Rhythm Measures and Language Endangerment: The Case of Deori

Krisangi Saikia[(✉)] and Shakuntala Mahanta [ID]

Department of Humanities and Social Sciences, Indian Institute of Technology Guwahati,
Guwahati 781039, India
{krisangisaikia,smahanta}@iitg.ac.in

Abstract. This paper presents a study on the rhythm in read speech for Deori (L1), a Tibeto-Burman language and Assamese, an Indo-Aryan language (L2) spoken by the Deori speakers in the state of Assam, India. This study aims to explore and analyze the rhythmic patterns exhibited in read speech for Deori, focusing on aspects such as syllable timing, and duration for both L1 and L2 read speech. To analyze the speech rhythm, rhythm measures such as %V, nPVI, rPVI, varco-V, varco-C, ΔV, and ΔC were calculated for the read speech. Regardless of the rhythmic class of Deori (L1), the results on read speech showed that Assamese (L2) are similar to Deori (L1) in terms of nPVI-V, rPVI-C, whereas it exhibits a shift towards the mora-timed class in terms of %V and ΔC. This study provides valuable insights into the complex interplay between L1 and L2 rhythm patterns. These findings highlight the significance of considering factors such as speech rate and prosodic structure when examining rhythmic differences in bilingual speech.

Keywords: Deori · Assamese (L2) · Rhythm · Speech Rate

1 Introduction

Deori is a Tibeto-Burman language spoken in the state of Assam, India. Deori, an endangered language, is currently in a highly vulnerable state, considering it almost moribund. However, a glimmer of hope emerges from a recent study where the research highlights that some young children have been observed learning the Deori language, indicating a potential avenue for extending its existence [1].

Since Deori is spoken in Assam, and almost the entire community is bilingual their speech repertoire comprises their native language as their first language (L1) and Assamese as their second language (L2). In this context, it is important to investigate whether speakers who have an Indo-Aryan language with trochaic prominence as their second language (L2) experience any impact on rhythm. This is particularly relevant when these L2 speakers come from a linguistic background that includes a highly vulnerable language within the Tibeto-Burman language group, which exhibits an iambic prominence pattern characterized by initial vowel lengthening [2]. This paper attempts to study the rhythmic patterns in Deori (L1) and L2 (Assamese). Both of them are compared with the previously analyzed prototypical stress, syllable and mora timed languages.

Rhythm is a significant prosodic characteristic that plays a crucial role in the naturalness of speech. Traditionally, spoken languages have been divided into three rhythmic categories, known as "stress-timed," "syllable-timed," and "mora-timed" [3–5]. The categorization is based on the concept of isochrony, which states that speech is divided into relatively equal units of duration: syllables in syllable-timing languages such as French and Italian, inter-stress intervals in stress-timing languages like English and German, and mora intervals in mora-timing languages such as Japanese [6].

However, there is no reliable acoustic evidence that proves the presence of isochronous units [7–9]. Isochrony is thus viewed as a more impressionistic trait that correlates with particular phonological features such as syllable structure, vowel reduction, and stress [7]. Recent research has shifted away from the primary focus on isochrony in favor of a more detailed study of the variability in the durations of consonantal and vocalic intervals for the acoustic perception of rhythmic distinctions. The standard deviation of consonant duration (ΔC), percentage of vocalic duration (%V) [8], and pairwise variability index (PVI) [9] for vocalic and consonant durations are all examples of such measures. Speech rhythm is usually divided into rhythmic classes, with languages being either stress-, syllable-, or mora-timed. So, the basic unit of rhythmic speech is either the foot (e.g., English), syllable (e.g., French), or mora (e.g., Japanese).

Stress-timed languages have complex syllable structure and vowel reduction in contrast to syllable-timed and mora-timed, they have simple syllable structure and avoid vowel reduction [7]. Temporal measurements, such as ΔC (standard deviation of consonantal intervals), ΔV (standard deviation of vocalic intervals), and %V (percentage of vocalic intervals in an utterance) were measured. Out of these three temporal measures, the combination of %V and ΔC was considered to best fit for distinguishing rhythm classes. The stressed-timed and syllable-timed languages cluster differently when %V and ΔC are plotted on an x-y plane [8]. Speaking rate affects measurements like %V, ΔV, and ΔC, making them less efficient in distinguishing rhythm classes. Thus, Pairwise Variability Index (PVI) were proposed to decrease the effect of speaking rate. This approach classifies languages based on durational variability of successive units of speech and can reflect normalized (npvi) or raw (rpvi) values [9]. Whereas Varcos, were developed to minimize the effect of speech tempo [10]. It is important to mention that some claims have been made in the literature suggesting that the existing rhythm metrics are not capable of adequately classifying languages into distinct rhythmic classes [11].

