PERCEPTION, PRODUCTION, AND THE IMPLEMENTATION OF PHONOLOGICAL OPACITY IN THE BANGLA VOWEL CHAIN SHIFT

Traci Christine Nagle

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Linguistics,
Indiana University
August 2020

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Doctoral Committee

_____
Kelly H. Berkson, Ph.D., chairperson

_____
Stuart Davis, Ph.D.

_____
Kenneth de Jong, Ph.D.

_____
Michael P. Adams, Ph.D.

July 9, 2020

for Ma,

and for her granddaughter, Tara

Acknowledgments

This project has accrued a globe-spanning arc of indebtedness. Foremost among those who have helped me complete this work are the members of my committee. Without Kelly Berkson this dissertation would never have been finished. I showed up out of nowhere to ask that she supervise my work, and it was the best decision I made in my entire graduate career. Encouragement, understanding, prodding, innovation, and advice—she offered it all, and always at just the right moments. Ken de Jong taught me phonetics, coached me in statistics, and seemed always to have faith in me even when I had little in myself. Stuart Davis has given me unfailingly wise counsel for many years and through many stages of my work. Michael Adams and I arrived at IU in the same year; he inspired me to study the linguistics and lexicography of English for my masters, a secondary line of research that continues to bring me joy (and publications). I thank each of them for their patience and support.

The research described herein began as a term paper in Advanced Phonology, taught by Dan Dinnsen and Ashley Farris-Trimble. I am grateful to both for inspiring me to pursue phonology and for pointing out this pattern in Bangla before I even knew what a chain shift was. After Dan retired, Michael Becker landed in Indiana, and in his brief time here he brought the buried programmer in me back to life and helped me craft the experimental dissertation that I had in mind.

Unsurprisingly, given that I began this project under Dan Dinnsen, my initial intention was to study how children acquired the vowel chain shift in Bangla verbs. I spent 5 months in India collecting recordings of sometimes-hilarious child speech with the invaluable help of Poulomi Paul and Soumi Banerjee. Thanks to them not only do I have many hours of child data for future

work, but I now recognize the advantages of the subway in Kolkata and can find College Street and Howra without a map. I promise I'll get to those child recordings soon.

At Indiana University, I benefited from the assistance of undergraduate researchers in the Phonetics and Phonology Lab, particularly Caroline Ehlhardt, Katie Blake, Chelsea Bonhotal, Abby Elston, Jacob Henry, Kristina Mihajlovic, and Max Nelson.

No project this long or this complex would have been possible without a village worth of friends and colleagues in the U.S., Canada, and Kolkata who helped in innumerable ways: the late Gautam Adhikari, Abhijit and Ilora Basu, Anumita Basu, Shohini Bhattasali, Nayana Bose, Becca Cambridge, Anamitraa Chakraborty, Swapan Chakravorty, Shibashis Chatterjee, Mina Dan, Anupam Das, Paramita Das, Suranjan Das, Stephanie Dickinson, Nandini Gupta, Walt Hakala, Abdul Hakim, Alisa Hendrix, Taj Islam, Ujjwal Jana, Ananda Lal, Shamsuddin Mahmud, Shivnath Mazumder, Shikha Mukerjee, Dipika Mukherjee, Sreemati Mukherjee, Rahul Mukherji, Joe Pater, Rinku Roy Chowdhury, Atanu Saha, Amitava Sarkar, Saurov Syed, and Rashad Ullah. I am also grateful to all the participants in my experiments, some of whom I know or got to meet, and others who remain anonymous even to me.

Various iterations of this work were presented as papers and posters at AMP (5th, New York University), LabPhon (15th, Cornell University), FASAL (6th, Massachusetts–Amherst), MidPhon (20th, Indiana–Bloomington; 19th, Wisconsin–Madison; and 13th, Ohio State), ASA (Pittsburgh), and at several departmental conferences and colloquia. I benefited there from audience questions and comments, and from stimulating conversation with like-minded scholars.

The travel required for data collection and conference presentations and graduate study in general, are costly but were made easier by funding from various entities, including the IU Linguistics Department, the IU Center for the Study of Global Change, the IU India Studies

Program, the IU Graduate and Professional Student Organization, the Kappa Alpha Theta Foundation, and the IU Credit Union.

