

UNDERSTANDING LEARNERS' INCONSISTENCIES WITH THE 'O' VOWEL AND 'O-KAR' IN BANGLA LANGUAGE

Mohd Twarique Ahsan, Mohammad Torikul Islam

Department of Foreign Languages

College of Arts and Humanities, Jazan University, Jazan, Saudi Arabia

E-mail: ahsan123k@yahoo.com

Received: 2025-02-23

Accepted: 2025-05-12

Published: 2025-06-26

Abstract

Bangla, an Indo-Aryan language of the Indo-European family, employs an orthographic system where vowels are represented both as independent letters and diacritical marks. This dual representation introduces inconsistency and complexity in spelling and reading, particularly for young learners acquiring foundational literacy skills. A persistent challenge arises from the use of the O-Kar diacritic, which modifies consonants to produce the vowel /o/ sound, but is often omitted in certain lexical contexts, leading to inconsistencies in Bangla pronunciation and spelling. This study investigates the difficulties faced by native Bangla-speaking children in spelling and pronouncing words containing the inherent consonant O, especially when the O-Kar diacritic is absent. Data were collected through written assessments and structured interviews with 52 native Bangla-speaking children aged 8–11 from Berhampore, West Bengal. Three separate sets of word—each with 15 target words featuring inherent O consonant—were presented both with and without the O-Kar diacritic in the initial, medial, and final positions. Findings reveal significant spelling inconsistency when the diacritic was omitted, with learners frequently deviating from standard orthographic conventions. Conversely, words including the diacritical were consistently spelled and pronounced more accurately. This inconsistency persisted across all word positions, suggesting that the optional application of O-Kar impedes overall reading and writing fluency. The study highlights the need for standardized pedagogical approaches to mitigate spelling inconsistencies and improve literacy outcomes in Bangla. These findings have broader implications for Bangla phonological and orthographic conventions and are practically relevant to lexicographers, language education policymakers, and teachers.

Keywords: *Bangla spelling; O-Kar diacritic; O vowel; literacy development; reading and writing inconsistencies*

1. Introduction

Language serves as a medium for individuals to communicate, express ideas, and share thoughts. It is “a purely human and non-instinctive method of communicating ideas, emotions, and desires by means of a system of voluntarily produced symbols” (Sapir, 1921).

It is primarily realized through two forms: speech and writing, each offering distinct ways to convey meaning. There is a notable distinction between verbal and written expressions. Every language has its own writing system, complete with a unique alphabet. For effective word recognition, listeners must associate speech signals with the phonological representations of these words stored in their lexical memory (Peereaman et al., 2009). However, inconsistencies often exist between how words are pronounced and how they are written. Many argue that written language should reflect spoken language, but this is rarely the case. Spelling conventions frequently take precedence over phonetic representation, resulting in various forms of the same word that are often classified as spelling errors. Bangla is no exception. Scholars have extensively discussed the factors contributing to spelling errors in Bangla and have attempted to establish rules to minimize these mistakes. However, it is often difficult for language users—especially those learning the language—to remember these rules as they frequently make spelling mistakes.

Across the globe, various writing systems exist, each with its own orthographic organization and employing different phonological and semantic units as the basis for individual graphemes. For example, logographic and ideographic systems, such as Chinese, represent entire words or concepts with a single grapheme. In contrast, Korean corresponds to morphemes, while both Japanese and Cherokee use syllabaries, where each grapheme represents a syllable rather than a phoneme (Mukerjee, 2023). Modern languages like English, German, and Finnish use alphabetic scripts that operate on a phonemic basis, with each grapheme typically corresponding to a single sound. The boundaries between these writing systems are not rigid; some languages employ modified or hybrid systems that do not fit neatly into established categories.

The Bangla orthographic system is based on an alphabetic script but follows an alpha syllabary structure. Such orthographic systems significantly influence how literate individuals perceive and process words (Mukherjee, 2023). Consequently, a lot of inconsistencies exist between Bangla phonology and orthography. She also notes that these inconsistencies can hinder reading aloud unless the phonological contexts for specific sound values are clearly defined. A particular source of inconsistency is the lack of a diacritic mark for the grapheme অ, which represents the inherent vowel in Bangla. As a result, this vowel may be pronounced as /o/, /ɔ/, or omitted entirely, depending on the phonological context. This ambiguity often poses challenges for learners, especially when the inherent vowel is not explicitly marked, making it difficult to predict pronunciation. As educators, we have observed that students frequently struggle with these complexities, particularly those related to the inherent vowel.

The acquisition of Bangla pronunciation and orthography presents unique challenges for young learners due to complex grapheme-phoneme correspondence and dual vowel representation system. Unlike transparent orthographies, Bangla's use of inherent vowel sounds, optional diacritics (e.g., O-Kar), and context-dependent letters results in ample inconsistencies. Consequently, Bangla pronunciation creates significant hurdles in developing reading fluency and spelling accuracy. Children must master not only the 11 vowel graphemes (including অ "o" and ও "ô"), but also their diacritical variants and positional allophones, while navigating frequent inconsistencies between written and spoken forms. Research indicates these orthographic-phonological complexities particularly affect early literacy development. The inconsistent use of the O-Kar diacritic, which alters consonants to generate the /o/ sound but is often left out in written form, leads to

inconsistency in spelling and pronunciation for young learners (Barman & Sircar, 2020; Khan, 2018).

Therefore, the inconsistent use of the O-Kar diacritic in Bangla leads to frequent spelling mistakes and disrupts reading fluency, making early literacy development more challenging for children. This slows literacy acquisition and diverts attention from comprehension to basic language decoding. For educators, these issues highlight the need for targeted teaching strategies that address orthographic inconsistencies and prevent entrenched spelling errors. Addressing these challenges is essential for improving learning outcomes, guiding effective curriculum design, and helping lexicographers create more accessible educational resources—ultimately supporting both learners and the broader Bangla education community.