In addition to rhythmic studies on native speech (L1), some studies have investigated rhythmic patterns in non-native speech (L2), such as English as a second language for Mandarin and Cantonese speakers with Mandarin or Cantonese as their first language [12]. Studies also explored the influence of the first language (L1) on the second language (L2) for Dutch, English, and Spanish speakers [13]. It is essential for computer-assisted language learning systems to be able to recognize rhythmic patterns in non-native speech. Some occurrences of rhythmic similarities between the L1 and L2 in non-native speech lend credence to the hypothesis of L1 transfer effects. In other cases, non-native speech shows rhythmic patterns nearly identical to either L1 or L2 [6].

2 Methodology

This work investigates the rhythm of read speech of Deori speakers in Assam. Assamese is the dominant language in Assam so, speakers of Deori are bilingual as they can speak Assamese as L2 and to some extent English L3, especially the younger generation [1]. In this work, we investigated the difference in rhythm of the speakers of Deori (L1) and Assamese (L2) reading the story "The North Wind and the Sun". Conventional rhythm measures, such as %V, nPVI, rPVI, varco-V, varcoC, ΔV and ΔC are calculated for read speech.

2.1 Participants

A total of eight participants, all native speakers of Deori (L1) and also proficient in Assamese (L2), took part in two production experiments. The age range of the participants was between (21 to 36 years), consisting of four male and four female speakers who recorded both languages. Each participant was asked to produce the story four times, ensuring a natural speech rate and intonation pattern. The best three repetitions produced by each speaker were considered for final analysis. The translated story comprises roughly 11 sentences for each language with varied syllable lengths (ranging between 6 to 12 syllables per sentence). The recorded speech data were annotated at the phoneme level in Praat 6.1.06 [14], delineating vocalic and consonantal intervals based on auditory and acoustic cues according to standard segmentation criteria [15].

2.2 Materials

The English version of "The North Wind and the Sun" was translated into Deori and Assamese [16]. Translation has been done by a native speaker of Deori language. Prior to recording, the data sets were given to them to familiarize themselves with the sentences and were allowed to rehearse a couple of times to avoid pauses and hesitations. Speakers were instructed to read the sentences on a sheet at their own pace and as naturally as they would in a conversation.

2.3 Procedure

After the data was recorded, it was annotated at the phoneme level in PRAAT [14]. The *Correlatore* program (version 2.3.4) [17] was used to extract different rhythmic metrics, including Cmean, Vmean, %V, ΔC, ΔV, Varcos (Varco-V, Varco-C), and the PVI (nPVI, rPVI) from the annotated speech data. The *speaking rate* also influences rhythm measures. The speech rate is calculated in terms of the time taken syllables per second and segments per second. The values of these matrices were plotted against each other using the *ggplot* package (Figs. 1 and 2, for example) in the R software (version 4.2.2 (R Core Team, 2022) [18].

3   Results

3.1   Syllable Structure of Deori

Deori typically employs the CV syllable type as the default or unmarked syllable type. This aligns with the moraic theory of syllable weight. Deori follows a canonical syllable structure of (C)V(C), where the onset (initial consonant) and coda (final consonant) are optional [2]. This is also true for Assamese (L2) [19] as can be seen in Fig. 1 for comparison. Deori syllables tend to resemble to French language. Deori (L1) CV- interval shows 75.1% among other syllable types. Whereas Assamese (L2) CV intervals shows 65.5% in the entire passage.
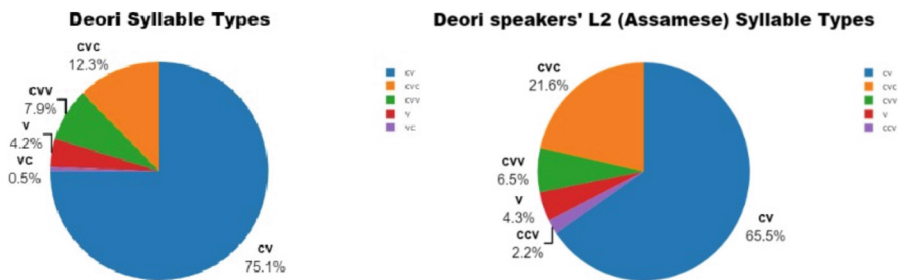

Fig. 1. Syllable types of Deori (L1) on the *left* and Assamese (L2) on the *right* of the read passages in the story *"The North Wind and the Sun"*.

