Jessica K. Ivani* and John Peterson

Introduction

<https://doi.org/10.1515/jsall-2021-2036>

The present special issue of the *Journal of South Asian Languages and Linguistics* is dedicated to the topic of language contact in South Asia. These following five articles were originally presented as talks at the special theme session “South Asia as a ‘Sprachbund’? Advances in the study of language contact in South Asia” at the 35th South Asian Languages Analysis Roundtable (SALA 35), at INALCO-Paris, on October 29–31, 2019, organized by Ghanshyam Sharma.

Language contact/convergence in South Asia now looks back at a long history. For example, as early as Caldwell’s comparative grammar of Dravidian languages (Caldwell 1875) a number of features were already identified as being found in different families of the subcontinent, specifically Indo-Aryan, Dravidian and, to a certain extent, Tibeto-Burman.

Bloch (1934) was, to our knowledge, the next author to comment on language convergence in South Asia. Although Bloch (1934)’s main focus is the presence and the description of non-Indo-European linguistic features in Indo-European, he mentions several phonological, lexical and morphosyntactic features which are shared across language families by a large number of languages in South Asia or by restricted sets of unrelated languages in the region. These features, which provided the basis for later South Asian areal studies, include, among others, the extensive presence of lexical items of Indo-Aryan origin borrowed into all Dravidian languages, the phonemic status of retroflexes in Indo-Aryan, Dravidian and Munda, the general absence of dual number (though attested in Sanskrit) and a strong tendency towards suffixes as marking strategies over prefixes and infixes.

Emeneau (1956) sets a turning point in the South Asian linguistic literature by postulating the existence of an Indian linguistic area. Emeneau calls attention to the spread of a number of features throughout the subcontinent, in addition to the ones postulated by Bloch, such as the *pan-Indic* (sic!) presence of echo-word constructions (i.e., melodic overwriting), general converbs (generally referred to as “conjunctive participles” in South Asian linguistics) and the use and functions of

*Corresponding author: Jessica K. Ivani, University of Zürich, Zürich, Switzerland,

E-mail: jessica.ivani@uzh.ch

John Peterson, University of Kiel, Kiel, Germany, E-mail: jpeterson@isfas.uni-kiel.de

classifiers in South Asia. Discussing linguistic data from Indo-Aryan, Dravidian and Munda languages, Emeneau concludes by noting that the end result of this heavy borrowing is that the Indo-Aryan languages of South Asia seem more akin to Dravidian than to other Indo-European languages (Emeneau 1956: 16).

Since Emeneau's work, an extensive literature on this topic has appeared, with different suggestions as to the choice of features to be compared throughout the subcontinent and how to define the respective variables. Further contributions include Andronov (1964), among others, which mentions additional features that are shared by both Indo-Aryan and Dravidian families, and also suggests the direction of the spread of these features across the two families; and Emeneau again, first by expanding the range of "pan-Indian" features to onomatopoeia (Emeneau 1969) and then exploring semantic domains and lexical items shared by Indo-Aryan and Dravidian (Emeneau 1974).

Many of these features are summarized in Masica (1976), a book-length study of South Asia as a linguistic area that examines primarily word order, causative verb morphology, general converbs, the so-called explicator compound verbs (forms deriving from lexical verbs which mark *Aktionsart* and often the passive), and the "dative construction", in which the stimulus appears in the nominative and the experiencer appears in another case, generally corresponding to the (accusative and) dative. Masica (1976) represents a landmark in the study of South Asia as a "linguistic area": instead of focussing only on the presence of shared linguistic features and on the processes that determined their convergence in South Asia, Masica's research questions aim to expand the scope of study and define the extent to which each of these features can be considered exclusively South Asian, and whether the South Asian linguistic area is part of a larger "Asian Sprachbund".

One of Masica's (1976: 185) suggestions for future work was to search for further linguistic criteria to begin to define subregions in South Asia, and this has become the main emphasis of research in contact linguistics in the subcontinent since then.

The present issue adds five new and highly innovative contributions to the growing amount of literature on language contact in South Asia. Contact dynamics are explored here from different perspectives and methods: from a bird's eye view of the South Asian subcontinent, such as in Borin et al., to northern-central India (Ivani et al.), to smaller regions characterised by high linguistic diversity (Paudyal and Peterson), more generally between different languages families of South Asia (Hock), and within and between South Asian neighbouring regions (Hindu-Kush-Karakorum area, Liljegren).

Hock presents three case studies that suggest that we rethink convergence phenomena from various angles. The first case study concerns retroflexion in Kuki-Chin languages. While retroflexion in Kuki-Chin can be claimed to be a

result of influence from Indo-Aryan, the author illustrates, based on geographical evidence, that retroflexes have developed in many other Tibeto-Burman languages independently, and are thus not necessarily of Indo-Aryan origin in Kuki-Chin. The second study is a continuation on the author's previous work which suggests an apparent geographical alignment in the development of geminate Dravidian alveolar stops and Indo-Aryan clusters of r + dental t based on a former contact context. According to the updated view of the author, his previous contact explanation is chronologically problematic, and thus he revises this alignment as most likely due to chance. In the third case study, Hock focusses on Krishnamurti et al.'s (2001) claim of Dravidian influence regarding the major changes in syllable structure that affected Indo-Aryan in its transition from Old to Middle Indo-Aryan and, again, from Middle Indo-Aryan to Modern Indo-Aryan. By invoking structural, historical and sociolinguistic evidence, the author suggests strengthening putative convergence phenomena with fine-grained typological and historical evidence.