This investigation aims at shedding light on the inconsistency that arises from the two different vowel sounds (অ /ɔ/, ও /o/) and the O-Kar diacritical mark. These features of Bangla transcript often pose a challenge for new and young learners. It is important to acknowledge that this challenge can be quite intimidating for those who are new to the language. Nevertheless, the study will endeavour to investigate potential recommendations and solutions.

The study investigates the possible inconsistency between the vowel diacritics, mainly the O-Kar, and the inherent /ɔ/ or /o/ sounds of the consonant letters in Bangla. The following are the two research questions (RQs) to help guide this study:

RQ 1 : Almost all the letters carry an inherent /ɔ/ sound which is pronounced like /o/. In that case, how the O-Kar and without the O-Kar (in the initial, medial and final positions) create inconsistency in Bangla literacy.

RQ 2 : Is the pattern of inconsistency in reading and writing the same or different?

2. Literature Review

Over the years, individual scholars and institutions have shared their views on spelling mistakes and prescribed correct spelling in Bangla. Almost a century ago, Bengali scholars and linguists such as Suniti Kumar Chatterji (1890-1977) and Haraprasad Shastri (1853-1931) collaborated to develop a set of guidelines aimed at addressing these spelling issues, focusing primarily on the etymological roots of words. This initiative received endorsement from Rabindranath Tagore (Mahalanabish, 1925). To further standardize Bangla spelling, Kolkata University Press (KUP, 1936) published a set of spelling rules in 1936 and released a revised third edition in May 1937.

Individual academics have also played a crucial role in standardizing the spelling of Bangla words, with some notable figures making perceptive observations, academic judgments, and logical deductions. Among them, Rajshekhar Basu's work stands out. In his monolingual dictionary (*Chalantika*), Basu (1962) wrote 30,000 Bangla words and gave orthographic and phonological guidelines about the Bangla language. He mainly explained the definitions and grammatical features of words and followed the Kolkata University rules (1936) for spelling.

2.1 linguistics Inconsistencies of Bangla

Interestingly, academic discussions have predominantly centred on the correct spelling of words rather than exploring the causes and solutions for spelling inaccuracies,

where Dash (2018) claimed that spelling errors arise from a combination of linguistic and non-linguistic factors. The following are some of the linguistics inconsistencies:

2.1.1 Inherent অ [ɔ] Sound in Bangla Orthography

Bangla consonants inherently carry a default vowel sound, typically realized as /ɔ/ or /o/ (Anderson, 1917). For instance, the grapheme 'ক' is pronounced /kɔ/ or /ko/ contextually. Ghosh (1994) attributes this challenge due to inherent Bangla vowel system. This inherent vowel remains active unless suppressed by a vowel diacritic or the *hasanta* (silencing mark) (Bagchi, 1996). Crucially, the unmarked presence of this vowel in writing—contrasting with its variable phonetic realization—creates a systematic divergence between spoken and written forms. This inherent vowel system exemplifies Bangla's orthographic complexity, necessitating rule-based pedagogical approaches for literacy development.

2.1.2 Inconsistencies in Phonology and Orthography

Research indicates that Bangla's spelling inconsistencies stem primarily from its writing system's inability to fully represent spoken language features (Atkins et al., 1992). Das (2006) attributed this to the inconsistent grapheme-phoneme correspondence, particularly regarding consonant pronunciation. Vowel graphemes often represent multiple phonetic forms: অ may be realized as /ɔ/ or /o/, আ as /a/ or /æ/, and এ as /e/ or /æ/. Conversely, a single sound may correspond to multiple graphemes such as ই/ঈ for /i/ and উ/ঊ for /u/, creating spelling challenges (Mukherjee, 2023). Consonant inconsistencies further compound this issue, with ন/ব representing /n/ and শ/ষ/স often denoting /ʃ/. Additionally, consonant clusters exhibit positional variability, where the second consonant (e.g., /j/, /m/, or /b/) influences pronunciation unpredictably. Mukherjee notes that these discrepancies impede fluent reading and spelling accuracy, underscoring the need for systematic orthographic instruction. This systematic divergence between sound and script warrants the need for pedagogical strategies addressing Bangla's orthographic complexity.

2.1.3 Allographic Variation in Vowel Representation

An allograph refers to an alternative written representation of the same letter or character that conveys the same sound or phoneme within a language. In Bangla, each vowel grapheme exists in two distinct forms: as a vowel letter (independent form) and as a diacritic (vowel sign/marker) (David, 2015). The Bangla language comprises 11 vowels, all vowel letters appearing at the beginning of words, while the initial position diacritic is not used.

Notably, the vowel অ is unique in that it does not have a corresponding diacritic; it is represented only as a letter in the initial position. In medial and final positions, অ appears as the inherent vowel sound that accompanies a consonant letter. With the exception of the vowels ই /i/ and ও /o/, all other vowels are represented exclusively by a diacritic in both medial and final positions. The vowels ই /i/ and ও /o/ can be represented as both a letter and a diacritic in all positions (initial, medial, and final) within words. This vowel complexity emphasizes the intricacies of Bangla orthography and points out the difficulties encountered in developing precise reading and spelling abilities.

2.1.4 Orthographic Variability in Bangla Vowel

The Bangla writing system exhibits notable variability in vowel marks, particularly through the optional omission of diacritics. While most vowels are explicitly marked by diacritics, their occasional absence forces pronunciation to rely either on the consonant's inherent vowel (/ɔ/ or /o/) or contextual phonological rules, such as লিখি (with diacritic) versus লিখি (without diacritic), which may be realized as either /likho/ or /likh/. This variability of diacritics introduces ambiguity, particularly for language learners, as it requires reliance on implicit phonological rules rather than explicit orthographic cues.

