

Comparative analysis of Negation in Marathi and Hindi in context of Translation

^[1] Shraddha Kalele, ^[2] Lenali Singh, ^[3] Ajai Kumar

^[1] ^[2] ^[3] C-DAC, Pune, Maharashtra, India

Corresponding Author Email: ^[1] shraddhak@cdac.in, ^[2] lenali@cdac.in, ^[3] ajai@cdac.in

Abstract— In a natural language, the expression of denying or negating something is expressed by Negative sentences. It allows speakers and writers to indicate the opposite or absence of a particular action, state, or proposition. Every language has its unique way of expressing negation. In this paper, we have conducted the comparative analysis of these negative sentence patterns in Hindi and Marathi by exploring the categorical, syntactical and morphological dimensions of negation, offering insights valuable to translators, linguists, and developers of Neural machine translation working on Hindi-Marathi translation.

Index Terms— Comparative analysis, Hindi, Marathi, Morpho-syntactic analysis, Machine translation, Negation, Negative marker, translation.

I. INTRODUCTION

Hindi and Marathi, two prominent Indo-Aryan languages spoken in India, despite their shared linguistic ancestry, these languages exhibit notable distinctions in terms of morphology, syntax and semantics. Negative sentence is one such area where these languages diverge significantly. In this study, we have conducted an analysis of negative sentence pattern of Hindi and Marathi language firstly to throw light on differences in category, sentence patterns and morphology, to assist translators by highlighting the complexities of negative sentence structures in both Hindi and Marathi, and to provide a set of examples in negation that can serve as a reference for creation of training as well as test data for developing and evaluating the machine translation systems.

II. RELATED WORK

In the past, various Marathi grammarians, such as Damle (1970) [1] and Walimbe (2010) [2], have delved into the subject of Marathi negation. Wali, K, and Dhongde, R. V. (2009) [3] have also explained negation. Pandharipande (1977) [4], extensively explored sentence negation and non-finite clause negation, offering numerous illustrative examples. Additionally, Pandharipande (1986) [5] examined the changes in negation within the future tense in the context of language contact between Nagpuri and Puneri Marathi. Mhaskar (2012) [6] delved into the study of negation in language contact scenarios by comparing spoken data from Nagpuri Marathi and Standard Marathi. Moon (2016) [7] contributed a morphological analysis of negation in Marathi. Likewise, in Hindi, Kachru (1966) [8] offers the initial linguistic description of negation, by briefly discussing the derivation of negative sentences. Her analysis delves into the relationship between the negative particle and other sentence

constituents in Hindi. Subsequently, numerous scholars have undertaken the study of negation in Hindi. Bhatia (1979) [9] provides a comprehensive account of the syntax and semantics of negation in six South Asian languages, including Hindi. Concise descriptions of negation in Hindi are also found in Mohanan (1994) [10] and Dwivedi (1991) [11]. Kachru (2006) [12] further delves into the intricacies of sentential and constitutional negation in Hindi.

These scholars have provided valuable insights into negation in Hindi as well as in Marathi and described the phenomenon of negation for individual language. In this work, we analyze negations in Hindi and Marathi in context of their role in the machine translation systems. For this, we have conducted comparative analysis of negation in Hindi and Marathi based on i) syntax of negative sentences, ii) position of negating token iii) part of speech category of negating token and iv) the morphological changes occurring in the preceding or following verb forms. These properties, if not handled correctly may lead to inadequate, syntactically or semantically ill formed sentences. In tern it serves as the benchmark pattern for test as well as training data for machine translation model.

III. MORPHO-SYNTHETIC COMPARATIVE ANALYSIS OF NEGATION IN HINDI AND MARATHI

Every natural language has ways at its disposal to reverse the value of truth in a sentence which is called negation. In the context of negation, the languages vary with respect to following

- Number and form of negative elements,
- Part of speech category,
- Syntactical structures employed for expressing negation, and,
- Position of negative elements

Keeping these facts in mind, the behavior of negating words and the sentences, in Hindi and Marathi have been analyzed and following observation were made:

- Number and form: Marathi employs a diverse set of seven words (नाही, नको, नये, नसू, नव्हे, न, ना) whereas Hindi employs only four words (नहीं, मत, न, नही). The form of these negative words in Marathi varies with respect to tense and the subject, where as in Hindi the form of negative element remains unaltered with respect to tense or subject.
- Part of speech: In Marathi, these negative words belong mainly two part of speech category namely verb and particle, whereas in Hindi, they exclusively function as particles.
- Syntactical patterns: Both the languages exhibit distinct syntactical patterns which are discussed below with relevant examples
- Position: In Marathi the negating token generally takes sentence final position while in Hindi it is followed by the verb token.

