

A LINGUISTIC ANALYSIS OF DRAVIDIAN AND INDO-ARYAN LANGUAGE FAMILIES IN INDIA

G. Saidulu

Research Scholar
Arni University
Himachal Pradesh

Dr. P. Suresh

Research Supervisor
Arni University
Himachal Pradesh

Dr. D. Sucharitha

Research Co-Supervisor
Arni University
Himachal Pradesh

ABSTRACT

This paper addresses the theme of the seminar from the perspective of historical linguistics. It introduces the construct of 'language family' and then proceeds to a discussion of contact and the dynamics of linguistic exchange among the main language families of India over several millennia. Some prevalent hypotheses to explain the creation of India as a linguistic area are presented. The 'substratum view' is critically assessed. Evidence from historical linguistics in support of two dominant hypotheses 'the Aryan migration view' and 'the out-of-India hypothesis' is presented and briefly assessed. In conclusion, it is observed that the current understanding in historical linguistics favours the Aryan migration view though the 'substratum view' is questionable.

Keywords: Aryan migration; historical linguistics; language family; Out-of-India hypothesis; substratum

INTRODUCTION

This thesis investigates the origin and spread of the dative-subject construction within the Indo-Aryan languages of South Asia. That this construction is present in many, if not most, of the Indo-Aryan languages as well as those of the genealogically unrelated but geographically contiguous Dravidian language family raises questions as to its status as a contact-induced feature. This thesis develops an account of this construction, arguing that its presence in Indo-Aryan is, in part, a product of contact with Dravidian. In this section, I present the relevant background for this study. I introduce South Asia and the languages that characterize this region, and provide a brief sketch of the region's prehistory. I then review some of the previous literature pertaining to historical Indo-Aryan/Dravidian language contact and the dative-subject construction, identify the shortcomings in this literature, and conclude with the point of departure for this thesis.

The Indo-Aryan group is one of the few groups of Indo-European languages, if not the only one, for which no classification based on rigorous genetic criteria has been suggested thus far. The cause of such a situation is neither lack of data, nor even the low level of its historical interpretation, but rather the existence of certain prejudices which are widespread among Indo logists. One of them is the belief that real genetic relations between the Indo-Aryan languages cannot be clarified because these languages form a dialect continuum. Such an argument can hardly seem convincing to a comparative linguist, since dialect continuum is by no means a unique phenomenon: it is characteristic of many regions, including those where Indo European languages are spoken, e.g. the Slavic and Romance-speaking areas. Since genealogical classifications of Slavic and Romance languages do exist, there is no reason to believe that the taxonomy of Indo-Aryan languages cannot be established. However, this scholarly pessimism does have some grounds. Certain approaches and methods used nowadays in Indological historical linguistics have proven to be inefficient, and without changes in research methodology, significant progress in the genetic classification of the

Indo-Aryan languages is hardly possible. This issue will be discussed at some length below. In the past Indologists had more than once attempted to classify the languages that they studied. In the late 19th and 20th centuries several alternative classifications were suggested. Since some of them still appear in modern Indological publications, I feel it necessary to dwell on them here. The scholar who seems to have been the first one to approach this problem is Rudolf Hoernle. He hypothesized that the Aryans migrated to the Indian subcontinent in two successive waves.

LITERATURE REVIEW

A. Mukhopadhyay et.al (2020) Ever since the discovery of Indus valley civilization, scholars have debated the linguistic identities of its people. This study analyzes numerous archaeological, linguistic, archaea genetic and historical evidences to claim that the words used for elephant (like, 'pīri', 'pīru') in Bronze Age Mesopotamia, the elephant-word used in the Hurrian part of an Amarna letter of ca. 1400 BC, and the ivory-word ('pīruš') recorded in certain sixth century BC Old Persian documents, were all originally borrowed from 'pīlu', a Proto-Dravidian elephant-word, which was prevalent in the Indus valley civilization, and was etymologically related to the Proto-Dravidian tooth-word '*pal' and its alternate forms ('*pīl'/*'pīl'/*'pel'). This paper argues that there is sufficient morphophonemic evidence of an ancient Dravidian '*pīl'/*'pīl'-based root, which meant 'splitting/crushing', and was semantically related to the meanings 'tooth/tusk'. This paper further observes that 'pīlu' is among the most ancient and common phytonyms of the toothbrush tree *Salvadora persica*, which is a characteristic flora of Indus valley, and whose roots and twigs have been widely used as toothbrush in IVC regions since antiquity.

Anastasia Krylova et.al (2019) In this paper, I discuss the issue of replacement of basic lexicon due to language contact, specifically focusing on Indo-Aryan, Dravidian, and Munda languages in South Odisha but with implications for Eastern India as a whole. Apart from direct borrowings, special attention is paid to cases of not particularly obvious substrate influence in Indo-Aryan languages, in which inherited lexicon is substituted by other inherited items as a result of semantic shift typical of a neighbouring language family. This type of phenomena should necessarily be taken into consideration in any quantitative studies involving lexicostatistical calculations.