2.2 Efforts to Resolve the Inconsistencies

Linguists, experts, scholars, and institutions in both West Bengal and Bangladesh continue to work towards standardizing the script to reduce these inconsistencies. They have recently taken steps to make the script more transparent, such as promoting clearer graphical forms of consonant clusters and addressing the opaque nature of vowel diacritics like the O-Kar in textbooks. However, implementation challenges persist due to dialectal variations and entrenched writing conventions (Mukherjee, 2023). As a result, these reforms are not yet uniformly adopted across all printed materials, requiring learners still need to recognize both traditional and reformed spellings. The following are a few of the harmonizing efforts:

2.2.1 Historical Reforms of Bangla Spelling

The history of spelling reform in the Bangla language has been marked by ongoing debates and proposals, but a universally accepted solution remains elusive. Institutions and individual scholars have produced guidebooks and dictionaries to promote orthographic consistency, yet discrepancies persist. In addition, recent research has focused on the relationship between pronunciation and orthography, highlighting the prevalence of spelling discrepancies and the challenges (Dash, 2005, 2006 & 2018; Mandal, 2004; Mukherjee, 2005; Biswas, 2005; Khan, 2005). Proposals have ranged from new spelling conventions to revised orthographic systems, reflecting a sustained scholarly effort to address these issues.

2.2.2 Harmonizing the Contribution of Bangla Literature

Analyses of modern Bangla literature reveal that spelling changes are primarily driven by efforts to align orthography with phonetic realities, reflecting a broader tendency to reduce the gap between written and spoken forms (Dash, 2011). This motivation is evident in the influence of various phonological processes—such as vowel harmony, nasalisation, de-nasalization, segment assimilation, and haplology—on the evolution of spelling conventions. These harmonizing processes help with the inconsistent nature of Bangla spelling as literateurs, writers and language planners attempt to capture pronunciation more faithfully in script.

In Bangla orthographic system, the use of O-Kar and without O-Kar is one of the most debatable topics. While some scholars argue that O-Kar is redundant (Ghosh, 1953), others document its functional role in marking medial/final /o/ sounds, e.g., ভালো / b^halo / (Chaudhuri, 1994). Mamud (2016) further noted context-dependent grammatical rules—such as the transformation of অ /ɔ/ to ও /o/ after certain consonants; he also pointed out that these rules are often too complex for learners and teachers to apply consistently. As a

result, the lack of consensus on the O-Kar usage has led to significant variation in spelling practices, with some advocating for its restriction to reduce ambiguity and others supporting its continued use for phonetic clarity. This persistent uncertainty underscores the central challenge posed by the interplay of inherent vowels and diacritical marks in Bangla orthography.

2.3 Research Gap

While existing studies have extensively analysed the orthographic-phonological challenges posed by Bangla's inherent vowel অ /ɔ/ and O-Kar diacritic, a critical gap remains in understanding how young learners navigate these complexities during early literacy acquisition. Not enough empirical research has systematically examined native Bangla-speaking children's ability to process words with and without the O-Kar, particularly their confusion between inherent /ɔ/ and diacritical mark /o/ sounds. This study addresses this gap by investigating whether these features contribute to confusion or inconsistency in young learners' early literacy development and aims to inform pedagogical strategies and material design to mitigate spelling inconsistencies and better support early literacy acquisition.

3. Research Method

At a young age, children learn to spell to develop their reading and writing skills in elementary school. The study deals with Bangla spelling, mainly O-Kar words. Therefore, the data was collected through reading and writing tests. The participants were targeted between the ages of 8 to 11 years (Class 3, 4 and 5). These young children were chosen because they do not know the words' conventional spelling as they are still at a developmental stage. Consequently, they were expected to spell the words they perceive and read according to the alphabetical structure.

3.1 Sample and Setting

For this study, data was collected from 52 native Bengali students from a small town, Berhampur, West Bengal, India. The participants aged 8 to 11, including both boys and girls, were from grades 3, 4 & 5. 7 participants were from class 3, 21 from class 4, and 24 from class 5. The statistical evaluation of the age group revealed a mean of 9.5, with both the median and mode at 10, and a standard deviation of 0.828. The data was collected from a private institute where students of classes 2 to class 10 were present. This purposive sampling considers 15 words with and without O-Kar in all three positions, i.e., beginning, middle, and end (3x15=45 words). The setting was selected keeping in mind the native Bangla language, which is dominant in that area.

3.2 Language Assessment Tools

Bangla contains a good percentage of words that lack the O-Kar. However, they are pronounced similarly due to the presence of inherent O vowel in consonants. Given the extensive number of potential confusing words, compiling and evaluating all of them is impractical. Therefore, the assessment tool consists of 15 Bangla words. These words have been chosen because they appear most frequently in upper elementary Bangla books.

In the writing segment, participants were instructed to listen to a recording and write the sentences from the recording to measure how they handled O-Kar in those sentences, specifically whether they included O-Kar in their written sentences. Following the writing assessment, a reading test was conducted. In this section, participants were asked to read the same sentences, both with and without O-Kar sounds. The interviewer observed the pronunciation of the targeted words and marked them with a check (✓) and noted how the participants pronounced them.

3.3 Procedure

For the writing test, all participants sat in a room. They were instructed to listen to the sentences from an audio file and write them in Bangla. In-person interviews were carried out with eight participants for the reading assessment. These eight students represented the entire sample, comprising two participants from class 3, two from class 4, and four from class 5, in accordance with the class size distribution. The interviewer asked them to read the sentences. When they read those sentences, the interviewer ticked (✓) if they pronounced them correctly.

3.4 Ethical Consideration

During data collection for this research, we conducted written evaluations and structured interviews to assess the Bangla pronunciation and writing skills of children. By responding to the questions, participants enabled us to evaluate their comprehension of pronunciation and writing. Although they consented to participate, they retained the right to withdraw at any time or decline to answer any questions without providing a reason or facing any consequences. Participation was entirely voluntary. To safeguard their identities, all information and responses were treated with strict confidentiality and anonymity. Their answers will be securely archived for five years for the purposes of this study, after which they will be destroyed. We made sure participants had read and understood the information provided to them. Additionally, they were encouraged to ask any relevant questions. As the participants were minors, both they and their parents or guardians provided signed consent for participation in this research.