3.2   Correlation of Rhythm Metrics

Several rhythm measures have been demonstrated to be directly or indirectly proportional to the rate of speech in the literature. It has been suggested that utterance length is another aspect to which rhythm metrics are particularly sensitive. It has also been demonstrated that the extent to which these factors influence rhythm measures varies from one language to another. As can be seen in Figs. 2 and 3. Pearson correlation was calculated for each text independently. The figures clearly show that the measurements' correlation varies by language. Rate of articulation, in terms of segments per second (sg/s), has a negative correlation on (L1) and nPVI-V, ΔV. However, there is a robust inverse relationship between Varco-C for both length and syllable per second as can be seen in Fig. 2. In the case of Assamese (L2), the impact of rate of articulation is highly significant across all seven rhythm measures investigated in this study. As seen in Fig. 3, segment/second is negatively correlated with all the rhythm measures.

3.3   Rhythm Measures

Deori (L1) and Deori speaking Assamese (L2) rhythm results are presented in Table 1, along with other languages [8]. This allows us to make direct comparisons to earlier findings. And to compare the results with the previous findings, we plotted the values
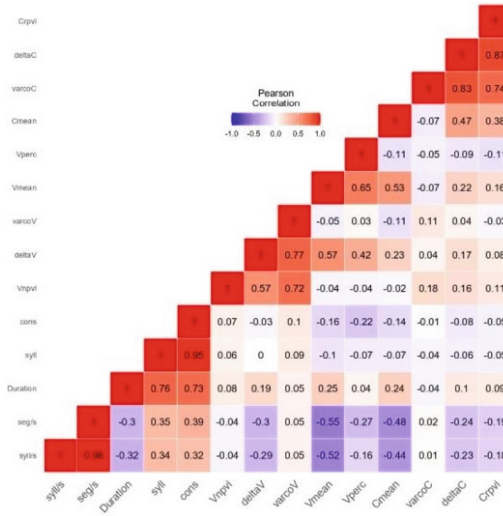
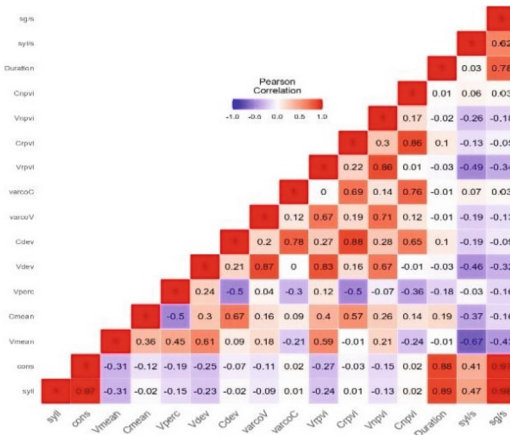**Fig. 2.** Pearson correlation matrix of measures for Deori.

Fig. 3. Pearson correlation matrix of measures for Deori L2 (Assamese).