Anton I. Kogan et.al (2016) Genetic relations among Indo-Aryan languages are still unclear. Existing classifications are often intuitive and do not rest upon rigorous criteria. In the present article an attempt is made to create a classification of New Indo-Aryan languages, based on up-to-date lexicostatistical data. The comparative analysis of the resulting genealogical tree and traditional classifications allows the author to draw conclusions about the most probable genealogy of the Indo-Aryan languages.

Arun Kumar et.al (2017) The Dravidian family is one of the most widely spoken set of languages in the world, yet there are very few annotated resources available to NLP researchers. To remedy this, we create DravMorph, a corpus annotated for morphological segmentation and part-of-speech. Also, we exploit novel features and higher-order models to achieve promising results on these corpora on both tasks, beating techniques proposed in the literature by as much as 4 points in segmentation F1.

Ashwini Deo et.al (2007) While New Indo-Aryan languages are a common example of morphological ergativity, the range of variation in ergative marking and agreement among these languages has not been examined in detail. The goals of this article are twofold. We first present a typology of ergative marking and agreement in Indo-Aryan languages, demonstrating that a progressive loss of ergative marking has occurred to varying degrees in different systems. This process is manifested in two distinct strategies of markedness reduction: loss of overt subject marking in the nominal domain and loss of marked agreement in the verbal domain. Using the framework of Optimality Theory, we account for the typology in terms of universal sub hierarchies of markedness.

Evidence for Dravidian as the source language family

Thomason's second diagnostic requires a source language to be identified for the particular feature under investigation. In the preceding section, I outlined some similarities between Indo-Aryan and Dravidian languages that started appearing in Indo-Aryan around the same time as that of the dative-subject construction. Although this concise enumeration of similar features suggests that transfer between these two families was taking place at the proper time, it still remains to be seen what the directionality of that transfer might have been. I propose that the source language family for the dative-subject construction, and for most, if not all, of the features mentioned in the preceding section for that matter, is Dravidian. Not only is there numerous convincing evidence of Dravidian contact-induced features entering Indo-Aryan at earlier stages, but several of the features noted above, possibly including the dative-subject construction (see Section 5.2), can plausibly be reconstructed for Proto Dravidian. It is also relevant to note that these innovations in early New Indo-Aryan were concurrent with an overall major restructuring of the Indo-Aryan grammatical system from more synthetic and inflectional to more analytic. The Dravidian languages, on the other hand, have changed relatively little structurally over the course of their written history. They preserve the case system and the robust agglutinating structure of their shared ancestral language, Proto-Dravidian. This provides further support for the idea that Indo Aryan was the recipient language family for many Dravidian-origin features.

Absence of feature in Indo-Aryan prior to contact with Dravidian

Thomason's third diagnostic requires that the feature under investigation must be shown to have been absent in the RL prior to contact with the SL. This is not entirely possible in the case of Indo-Aryan and Dravidian, as contact between these families has been hypothesized by most to have begun even before the composition of the Vedas, the oldest attested form of Indo-Aryan. However, it can be shown that the dative-subject construction as it appears in the modern Indo-Aryan languages is not simply a development from an earlier construction in OIA, but that it is a completely novel construction which only began appearing in the early New Indo-Aryan languages, and thus was not feature of Indo-Aryan before the period of contact with Dravidian which took place as the NIA languages were emerging.

examines several instances of verbs in Sanskrit which in some cases take oblique subject-like arguments, but ultimately concludes that Sanskrit did not have a dative subject construction., on the other hand, believe that it did, and that the construction can moreover be reconstructed for Proto-Indo-European. Regardless of the position one takes, however, it is clear that these

verbs which took oblique subject like objects were few and far between in Sanskrit. Moreover, these verbs are not the same one's which develop dative subject-marking in later stages. For this reason, it is the general consensus that the dative-subject construction in the modern languages is not a continuation of a pre-existing construction in earlier Indo-Aryan, but that it developed independently.

Presence of feature in Dravidian prior to contact with Indo-Aryan

Thomason's fifth criterion is a bit more difficult to satisfy. It requires that evidence of the feature in question be found in the proposed SL prior to contact with the RL. This cannot be done through textual data in the case of the dative-subject construction, considering that the earliest textual evidence for Dravidian dates back only to 254 BCE which is, on most accounts, quite a while after Dravidian speakers would have first come in contact with Indo-Aryan speakers.

METHODOLOGY

This chapter deals with the methodology that is employed in the current study, which aims to investigate linguistic variations between speakers of Dravidian and Indo-Aryan languages within the diverse linguistic landscape of Uttar Pradesh (Hindi), West Bengal (Bengali), Tamil Nadu (Tamil), and Andhra Pradesh (Telugu). The focal point of this chapter is to provide a clear and systematic account of the research design, data collection procedures, and analytical framework utilised in pursuing our research objectives. The study encompasses four linguistically distinct states, namely Uttar Pradesh, where Hindi is predominantly spoken; West Bengal, where Bengali is the primary language; Tamil Nadu, characterised by the prevalence of Tamil; and Andhra Pradesh, with Telugu as its dominant language. These states collectively constitute our research universe and represent regions for investigating linguistic variations. Within each state, we have selected four districts as our specific areas of study to ensure a balanced representation of urban and rural settings. The selection of districts and individual samples has been guided by specific inclusion and exclusion criteria, considering population size, linguistic diversity, geographical location, age, gender, educational background, and language proficiency. Rigorous sampling techniques have been employed to maintain statistical significance, and pilot studies were conducted in each state to refine data collection methods. The primary data source consisted of language samples obtained through picture-based written tasks.