3.5 Analysis Technique

The data collection process is aimed at addressing two main domains: reading and writing abilities. Data analysis was conducted using percentage metrics. Writing skills were evaluated in two distinct groups: percentage of participants utilizing diacritical O-Kar during listening activities (O-Kar users) and percentage of participants who did not use diacritical O-Kar (non-O-Kar users). Reading skills were further categorized into groups based on O-Kar usage and the nature of the words. The diagram groups are as follows:

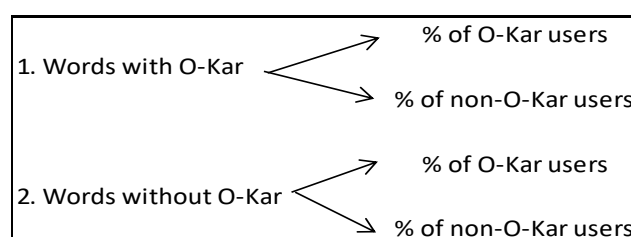


Figure 1. Groups of O-Kar and non-O-Kar Readers

4. Results and Discussion

4.1 Findings of Using O-Kar in the Initial Position of the Writing Test

Participants were instructed to listen and note down words that begin with the O-Kar sound. Table 1 underneath presents only the relevant words, illustrating the percentage of users who employed O-Kar and non-O-Kar sounds at the beginning of words.

Var.	O-Kar words	Non-O-Kar words	O-Kar users in %	Non-O-Kar users in %	Pronunciation change	Meaning change
1	কোনো kono	কনো kno	83.84	16.16	X	X
2	কোহিনুর Kohinur	কহিনুর kchinur	76.92	23.08	√	√
3	গোলাপ golap	গলাপ glap	92.3	7.69	X	X
4	লোকসান loksān	লকসান lkksan	90.38	9.61	√	√
5	পোকামাকড় pokamākor	পকামাকড় pckamakor	71.15	28.84	X	X
6	ধোকাবাজ dhokabaz	ধকাবাজ dhckabaz	80.76	19.23	X	X
7	বোতাম botam	বতাম btam	90.38	9.61	X	X
8	কোটোয়াল kotoyal	কতোয়াল ktoyal	46.15	53.84	X	X
9	কোণঠাসা konthasa	কণঠাসা knthasa	88.46	11.53	X	X
10	পয়োবারো poyabaro	পয়াবারো pyabaro	78.84	21.15	X	X
11	পোশাক posak	পশাক psak	86.53	13.46	X	X
12	ঘোমটা ghomta	ঘমটা ghmta	73.07	26.92	√	X
13	গোলমলে golmele	গলমলে glmele	69.23	30.76	√	X
14	সোনালী sonali	সনালী snali	80.76	19.23	X	X
15	ক্সোদতি khodito	ক্খদতি khcdito	92.3	7.69	X	X

Table 1. Results of using O-Kar spelling in the initial position

Table 1 addresses Research Question (RQ 1), detailing the percentage of inconsistency observed when participants wrote without O-Kar in initial positions. The data indicates that overall inconsistency was minimal, as most students consistently applied O-Kar wherever they perceived an "O" sound. Notably, many learners extended its use to the inherent /ɔ/ sound (অ), which traditionally does not require O-Kar. However, a minority omitted O-Kar for this inherent sound, suggesting variability in their grasp of orthographic rules.

The confusion arises from the subtle acoustic distinction between /ɔ/ (অ) and /o/ (ও). The minimal difference in pronunciation complicates learners' ability to discern when O-Kar is necessary, particularly for children still developing phonetic awareness. This overlap likely contributes to inconsistent application, as students default to associating the "O" sound with O-Kar regardless of contextual rules.

4.1.1 Findings of Using O-Kar in the Middle Position

Like the initial position, participants were asked to listen and write sentences with words which have the O-Kar sounds in the medial position. This partially answers RQ 1, i.e., percentage of inconsistency while writing without the O-Kar words in the medial position.

Var.	O-Kar words	Non-O-Kar words	O-Kar	Non-O-	Pronun	Meaning
------	-------------	-----------------	-------	--------	--------	---------

			users in %	Kar users in %	ciation change	change
1	কুমোর kumor	কুমর kumər	92.3	7.69	X	X
2	ঘাটোয়ালী ghatoyali	ঘাটয়ালী ghatyali	92.3	7.69	√	X
3	ছোটোলক chotolok	ছোটলক chotolok	78.84	21.15	X	X
4	তপোবন topobon	তপবন tɔpbɔn	88.46	11.53	√	√
5	লখিচিহ্ন লখিচিহ্ন likhochittro	লখিচিহ্ন likhchittro	90.38	9.6	√	X
6	পটোল potol	পটল pɔtol	7.69	92.3	X	X
7	ততখিক totodhik	ততখিক tɔtdhik	86.53	13.46	√	√
8	জগমোহন jgmmohon	জগমহন jgmmɔn	46.15	53.84	X	X
9	বারোয়ারী baroyari	বারয়ারী baryari	92.3	7.69	√	√
10	বক্সোভ bikhov	বক্সভ bikɔv	9.61	90.38	X	X
11	বিনোদী binodi	বিনদী binɔdi	83.84	16.16	X	X
12	বালভোগ balvog	বালভগ balvɔg	80.76	19.23	√	√
13	শরিনোম shironam	শরিনাম shirnam	92.3	7.69	√	√
14	দোমোট dhomot	দোমট dhomɔt	86.53	13.46	X	X
15	মনোভূমি monovumi	মনভূমি mɔnvumi	86.53	13.46	√	√

Table 2. Results of using O-Kar spelling in the initial position

Similar to the cases observed in the initial position, the pattern of inconsistency among students in the medial position is nearly identical. In this instance, students consistently employed O-Kar for all O sounds, except in variations 6 and 10, where they did not apply O-Kar to the inherent O sound in the medial position.