of %V and ΔC rhythm metrics of Deori (L1) and (L2). As we can see in Fig. 4, the ΔC values for Deori (L1) are close to the other three syllable-timed languages (French, Spanish and Catalan), which makes it clear that Deori should be categorized as syllable-timed while Deori-speaking Assamese (L2), is more a mora-timed language which tends to cluster with Japanese. Whereas in Fig. 5 we plot the values of nPVI-V and rPVI-C with other languages [9]. The results are presented in Table 2. It can be seen that both Deori (L1) and Assamese (L2) the nPVI-V is similar to that of Japanese, but rPVI-C for Deori (L1) showing tendency of shifting towards syllable-timed language and can be seen clustering with French.
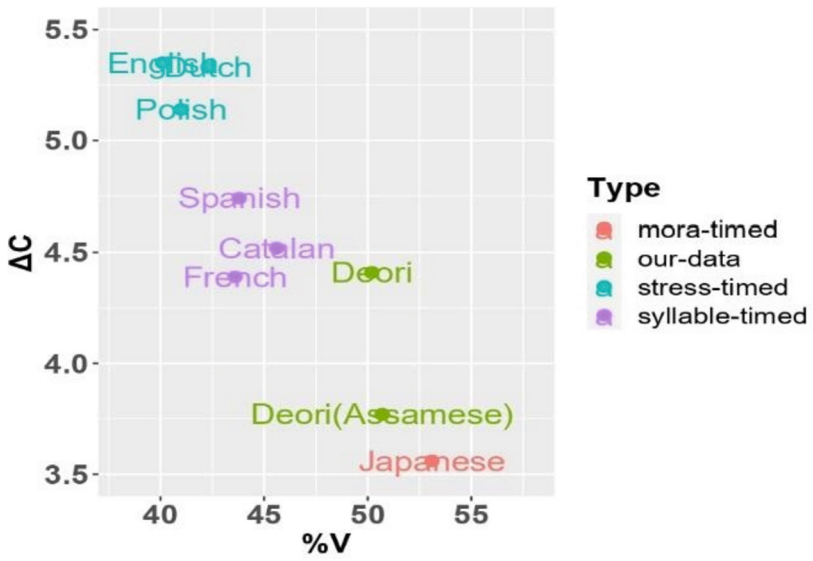
Fig. 4. ΔC and %V for Deori L1 and L2 speech with different languages [8].
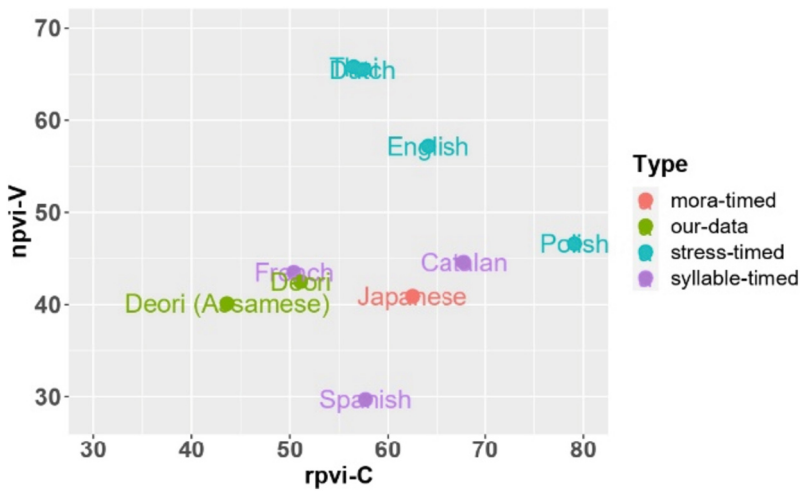

Fig. 5. nPVI-V and rPVI-C for Deori L1 and L2 speech with different languages [9].

Table 1. The values of Rhythm correlates for [British English, Polish, Dutch (Stressed-timed), Spanish, French, Catalan (Syllable-timed) and Japanese (Mora-timed)] along with their standard deviation as proposed by [8] are compared with the values for Deori (L1) and Deori Assamese (L2).

Language	%V(SD)	ΔV(SD)	ΔC(SD)
English	40.1 (5.4)	4.64 (1.25)	5.35 (1.63)
Polish	41.0 (3.4)	4.23 (0.67)	5.33 (1.18)
Dutch	42.3 (4.2)	3.11 (0.93)	5.37 (1.5)
Spanish	43.8 (4.0)	3.32 (1.0)	4.74 (0.85)
French	43.6 (4.5)	3.78 (1.21)	4.39 (0.74)
Catalan	45.6 (5.4)	3.68 (1.44)	4.52 (0.86)
Japanese	53.1 (3.4)	4.02 (0.58)	3.56 (0.74)
Deori	50.2 (5.7)	3.77 (1.64)	4.41 (0.75)
Deori (Assamese)	50.7 (5.8)	3.61 (1.2)	3.77 (0.8)

Table 2. The values of Rhythm correlates for [British English, Thai, Dutch, Polish (Stressed-timed), Spanish, French, Catalan (Syllable-timed) and Japanese (Mora-timed)] as proposed by [9] are compared with the values for Deori (L1) and Deori Assamese (L2).

Language	%V	nPVI-V	rPVI-C	ΔC	ΔV
British English	41.1	57.2	64.1	56.7	46.6
Thai	52.2	65.8	56.5	46.1	74.8
Polish	42.3	46.6	79.1	71.4	44.9
Dutch	44.9	65.5	57.4	53.7	48.4
Spanish	50.8	29.7	57.7	47.5	20.7
Catalan	43.6	44.6	67.8	62.1	33.9
French	50.6	43.5	50.4	42.4	35.5
Japanese	45.5	40.9	62.5	55.5	53.0
Deori	50.2	42.5	51.1	44	37.7
Deori (Assamese)	50.7	40.1	43.6	37.7	36.1

4   Conclusion

We have analyzed the differences in rhythmic patterns of speech of Deori (L1) and Assamese (L2) speakers who were raised speaking Deori as their first language. The rhythmic patterns of read speech were evaluated using nine different rhythm measures. In terms of rate-normalized measures such as nPVI-V and rPVI-C values it was found that L2 speakers gave a rhythmic mode relatively comparable to Deori (L1), and in terms of %V and ΔC values, Assamese (L2) tends to cluster with Mora timed languages,

regardless of the rhythmic class of (Deori) L1 which tends to cluster with syllable-timed languages viz., (French, Spanish) as can be seen in Fig. 4. This matches our subjective auditory impression of L2 speech, in which the perceived rhythm may not fit neatly into any of the rhythm class categories. Possible replacement metrics are needed, and the link between rhythmic metrics and other measures of fluency and naturalness must be explored.

