativity, and the Duke-of-York Gambit. In C. Féry & R. van de Vijver (Eds.), *The syllable in optimality theory* (pp. 23-76). Cambridge: Cambridge University Press.

McCarthy, J. J. & A. Prince. (1985). Dissecting the skeleton. Paper presented at the *1985 West Coast Conference on Formal Linguistics*.

McCarthy, J. J. & A. Prince. (1993a). Generalised alignment. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1993* (pp. 79-153). Doedrecht: Kluwer.

McCarthy, J. J. & A. Prince. (1993b). *Prosodic morphology: constraint interaction and satisfaction*. Ms., University of Massachusetts, Amherst and Rutgers Centre for Cognitive Science (RuCCS) technical report 3. ([ROA-482-1201, Rutgers Optimality Archive, URL: http://roa.rutgers.edu])

McCarthy, J. J. & A. Prince. (1994). The emergence of unmarked: optimality in prosodic morphology. In M. Gonzalez (Ed.), *Proceedings of the North East Linguistics Society 24* (pp. 333-379). Amherst, Massachusetts: GLSA.

McCarthy, J. J. & A. Prince. (1995). Faithfulness and Reduplicative Identity. In J. Beckman, S. Urbanczyk & L. Walsh Dickey (Eds.), *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory* (pp. 249-384). Amherst, Massachusetts: GLSA.

McCarthy, J. J. & A. Prince. (1999). Faithfulness and identity in prosodic morphology. In R. Kager, H. van der Hulst & W. Zonneveld (Eds.), *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press.

Murray, R. & T. Vennemann. (1983). Sound change and syllable structure in Germanic phonology. *Language 59*, 514-528.

Noyer, R. & H. Harley. (1999, April). Distributed morphology. *Glot International, 4,* 3-9.

Ohala, J. (1983). The origin of sound patterns in vocal tract constraints. In P. MacNeilage (Ed.), *The Production of Speech* (pp. 189-216). New York: Springer Verlag.

Oostendorp, M. v. (2007). An exception to final devoicing. In v. d. Torre, E. Jan & J. v. d. Weijer (Eds.), *Voicing in Dutch*. Amsterdam: John Benjamins.

Paradis, C. (1988). On constraints and repair strategies. *The Linguistic Review 6*(1), 71-97.

Paradis, C. (1990). "Focus" in Gere configurational constraints. In J. P. Hutchison & V. Manfredi (Eds.), *Current Approaches to African Linguistics* (Vol. 7). Dordrecht: Foris.

Paradis, C. & J. F. Prunet (Eds.). (1991). Coronals, Consonant Clusters and the Coda Condition. San Diego: Academic Press Inc.

Pater, J. (1999). Austronesian nasal substitution and other NC effects. In R. Kager, H. van der Hulst & W. Zonneveld (Eds.), *The Prosody-Morphology Interface* (pp. 310-343). Cambridge: Cambridge University Press.

Prince, A. (1984). Phonology with tiers. In M. Aronoff & R. T. Oehrle (Eds.), *The Language Sound Structure* (pp. 234-244). Cambridge: MIT Press.

Prince, A. & P. Smolensky. (1993). *Optimality theory : constraint interaction in generative grammar*. Ms., Rutgers University, New Brunswick and University of Colorado, Boulder. Later published as in 2004 from Blackwell Publishing, Malden, MA.

Prince, A. & P. Smolensky. (1997). Optimality: from neural networks to universal grammar. *Science 275*, 1604-1610.

Radice, W. (1994). Teach yourself Bengali. London: Hodder & Stoughton.

Ray, P. S., M. A. Hai & L. Ray. (1966). Bengali language handbook. Washington, D.C.: Center for Applied Linguistics.

Ringen, C. (1975). Vowel Harmony: Theoretical Implications. Indiana University. *PhD dissertation.*

Rose, S. (2000). Epenthesis positioning and syllable contact in Chaha. *Phonology 17*, 397-425.

Rosenthall, S. (1994). Vowel/glide alternation in a theory of constraint interaction. University of Massachusetts, Amherst. *PhD dissertation.*

Rubach, J. (1989). Final devoicing and cyclic syllabification in German. *Linguistic Inquiry 21*(1), 79-94.

Rubach, J. (2004). Derivation in Optimality Theory: A Reply to Burzio. *Linguistic Inquiry 35*(4), 656-670.

Sarkar, P. (1985). Bangla diswar dhoni. Kolkata: Bhasa.

Sarkar, P. (1986). Aspects of Bengali syllables. Paper presented at the *National Seminar on the Syllable in Phonetics and Phonology, Hyderabad.* Osmania University.