In variation 3, the term /chotolok/ appears twice with O-Kar in the medial position: /to/ and /lo/. When participants encountered this word, the majority opted to use O-Kar in the second medial position /chɔtɔlok/ /lo/, while some variation was also noted in the first medial position [to], instead of using /chotolok/.

4.1.2. Findings of Using the O-Kar in the Final Position

Child participants were asked to listen and write sentences with words, which have the O-Kar sounds in the final position. This partially answers RQ 1, i.e., percentage of inconsistency while writing without the O-Kar words in the final position.

Var.	O-Kar words	Non-O-Kar words	O-Kar Users in %	Non-O-Kar Users in %	Pronunciation change	Meaning change
1	আঁকোঁ aako	আঁক aak	100	0	√	√
2	লখিঁ likho	লখি likh	100	0	√	√
3	জাগোঁ jago	জাগ jag	100	0	√	√
4	বাঁচোঁ bacho	বাঁচ bach	100	0	√	√
5	আছোঁ accho	আছ acch	100	0	√	√

6	ছোটো choto	ছোট chot	100	0	√	√
7	ততো toto	তত tot	100	0	√	√
8	কাঁদো kado	কাঁদ kad	100	0	√	√
9	পুরানো purano	পুরান puran	100	0	√	√
10	খাবো khabo	খাব khab	100	0	√	√
11	কাঠামো katham	কাঠাম katham	100	0	√	√
12	করো koro	কর kor	100	0	√	√
13	ছলি চিলো chilo	ছলি chil	100	0	√	√
14	পালতি palito	পালতি palit	100	0	√	√
15	ভালো valo	ভাল val	100	0	√	√

Table 3. Results of using O-Kar spelling in the initial position

The findings in Table 3 show that all participants used the O-Kar in the final position, even though the standard spellings of these words do not include the O-Kar. They wrote /toto/ with the O-Kar instead of following the conventional spelling, which omits it. This pattern indicates that the participants did not understand the concept of the inherent /o/ sound. They inserted the O-Kar in almost all variations and some of which diverged from established spelling conventions.

In addition, words without the O-Kar are pronounced differently. For instance, in variation number 2, /likho/ (with O-Kar) and /likh/ (without O-Kar) have distinct pronunciations. Although both forms share the same meaning, their usage varies by context: /likho/ with O-Kar is suitable for formal address, while /likh/ without O-Kar is typically used among close acquaintances or in casual settings.

Overall, these findings suggest a lack of awareness among participants regarding the inherent vowel, leading to non-standard spelling and contextually inappropriate usage. The distinction in pronunciation and context further highlights the functional importance of the O-Kar sound.

4.2 Findings of the Use of O-Kar in the Reading Test

Like the writing test, reading tests were also conducted to observe how the participants dealt with or without the O-Kar sounds, while they read the O-Kar words or without the O-Kar, which have inherent অ /ɔ/ or ও /o/ sound. They read the O-Kar words without any difficulties, even though there were no differences or variations. All the participants read the O-Kar for all the variations in all three positions (initial-medial-final). However, participants faced difficulties with inherent অ /ɔ/ or ও /o/ sounds and spelt without the O-Kar. For example, the word লখি (with diacritic O-Kar) is pronounced as it is /likho/, but the word লখি (without the O-Kar) is pronounced both /likh/ and /likho/. The following bar graph uses two different pronunciations of the same word (without the O-Kar) in different positions and shows the extent of participants' inconsistency.

4.2.1 O-Kar Pronunciation in the Initial Position of the Reading Test

This section describes RQ 1 and the extent of participants' inconsistency while reading without O-Kar words in the initial position.

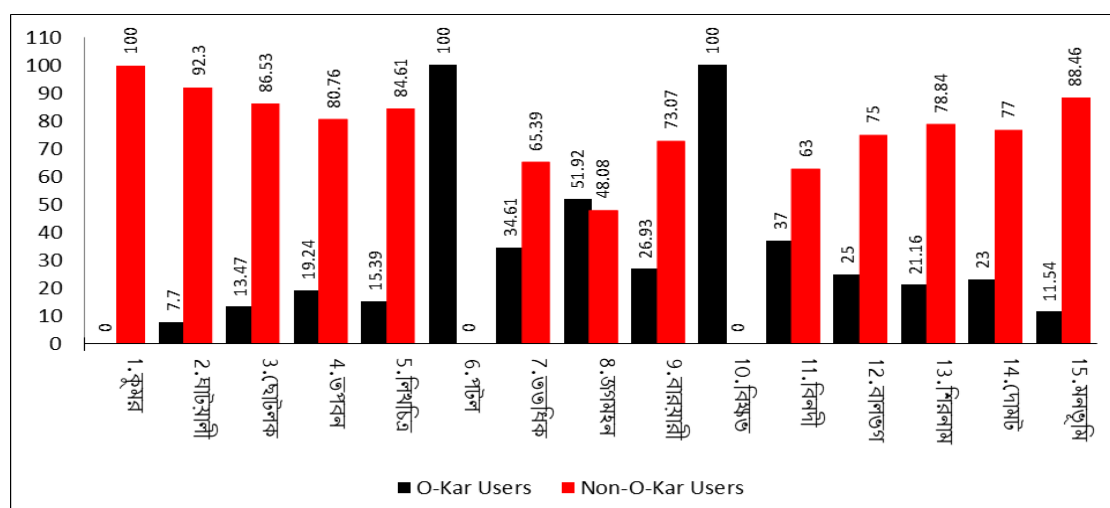


Figure 2. Pronunciation of without the O-Kar words in the initial position in %

As outlined in the method section, participants read sentences lacking O-Kar in the initial position. The findings (Fig.1) reveal divergent responses to the inherent অ /a/ and ও /o/ sounds. During the reading assessment, many participants struggled with monosyllabic words containing prefix sounds. Some pronounced the initial letter with /a/, paused, and continued without the O-Kar. For example, in Variation (Var.) 3, 80.76% segmented গ /ga/ (first syllable) and লাগ /lap/, omitting the O-Kar. This segmentation pattern persisted in: Var. 4: /lɔk+san/, 5: /pɔ+ka+makar/, 6: /dɔ+ka+baz/, 7: /bɔ+tam/, 9: /kɔn+thasa/, 12: /ghɔm+ta/ and 13: /gɔl+mele/.

