Vowel epenthesis in Bengali: An Optimality Theory analysis

Khaled Karim
Ph.D. Candidate, University of Victoria Linguistics
khaledk@uvic.ca

This paper examines the occurrence of epenthetic vowels before and between the initial consonant clusters in Bengali speakers of English, and provides an Optimality Theory (OT) analysis to account for this phenomenon. Native Bengali words disallow initial consonant clusters, and many word-initial consonant clusters in loan words are simplified according to these phonotactics. The maximum syllabic structure is CVC in Bengali and speakers often carry this restriction over to loan words. For example, geram (CV.CVC) instead of gram (CCVC) for the Sanskrit loan word "village", or iskul (VC.CVC) instead of skul (CCVC) for the English word "school" (Kar, 2009). I argue that in rising sonority clusters, a vowel is inserted between the two consonants and in falling sonority clusters (i.e., [s]-stop clusters) the vowel is inserted before the consonant cluster. I also explain that the sonority sequencing constraint SYLLABLE CONTACT treats [s]-stop clusters differently from obstruent-sonorant clusters, and the differing epenthesis pattern can be explained properly if it is considered an effect of SYLLABLE CONTACT – the preference of sonority to fall across a syllable boundary, which was proposed by Murray and Venneman (1983) and also supported by Gouskova (2001). With tableaux, I demonstrate that the epenthesis in consonant clusters is caused by the prohibition on consonant clusters in Bengali and the site of epenthesis is determined by SYLLABLE CONTACT (Gouskova, 2001). I also demonstrate that the constraint that prefers epenthesis before the [s]-stop cluster is CONTIG-IO (Kager, 1999). Furthermore, I propose that apart from SYLLABLE CONTACT, two other constraints *OO and *OR can also account for the vowel epenthesis in Bengali.

1 Introduction

In this paper, I examine the occurrence of epenthetic vowels before and between the initial consonant clusters in Bengali speakers of English. Native Bengali words disallow initial consonant clusters and many word-initial consonantal clusters in loan words are simplified according to these phonotactics. The maximum syllabic structure is CVC and speakers often carry this restriction over to loan words. For example, *geram* (CV.CVC) instead of *gram* (CCVC) for the Sanskrit loan word "village" or *iskul* (VC.CVC) instead of *skul* (CCVC) for the English word "school" (Kar, 2009). I argue that this epenthesis is sensitive to sonority. Clusters that rise in sonority are broken up by an epenthetic vowel, and clusters that fall in sonority are resolved by placing an epenthetic vowel before the cluster. In addition, I find that obstruent clusters involving [s] pattern differently from obstruent-sonorant clusters.

The structure of the paper is as follows. First, I present notable features of Bengali language and two sets of data for two vowel epenthesis patterns in Bengali. I then provide an analysis of the data within the framework of Optimality Theory (Prince and Smolensky, 1993). Finally, I demonstrate the constraints in tableaux and present my alternative analysis and conclusion.

2 Notable features of Bengali

Bengali is an eastern Indo-European language with approximately 211 million speakers in Bangladesh and the Indian state of West Bengal. Bengali emerged as a new Indo-Aryan language by 900–1000 AD through Magadhi Apabrangsa and Abahatta, two stages of Magadhi Prakrit (600 BC – 600 AD), along with two other Indo-Aryan languages, Oriya and Assamese (Chatterji, 1926). Until the 14th century, there was little linguistic difference between Bengali and Assamese. Bengali has two literary styles: one is called *Sadhubhasa* (elegant language) and the other *Chaltibhasa* (current language). The former is the traditional literary style based on Middle Bengali of the sixteenth century, while the later is a 20th century creation and is based on the speech of educated people in Calcutta (Sahidullah, 1965). The differences between the two styles are not huge and involve mainly forms of pronouns and verb conjugations (Sahidullah, 1965).

The Bengali alphabet is a syllabic alphabet in which consonants all have an inherent vowel. Vowels can be written as independent letters, or by using a variety of diacritical marks which are written above, below, before or after the consonant they belong to. Word order in Bengali is SOV (Kar, 2009), for example, *ami (I) bhat (rice) khai (eat)* instead of “I eat rice” in English.

3 The problem and related data

The restrictions on word-initial consonant clusters in native Bengali carry over to the pronunciation of English words by Bengali speakers learning English as a foreign language. These learners use a strategy of vowel epenthesis to break up initial consonant clusters.