These above-mentioned points are discussed below with the common and language specific sentence pattern along with the examples from both or relevant language/s..

IV. SYNTACTIC ANALYSIS

This section briefly discusses the diverse negative sentence pattern found in Hindi and Marathi and provides a comparative analysis by focusing on specific characteristics such as tense, position of negative token, it's part of speech category etc

Sentence pattern:1.: This is the usual type of sentence pattern found in both the languages. In this pattern, in Hindi, the position of the negative marker is always before the verb, in all the three tenses (present, past and Future) (Kumar 2006 [13], Bhatia 1978 [9], Kachru 1966 [8]). whereas the negative marker occurs in sentence final position in Marathi. This pattern is generally seen in the sentences with nonverbal predicate (Table I).

Table I

Language	Present	Past	Future
Hindi	मैं विद्यार्थी नहीं हूँ।	मैं विद्यार्थी नहीं था।	मैं विद्यार्थी नहीं होऊंगा। ¹
Marathi	मी विद्यार्थी नाही।	मी विद्यार्थी नव्हतो.	मी विद्यार्थी नसेन.

As we can see that in Marathi the form of the negating element changes with respect to tense as नाही in present,

नव्हतो in past, and नसेन in future. Contrast to this the negative marker remains the same irrespective of tense in Hindi. These 3 variants of negating token, seen in 3 tenses in Marathi are formed from different root. नाही from नाह or नाहीये from नाही+आहे (present), नव्हतो from न+हो, नसेन from न+अस.

Similar sentence pattern is seen in imperative sentences in both the languages. It is interesting to note that Marathi uses नको (don't want), a different form of negating token, in imperative sentences that too in sentence final position, while Hindi uses the same नहीं in preverbal position with चाहिये as verb. (Table II).

Table II

Marathi	Hindi	Marathi	Hindi
नको (v)	नहीं चाहिये	मला चपलानको.	मुझे चपल नहीं चाहिये।
नको (v)	नहीं चाहिये	मला खाणेनको.	मुझे खाना नहीं चाहिये।

Sentence pattern:2. This is a usual type of sentence pattern which occurs in both Hindi and Marathi. This shows similarity with pattern1 in the position of negating token as preverbal in Hindi and sentence final in Marathi. The difference is that this negating token is preceded by the object of verb and followed by the verb in Hindi while preceded by the verbal Noun (krudant [14]) in Marathi. (Table:III).

Table III

Language	Present	Past	Future
Hindi	वह काम नहीं करता.	वह काम नहीं करता था।	वह काम नहीं करेगा।
Marathi	तो काम करत नाही.	तो काम करायचा नाही.	तो काम करणार नाही.

This krudant form is discussed in detailed in the Variation in Morphology section below.

The sentence pattern 2 is also common in imperative sentences showing instruction, request or advice. Here, Marathi uses नको as well as नये (not used with nonverbal predicates in sentence pattern1) and restricts the use of नाही and Hindi uses न and ना and नहीं ... चाहिये in place of नहीं. (Table:IV)

Table IV

Hindi	Marathi
यह काम न/ ना करें।	हे काम करू नका/ नये.
खाना देरी से नहीं खाना चाहिये।	जेवण उशीरा करू नये.

Sentence pattern:3. This pattern is similar in both Marathi and Hindi, where the negative marker act as conjunction and joins two nouns, adjectives or even sentences. (Table V)

¹ The example is not correct as in Hindi the bare future form is not used. It is given just to show equivalence for bare future form in Marathi.

² नाही and नाहीये are different words. नाही is a verb and is derived from नाह dhatu and has forms like नाहीस, नाहीत while नाहीये is made due to sandhi of नाही+आहे.