The support and encouragement of my family—Anirban, Ananya, Arko, Mom, and Lou—sustained me all these years. I only wish my *shashuri* Nandini were still with us to see that I finally finished this. It was because of her son, Sumit, that I first started to study Bangla and thanks to him that I was able to get through to this finish line. Lastly, helping Tara grow into the intelligent, kind, and accomplished person she is took a lot of time away from this dissertation, but it was worth every minute.

Traci Christine Nagle

PERCEPTION, PRODUCTION, AND THE IMPLEMENTATION OF PHONOLOGICAL OPACITY IN THE BANGLA VOWEL CHAIN SHIFT

This dissertation explores the productivity of a phonologically opaque synchronic chain shift in Bangla, a language spoken in India and Bangladesh by 265 million people. Bangla exhibits a vowel-harmony chain shift in which certain verb-stem vowels alternate [(i ~ e), (e ~ æ), (u ~ o), and (o~ ɔ)] based on the identity of a vowel in the neighboring suffix. This pattern is called a chain shift because harmonization with affix vowels causes stem vowels to move only one step along an ordered scale of vowel heights. Previous work has suggested that opaque patterns are difficult to learn and may not be extended to nonce words (e.g., Zhang 2016). Using production and perception experiments conducted with native speakers of Bangla, this dissertation investigates the nature of this alternation pattern in the grammar of native Bangla speakers. The central question is whether speakers can extend this chain-shift pattern to nonce words.

Though the chain shift pattern is exceptionless in real verbs, the results indicate that Bangla speakers do not freely extend this pattern to nonce words. Instead, they apply the chain shift only about half of the time. Analyses conducted to probe possible acoustic explanations for these results (e.g., incomplete neutralization, vowel-to-vowel coarticulation) suggest that the patterns observed cannot be explained by phonetics alone. Rather, they seem to point to underlearning (Becker, Ketrez, & Nevins 2011) of the chain shift pattern.

As Bangla remains under-examined in modern theoretical linguistic and phonological literature, this dissertation contributes to the general literature on Bangla. Its findings also carry implications for formal models of phonology and language learning, and for the

conceptualization and modeling of synchronic chain shifts. The results indicate that chain shifts, and possibly opaque patterns more generally, may operate very differently in cognitive systems than some theoretical models have assumed.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> This subject has hitherto been utterly disregarded in Europe; and it is scarcely believed that Bengal ever possessed a native and peculiar dialect of its own, distinct from that idiom which, under the name of *Moor's,* has been supposed to prevail over all India.
>
> (Halhed, 1969 [1778], p. ii)

1.1 Overview

This dissertation explores the productivity of a vowel harmony pattern in Bangla that has been analyzed in prior research as a synchronic chain shift (e.g., Mahanta, 2007, 2012). Chain shifts are conceptualized in phonological theory as instances of opaque alternations, called that because the linguistic principles that govern the alternations are not fully transparent in surface realizations. McCarthy (2007, p. 2) has called phonological opacity "one of the central themes of generative phonology." From a theoretical standpoint, opaque interactions can be challenging to model in adult grammars (see, e.g., discussion in McCarthy, 2002, § 3.3.3.1), child learning trajectories (see, e.g., introductory discussion in Kager, Pater, & Zonneveld, 2004, pp. 5-6), or both. From an experimental standpoint, some opaque interactions that have persisted in spoken languages for generations seem to resist generalization to novel lexical items (e.g., Hsieh, 1970; Sanders, 2003; Zhang, 2019).

This lack of productivity is puzzling, for if a pattern has been passed from generation to generation, it must be learnable and extendable to not-previously-encountered words as children acquire an adult vocabulary. Indeed, formal computational modeling has shown opaque patterns to be learnable by machine subjects (e.g., Nazarov & Pater, 2017; Pater, 2014), and artificial-grammars that mimic opaque patterns in real language have proved learnable by human subjects (e.g., Ettlinger, 2008). And yet multiple experiments with speakers of languages such as Polish

(Sanders, 2003), Taiwanese (Zhang et al., 2006), and Hebrew (Sumner, 2003) have cast doubt on the productivity of opaque phonological patterns and, in the case of Sumner's work on Hebrew, on phonologists' ideas about how speakers represent these phenomena in their grammars.