Sometimes vowel epenthesis occurs between the two consonants of the consonant clusters. For example:

(1)	ENGLISH		BENGALI
a.	Front:	/frʌnt/	/fərʌnt/
b.	Flat:	/flæt/	/fəlæt/
c.	Cream:	/krim/	/kərim/
d.	Group:	/grup/	/gərup/
e.	Floor:	/flɔr/	/fəlor/

(adapted from Islam, 2004)

In some cases epenthesis occurs before the initial consonant clusters. For example:

(2)	ENGLISH		BENGALI
a.	Special:	/'speɪʃl/	/ɪspeɪʃal/
b.	Spain:	/'speɪn/	/ɪspeɪn/
c.	Station:	/'steɪʃn/	/ɪsteʃon/
d.	School:	/sku:l/	/ɪskul/

(adapted from Islam, 2004)

But, when a consonant cluster occurs between two vowels, epenthesis does not occur. For example:

(3)

a. Astonish: /əstɒn ɪʃ/
b. Continue: /kəntɪn yu/
c. Monday: /mʌn deɪ/
d. April: /eɪ prəl/

(data source: author)

4 Analysis

The data sets in section 3 illustrate two different strategies to break the consonant clusters. When the words start with obstruent and resonant, the vowel insertion occurs in between obstruent and resonant. And when the words start with obstruent [s] followed by a stop, then epenthesis occurs word initially. In this section an analysis of the place of epenthetic vowel[1] will be considered for analysis within the framework of Optimality Theory (McCarthy & Prince 1993; Prince & Smolensky, 1993).

In the Optimality Theory (henceforth OT) structure phonological constraints are ranked and violable (Prince & Smolensky, 1993). These constraints are minimally violated by potential surface forms (possible set of candidates) and the one which violates the lowest ranked constraints wins. The seriousness of a violation depends on the hierarchies of constraints and the violations of higher-ranked constraints are most serious. There are two types of constraints: markedness and faithfulness constraints (Prince & Smolensky, 1993). Markedness constraints enforce well-formedness of the output candidate, prohibiting structures that are difficult to produce or comprehend, such as consonant clusters (Prince & Smolensky, 1993). These constraints usually impose restrictions on the occurrence of certain segments. Examples of such markedness constraint are: syllables must not have codas (NOCODA); syllables must have onsets (ONSET); and obstruents must not be voiced (*VDOBS) (Kar, 2009). Faithfulness constraints, on the other hand, impose similarity between input and output. For instance, all morphosyntactic features in the input to be overtly realized in the output (Kar, 2009). Some of the faithfulness constraints are: the output must present all segments present in the input (DEP-IO); elements adjacent in the input must be adjacent in the output (CONTIGUITY); and input segments must have counterparts in the output (MAX-IO) (Kar, 2009).

The data in (1) and in (2) illustrate that there is a restriction against word initial consonant clusters and there is a different epenthesis site for [s]-obstruent clusters. This kind of restriction can be translated into an OT constraint. The constraint is called *CC$_{\text{ONS,}}$ which assigns a violation mark to words with consonant clusters (Kager, 1999). For example, outputs like ‘special’ or ‘front’ will not be allowed. An output with word-initial vowel epenthesis to break the consonant cluster is a possible solution. For example: /ɪspeɪʃal/. Another solution is output with vowel epenthesis between consonants of the initial cluster. For example: /fərʌnt/. But these will be possible only at the cost of violation of the faithfulness constraint DEP-IO, which assigns a violation mark to words with epenthetic vowels. Other possible solutions (given example inputs /speɪʃal / and /frʌnt/) are shown overleaf.

[1] The epenthetic vowel, I argue, is a central schwa-like vowel and this schwa can take colour from surrounding segments. But the analysis of the quality or different kinds of vowel is beyond the scope of this paper.

Solutions:	Possible output candidates:
Do nothing (retain 2 adjacent obst. clusters)	[speɪʃal] & [frʌnt]
Appear with one deleted segment	[peɪʃal] & [fʌnt]

Candidates:	Violations:
[speɪʃal] & [frʌnt]	*CC_{ONS}
[peɪʃal] & [fʌnt]	MAX-IO

The constraint MAX-IO assigns a violation mark to words with deleted segments (Kager, 1999). Therefore, outputs like [peɪʃal] and [fʌnt] with one deleted obstruent will be violations of MAX-IO. The constraints, as we see so far, for Bengali data are: *CC_{ONS}, DEP-IO and MAX-IO.