Table V

Marathi	Hindi
हा सिनेमा लहानासाठी नसून मोठ्या मुलांसाठी आहे	पिक्चर छोटे नहीं बड़े बच्चों के लिए है।
मी घरी जात नसून, बाजारात जात आहे.	मैं घर नहीं बाजार जा रही हूँ।

In the first example (Table:5) it is joining 2 adjectives (लहानासाठी नसून मोठ्या मुलांसाठी, छोटे नहीं बड़े बच्चों) while in second case it is joining 2 sentences (१. मी घरी जात नाहीये., २. बाजारात जात आहे.).

Some exception to this sentence pattern may occur in Hindi with the use of *तो* and *इसलिए* (with negation) but in those examples as well Marathi continues to use negative marker as conjunction as shown in (Table:VI)

Table VI

Marathi	Hindi	Marathi sentences	Hindi sentences
हे उपलब्ध नसल्यास, ते स्वच्छ बांबूच्या चटईवर देखील वाळवले जाऊ शकतात.	यदि ये उपलब्ध नहीं हों तो बाँस की साफ चटाई पर भी इसे सुखाया जा सकता है	1. हे उपलब्ध झाले नाही 2. त्याला स्वच्छ बांबूच्या चटईवर देखील वाळवले जाऊ शकते	1. ये उपलब्ध नहीं हैं। 2. इसे बाँस की साफ चटाई पर भी सुखाया जा सकता है।

Sentence pattern 4: This sentence pattern is similar in both Hindi and Marathi where the negative element occurs in the coordinator position in Marathi [4] as well as in Hindi. The examples given below (table VII), shows that in both the languages, the negative marker is kept before each constituent which is to be negated

Table VII

Marathi	Hindi
त्याने ना आंबे आणले ना पेरू.	उसने ना आम लाए ना अमरूद [15]
हा पदार्थ जाळल्यावर ना वितळतो ना कुजतो	यह ऐसा पदार्थ है जो जलाने पर ना गलता है ना सड़ता है।

Sentence pattern: 5: This sentence pattern illustrates the negation of non-finite verb from i.e., participle, gerund, and infinitive. It has several sub patterns. Let us discuss them one by one

Participle: To negate a participle, Marathi uses 2 negating markers *न*, and *नको*. These two markers are used in different sub patterns

The negating particle *नको* is used, when negating a participle at sentence initial position particularly formed from verb root 'अस' (to be, to exist) or 'वाट' (to feel) example:

- *नको असलेल्या जाहिराती, पॉप-अप आणि मालवेअर काढून*

³ In very rare case, n Marathi this 'नकोशी' - a form of 'नकोसा'- is used as Noun in the sense of नको असलेली मुलगी. Examples: 'काही तासांपूर्वीच जन्मलेली ही 'नकोशी' सुखरूप आहे.' [16]

टाकणे participle (perfective)

- *नको वाटणाऱ्या व्यक्तीचा त्रास होतो.* Participle (Habitual). This sentence pattern is missing in Hindi.

Table VIII

Non-finite	Marathi	Hindi
Participle (Habitual)	उपवासात न चालणारा पदार्थ कोणता?	व्रत में ना खाया जाने वाला व्यंजन कौन-सा है।
Participle (perfective)	मला तु न वाचलेले पुस्तक दे.	तुमने जो किताब ना पढी हो वो मुझे दो। तुमने जो किताब पढी ना हो वो मुझे दो.

The negating particle *न* is used, when negating the participle derived from verb root other than 'अस' (to be, to exist) or 'वाट' (to feel). In both the above cases Hindi uses *न* to negate the non-finite clauses irrespective of verb root from which it is formed. (Table VIII)

Gerund: For negating gerund in both the languages *न* particle is places just before the gerund which is to be negated. (Table IX)

Table IX

Non-finite	Marathi	Hindi
Gerund	रात्री वेळेवर न झोपल्याने आजार उद्भवतात.	रात को समय पर न सोने के कारण बिमारियां होती है।

Infinitive: For negating gerund, Marathi particle *नको* in is used in preverbal position which is similar to Hindi which uses particle *नहीं* in preverbal position. (Table X)