This dissertation investigates the productivity of another opaque phonological pattern, using perception and production experiments and other measures to gain an understanding of how this process works in the grammars of speakers. The central question it poses is whether speakers can extend this chain-shift pattern to nonce words. The pattern in the language explored here (Bangla) differs from the opaque patterns in other languages just mentioned in several potentially critical ways. In both the Polish opacity(s) studied by Sanders and the Hebrew opacity studied by Sumner, the patterns occurred in only certain words within the lexical category under study. In the Taiwanese circular tone shift studied by Zhang at al., parts of the pattern defied notions of naturalness and markedness because a circular chain-shift analysis involves the changing of a less-marked form into a more-marked form. In each of those cases, it could be argued that these factors worked against the extension of the opaque pattern to nonce words. Unlike these previously studied cases, the opaque pattern in Bangla is phonetically grounded and is exceptionless—no variation affects the frequency of the alternation. Thus, if variation in the relevant portion of the grammar and an increase in markedness do result in a reduced pressure for implementation of a phonological pattern, then Bangla may be a more ideal language for the investigation of opaque alternations.

Nevertheless, the results of the study presented here, like those of the prior studies mentioned above, indicate that Bangla speakers do not freely extend this pattern to nonce words. Instead, they apply the chain shift only about half of the time. Analyses conducted to probe possible acoustic explanations for the pattern (e.g., incomplete neutralization, vowel-to-vowel

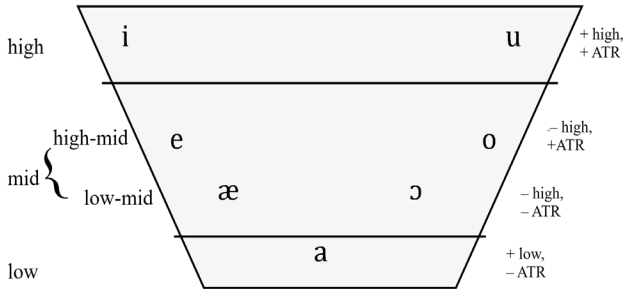
coarticulation) suggest that the patterns observed cannot be explained by phonetics alone. As will be discussed in later chapters, that fact that speakers do implement the pattern in nonce words at rates greater than chance, but not at rates that mirror their productions for real words, suggest underlearning (Becker, Ketrez, & Nevins 2011) of the chain shift pattern. More investigation addressing some potential shortcomings of the experiments presented here may further illuminate the influences on implementation rates.

The rest of this chapter will present a description of the opaque interaction in Bangla and then will address the relevance of opacity to linguists' understanding of natural languages.

1.2 The vowel chain shift in Bangla verbs

Bangla (also known as Bengali) contains seven oral monophthongs, as shown in Figure 1. Six of these participate in an alternation phenomenon generally described as a chain shift. This chain shift occurs in monosyllabic verb stems, as can be seen in the present-tense inflections of the sample of Bangla verbs shown in Table 1. There we see the following alternations: æ ~ e in the verb stem /dæk^{h}-/ 'see'; e ~ i in the verb stem /ʃek^{h}-/ 'learn'; ɔ ~ o in the verb stem /kɔr-/ 'do'; and o ~ u in the verb stem /oʈ^{h}-/ 'rise'. The higher alternating stem vowel in each pair occurs when a high vowel is present in the neighboring affix, and the lower alternating stem vowel

Figure 1: Vowel phonemes of Bangla in traditional vowel-space schematization, with relevant feature values



Sources: Chatterji (1986 [1928]), Ghosh (1997), Thompson (2010).

Table 1: Vowel alternations in the present indicative of Bangla verbs

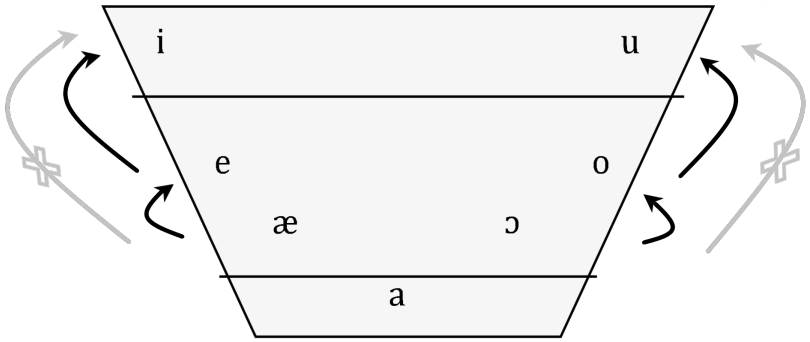
Affix		/dækʰ-/ 'see'	/ʃekʰ-/ 'learn'	/kɔr-/ 'do'	/oʈʰ-/ 'rise'	/rakʰ-/ ''
1p.	/-i /	[d**e**kʰi]	[ʃ**i**kʰi]	[k**o**ri]	[**u**ʈʰi]	[r**a**kʰi]
2p. (fam.)	/-iʃ /	[d**e**kʰiʃ]	[ʃ**i**kʰiʃ]	[k**o**riʃ]	[**u**ʈʰiʃ]	[r**a**kʰiʃ]
2p. (casual)	/-o /	[d**æ**kʰo]	[ʃ**e**kʰo]	[k**ɔ**ro]	[**o**ʈʰo]	[r**a**kʰo]
3p. (casual)	/-e /	[d**æ**kʰe]	[ʃ**e**kʰe]	[k**ɔ**re]	[**o**ʈʰe]	[r**a**kʰe]