CONCLUSION

The conclusion drawn from the extensive analysis of typological error patterns in written English among speakers of the Indo-Aryan and Dravidian language families, as detailed in Chapter 5, underscores the intricate relationship between learners' linguistic backgrounds and their learning of English morpho-syntactic structures. The comprehensive examination of errors in inflectional morphemes across these linguistic groups reveals significant challenges that learners face, particularly in applying rules of verb tense, plurality, adjective forms, and the correct use of inflectional morphemes such as third-person singular present tense, superlative forms, and others. The data suggest that these challenges are not merely isolated incidents but form a consistent pattern across learners from different linguistic backgrounds, indicating a profound impact of native language structures on learning English.

REFERENCES

1. A. Mukhopadhyay et.al (2020), "Ancestral Dravidian languages in Indus Civilization: ultra conserved Dravidian tooth-word reveals deep linguistic ancestry and supports genetics", *Humanities & social sciences communications*, ISSN:2662-9992,vol.8,<https://doi.org/10.1057/s41599-021-00868-w>
2. Anastasia Krylova et.al (2019), "Some areal features of the 110-item wordlist for the Indo-Aryan languages of South Odisha", *Journal of Language Relationship*, ISSN: 2219-4029,vol.17,issue.4,Pp.285–296
3. Anton I. Kogan et.al (2016), "Genealogical classification of New Indo-Aryan languages and lexicostatistics", *Journal of Language Relationship*, ISSN: 2219-4029, vol.14,issue.4,Pp. 227—258
4. Arun Kumar et.al (2017), "Morphological Analysis of the Dravidian Language Family", *Association for Computational Linguistics*, ISSN:1530-9312,Volume 2, pages.217–222
5. Ashwini Deo et.al (2007), "Typological variation in the ergative morphology of Indo-Aryan languages", *Linguistic Typology*, ISSN: 1613-415X, Volume. 10, Issue.3,<https://doi.org/10.1515/LINGTY.2006.012>
6. Ayaz ud Din et.al (2019), "'Phonemic Inventory of Khovar Language: An Acoustic Analysis," *Global Regional Review*, ISSN: 2616-955X ,vol.4,issue.(4), pages 369-380
7. Bharathi Raja Chakravarthi et.al (2022), "Dravidian Code Mix: sentiment analysis and offensive language identification dataset for Dravidian languages in code-mixed text", *Language Resources and Evaluation*, ISSN: 1574-0218, Volume. 56, pages.765–806
8. Bryang.levman et.al (2023), "the influence of proto-dravidian on indo-aryan phonology morphology and syntax", *International Journal of Dravidian Linguistics*, ISSN: 0378-2484,Vol.52,issue.1
9. Chundra A. Cathcart et.al (2020), "Disentangling dialects: a neural approach to Indo-Aryan historical phonology and subgrouping", *Association for Computational Linguistics*, ISSN:1530-9312, pages.620–630,<https://doi.org/10.18653/v1/P17>
10. Claus Peter Zoller et.al (2016), "Outer and Inner Indo-Aryan, and northern India as an ancient linguistic area", *Acta Orientalia*, ISSN: 1588-2667, vol.77, pages. 71–132
11. Clyde Winters et.al (2008), "Origin and Spread of Dravidian Speakers", *Int J Hum Genet*, ISSN :2456-6330,vol.8,issue.(4),pages.325-329
12. Dr. T.Balasubramanian et.al (2019), "dravidians culture and aryan settlement in tamilnadu analysis", *global journal for research analysis*,ISSN:2277 - 8160, volume-8, issue-8,
13. Henrik Liljegren et.al (2017), "Profiling Indo-Aryan in the Hindukush-Karakoram: A preliminary study of micro-typological patterns", *journal of South Asian Languages and Linguistics*,ISSN:2196-078X,Volume.4, Issue.1, <https://doi.org/10.1515/jsall-2017-0004>
14. Jasleen Kaur et.al (2014), "A Study and Analysis of Opinion Mining Research in Indo-Aryan, Dravidian and Tibeto-Burman Language Families", *International Journal Of Data Mining And Emerging Technologies*, ISSN: 2249-3220, vol.4, issue.(2),pages.53-60,DOI:10.5958/2249-3220.2014.000 02. 0
15. Jessica K. Ivani et.al (2021), "Indo-Aryan – a house divided? Evidence for the east–west Indo-Aryan divide and its significance for the study of northern South Asia", *Journal of South Asian Languages and Linguistics*, ISSN 2196-078X, Volume.7, Issue. 2,<https://doi.org/10.1515/jsall-2021-2029>