Table X

Non-finite clause	Marathi	Hindi
Infinitive	मला जाणे नको वाटते.	मुझे जाना नहीं है। मुझे जाने की इच्छा नहीं है।

Sentence pattern:6. This is a variation of sentence pattern 2 which depicts a special phenomenon where the negating particle changes its position with the attachment of emphatic marker in both the languages. Therefore, it is kept under separate sentence pattern. In sentence pattern 2 we have shown that the negating particle takes sentence final position in Marathi and preverbal in Hindi. But this structure changes with the introduction of emphatic marker. It causes the change in position of negating particle from sentence final to preverbal position in Marathi and sentence final position in Hindi. So when the emphatic marker is attached to negating particle in Marathi, it causes the shift of negating particle from sentence final position to preverbal position. Though in this case the Hindi follows the sentence pattern 2 (with no relocation of negating particle) even after addition of *ही* to negating token (Table: XI).

Table XI

Marathi	Hindi
तो घरी नाहीच जाणार.	वह घर नहीं ही जाएगा।

But when this emphatic marker to added to verb form in Hindi. It causes the shifts of negating token to sentence final position but even after the attachment of emphatic marker to verb form Marathi continues to follow sentence patten 2 with negating token in sentence final position. Examples are given below. (Table XII)

Table XII

Marathi	Hindi
तो घरी जाणारच नाही.	वह घर जाएगा ही नहीं।

Sentence pattern 7.: This pattern describes double negation in Hindi and Marathi. In a conversational situation, where double negation is most common, the first negation (sentence initial position) is answer to question and remaining part of the sentence is a statement with negation in both the languages. This is used to reinforce certain action. The examples are given in table XIII.

Table XIII

Marathi	Hindi
नाही, तो जेवणार नाही.	नहीं, वह खाना नहीं खाएगा।

Sentence pattern:8. This sentence pattern is unique to Marathi only where the negating particle is attached to verb which shows the absence of willingness or ability. These particles include -ना, -नासा। Example:

- त्याच्याकडून हे काम होईना.
- तो हे काम करीनासा झाला.

V. VARIATION IN MORPHOLOGY

In sentence pattern (2) of Marathi, an interesting observation is made regarding the Krudant form of the verb and the negating token. In this context, the Krudant form conveys tense information through its krudant suffixes. From the examples given in Table 3, 'करत' signifies the present tense, 'करायचा' denotes the past tense and 'करणार' indicates the future tense, Notably, this Krudant form remains unaltered with respect to agreement markers for person and number. In these sentences, negative marker 'नाही' agrees with the subject (or object) and consequently undergoes alterations. In same examples, in Hindi, the verb form 'करता' not only conveys tense information but also agrees with the subject or object and undergoes alterations.

The following table illustrates the variations of 'नाही' based on person and number in Marathi and that of 'कर' in Hindi. (Table XIV)

Table XIV

Person	Marathi		Hindi	
	Singular	Plural	Singular	Plural
1 st	करत नाही	करत नाही	नहीं करता	नहीं करते
2 nd	करत नाहीस	करत नाही	नहीं करते	नहीं करते
3 rd	करत नाही	करत नाहीत	नहीं करती/नहीं करता	नहीं करते /नहीं करतीं

Similar morphological phenomenon is observed in the sentence's construction with 'नको' and 'नये' in sentence pattern 2.

VI. CATEGORY VARIATION

Hindi and Marathi show difference in the Part of speech category for the negative marker. In Hindi, it is always a particle while in Marathi it falls in following categories verb, particle and sometimes coordinator and conjunction.