In their contribution to this issue, Borin et al. bring computational tools into areal and genealogical debates in South Asian linguistics. Using computational methods, they focus on the subclassifications of Indo-Aryan, by processing and plotting the distance metrics of 63 features from 200 languages, representing five language families and one isolate, extrapolated from Grierson's monumental *Linguistic Survey of India*, in digitised format. The dataset has undergone three different processes of computer processing: 1) phylogenetic algorithms with network visualisations (using SplitsTree), 2) multiple component analysis (MCA) for data exploration and for the emergence of characteristic principal components in relation to language groupings, both genealogically and geographically, and 3) visualisation tools, such as maps and graphical representations of the distributions of features (or sets of features) in South Asia. The combined analysis of these three methods provides further support to previous research and intriguing results on the linguistic diversity of South Asia; the intertwining of areality and genealogy also further confirms the presence of an East-West divide in Indo-Aryan languages, which is also discussed in Ivani et al. (this issue).

Liljegen explores areality and contact in the Hindu-Kush-Karakorum region, a typologically underdescribed area that includes languages from at least six different families. The author explores first-hand collected data from 59 different languages in an overall structural analysis, which includes domains such as phonology, lexicon, semantics and morphosyntax, defined by 80 binary structural features. In his results, performed through various distance-based visualisation methods such as COG and SplitsTree, the author illustrates how on the one hand the cross-linguistic comparison of lexical and phonological features lines up with the established genealogical classification of the region and within the Hindu-Kush-Karakorum

area, while on the other hand the morphosyntactic domains often cut across phylogenetic boundaries and share macro-areal trends of neighboring geographic regions, such as South Asia and Central-West Asia. The author touches upon the presence of layers of areality, recognizing six micro-areas which align with the ethno-history of the region and that reflect a high linguistic diversity.

Ivani et al. investigate the claim of the presence of an east–west divide in the central northern Indo-Aryan languages attested in previous literature such as Peterson (2017) with the further hypothesis that this divide may be linked to the influence of the Munda languages in the East. Working with 217 fine-grained variables on a sample of 27 Indo-Aryan and Munda languages, the authors test the presence of a geographical divide within Indo-Aryan using computational methods such as cluster analysis in combination with visual statistical inference. Having confirmed the presence of a geographical divide for the whole dataset and most of the individual features, they proceed to compute the degree of similarity between the Indo-Aryan languages and Munda, using a Bayesian t-test. The results for most features support the claim that the languages identified in the eastern cluster are more similar to Munda, opening up further research scenarios for the history of the region.

In the last contribution to the issue, Paudyal and Peterson explore language contact scenarios in the state of Jharkhand in eastern central India. The authors specifically investigate four Indo-Aryan languages spoken in the region: Sadri/Nagpuri, Khortha, Kurmali and Panchparganiya, which are collectively referred to as “Sadani” by their speakers and which are traditionally considered by the international linguistic community to be dialects of other predominant languages of the region, such as Bhojpuri, Magahi and Maithili. Using the software COG from the Summer School of Linguistics to explore lexical and phonetic similarity, and discussing in detail other specific features such as lexical items, ergativity, polypersonal marking, alienable/inalienable possession and the “narrative” verbal marker with respect to each of the four varieties, the authors show that these four languages form a compact genealogical group within the Magadhan group of Indo-Aryan, which they refer to as Sadani, and show how this preliminary comparison might offer further insights into earlier stages of Sadani. The authors conclude that the traditional classification of these languages as dialects of other languages appears to be due to morphosyntactic differences which have arisen through the respective contact situations in which these four languages are found.

Taken together, these five contributions add considerably to our current state of knowledge while at the same time illustrating the use of innovative methods to take the field of contact linguistics in South Asia in new directions. This in turn will provide us with a clearer picture of past contact scenarios in South Asia as well as a better understanding of the exact mechanisms involved in language contact in general.

References

- Andronov, Mikhail. 1964. On the typological similarity of new Indo-Aryan and Dravidian. *Indian Linguistics* 25. 119–126.
- Bloch, Jules. 1934. *L'indo-aryen du Veda aux temps modernes*. Paris: Adrien-Maisonneuve.
- Caldwell, Robert. 1875. *A comparative grammar of the Dravidian or South-Indian family of languages*. London: Trübner.
- Emeneau, Murray B. 1956. India as a linguistic area. *Language* 32(1). 3–16.
- Emeneau, Murray B. 1969. Onomatopoeitics in the Indian linguistic area. *Language* 45(2). 274–299.
- Emeneau, Murray B. 1974. The Indian linguistic area revisited. *International Journal of Dravidian Linguistics* 3(1). 92–134.
- Krishnamurti, Bhadriraju. 2001. *Comparative Dravidian linguistics: Current perspectives*. Oxford: Oxford University Press.
- Masica, Colin. 1976. *South Asia as a linguistic area*. Chicago: University of Chicago Press.
- Peterson, John. 2017. Fitting the pieces together – Towards a linguistic prehistory of Eastern-Central South Asia (and beyond). *Journal of South Asian Languages and Linguistics* 4(2). 211–257.