occurs when the vowel in the affix is a mid vowel (or a low vowel, though none are shown in the mini-paradigm here). Monosyllabic verbs with an [a] in their stem do not exhibit alternations in the present indicative, as shown for the stem /rakʰ-/ in the rightmost column of the table. (In the present tense shown in Table 1, the only affix present in the surface representation is the person marker; the Bangla verb system includes other affixes for tense and aspect that intervene between the stem and the person marker, as is shown in the full verb paradigm in Appendix A. The distinction between *sadhu bhasha* and *cholit bhasha* shown in that appendix will be explained in Chapter 2.)

A number of scholars who have analyzed these alternations (e.g., Dasgupta, 1982; Dey, 1979; Lahiri, 2000b; Nagle, 2008; Paul, 1986; Sarkar, 2004; Singh, 1980) have described the pattern as one of regressive vowel height harmony or assimilation, triggered by the presence or absence of a [+high] or [+ATR] feature value in the vowel in the inflectional affix closest to the stem.[1] This assimilation is not complete, however. It takes the form of a chain shift, in which feature values change only one step along an ordered scale toward the harmonization trigger, as shown in Figure 2, where the black arrows show that æ → e → i and ɔ → o → u. Importantly,

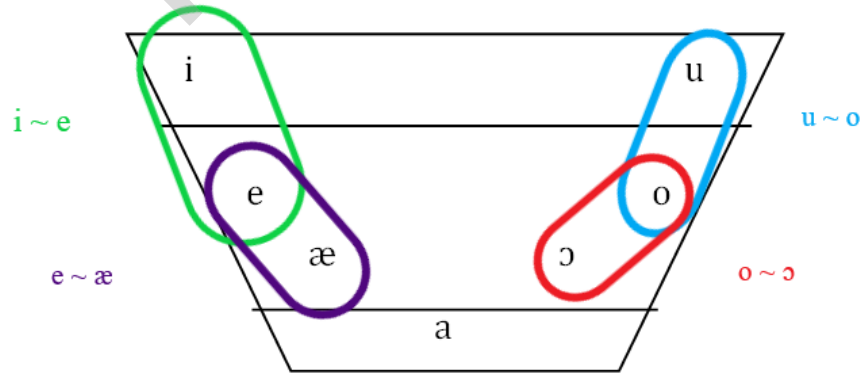
[1] The data presented in Table 1 is shown as a raising pattern (with a lower vowel in the underlying stem form) for the sake of simplicity of exposition. Whether the assimilation is one of raising caused by a [+high] affix vowel or lowering caused by a [-high] affix vowel is a matter of interpretation and does not materially affect the analysis presented here. The experimental results in Chapter 4 and Chapter 5 suggest the possibility that neither is correct.

Figure 2: The vowel chain shift in Bangla verbs



though, there is not a complete neutralization of height contrast between the focus (stem) vowel and the trigger (affix) vowel—that is, *æ → i and *ɔ → u, as shown by the gray arrows in Figure 2 (see, e.g., Lahiri, 2000a; Mahanta, 2007; Nagle, 2008; Schmidt, 1996 describes a markedly similar pattern in the Bantu language Basaa). This alternation pattern is further illustrated in Figure 3, which groups each set of alternations and also shows clearly that the high-mid vowels [e] and [o] are each part of two alternations: each is the higher alternant when paired

Figure 3: Vowel alternations in present indicative Bangla verbs



a

(does not alternate in the present indicative)

with a low-mid vowel, and each is also the lower alternant when paired with a high vowel. These dual roles are important to note and will be discussed further in later chapters.