Sarkar, P. (2006). Bangla balo. Kolkata: Dey's Publishing.
First published in 1990.

Scobbie, J. (1991). Attribute value phonology. University of Edinburgh, Edinburgh. *PhD dissertation.*

Selkirk, E. (1978). The syllable. In H. van der Hulst & N. Smith (Eds.), *The Structure of Phonological Representations, Part - II (1982)* (pp. 337-383). Dordrecht: Foris Publications.

Selkirk, E. (1980). Prosodic domains in phonology: Sanskrit revisited. In M. Aronoff & M.-L. Kean (Eds.), *Juncture*. Saratoga, CA: Anma Libri.

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of Phonological Representations*. Dordrecht: Foris Publications.

Sen, S. (1993). Bhashar itibritto Kolkata: Ananda Publishers Private Limited.
First published in 1932 (Bengali 1339).

Shepherd, M. A. (2003). Constraint interactions in Spanish Phonotactics: an optimality theory analysis of syllable-level phenomena in the Spanish Language. California State University, Northridge, California. *MA dissertation.*

Shinohara, S. (2002). Emergence of Universal Grammar in foreign word adaptations. Paper presented at the *CNRS UPRESA 7018, Université Paris 3.*

Silverman. (1992). Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology 9*, 289-328.

Simmler, F. (1974). Die westgermanische konsonantengemination im Deutschen unter besonderer berücksichtigung des Althochdeutschen. München: Fink.

Singh, U. N. (1980). Comments on 'On rule ordering in Bengali phonology'. *Indian Linguistics 40*(2), 91-101.

Smolensky, P. (2005). Optimality in phonology II: Harmonic completeness, local constraint conjunction, and feature-domain markedness. *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*, 590-716.

Spencer, A. J. (1994). Syllabification in Chukchee. In R. Weise (Ed.), *Recent Developments in Lexical Phonology* (pp. 205-226).
Theorie des Lexikons. Arbeiten des Sonderforschungsbereichs 282, Nr. 56.
Düsseldorf: Heinrich Heine University.

Steriade, D. (1979). Vowel harmony in Khalkha Mongolian. In K. Safir (Ed.), *Papers on Syllable Structure, Metrical Structure and Harmony Processes* (pp. 25-50).
Cambridge: MITWPL.

Steriade, D. (1982). Greek prosodies and the nature of syllabification. MIT, Cambridge.*dissertation.*

Steriade, D. (1995). Underspecification and markedness. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 114-174). Oxford: Basil Blackwell.

Taylor, M. (1985). Some patterns of geminate consonants. *The University of Chicago Working Papers in Linguistics 1*, 120-129.

Trubetzkoy, N. S. (1939). Grundzüge der Phonologie (*C. A. M. Baltaxe, Trans.*).
Berkeley: University of California Press.

Truckenbrodt, H. (2002). Variation in p-phrasing in Bengali. *Linguistic Variation Yearbook 2*, 259-303.

Vennemann, T. (1988). Preference laws for syllable structure and the explanation of sound change. Berlin: Mouton de Gruyter.

Vennemann, T. & R. Murray. (1972). On the theory of syllabic phonology. *Linguistische Berichte 18*, 1-18.

Vijayakrishnan, K. G. (2003). Weakening processes in the optimality framework. In J. M. van de Weijer, V. van Heuven & H. van der Hulst (Eds.), *The Phonological Spectrum* (pp. 241-255). Amsterdam: John Benjamins.

Weber, G. (1997). Top languages. *Language Today*(3), 12-18.

Wheeler, M. (2005). Cluster Reduction: Deletion or Coalescence? *Catalan Journal of Linguistics 4*.

Wiese, R. (1996). The phonology of German. Oxford: Clarendon Press.

Yavaş, M. (2003). Role of sonority in developing phonologies. *Journal of Multilingual Communication Disorders 1*(2), 79-98.

Yip, M. (1991). Coronals, Consonant Clusters and the Coda Condition. In C. Paradis & J. F. Prunet (Eds.), *The Special Status of Coronals* (pp. 61-78). San Diego: Academic Press Inc.

Zoll, C. (1996). Parsing below the segment in a constraint based framework. University of California, Berkeley, CA. *PhD dissertation.*

Zoll, C. (1998). *Positional Asymmetries and Licensing*. Ms., ROA#282.

Zuraw, K. R. (2003). Optimality theory in linguistics. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (2 ed., pp. 819-822). Cambridge, Massachusetts: MIT Press.