Research conducted on Deori phonology reveals that it exhibits an iambic pattern, with a notable lengthening of the second syllable [2]. In contrast, study on Assamese indicates that it displays a trochaic pattern with a preference for heavy syllables [19]. These findings support our observations, considering the vulnerability of Deori and the ongoing language shift observed among Deori speakers. Our research suggests that Deori speakers can attain a high level of proficiency in bilingualism, despite their native language being a Tibeto-Burman language with distinct characteristics such as iambic prominence and remnants of tonal features.

The results of this study contribute to the understanding of rhythm in Tibeto- Burman languages and provide a foundation for further research in the field. The findings show that speakers with extensive language shift to a dominant L1 may be proficient in the subtle feature of the rhythm properties of the L2. This again validates our concern that language endangerment is a gradual process – it starts with gradual bilingualism, extensive proficiency in the L2 and finally acquiring the L2 with great sophistication leading to the complete replacement of the L2 with the L1.

Further, subjective listening tests will be conducted to analyze rhythmic patterns in L1 (native language) and L2 (second language) speech to see if L2 speakers exhibit significant deviations from the rhythmic patterns of their L1. These tests can reveal how speakers, perceive the linguistic features of their native language. If these tests show that individuals have difficulty perceiving or identifying these features, it may indicate that the language is undergoing a shift, and are no longer fully attuned to its linguistic nuances.

References

1. Acharyya, P., Mahanta, S.: Language vitality assessment of Deori: an endangered language. Lang. Doc. Conserv. **13**, 514–544 (2019)
2. Acharyya, P.: Phonology of Deori: an 'endangered' language. Ph.D. dissertation. Indian Institute of Technology, Guwahati, Assam (2019)
3. Abercrombie, D.: Elements of General Phonetics. University Press, Edinburgh (1967)
4. Ladefoged, P.: A course in Phonetics. Harcourt Brace Javanovich, New York (1975)
5. Pike, K.L.: Intonation of American English. University of Michigan Press, Ann Arbor (1945)
6. Gu, W., Hirose, K.: Rhythmic patterns in native and non-native Mandarin speech. In: Social and Linguistic Speech Prosody: Proceedings of the 7th International Conference on Speech Prosody, pp. 587–519 (2014)
7. Dauer, R.M.: Stress-timing and syllable-timing reanalyzed. J. Phon. **11**, 51–69 (1983)
8. Ramus, F., Mehler, J.: Language identification based on supra-segmental cues: a study based on resynthesis. J. Acoust. Soc. Am. **105**(1), 512–521 (1999)
9. Grabe, E., Low.: Durational variability in speech and rhythm class hypothesis. Pap. Lab. Phonol. **7**(515–546) (2002)

10. Dellwo, V.: Rhythm and speech rate: a variation coefficient for deltaC. In: Language and Language-Processing, pp. 231–241. Peter Lang, Frankfurt am Main (2006)
11. Arvaniti, A.: The usefulness of metrics in the quantification of speech rhythm. J. Phon. **40**(3), 351–373 (2012)
12. Mok, P., Dellwo, V.: Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. In: Proceedings of Speech Prosody, Campinas, Brazil, pp. 423–426 (2008)
13. White, L., Mattys, S.L.: Calibrating rhythm: first language and second language studies. J. Phon. **35**, 50–52 (2007)
14. Boersma, P., Weenink, D.: Praat: doing phonetics by computer (version 5.4). [Computer program]. The Netherlands: University of Amsterdam, Amsterdam (2014). http://www.praat.org
15. Frota, S., Vigário. M.: On the correlates of rhythmic distinctions: the European Brazilian Portugues case, vol. 13, no. 2, pp.247–275 (2001)
16. Mahanta, S.: Assamese. J. Int. Phon. Assoc. **42**(2), 217–224 (2012)
17. Mairano, P.: Correlatore. 2.1 (2009). https://www.lfsag.unito.it/correlatore/downloadenhtml
18. R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2022). https://www.R-project.org/
19. Mahanta, S.: Some aspects of prominence in Assamese and Assamese-English, MPhil thesis, Central Institute of English and Foreign Languages (2001)