Chain-shift patterns are phonologically opaque, in the sense that speakers must form a generalization about the phonology that is not always true in the utterances of speakers (i.e., not "surface-true"). For instance, the generalization necessary to produce the first-person forms for the Bangla verb meaning 'rise' (in the rightmost column of Table 1—that [u] is preferred to [o] in the stem when preceding the high vowel [i] in the affix—does not hold true in the case of the first-person, familiar second-person, or formal forms of the verb meaning 'do' (also shown in Table 1), where [o] rather than [u] appears to be the preferred stem vowel in the high-affix context. The generalization "[u] is preferable to [o] before a high-vowel affix" that accounts for the stem vowel patterning in the verb 'rise' is not true, apparently, in the case of the verb 'do'. A similar inconsistency arises when generalizing about the surface forms of the front stem vowels /æ/, /e/, and /i/. In a verb like that meaning 'learn', it appears that [i] is preferable to [e] before the high affix vowel [i]. And yet in a verb like the one meaning 'see', such a generalization is untrue, for there [e] is the vowel that appears in the syllable before that same high affix. In the case at hand, therefore, we can say that the generalizations are "not surface-true." We will see in Chapter 2 that this surface-inconsistent vowel pattern is replicated in all monosyllabic verbs in Bangla, meaning that speakers of Bangla are exposed to a large amount of information that, on the surface, does not appear to be consistent. In the terminology used to describe opacity, this is a case of an "underapplication" of the generalization that [u] or [i] are preferable to [o] or [e], respectively, before a high vowel in the neighboring affix. In derivational terms, this pattern is described as "counterfeeding."

1.3 Why does opacity matter?

Synchronic chain shifts are found in diverse languages (a partial compendium is provided in Moreton & Smolensky, 2002), and they are just one type of phonological opacity as it has been documented and analyzed in many languages and in several analytical frameworks (for a sampling, see, e.g. Anttila, 2006; Baković, 2007; Kiparsky, 2000; McCarthy, 2007; Moreton, 2004). McCarthy (2007, p. 2) has called phonological opacity "one of the central themes of generative phonology." It initially presented a challenge to derivational frameworks with systems of simultaneous rules (e.g., Chomsky & Halle, 1968)—a challenge that was met with the development of models that included ordered rules (e.g., Kiparksy, 1968). But as many scholars in the field of phonology moved away from rule-based analyses and toward surface-oriented constraint-based models such as Optimality Theory (OT; Prince & Smolensky, 1993/2004) grounded in terms of typologically attested patterns (= markedness) and principled departures from it (= faithfulness), the challenges presented by opacity resurfaced. Certain phenomena, including chain shifts, pose difficulties for constraint-based models, since these patterns seem to arise from something other than just the interactions of markedness and faithfulness constraints (see, e.g., McCarthy, 2007, § 1.1; Mielke, Armstrong, & Hume, 2003, § 2). Numerous interpretations of and adjustments to constraint-based models have been proposed to account for the surface facts of chain shifts and other opaque patterns. Some re-introduce elements of sequenced derivation; others assign weights or probabilities to constraints and their ranking. These adjustments supplement the constraint interactions and thereby accommodate language patterns, such as opaque alternations, that seem to defy the basic idea of strictly ranked, permutable markedness and faithfulness constraints. Some of these ideas will be discussed in section 2.2.1.

Language patterns such as phonological opacity offer both a challenge and an opportunity to linguists, psycholinguists, and cognitive scientists attempting to understand and model the nature of "language" in the human mind. These patterns implicate some of the most fundamental questions being asked about human language today, such as which elements of language are stored in memory and in what form they are stored, how stored elements of language are accessed during the processes of language perception and production, and the means by which language-specific patterns are learned and applied both to items retrieved from storage and to newly encountered lexical items.

1.4 Opaque phonology, productivity, and natural language

Experimental investigations of phenomena in natural languages are crucial to the process of revising and refining linguistic theories, and they also offer fresh insights into the nature of much-described and much-analyzed linguistic phenomena. For instance, in their "islands of reliability" study of the alternations involving diphthongization in Spanish, Albright, Andrade, and Hayes (2001) challenged the assumptions of numerous prior proposed analyses that Spanish speakers must distinguish between alternating and non-alternating verbs in their stored lexical representations. By studying the behavior of Spanish speakers encountering nonce verbs, these authors showed that speakers have intuitive knowledge of phonotactic conditioning of the alternation pattern that makes such storage details unnecessary. Although the alternations examined by Albright et al. in Spanish do not exhibit opacity, their study and others like it demonstrate the importance of testing theoretical assumptions against speakers' knowledge and implementation of patterns in their language, because our theory does not always imagine what actual speakers will do.