Conversely, other participants read initial segments as syllables with the O-Kar, inserting it even when absent in spelling. Examples include: Var. 1: /kono/, 3: /kohi+nur/, 7: /botam/, 8: /kto+yal/, 10: /poya+baro/, 11: /posak/, 14: /sonali/ and 15: /khodito/.

These results demonstrate that syllabic vs. non-syllabic structures influence O-Kar application, particularly when the word is not spelled with it. This confusion is created due to the inherent vowel, which is pronounced অ /a/ and ও /o/; it aligns with Ghosh (1994) and Dash (2006).

4.2.2. O-Kar Pronunciation in the Middle Position

This part explains RQ 1 and how students become confused when reading without the O-Kar terms in the middle place.

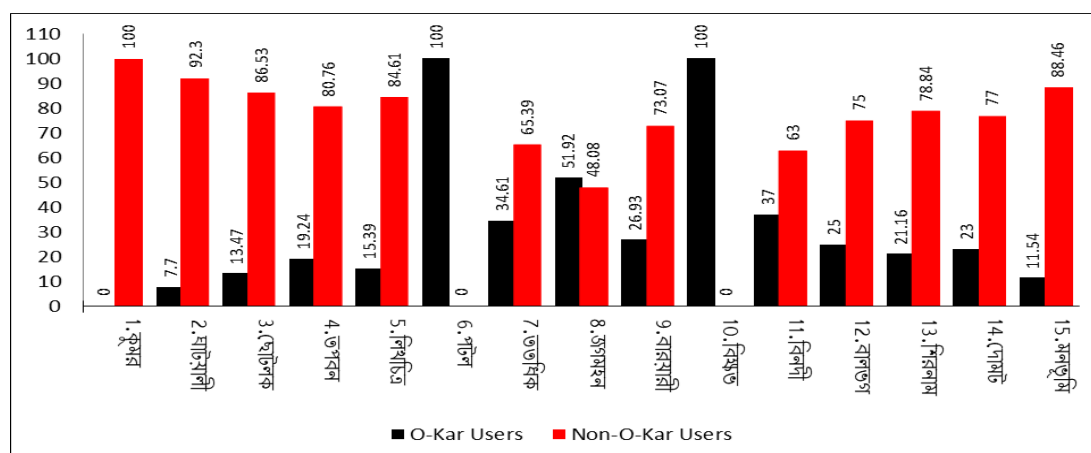


Figure 3. Pronunciation of without the O-Kar words in the medial position in %

The above graph indicates that participants' use of the O-Kar pronunciation is dependent on syllabic structure and pauses. In bi/multisyllabic words with prefixes followed by pauses, majority of the participants omitted the O-Kar in var. 1: /ku+mɔr/, 11: /bl+nɔdi/, and 14: /do+mɔt/.

Similarly, pauses between syllables also led to O-Kar absence in Var. 2: [gʰat+ali], 3: /tʃʰot+lɔk/, 4: /tɔp+bɔn/, 5: /likh+tʃitro/, 7: /tɔt+dhik/, 8: /dʒɔg+mɔhon/, 9: /bar+ari/, 12: /bal+vhɔg/, 13: /lɪr+nam/ and 15: /mɔn+vhumi/.

All participants used monosyllabic words with the diacritic O-Kar, diverging from standard spelling conventions in Var. 6: /pɔtol/ and 10: /bikhob/.

It is observed from the results that participants did not use O-Kar when pauses occurred after prefixes or between syllables. The O-Kar also appeared in uninterrupted speech (e.g., non-paused contexts).

This suggests the O-Kar functions as a phonetic feature tied to prosodic continuity rather than lexical structure. This different pronunciation pattern creates the discrepancy with standard spelling and highlights potential mismatches between orthography and spoken phonetics in the studied language as Dash (2005) observed.

4.2.3. O-Kar Pronunciation in the Final Position

The following section explains how learners become confused when reading without the 'O-Kar' terms in the last position of the words mentioned in RQ1.

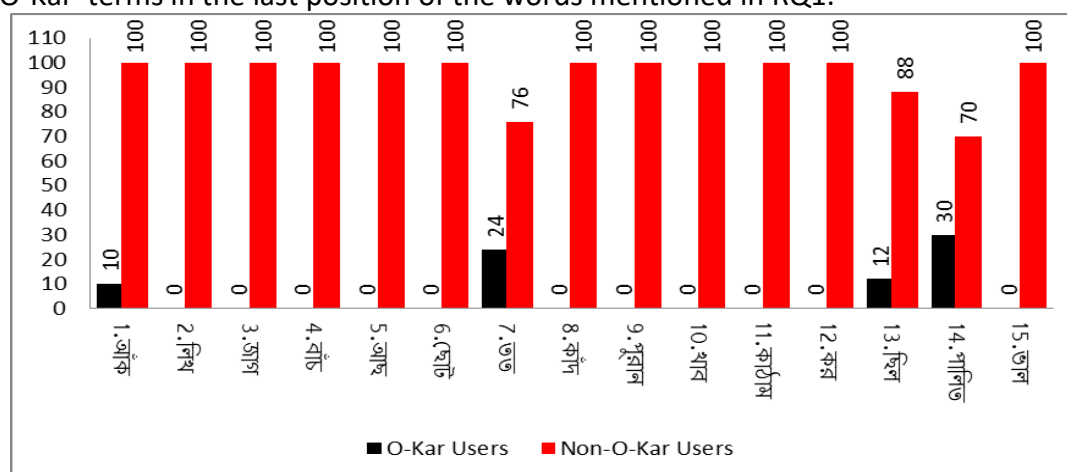


Figure 4. Pronunciation of without the O-Kar words in the final position in %

Nearly all participants refrained from using the O-Kar in the final position. Notably, the inherent sound of the letter **ꦱ** /o/ in certain words is typically represented in conventional spellings without the O-Kar. However, the participants failed to recognize these inherent sounds and omitted the O-Kar sound when pronouncing the words. For instance, the conventional spellings for variations 2, 7, and 12 are /litho/, /toto/, and /koro/—none of which include the O-Kar—yet participants pronounced them as if the O-Kar were present. In contrast, all participants articulated these words—variations 2, 7, and 12—without the O-Kar, resulting in the pronunciation of /likh/, /tot/, and /kor/. These differ significantly both in meaning and pronunciation from the original words. Participants consistently read all variations without the O-Kar, pausing at the end of each word, with only minor exceptions for variations 14, 7, 12, and 10.