VII. ROLE OF SENTENCE PATTERN IN EVALUATION OF MACHINE TRANSLATION SYSTEM AND CREATION OF TRAINING DATA

To understand the extent of the knowledge acquired by language models within machine translation systems, it is necessary to comprehend the intricacies of language behavior and construct benchmark sentences that contains the unique linguistic attributes of the language under evaluation. The study of negative sentence structures detailed in this paper aims at shedding light on the distinctive and common characteristics of Hindi and Marathi in handling negation. These identified sentence structure can act as the pilot patterns to understand the capability of language model in handling translation of negation. Thus, these identified patterns were tested on 4 Machine translation (MT) system, and the outputs of these MT systems were assessed to gauge their performance in handling these patterns. The observations are provided below:

When translating sentence pattern (3), from Hindi-Marathi where the negating token is acting as conjunction between two nouns, it is observed that the Marathi output of all the MT system retains negation but none has used negating token as conjunction. For example, the MT_1 and MT_2 and MT_4 break the sentence into two sentences and joined it with *तर* instead of *नसून*. in Marathi output. The output of MT_3 has issue of adequacy as the translation of "बड़े बच्चों के लिए है" is missing in MT output (Table:XV)

Table XV

Input sentence	पिक्चर छोटे नहीं बड़े बच्चों के लिए है। (Hindi)
Human translated	हा सिनेमा लहानांसाठी नसून मोठ्यासाठी आहे. / हा सिनेमा लहानांसाठी नसून मोठ्या मुलांसाठी आहे. (Marathi)

Input sentence	पिक्चर छोटे नहीं बड़े बच्चों के लिए है। (Hindi)
MT_1	चित्र लहान मुलांसाठी नाही तर मोठ्या मुलांसाठी आहे.
MT_2	हे चित्र लहान मुलांसाठी नाही तर मोठ्या मुलांसाठी आहे.
MT_3	पिक्चर लहान मुलांसाठी नाही.
MT_4	पिक्चर लहान नाही तर मोठ्या मुलांसाठी आहे.

Similarly, the translation of sentence pattern (6) to Marathi, where in Hindi has emphatic marker 'ही' with negation, it is observed that, all the four MT system generates negative output which is without the emphatic marker 'चै' in Marathi. (Table XVI)

Table XVI

Hindi-Marathi	
Input sentence	वह घर नहीं ही जाएगा। (Hindi)
Human translated	तो घरी नाहीच जाणार. (Marathi)
MT_1	तो घरी जाणार नाही.
MT_2	तो घरी जाणार नाही.
MT_3	तो घरी जाणार नाही.
MT_4	तो घरी जाणार नाही.

Similar behavior as above is observed when translating the same pattern (6) from Marathi to Hindi. Here also there is an issue of adequacy since 'चै' the Marathi emphatic particle is not translated in any of the Hindi MT output. Consequently, the original intent of the sentence is not conveyed. (Table XVII)

Table XVII

Marathi-Hindi	
Input sentence	तो घरी नाहीच जाणार. (Marathi)
Human translated	वह घर नहीं ही जाएगा। (Hindi)
MT_1	वह घर नहीं जायेगा.
MT_2	वह घर नहीं जायेगा.
MT_3	वह घर नहीं जायेगा.
MT_4	वह घर नहीं जायेगा.

Sentence pattern (7), is also tested as given in table 19, shows that in MT_1 output is not adequate as आम is missing and verb आणले (to bring) is wrongly translated to लगाए (to plant), in MT_2 output bring is correctly translated to लांना and negation is retained but the syntax is not correct, in MT_3 instead of negating nouns it the verbs are negated, and finally MT_4 generated no output. (Table XVIII)

Table XVII

Marathi-Hindi	
Input sentence	त्याने ना आंबे आणले ना पेरू.
Human translated	उसने ना आम लाए ना अमरूद
MT_1 output	ये अमरूद हमने नहीं लगाए.
MT_2 output	वह आम या अमरूद नहीं लाता था।
MT_3 output	उसने न तो सिरके लाए और न ही बोए।
MT_4 output	NO output

Apart from these pattern MT systems are also tested on language specific patten where we have tested both (1) and (2) sentence from pattern (8), to generate the Hindi output. For sentence (1) it is observed that MT_1 output is adequate since the ability verb 'होना' is used here but the other MT outputs (MT_2, MT_3 and MT_4) have adequacy issue as the meaning of absence of willingness or ability (ना होना) is not conveyed.