There is an interesting difference in the types of consonants involved in the kinds of consonant clusters which get broken up by internal epenthesis, versus the kinds of consonant clusters which get resolved by initial epenthesis. The first type involves sonorant and obstruent segments and the second type involves obstruents, specifically [s] followed by a stop. The different epenthesis process for '[s]-stop' clusters can be explained by the fact, which Gouskova (2001) correctly observed, that a sonority sequencing constraint such as SYLLABLE CONTACT treats '[s]-stop' clusters differently from obstruent-sonorant clusters. According to Gouskova (2001), the split epenthesis pattern (also evident in Hindi, Sinhalese, Wolof and Uyghur) can be explained properly if it is considered an effect of SYLLABLE CONTACT-the preference of sonority to fall across a syllable boundary, proposed by Murray & Venneman (1983). The epenthesis in consonant clusters is caused by the prohibition on consonant clusters in Bengali but the site of epenthesis is determined by SYLLABLE CONTACT (Gouskova, 2001). Vowel epenthesis occurs before the cluster whenever the first consonant is of higher sonority than the second consonant (e.g., speɪʃal→ ɪspeɪʃal) and on the other hand, the epenthesis is inside the two initial consonants whenever the first consonant is of lower sonority than the second consonant (e.g., frʌnt→ fərʌnt).

The constraint that prefers epenthesis before the cluster is CONTIGUITY-IO (Kager, 1999). This constraint ensures the epenthesis before the consonants in [s]-obstruent clusters when SYLLABLE CONTACT is not at stake. So, the candidates [sɪpeɪʃal], [fərʌnt] and [əfrʌnt] will have the following violations:

Candidates:	Vioaltions:
[speɪʃal] and [fərʌnt]	CONTIG-IO & DEP-IO
[əfrʌnt]	SYLLABLE CONTACT

While the candidate [əfrʌnt] will have the violation of SYLLABLE CONTACT, the candidate [ɪspeɪʃal] will not violate this constraint.

4.1 Constraint definitions

$*CC_{ONS}$:	No consonant cluster in the onset.
MAX-IO:	Input segments must have output correspondents (No deletion).
DEP-IO:	No epenthesis.
CONTIGUITY-IO:	No medial epenthesis or deletion of segment.
SYLLABLE CONTACT:	Sonority must not rise across a syllable boundary (Murray & Vennman, 1983; Gouskova, 2001).

4.2 Crucial ranking

The markedness constraint needs to be crucially ranked higher than faithfulness constraint to select an optimal candidate which shows an alternation over other possible candidates which do not. The optimal candidate will violate faithfulness constraints, therefore DEP-IO needs to be ranked lower than $*CC_{ONS}$ to allow epenthesis. This faithfulness constraint DEP-IO is crucially ranked[2] to select the optimal candidate. Another faithfulness constraint CONTIG-IO also needs to be ranked lower than $*CC_{ONS}$ and SYLLABLE CONTACT to determine the optimal candidate. The markedness constraint SYLLABLE CONTACT needs to be ranked above CONTIG-IO to ensure the epenthesis site. Faithfulness constraint MAX-IO needs to be ranked above DEP-IO and CONTIG-IO but below $*CC_{ONS}$ to account for the optimal candidate. But the constraints MAX-IO and SYLLABLE CONTACT[3] are not crucially ranked with respect to each other. So, the ranking of the constraints to account for the Bengali data is as follows in (4):

(4)

$*CC_{ONS}$ >>SYLLABLE CONTACT, MAX-IO>> CONTIG-IO>>DEP-IO.

[2] As it represents the crucial ranking, it needs to be indicated with solid lines in tableaux.
[3] MAX-IO and SYLLABLE CONTACT can be kept in dashed lines as the order of their ranking would provide the same result.

4.3 Tableaux

Tableau 1

/frʌnt/	*CC_{ONS}	SYLLABLE CONTACT	MAX-IO	CONTIG-IO	DEP-IO
a. ☞ fərʌnt				*	*
b. əfrʌnt		*!			*
c. frʌnt	*!				
d. fʌnt			*!		