Ghosh (1953) discussed this ambiguity in the pronunciation of the final position of the words and noted that this ambiguity is influenced by pauses and the presence of vowels, depending on the context. He illustrated this issue using homonyms, showing that identical spellings can represent different meanings in different contexts. This poses challenges for young learners who are still acquiring language literacy, as they often struggle to understand these contextual nuances. Therefore, the study indicates that such learners do not experience confusion when the O-Kar is included in the final position, a finding that Ghosh contested. The inclusion of the O-Kar can facilitate the learning process and clarify distinctions in meaning, as suggested by Mamud (2016). This would help young learners better comprehend and distinguish between homonyms, reducing ambiguity and supporting language acquisition.

	Initial		Medial		Final	
	O-Kar Users	non-O-Kar Users	O-Kar Users	non-O-Kar Users	O-Kar Users	non-O-Kar Users
Mean	41.6	10.466	39.933	12.066	52	0
Median	42	10	45	7	52	0
Mode	48	4	48	4	52	0
Standard Deviation	6.276	6.334	12.555	12.555	0	0

Table 4. Statistical analysis of O-Kar and non-O-Kar users in writing (n=52)

	Initial Position		Meddle Position		Final Position	
	O-Kar Users	non-O-Kar Users	O-Kar Users	non-O-Kar Users	O-Kar Users	non-O-Kar Users
Mean	26.2	25.8	20.333	31.666	2.333	49.666
Median	18	34	13	39	0	52
Mode	52	0	52	0	0	52
Standard Deviation	17.1097	17.109	17.360	17.360	5.205309	5.205

Table 5. Statistical analysis of O-Kar and non-O-Kar users in reading (n=52)

The two tables above (table 4 & 5) clearly illustrate the students' struggle with inconsistencies in spelling and pronouncing words containing the O sound. The perplexity associated with composing and interpreting words with the letter O was particularly distinct. When comparing the confusion experienced in writing versus reading, it becomes evident that the types of inconsistency differ between these two activities.

Typically, variations among students are more noticeable in writing; however, in this instance, greater discrepancies were found in reading. In writing, the majority of students spelt the word with O-Kar rather than using the letter O or the inherent O sound. Conversely, in reading, students tended to segment the word, sometimes pronouncing the prefix or the root independently. This led to variations in the application of O-Kar. In the writing segment, nearly all students appended O-Kar to the ends of words due to their recognition of the "O" sound. However, during the reading of words ending with the O sound, they refrained from using O-Kar as they paused to comprehend the text.

Analysis of their reading behaviour suggests that if the students composed the word independently without auditory input, more differences would have emerged in their writing. Moreover, this section explains how the pattern of misperception in reading is different from that in writing (RQ 2).

5. Conclusion

Bangla holds immense linguistic and cultural significance as both the native and official language of Bangladesh and the second most spoken language in India alongside one of the official languages. It ranks as the seventh most commonly spoken language globally and is the fifth most prevalent Indo-European language. This language not only connects with the global Bengali diaspora but also continues to grow in the digital and academic spheres, with increasing demand for Bangla-language technologies and educational resources. This widespread usage and recognition of cultural values pose challenges in Bangla literacy acquisition, particularly the orthographic complexities surrounding the O-Kar diacritic.

The findings of this study emphasize the challenges posed by the dual representation of the /o/ vowel sound in Bangla orthography, particularly the inconsistent use of the O-Kar diacritic. The persistent spelling and pronunciation errors observed among young learners when encountering words without the diacritical mark highlight a fundamental gap between the writing system's conventions and children's phonological processing. The inherent vowel **ও** [ɔ] and its overlap with the O-Kar mark **ঔ** [o] create a cognitive burden, as child learners struggle to reconcile implicit phonological rules with explicit orthographic representations. This inconsistency is not limited to specific word positions but persists across initial, medial, and final contexts, suggesting systemic issues in Bangla literacy instruction. This research advocates representing the inherent O sound with the O-Kar, as it does not alter the meaning or pronunciation of the original term. The findings indicate that when the inherent vowel is not represented with the O-Kar, students tend to pronounce words based on syllabic divisions, leading to spelling mistakes. Students typically articulate prefixes and syllables in isolation, often pausing at the end of syllables, which results in the omission of the O sound. This issue persists even in multi-syllable words, too.

For lexicographers, these findings advocate standardized diacritic usage in learner dictionaries and textbooks, reducing ambiguity for young readers. Moreover, policymakers might consider revising the Bangla language curriculum to explicitly address the O-Kar's variability and ensure alignment across instructional materials. Teachers, meanwhile, can adopt diagnostic assessments to identify learners' specific inconsistent points (e.g., medial-position omissions) and tailor instructional pedagogy accordingly. Digital tools—such as interactive apps that gamify diacritic placement—could also reinforce these skills. Eventually, addressing the O-Kar's inconsistencies requires a collaborative effort—one that

balances linguistic accuracy with pedagogical pragmatism. By demystifying Bangla's orthographic quirks early on, educators can foster greater reading fluency and writing confidence, laying a stronger foundation for child literacy. Future research should explore the longitudinal impacts of such interventions and their applicability to non-native learners, ensuring equitable access to effective Bangla language education.