In case of sentence (2), the output generated by MT_1, MT_2 and MT_3 are not correct, since it is not adequate enough to convey the meaning intended in Marathi sentence

Table XVIII

	Marathi-Hindi	Marathi-Hindi
Input sentence	त्याच्याकडून हे काम होईना. (Marathi) (1)	तो हे काम करीनासा झाला. (Marathi) (2)
Human translated	उससे यह काम नहीं होगा। नहीं हो पा रहा था (Hindi)	उससे यह काम नहीं हो पा रहा था। (Hindi) / वह यह काम कर नहीं पा रहा था।
MT_1	ऐसा उससे नहीं होगा.	ये काम करीना ने किया था.
MT_2	यह उसके लिए काम नहीं करता था।	उसने काम किया।
MT_3	उसका काम नहीं हुआ।	वह काम नहीं कर रही थी।
MT_4	वह काम नहीं करता था।	No output

The significance of analyzing negative sentence structures becomes evident through the provided examples and observations. These insights not only aid in comprehending the capabilities of machine translation systems through the creation of benchmark data for testing but also assist in identifying the necessary enhancement in training data sets, particularly when examining variations in specific sentence structures across languages

VIII. CONCLUSION

This comparative morpho-syntactic analysis of negative sentences in Marathi and Hindi underscores the remarkable differences in morphology and syntax within these two prominent Indo-Aryan languages, despite their common linguistic lineage. Additionally, this study has established benchmark negation patterns that serve the dual purpose of

evaluating machine translation system output and facilitating the generation of sentences using these patterns to enhance training data sets and improve the quality of Machine translation outputs. [16] Maharashtra times

IX. FUTURE WORK

The current study is limited to the contrastive Morpho-syntactic analysis of Negation in Hindi and Marathi covering most of the negative sentence structure including negation of finite and non-finite clauses, double negation, and negation as conjunction etc. The comparative study may also be extended to negation with subordination, multiple negation, negation with interrogative sentences etc.

Ms. Shraddha Kalele would like to thanks Mr. Sushant Devlekar (Rajya Marathi Vikas Sanstha, Mumbai) for technical discussion.

REFERENCES

- [1] Damle Moro Keshav: Shastriya Marathi vyakarana: 1st edition: Mauj Printing bureau: (1970)
- [2] Walimbe, M.R: Sugam Marathi vyakaran Lekhan: Nitin Prakshan: Pune: (2010)
- [3] Ramesh V. Dhongde and Kashi Wali: Marathi: John Benjamins Publishing company: North America: (2009)
- [4] Rajeswari Pandharipande: Marathi: London and New York: (1977)
- [5] R Pandharipande: Language Contact and Language Variation: Nagpuri Marathi Linguistics: South Asian Languages: Structure, Convergence, and Diglossia, 1986 4.
- [6] Rahul N. Mhaiskar: Change in Progress: Phonology of Marathi-Hindi contact in eastern Vidarbha; International journal of English language, literature and Humanities
- [7] Moon, Swapnil D.: A Computational Morphological Analysis of Marathi Negative Markers: HindiTech: A Blind Double Peer Reviewed Bilingual Web-Research Journal, 7 (8), 87-103 ((2016).
- [8] Kachru, Y.: An introduction to Hindi syntax.: IL: University of Illinois Press: Urbana (1966)
- [9] Bhatia, T.K.: Negation in south Asian Languages. Urbana-Champaign, South Asian Languages Analysis Vol.1 Depart. of Linguistics (1979)
- [10] Mohanan, T.: Argument structure in Hindi.:CA: CSLI Publications: Stanford (1994)
- [11] Dwivedi, V.: Negation as a functional projection in Hindi. Proceedings of the western conference on linguistics, ed. by K.Hunt, T. Perry, and V. Samian, 88-101. Fresno: California State University Press (1991)
- [12] Kachru Y.: Hindi: University of Illinois at Urbana-Champaign: – John Benjamins B.V 2006
- [13] Kumar R.: NEGATION AND LICENSING OF NEGATIVE POLARITY ITEMS IN HINDI SYNTAX, Routledge Taylor & Francis Group (2006)
- [14] Ganesh Bhosale, Subodh Kembhavi, Archana Amberkar, Supriya Mhatre, Lata Popale and Pushpak Bhattacharyya, Processing of Participle (Krudanta) in Marathi, International Conference on Natural Language Processing (ICON 2011), Chennai, December, 2011.
- [15] COSH Conc