In Tableau 1, candidate (a) [fərʌnt] is the winning candidate because, although it violates two lower ranked constraints CONTIG-IO and DEP-IO, it satisfies higher ranked constraints. Candidate (b) [əfrʌnt] violates lower ranked candidate DEP-IO but gets ruled out for violating higher ranked constraint SYLLABLE CONTACT. Candidate (c) [frʌnt] loses for violating the higher ranked constraints*CC_{ONS}. The last candidate (d) [fʌnt] also loses for violating MAX-IO, a higher ranked constraint.

Tableau 2

/speɪʃal/	*CC_{ONS}	SYLLABLE CONTACT	MAX-IO	CONTIG-IO	DEP-IO
a. ☞ ɪspeɪʃal					*
b. sɪpeɪʃal				*!	*
c. speɪʃal	*!				
d. peɪʃal			*!		

In Tableau 2, candidate (a) [ɪspeɪʃal] is the winning candidate because it has no fatal violation. Although it violates the lower ranked constraints DEP-IO, it satisfies the higher ranked constraints. Candidate (b) [sɪpeɪʃal] violates CONTIG-IO and DEP-IO but gets ruled out for violating high ranked constraint CONTIG-IO. Candidate (c) [speɪʃal] violates the highest ranking constraint *CC_{ONS} and thus gets ruled out. The last candidate (d) [peɪʃal] violates only MAX-IO but gets ruled out for violating this constraint as this is also a higher ranked constraint.

5 Alternative analysis

As an alternative analysis, I propose that, apart from SYLLABLE CONTACT, two other constraints *OO and *OR[4] can also be used to account for the vowel epenthesis in Bengali. The constraint *OO does not allow two adjacent obstruents in a word, i.e., /speɪʃal/ will not be allowed. The other constraint *OR does not allow obstruents followed by a resonant in a word, therefore /frʌnt/ will not be allowed. It is noteworthy that, CONTIG-IO will also be required to account for the different vowel epenthesis process, as this constraint prefers epenthesis before the consonant clusters (i.e., [s]-obstruent clusters) (Gouskova, 2001). The ranking of the constraint will make sure that the optimal candidates win. The constraints used in the alternative analysis are defined as follows:

*OO:	Two adjacent obstruents are not allowed in a word.
*OR:	Obstruents followed by resonants are not allowed in a word.
CONTIG-IO:	No medial epenthesis or deletion of segment.
*CC_{ONS}:	No consonant cluster in the onset.
MAX-IO:	Input segments must have output correspondents (no deletion).
DEP-IO:	No epenthesis.

5.1 Constraint ranking

The markedness constraint needs to be crucially ranked higher than faithfulness constraint to determine the optimal candidate from the possible candidates. The optimal candidate violates faithfulness constraints, therefore CONTIG-IO and DEP-IO must be ranked lower than *CC. The faithfulness constraint DEP-IO is crucially ranked to determine the optimal candidate. Another faithfulness constraint CONTIG-IO also needs to be ranked lower than *CC_{ONS}, *OR and MAX-IO to determine the optimal candidate. The markedness constraint *OO[5] also needs to be ranked lower to determine the optimal candidate. So, the ranking of the constraints to account for the Bengali data is as follows in (5) overleaf.

---

[4] These two constraints were proposed by me and used after consultation with Dr. Marion Caldecott, who was the instructor for the course which became the genesis of this paper.

[5] *OO needs to be ranked crucially (solid line) in the case of the candidates in data set 2 to win, as the optimal candidate will violate this constraint. Rest of the constraints can be kept in dashed lines as the order of their ranking would give the same result, i.e., the candidates other than optimal ones would be eliminated if the ranking was different for *OR, CONTIG-IO and MAX-IO; but CONTIG-IO needs to be ranked lower than *CC_{ONS}, *OR and MAX-IO to account for the optimal candidate from dataset 1, as the optimal candidate violates this constraint and it should be kept in solid line.

(5) $*CC_{ONS,}$ *OR, MAX-IO>> CONTIG-IO>>*OO>>DEP-IO

Tableau 3

/frʌnt/	$*CC_{ONS}$	*OR	MAX-IO	CONTIG-IO	*OO	DEP-IO
a. ☞ fərʌnt				*		*
b. əfrʌnt		*!				*
c. frʌnt	*!	*				
d. fʌnt			*!			