The study's implications extend beyond theoretical linguistics, offering actionable insights for refining literacy instruction. The consistent errors observed—such as substituting অ for ও or omitting the diacritic entirely—demonstrate that current teaching methods may not adequately address the cognitive load imposed by Bangla's orthographic complexity. A structured, rule-based approach to the O-Kar instruction could mitigate this, such as explicitly teaching contextual rules (e.g., when অ transforms to ও after specific consonants) or prioritizing high-frequency words where diacritic mark is most variable. The development of differentiated reading materials that progressively introduce the O-Kar variability, paired with repetitive spelling drills, might further help. Finally, these strategies should be embedded in a broader literacy framework that emphasizes metalinguistic awareness, helping learners understand why certain spellings deviate from pronunciation norms.

While this study provides valuable insights into Bangla learners' challenges with the O-Kar diacritic, several limitations should be acknowledged. The sample size, for example, was restricted to 52 native-speaking children from a single district, which may not fully represent the diverse dialectal variations across Bangla-speaking regions. Moreover, the study did not account for differences in instructional methods across schools, which could significantly impact learners' diacritic recognition skills. Finally, the short-term nature of the research design limits our understanding of how these orthographic and phonological challenges persist or evolve over time. Future research could address this gap through longitudinal studies.

References

- Anderson, J. D. (1917). The Phonetics of the Bengali Language. *Bulletin of the School of Oriental Studies*, University of London, 1(1), 79–84.
- Atkins, S., Jereme, C., & Nicholas, O. (1992). Corpus Design Criteria, *Literary and Linguistic Computing*, 7(1), 1-16. <https://doi.org/10.1093/llc/7.1.1>
- Bagchi, T. (1996). Bengali Writing. In Peter Daniels and William Bright (eds.), *The World's Writing Systems*, New York: Oxford University Press, 399–403. <https://doi.org/10.7202/1084044ar>
- Barman, B., & Sircar, S. (2020). The Role of Diacritics in Bangla Word Recognition: Evidence from Young Readers. *Journal of Child Language Acquisition and Development*, 8(2), 45–62.
- Basu, R. (1962). *Chalantika: Adhunik Bangla Bhasar Abhidhan* [Chalantika: Dictionary of Modern Bengali Language]. Kolkata: M.C. Sarkar.
- Biswas, N. (2005). Bangla Banan O Uchharan: Samasya O Samadhaner Upay [Bangla Spelling and Pronunciation: Problems and Solutions]. *Bangla Bhasa Matri Bhasa [Bangla Language Mother Tongue]*, 79-86.
- Chaudhuri, J. (1994). *Bangla Academy Bangla Banan Abhidhan [Bangla Academy Bangla Spelling Dictionary]*. Dhaka: Bangla Academy Press.

- Dash, N. S. (2005). Methods in the Madness of Bangla Spelling: A Corpus-Based Empirical Investigation. *South Asian Language Review*, 14(2), 63-92.
- Dash, N. S. (2006). *Bahurupi Bangla Banan [Multifaceted Bangla Spelling]*. Kolkata: Daksha Bharati.
- Dash, N. S. (2011). *A Descriptive Study of the Modern Bangla Script*. Saarland: Lambert Academic Publishing.
- Dash, N.S. (2018). Linguistic and Extralinguistic Factors behind Spelling Variation of Bangla Words. *Language Forum*, 44, 75-93.
- David, A. (2015). *Descriptive Grammar of Bangla*. Berlin, München, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9781614512295>
- Ghosh, D. (1953). *Bangla Bhasa O Banan [Bengali Language and Spelling]*. Kolkata: Modern Book Agency.
- Ghosh, M. K. (1994). *Bangla Banan [Bangla Spelling]*. Kolkata: Dey's Publishers.
- Khan, M. H. (2018). *Bangla Orthography and Literacy Development: A Psycholinguistic Perspective*. Dhaka: Dhaka University Press.
- Khan, K. (2005). Bangla Bhasar Banan Samskar: Agun Phele Anc!? [Spelling Reform of Bangla Language: Oven Without Fire?]. In Chakrabarti, Rabi and Kalim Khan (Eds.) *Bangla Banan Bangla Bhasa [Bangla Spelling Bangla Language]*. Kolkata: Kaurab Prakashani, 13-58.
- Kolkata, U, P. (1936). *Bangla Bananer Niyam [Rules of Bengali Spelling]*. Retrieved from <https://dsal.uchicago.edu/dictionaries/biswas-bangala/appendices/appendices.pdf>
- Mahalanabish, P. (1925). *Chalti Bhasar Banan [Spelling of Colloquial Speech]*. Kolkata: Prabasi.
- Mamud, H. (2016). *Bangla Lekhar Niyamkanun [Rules of Writing Bengali, Reprint]*. Dhaka: Protik Prakasana Sanstha.
- Mandal, K. (2004). Ucchrinkhal Banan: Shrinkhal-Bhanga Shrinkhala (Indiscipline Bangla Spelling: Rule-Breaking Rules). *Thirabijuri*, 4(7), 97 101.
- Mukherjee, A. (2005). *Banglay Yuktabarna O Kichu Prasangik Prashna (Bangla Consonant Clusters and Some Relevant Questions)*. Bangla Bai (Bangla Book), 1-12.
- Mukherjee, M. (2023). The Influence of Orthography on Spoken Word Recognition in Bangla. *Proceedings of Formal Approaches to South Asian Languages* 12(1), 169-182. <https://ojs.ub.uni-konstanz.de/jsal/index.php/fasal/article/view/291>
- Peereman, R., Dufour, S., & Burt, J. S. (2009). Orthographic Influences in Spoken Word Recognition: The Consistency Effect in Semantic and Gender Categorization Tasks. *Psychonomic Bulletin & Review*, 16(2), 363-368. <https://doi.org/10.3758/pbr.16.2.363>
- Sapir, E. (1921). *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace & World.