In Tableau 3, candidate (a) [fərʌnt] is the winning candidate because, although it violates two lower ranked constraints CONTIG-IO and DEP-IO, it satisfies higher ranked constraints. Candidate (b) [əfrʌnt] violates the lower ranked DEP-IO and gets ruled out for violating the higher ranked constraint *OR, which is a fatal violation. Candidate (c) [frʌnt] violates two higher ranked constraints$*CC_{ONS}$ and *OR and gets ruled out for violating the highest ranked constraint $*CC_{ONS}$. The last candidate (d) [fʌnt] has only one violation, i.e., MAX-IO, but gets ruled out as this is a higher ranked constraint.

Tableau 4

/speɪʃal/	$*CC_{ONS}$	*OR	MAX-IO	CONTIG-IO	*OO	DEP-IO
a.☞ ɪspeɪʃal					*	*
b. sɪpeɪʃal				*!		*
c. speɪʃal	*!				*	
d. peɪʃal			*!			

In Tableau 4, candidate (a) [ɪspeɪʃal] is a winning candidate because, although it violates two lower ranked constraints *OO and *DEP-IO, it satisfies the high ranked constraints. Candidate (b) [sɪpeɪʃal] violates CONTIG-IO and DEP-IO and gets ruled out for violating high ranked constraint CONTIG-IO. Candidate (c) [speɪʃal] violates a lower ranked constraint *OO and also violates the highest ranking constraint *CC and thus gets ruled out. And the last candidate (d) [peɪʃal] violates only MAX-IO but gets ruled out for violating this higher ranked constraint.

6 Conclusion

In this paper, I have provided an OT analysis to account for the vowel epenthesis in Bengali language. I have shown that the primary analysis properly explains the reason why Bengali has a split epenthesis pattern, i.e., the different epenthesis process for [s]-stop clusters than other obstruent clusters (i.e., clusters with obstruent and resonants). I have argued that in rising sonority clusters, a vowel is inserted between the two consonants and in falling sonority clusters (i.e., [s]-stop clusters), the vowel is inserted before the consonant cluster. I have also explained that, the sonority sequencing constraint SYLLABLE CONTACT treats [s]-stop clusters differently from obstruent-sonorant clusters, and the differing epenthesis pattern can be explained properly if it is considered an effect of SYLLABLE CONTACT, the preference of sonority to fall across a syllable boundary, which was proposed by Murray & Venneman (1983) and also supported by Gouskova (2001). With tableaux I have demonstrated that the epenthesis in consonant clusters is caused by the prohibition on consonant clusters in Bengali but the site of epenthesis is determined by SYLLABLE CONTACT (Gouskova, 2001). I have also demonstrated that the constraint that prefers epenthesis before the [s]-stop cluster is CONTIG-IO (Kager, 1999). Furthermore, I have proposed an alternative analysis, where I have demonstrated with tableaux, that instead of SYLLABLE CONTACT, two other constraints *OO and *OR can also account for vowel epenthesis in Bengali.

Acknowledgements

I am truly grateful to Janet Leonard for providing valuable insights into my data analysis. I am also grateful to my instructor Dr. Marion Caldecott for her valuable suggestions and feedback on several occasions and above all for her care and understanding. Also, thanks to three anonymous reviewers for their useful suggestions.

References

Chatterji, S.K. (1926). *The origin and development of the Bengali language.* Calcutta University, Calcutta.

Gouskova, M. (2004). Relational hierarchies in Optimality Theory: The case of syllable contact. *Phonology* 21, 201–250.

Islam, A. (2004). *L1 influence on the spoken English proficiency of Bengali speakers*. C-Essay in English, Högskolan Dalarna University.

Kager, R. (1999). *Optimality Theory.* Cambridge: Cambridge University Press.

Kar, S. (2009). *The syllable structure of Bangla in Optimality Theory and its application to the analysis of verbal inflectional paradigms in distributed morphology.* Ph.D. dissertation, Tübingen University.

McCarthy, J.J. & Prince, A. (1993). Generalised alignment. In: G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1993*. Doedrecht: Kluwer, 79–153.

Prince, A. & Smolensky, P. (1993). Optimality Theory: Constraint interaction in Generative Grammar. Ms., Rutgers University, New Brunswick and University of Colorado, Boulder.

Shahidullah, M. (1965). *Bangla bhasar itibirtto*. Dhaka: Mowla Brothers.